Healthcare Informatics Online

Part 1: When a PACS Archive Reaches the End

Switching vendors is more problematic than it sounds.

by Stephen S. Boochever and Jeffrey D. Older

The technology for managing medical images has advanced dramatically in recent years, producing picture archiving and communications systems (PACS) that are faster, better and more sophisticated than ever before. Many healthcare organizations, especially large academic medical centers, are in the process of acquiring their second or higher-generation PACS. This presents the challenges of migrating image data to the new system.

Data volumes

A PACS archive's life cycle is typically four to eight years. Organizations that produce 500,000 radiology exams annually may well acquire more than 100 terabytes of data in that time. Obviously, such volumes can't be moved overnight. Plus, a cleaning process should be carried out during migration, since archives tend to become pools of image sets--some with incorrect, incomplete or missing identifiers and some archived multiple times.

The average size per exam across all modalities, estimated at 30 megabytes, is expected to more than double over the next few years with proliferation of new cross-sectional scanners and related exam protocols, digital mammography and direct radiography.

Archive technologies

The base technologies found in a PACS archive are magnetic disks, optical disks and magnetic tape, which come in myriads of derivatives. Differentiators are form factor, price, speed, capacity and life cycles. Price per data unit of storage, speed, and capacity are highest for magnetic disks, followed by optical and tape media.

Magnetic disks are grouped in arrays, such as a storage area network (SAN), in which a disk array is attached via fibre channel, or network attached storage (NAS), in which a disk array is attached to the local area network by means of a server. Optical and tape media are organized in cabinets; data rates are increased by using multiple drives.

Price ratios for identical storage capacity between fast magnetic disk subsystems and tape subsystems are about 20:1. Speed ratios are more difficult to compare. Tapes have delays of 30 to 120 seconds for the robot to put the tape into a drive and wind/rewind it to the desired image file, but once positioned, data can be streamed at speeds similar to magnetic disks. Magnetic disks reach a capacity of 300 gigabytes, optical disks 30 gigabytes, and tapes 200 gigabytes.

Archive organization

Archiving images according to retrieval immediacy is most practical. When reading new exams, radiologists need immediate access to earlier images. In only 10 percent of cases do they look at anything older than two years, but most states require that images of adults be retained for seven years and those of adolescents for 21 years.

A typical image archive is built in two storage blocks: a short-term, fast archive using magnetic disks and long-term archive using slower/more affordable magnetic disk, optical or tape media. Archives installed a few years ago have a comparatively small SAN/NAS component (some with only a two-week capacity for active patients) and a large tape/optical disk storage block. The trend for modern archives is to depend largely on SAN/NAS, with optical/tape media limited to backup. Depending on disaster recovery requirements, SAN/NAS might be used in backup archives as well.

Compression algorithms allow files to be condensed into a smaller format compared to the raw data acquired. Bit-preserving (or lossless) compression is the method of choice and averages 2:1 compression ratios across all radiology image classes.

The DICOM myth

"We store DICOM on the archive" is a common claim of PACS vendors, giving uninitiated buyers a false sense of security. The only standardized way to exchange DICOM images with any PACS archive is through the PACS server and the network. The DICOM standard has no guidelines for storing images on the archive subsystem.

There is one exception for interchangeable storage media (with capacities up to a few gigabytes). DICOM Interchange Media, parts 10 to 12, defines a DICOM directory for removable media and is helpful for modalities (e.g., computed tomography, magnetic resonance imaging) that allow digital archiving of images. For facilities that add PACS later, these media could be used to populate the archive if the new PACS is equipped with drives/drivers capable of reading them.

Given the speed of innovation, this approach is often impractical. Modalities have a seven-year life cycle and IT devices a one- to two-year generation cycle. By the time a PACS is acquired or a switch is made to a new PACS, the drives/drivers needed to read earlier data are likely to be off the market. Vendors generally cannot offer backwards compatibility; they have to live by rules of the vastly larger IT market.

Thus, adopting DICOM Interchange Media for large-scale archiving and data migration has limited merits. Even if the DICOM committee attempted to adjust parts 10 to 12 to key products only, the task would be impossible due to the plethora of storage media on the market. No attempt has been made by either the DICOM Standards Committee or the IHE to standardize modern archives that depend on noninterchangeable media.

The only guaranteed way to exchange data via DICOM is on the server level, through the application layer using DICOM protocol. Faster migration or even portation is possible if all three storage layers (data, media, and physical media format) are identical in the two PACS. This applies to some removable media, not an entire archive.

The logical layers that are standardized by DICOM are the protocol (application) and the file format, which is optional. The application layer, including structure of the database, is proprietary to the vendor, and except for some removable media, vendors store even DICOM-formatted files in different ways.

Metadata, headers and pixel data can be stored in separate files and in different directories. Some vendors switch directories for each image; others store contiguously. Some vendors are on their third or higher generation of image-formatting strategy, so sometimes migration is required even when upgrading with the same vendor.

Part 2: When a PACS Archive Reaches the End

Switching vendors is more problematic than it sounds.

by Stephen S. Boochever and Jeffrey D. Older

Part 1 of this article, describing organization of and retrieval from PACS archives, appeared in the Jan. 19 issue of *Healthcare Informatics Extra.*

For the initial period of data migration to a new PACS system, the existing PACS should support both daily production (foreground load) without degradation and migration of data to the new system (background load). Most PACS are designed to process, when initially installed, 130 percent to 150 percent of the daily load. But due to growth of data volume, most of that safety margin will be gone toward the end of the life cycle, when migration is about to begin.

Thus, migration speeds are comparatively low, often averaging less than the equivalent of one archived day/day (e.g., migration of four years' worth of images on the archive would require at least four years). Limiting factors vary. One may be the bandwidth of the server or storage subsystem, especially when based on optical disks or slower tapes.

To speed up the process, it is important to forward new images to both PACS. Only prior exams are migrated. The new system can be put in production as soon as it has access to 12 (or better, 24) months' worth of image data. In most cases, that means switching from the old to the new system in 6 to 18 months.

Cutting over might occur sooner, if the old system is fast enough to transfer prior exams that radiologists need to the new system (usually twice the amount of new exams/day, prefetched during the night). Freed up from the foreground load, the old system will be able to support migration at higher speeds.

There are two caveats: The modalities should transmit only images that have been through the quality-assurance process (site specific), and changes the radiologist made to images (e.g., window/level) while reading on the old system will not propagate to the new system. It's the radiologist's call if that's acceptable.

Archived images are retrieved via DICOM query, issued by the new PACS, the RIS (if possible) or a specialized product designed for migration purposes. Priority is given to the most recent exams ("last in, first out"). Ad hoc queries by the user on the new system should be supported.

How long will migration take?

Take, for example, a system with 48 months worth of images on the archive supporting a migration speed of one-half day/day with foreground load, and two days/day without. Within 16 months, the new archive will have accumulated 24 months worth of recent exams--eight months of prior images and 16 months of new images. To migrate the remaining 40 months will take another 20 months, for a migration process of 36 months.

As mentioned, a cleaning process should be run, quasi-simultaneously with data transfer. Thus, users should plan for keeping the old PACS around for some time, including service coverage and upgrades if the increasing volume exceeds inherent limitations (e.g., database size) or support for parts expires.

Typically, migration between storage subsystems of the same vendor is faster, with little impact on the foreground process, because vendors usually provide migration tools allowing data transfer on the storage level.

Planning for migration

In considering a new PACS, the following issues should be addressed:

- **Sustained migration speeds of old and new systems, with and without foreground load, with minimal degradation of production**
- **Speed of transferring prior exams ad hoc or by prefetching to the new system**
- **Integrity of archived data and effort for the cleaning process**

This information will enable estimation of the time needed to complete migration and switch production. Based on the estimate, the old system vendor should be asked to quote equipment service and upgrades to cover the archive and its RIS interface for the migration period and the other devices until switch-over. Cost and effort for migration when the new system reaches the end of its cycle should be negotiated up front with the new vendor.

Is there a better way?

The current status of DICOM is clearly unsatisfactory for migrating images from one PACS to another. Copying the content of a large storage subsystem using the established DICOM protocol on the server level can take years, and substantial efforts are required to keep the archiving part of the old system up during the migration process. Even if both PACS had identical storage subsystems, the different formatting would prohibit simply connecting storage of the old PACS to the new server.

Considering the trend to store all images on high-capacity SAN/NAS, the DICOM Committee would be well advised to define a methodology for storing images interchangeably on such subsystems. As long as the new PACS is capable of connecting a storage subsystem on the operating system level, there should be a way, at least by means of an import/export process, to read the data.

Users then would be able to discontinue the old PACS within weeks of installing the new one. The new vendor would have to commit to migrating the (now readable) data as soon as the hardware for the old storage subsystem is nearing obsolescence.

Stephen S. Boochever is chairman and Jeffrey D. Older is chief technology officer, Innovative Health Strategies L.L.C. info@healthstrategies.com, Albany, N.Y.