EMC Isilon: Making Healthcare Data an Asset and Not a Burden

Introduction

Think about your last visit to the doctor. Maybe you had some blood work done, a CT scan or a mammogram. Not so long ago, those records and the resulting data would have been logged in a paper chart or captured on X-ray film and stored in a medical records library. But physical records can become lost, damaged or illegible, and retrieving and sharing them isn't easy. Plus, physical records require space to hold them.

Today, physicians have gone digital, with paper charts giving way to electronic medical records and film supplanted by digital imaging. The resulting modernization promises benefits for both healthcare providers and patients alike, including improved patient care, reduced duplication of testing and remote consultation.

But just because the data have been converted into 1s and 0s, storage issues haven't disappeared. Even electronic records must be stored somewhere. Increasingly, clinicians in small medical practices, hospitals or regional networks are implementing sophisticated systems to make that happen.

The volumes at play can be enormous. Miami Children's Hospital has more than a petabyte of total storage in its data center—that's a million gigabytes. And the Northern and Eastern Ontario Diagnostic Imaging Network (NEODIN), which archives medical images for a region of Canada that's roughly the size of France, has some 9 million medical images in its repository, growing at 2.5 million per year—some 100 to 150 terabytes of growth annually.

Typical data storage systems can't keep up with such voracious consumption. In most server-based storage designs, a controller node directs several slaves, each of which has some maximum amount of allowed storage.

But at some point, the system hits a silicon ceiling and all the data must be migrated to a larger system—an expensive and time-consuming proposition.

Physical records can become lost, damaged or illegible, and retrieving and sharing them isn't easy.

Enter EMC Isilon. Isilon is a "scale-out network-attached storage" (NAS) system capable of holding 20 petabytes or more of information and delivering it to users at high speed. But users needn't know how large a system they need at the outset—Isilon can grow with their business. As existing drives fill up, simply plug new storage into the server rack and go.

Indeed, EMC Isilon has become a key piece of IT infrastructure for users across the healthcare spectrum, helping to convert data from archive to asset. To find out how, we spoke with clients representing five diverse environments: a children's hospital, a regional health information exchange, an academic health system, an independent medical oncology/hematology practice and a university hospital

Here are their stories. ■

IT Consolidation at Miami Children's Hospital

The Miami Children's Hospital health system extends from Palmetto Bay, south of Miami to West Palm Beach, nearly 100 miles to the north. Their system includes nine facilities anchored by a 190-bed inpatient hospital on the main campus.

Until several years ago, Miami Children's had no unified system to handle clinical data, says David Bratt, the hospital's director of technology services. The far-flung clinicians and departments were saving their data—including patient records, private files and medical images—on an eclectic mix of IT infrastructure and file systems. Keeping pace wasn't easy. "Managing several different types of storage structures on several different types of arrays just took a lot of time and [was] very cumbersome," Bratt says. Performance was never quite optimal either.

So, about three years ago, the hospital installed an Isilon storage device, with its unified OneFS file system, and initiated the work of consolidating everything under its umbrella.

How much data? In total, the hospital system exceeds a petabyte in total storage, of which the Isilon system accounts for "hundreds of terabytes," including 30 to 40 TB added just within the past few months. According to Bratt, the Isilon is used mostly for unstructured data storage (file shares and home directories), medical imaging and analytics, and the network includes more than a hundred data-generating systems. But even security video data are stored on Isilon nodes. "Those videos don't change, they just need a place to live for an extended amount of time."

"Managing several different types of storage structures on several different types of arrays just took a lot of time and [was] very cumbersome,"

DAVID BRATT, DIRECTOR OF TECHNOLOGY SERVICES, MIAMI CHILDREN'S HOSPITAL

Overall, the system is doubling in size annually. But given its location in hurricane-prone south Florida, the Isilon system is actually doing so twice, as the storage array is based in two physical locations, data constantly syncing between them. "In the event the primary unit does completely fail—or connectivity to it fails—we have a secondary copy that's fully accessible with every single file that's in the primary," he says.

Bratt, who was involved in the Isilon acquisition, says Miami Children's Hospital had several key requirements in selecting a new storage solution: Performance and reliability, of course, but also scalability, cost of ownership, and ease of management. Instead of having to be expert in multiple backup systems, for instance, IT managers using Isilon need only navigate a single web site to control and monitor the system. "The single pane of glass into this massive storage array helps my guys manage the array efficiently," he says, adding that it takes only "minimal staff" to manage the Isilon infrastructure.

Today, Bratt says he anticipates growth to accelerate due to both analytics services and genome sequencing, including a planned Hadoop database for big-data processing. That dovetails nicely with Isilon's easy scalability. But it's not just size that counts, Bratt notes. With 10-gigabit data transfer capabilities, for instance, the system can accommodate multiple researchers simultaneously; key for handling—and manipulating—today's massive genomics datasets. Slow data transfer specs can cause some reporting and analytics to drag on for days if not weeks, Bratt says. Today, they can take hours. "It's like having a six-lane highway versus just a two-lane street," he says. ■

The NEODIN Medical Image Repository

Suppose a child goes to the hospital in Ontario with a possible broken arm. His digital X-ray is analyzed by radiologists on site at the ER. But there's no gray-scale film to be stored in a dusty medical archive. Instead, the data are sent to the Northern and Eastern Ontario Diagnostic Imaging Network (NEODIN), a regional health information exchange supporting 64 hospitals in a region of Canada roughly the size of France.

According to IT director Pat Harkins, NEODIN harvests all medical imaging (DICOM) data collected at any of these facilities, whether it be an X-ray, PET scan, echocardiogram or mammogram. Some 9 million medical images have been stored to date, and the archive is growing at 2.3 million per year.

To store that data, NEODIN has just under a petabyte of storage capacity in total split across two synced Isilon systems located in Sudbury and Thunder Bay, 1200 km apart. "It's an active-active solution," he explains. "Some facilities archive to one data center and the rest archive to the other, and they sync in the background."

NEODIN's Isilon system was purchased two years ago. "We've seen about a 33 percent increase in performance and the uptime on the system has exceeded our specs." Harkins adds that NEODIN has experienced no unscheduled downtime since the Isilon system went live.

That's good news for the clinicians who use it, of course. These practitioners have the ability to interact with the DI-r (diagnostic imaging repository) viewer directly, Harkins says, or through their local PACS (picture archiving and communication system), which, as far as the software is concerned, is just another network-attached drive.

"We've seen about a 33 percent increase in performance and the uptime on the system has exceeded our specs."

PAT HARKINS, DIRECTOR OF TECHNOLOGY SERVICES, NEODIN

The software, he adds, is intelligent enough to match patient records across tests and facilities. So, if that child with the possible arm fracture also had had additional X-rays in the past, the radiologist will see them all when reviewing the chart. The physician, he explains, "would see all my history of what is known within the DI-r, with NEODIN."

Harkins says NEODIN data is growing at 100 to 150 TB a year, meaning the organization should be able to go nearly 20 years without having to migrate its data to a new platform. That translates directly into cost-savings, he adds, that's "probably the equivalent of two FTEs every five years working six months on a migration."

Other savings include falling storage costs and reduced staffing. Of eight FTEs at NEODIN, just 1.5 FTE are dedicated to the Isilon, down from three. "It's actually a simpler system from our perspective," Harkins explains, and the decreased downtime and simpler remote management means his team can accomplish more with less. Adding new storage, for instance, is as easy as plug-and-play. "It really is that simple," he says. ■

Adding new storage, for instance, is as easy as plug-and-play. "It really is that simple,"

PAT HARKINS, DIRECTOR OF TECHNOLOGY SERVICES, NEODIN

Supporting a VNA at the Ochsner Health System

The Ochsner Health System in southern Louisiana comprises some 10 hospitals, 40 health centers, and 16,000 employees. Medical imaging is done at many of those locations—everything from PET scans and X-rays to mammograms and MRIs. The problem is, until relatively recently, images collected at one location couldn't be viewed in another, because the different centers saved their images in different PACS, which didn't communicate with one another.

Plus, those databases are enormous, and getting larger. "As imaging advances, the two things technically that happen to new studies is they have more images, and each image has more resolution," explains Willy Schley, director of technology at Ochsner. "So we have a rapidly growing need for storage."

About seven years ago Ochsner integrated those PACS into a single vendor-neutral archive (VNA) from Agfa. The VNA—providing access to X-rays, PET, cardio "and other kinds of 'ologies' that generate image data" was stored on a scalable EMC storage system called Centera, Schley says. But Centera wasn't ideal, because linking it to the Agfa VNA required an additional piece of software, "and that added complexity and cost."

So, two years ago, the company migrated again—to EMC Isilon. "Isilon is perfect for what we're doing," Schley says, "because it's a scale-out platform, so we can continue to grow it without having to replace any core component … And we don't need the middleware."

"As imaging advances, the two things technically that happen to new studies is they have more images, and each image has more resolution,"

WILLY SCHLEY, DIRECTOR OF TECHNOLOGY SERVICES, OCHSNER HEALTH SYSTEM

According to Radiology Informatics Director Brian Deshotel, Ochsner's VNA currently has a capacity of 519 TB. With consumption of about 5 TB/month and a million images per year, he anticipates that storage should last another eight and a half years—even factoring in anticipated increased usage to 7 TB/month due to the hospital network's continuing "expansion and acquisition of other facilities and services."

Ochsner also uses Isilon to provide storage for each of the company's employees—a networked drive they can

access from any authorized computer. A full duplicate is maintained in Tennessee to prevent data-loss in the event of another major storm like Hurricane Katrina.

Isilon is "a system that scales without adding complexity."

WILLY SCHLEY, DIRECTOR OF TECHNOLOGY SERVICES, OCHSNER HEALTH SYSTEM

Isilon provides the hospital network with significant cost savings, Schley explains, especially by obviating the need to periodically migrate data to larger systems. "Normally you try to design a system to support three years' growth, and then in three years you'd replace it." But migrating a VNA is more complex than a large-scale file copy. VNA technical reps are required to ensure the database isn't corrupted.

"It would be a six-figure engagement for us at this point," Schley says. "If we got much bigger, it could get to be a seven-figure engagement." Instead, as capacity fills, Ochsner's IT managers can just plug in a new node and carry on.

Simplicity brings additional savings, as well. Schley says maybe one-quarter to one-half of an FTE is dedicated to the Isilon system and IT staffing levels have been flat for five years. In two years, he notes, the system has never gone down (though individual components have failed). Budgetary needs, therefore, are simpler to predict. Unlike some computer systems that become more complicated the larger they get, Isilon is "a system that scales without adding complexity."

But perhaps most importantly, patients benefit, Deshotel says. Because clinicians can see every imaging study performed on a patient from a single unified system (as opposed to having to log into and check multiple databases), unnecessary duplicate studies are minimized, efficiency increases, and overall patient care improves. "You're able to treat a patient faster because you have access to the total record," he says. ■

Using Syncplicity at the Florida Cancer Specialists & Research Institute

The Florida Cancer Specialists & Research Institute bills itself as "the largest independent medical oncology/hematology practice in the United States," with more than 170 physicians and 80 locations serving the state from Naples to Tallahassee. But it isn't resting on its laurels: The network adds 10 to 15 new practices and 20 to 40 new physicians annually, says Corey Reynolds, manager of IT Infrastructure at FCS. The clinicians are attracted by the efficiencies inherent in upsizing, such as better drug pricing, accounting, human resources and more.

Naturally, managing such a disparate and dynamic network presents problems, not the least of which is scalability. In 2013, Reynolds says, FCS's "data footprint" was growing 800 percent per year as new physicians entered the network, each with their own set of legacy data. To manage those needs, FCS used an EMC VNX array, including tiered storage, virtual machines and a "sprawl of Windows file servers" running different versions of the operating system. But managing that, Reynolds concedes, was becoming quite a headache—especially when it came to the Windows fileservers, of which the company had about 15.

One of those Windows servers included a PACS system for storing medical DICOM images, which had been outsourced at a cost of $200,000 per year and that the company had recently brought in-house to save money. But the rate of growth of that system—more than 20 TB/year—strained the VNX hardware, so Reynolds and his team turned to Isilon to ease the system's workload. "It fit what we needed," he explains.

Naturally, managing such a disparate and dynamic network presents problems, not the least of which is scalability.

Today, those Windows fileservers have been offloaded to Isilon, which comprises three X400s totaling 90 TB. "There's not a single Windows file server still running in the environment," Reynolds says. So, too, has much of the virtual machine workload. To date, Reynolds says, the company has experienced no unscheduled down time.

Among the company's plans for its Isilon is EMC Syncplicity. Dropbox isn't allowed on FCS networks, Reynolds explains—its lack of encryption in the cloud and potential vulnerability is simply too risky for patient records. Syncplicity provides a Dropbox-like experience that, among other things, supports data

encryption and the ability to decide whether to keep data on local servers or in the cloud—a key benefit when dealing with private data and federal regulations. "Syncplicity allows us to maintain our data in-house but still offers the hybrid cloud approach where we do have space in the cloud if we want it."

The implementation plan, Reynolds says, is to launch Syncplicity with FCS's marketing team, then spread the word via "viral marketing" throughout the organization to increase adoption.

Between Syncplicity, VNX, and Isilon, FCS is a heavy user of EMC solutions. Reynolds says that's down to customer service. "I have a really great EMC rep who, honestly, is the primary reason I continue to stay with EMC," he says. "He's phenomenal." In fact, Reynolds says he went so far as to pen a note to EMC vice president Paul Normark: "You need to understand that this product install is without a doubt the smoothest product install I've ever done for any product in my entire career, especially a storage product," he wrote. "Isilon was literally a turnkey product; the closest thing to a turnkey product I've ever seen in a data center." ■

"Isilon was literally a turnkey product; the closest thing to a turnkey product I've ever seen in a data center."

COREY REYNOLDS, MANAGER OF IT INFRASTRUCTURE, FCS

DNA Sequencing at VCMC

Just a decade removed from the advent of "next-generation DNA sequencing," data output has reached staggering proportions. Illumina's ultra-high-throughput HiSeq 2500 produces nearly a trillion bases of DNA sequence data per week—and the clinical genetics laboratory at VU University Medical Center in Amsterdam has two of them. Used for genetic testing both within VUmc and for external customers, these instruments crank out As, Cs, Gs and Ts to the tune of 4 TB per week, says bioinformatician Daphne van Beek.

VUmc's clinical genetics lab performs three basic types of sequence-based testing: non-invasive prenatal testing (NIPT), targeted gene panel sequencing, and whole-exome sequencing, complemented with the occasional research sample. Workload varies week to week, but on average, van Beek says, the sequencers complete some 16 NIPT samples per week as well as a "couple hundred" gene panels and one thousand exomes—and climbing—per year.

The resulting datasets are enough to choke a server: three gigabytes (GB) per NIPT and 20 to 30 GB per exome. Naturally, storing them is a challenge, as there's both raw and processed data to consider. So, three years ago the lab acquired—at Illumina's suggestion and under its service contract—a 25-TB Isilon storage system. The lab recently boosted that to 50 TB to accommodate its second HiSeq.

"I was really having trouble with storing it all," van Beek says. "The drives were around 90 percent full."

To keep up, van Beek implemented a "strict" back-up strategy—and she sticks with it now even despite having more breathing room. After all, at 5 TB per week, 50 TB could fill up in just three months. "We have a policy of keeping [data] one month on the system, and that's just to make sure that all our customers and we also ourselves can run our analyses and check everything and make sure that everything's correct … And then we have to delete it."

The lab's Isilon system offers two primary advantages. The first is the ease with which it can be scaled up; the second is speed.

DAPHNE VAN BEEK, BIOINFORMATICIAN, VU UNIVERSITY MEDICAL CENTER

The lab's Isilon system offers two primary advantages, van Beek says. The first is the ease with which it can be

scaled up; the second is speed. Bioinformatics analyses take time, especially given the size of the datasets. If file transfer is slow, "you have a lot of time where the compute nodes are just waiting to get the data, and that really slows down your process." Conversely, fast data access means analyses finish faster—meaning data can be shunted off the servers more rapidly.

Another plus is the system's internal redundancy, which reduces the likelihood of data loss even in the event of a failure. "Luckily, I have never had to test this in the system," she says.

Of course, when it comes to clinical data, it's all about redundancy, and the VUmc clinical genetics lab is no exception. In addition to its Isilon system, the lab also keeps copies of patient reports on separate, larger hospital servers. Many files must be maintained for 15 years, but some have to be retained for three generations, or 115 years, van Beek says. Needless to say, in an era when even 10-year-old media frequently are unusable, such long-term storage presents significant technical hurdles.

"Are you going to save the raw data, or are you going to save the DNA? That's the big discussion in Holland," van Beek says. "Because you can also resequence it, of course, [and] maybe that's cheaper than actually storing everything." ■