



► Five-Level Redundancy

By Richard J. Cichelli, SCS President

The guiding philosophy underlying SCS's storage management configurations is this: Lots of Copies Keep Stuff Safe (LOCKSS).

The bulk of storage managed by SCS/Track™ holds ad documents. These are what are used to generate the images of the ads. They are usually created with desktop tools like InDesign®, QuarkXPress®, Multi-Ad Creator® and PhotoShop®. Most mid-market newspapers require 100 to 200 gigabytes of storage to hold ad documents during production and publishing.

The data space required to maintain the information *about* ads is often less than one-hundredth the size of the ad documents themselves. SCSTrack maintains this descriptive information in its database and uses pointers within the database to identify where ad documents are stored.

By not storing actual ads in data processing oriented databases, neither database corruption nor slow recovery times impact ad production. You can easily get to your ads in SCS/Track using the native file system. Ad documents in SCS systems are free from internal pointers. Graphic images used in an ad are stored in the ad's folder. The complexity and unreliability of linking is avoided by simply using cheap disk storage effectively.

SCS/Track distinguishes between active information and archived information. The active information is dynamic and includes the SCS/Track database and what we call "current ads." Current ads are the ad documents that are being worked on or are still running in the paper or that haven't reached their post-publication expiration date.

Archived ads are those ads which have reached their expiration date and have gone through the archival process. They are unchanging static data. As with current ads, archived ads have their location known by the SCS/Track database.

Software Consulting Services, LLC

630 Selvaggio Drive, Suite 420

Nazareth, PA 18064

Sales: 1-800-568-8006

Fax: 610-746-7900

E-mail: sales@newspapersystems.com

www.newspapersystems.com

Archived data is purged from online storage on a least recently used basis. Should you wish to keep ads online no matter how old they are, additional servers can be purchased to add to the storage network. The cost is low and the integration trivial.

You can always find the ad you are looking for in an SCS/Track archive. SCS/Track transparently integrates both relational database lookups on ad metadata (advertiser, publication date, etc.) and full-text retrieval on ad content (furniture, dining, etc.). The storage hierarchy is also transparent to SCS/Track users. Current, expired and archived ads are all retrieved in a unified and consistent manner.

The daily maintenance of the storage asset management system involves a quick, unattended incremental disk backup (which is likely to take less than 10 minutes) and an unattended writing of a DVD or Blue-Ray disk. All the rest of the storage management is totally automated.

Our recommended server configuration is a pair of Dell departmental servers with two to four terabytes each. We custom configure the Dell servers to support any size ad production department from the smallest to one serving several hundred ad builders.

The servers are paired in a simple "shared-nothing" architecture. Neither server depends on the other or any separate, costly, complex storage subsystem.

With the pair of servers, one is active and the other is on active standby. That is, the second server can be used to do useful work such as backing up databases or running analytic reports, but its job is to be a data replication repository for the changing information of the primary server. It does this with a real-time feed over a gigabit Ethernet connection between the servers. The data replication scheme avoids the liabilities of low-level mirroring solutions because failures on the primary server do not corrupt the secondary server database. Each replication transaction is revalidated before it is applied to the secondary server.

Last update: February 22, 2010

SCS BUILDS TRUSTED NEWSPAPER SYSTEMS.



While nice, it is not essential that the servers communicate by fast Ethernet. They operate asynchronously and in some installations are at different sites and communicate via the Web.

5-Level Redundancy works like this:

Level 0 is the storage that you would get if you had no redundancy in the system whatsoever. Unfortunately, a significant fraction of newspapers run their ad production operations at level 0. Many find themselves losing, misplacing or deleting critical ad data.

Level 1 redundancy is easily provided by using a RAID controller card to assure data availability even in the context of a single drive failure.

Level 2 redundancy is provided by having the second standby server. This server should have the same or greater capacity than the primary server. (The greater capacity could be used to replicate other applications.)

Level 3 redundancy is provided by the RAID architecture in the second server. With these levels of redundancy, both servers would need to fail simultaneously to take the application offline. Even the failure of the SCS/Track application itself wouldn't prevent you from accessing and building ads since there is easy and transparent access to the ad folders themselves within the storage hierarchy of the SCS/Track file management architecture. The servers also feature redundant fans and redundant power supplies to ensure further long-term, robust performance.

Removable large capacity hard disks are used for backing up dynamic data such as current ads, application databases and the systems' images. Incremental backups can be done daily using rsync. If two copies are made, then you can keep one onsite and one offsite. This provides a 4th and 5th level of redundancy for dynamic data.

The 4th and 5th levels for archive storage are the optical media, one for keeping onsite and the second copy for offsite storage. Since, by definition, the archive is for static data, a backup for it is built in small increments day by day.

Here, in more detail, is a typical backup/redundancy workflow for an SCS/Track system.

- Data replication is performed in real-time throughout the day. As transactions are updated on the primary server, transactions are queued to be updated on the backup server. The ads are copied over as well.
- If an ad is deleted from the primary by mistake, it can still be recovered. Deleted ads and any deleted component files can be retrieved off the backup server up until the nightly server mirroring.
- A synchronization occurs nightly for critical data and ad archives. Only changed data is copied then.
- A backup to removable media is also performed. SCS/Track copies all critical data including the application and its databases, current ads and staged archives waiting to be written to DVD (or other write-once media like Blue-Ray.) The production environment can be reestablished from the removable media.
- DVD/Blue-Ray incremental backups for ads being archived occur nightly. Such archives are self-defining and independent of each other. SCS/Track also makes a compressed archive containing the same ad data for fast, online access.
- Every three to six months, an entire system recovery image is created.
- Weekly, an entire system backup is saved to the standby server, making it easy to recover either of the servers in the pair from its partner.

The importance of offsite storage was brought home to us very clearly when one of our customer's office facilities burned to the ground. A backup copy and an emergency delivery of a server saved this newspaper from missing publication. By building extremely robust systems, SCS earns the loyalty, trust and admiration of its customers.