Whitepaper: Healthcare Disaster Recovery

Why Disaster Recovery is Different in Healthcare
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executive Summary</td>
<td>3</td>
</tr>
<tr>
<td>Introduction</td>
<td>4</td>
</tr>
<tr>
<td>Disaster Recovery is Still Job One for Healthcare IT</td>
<td>5</td>
</tr>
<tr>
<td>PACS and EHRs: An Explosion of Digital Data</td>
<td>5</td>
</tr>
<tr>
<td>Digital imaging goes beyond radiology</td>
<td>5</td>
</tr>
<tr>
<td>EHRs result in significant data growth</td>
<td>5</td>
</tr>
<tr>
<td>International focus on digitizing records</td>
<td>6</td>
</tr>
<tr>
<td>Back to Basics: Why Backup and Disaster Recovery Matter More Than Ever</td>
<td>6</td>
</tr>
<tr>
<td>The Healthcare Data Profile</td>
<td>7</td>
</tr>
<tr>
<td>Clinical systems</td>
<td>7</td>
</tr>
<tr>
<td>Administrative systems</td>
<td>7</td>
</tr>
<tr>
<td>The Data: Where, What and How Much?</td>
<td>8</td>
</tr>
<tr>
<td>Structured data</td>
<td>8</td>
</tr>
<tr>
<td>Unstructured data</td>
<td>8</td>
</tr>
<tr>
<td>Semi-structured data</td>
<td>8</td>
</tr>
<tr>
<td>Best Practices</td>
<td>8</td>
</tr>
<tr>
<td>RPO and RTO</td>
<td>8</td>
</tr>
<tr>
<td>HIS and EHR</td>
<td>9</td>
</tr>
<tr>
<td>PACS and ECM</td>
<td>9</td>
</tr>
<tr>
<td>Enter Archiving</td>
<td>10</td>
</tr>
<tr>
<td>A Holistic Approach</td>
<td>10</td>
</tr>
<tr>
<td>Conclusion</td>
<td>10</td>
</tr>
</tbody>
</table>
Executive Summary

The protection of data is important for the survival of any organization. However, in healthcare, it’s not just important – it is vital. Healthcare is quite different from other sectors and, as such, backup and disaster recovery (DR) need to be considered in a different way.

30% of the world’s storage can be found in healthcare\(^1\). This is largely due to the existing, and rapidly growing, volume of patient and administration data, much of which is considered mission critical. This upsurge comes as a result of the prevalence of digital medical image technologies (and resulting files), across multiple diagnostic departments. Additionally, many hospitals are committed to the adoption of the Electronic Health Record (EHR) and some are also moving towards “paperless” environments (including the retroactive scanning and archiving of patient information) – both of which generate vast amounts of data. Few healthcare organizations were prepared for this huge increase in digital information and, consequently, many infrastructures are under a tremendous amount of strain.

Furthermore, there is a large amount of pressure being exerted by governments, regulatory bodies and other influencing organizations (in the form of enacted laws, codes of practice and guidelines) as to how healthcare data should be protected and secured in the event of system outages, natural disasters, loss and even theft.

So, imagine the increased burden on healthcare IT professionals when attempting to create and manage a robust, ‘working’ backup and disaster recovery strategy for the protection of such large volumes of critical information – especially in the restricted time windows available. While this presents a very real challenge for healthcare IT teams, it is not insurmountable.

The key to understanding how to create a ‘Healthcare Disaster Recovery’ strategy is to understand the nature and characteristics of the data you need to protect and the storage environment. Firstly, in healthcare IT, we can typically distinguish applications into two categories:

- **Clinical systems**, including the likes of PACS, HIS, RIS, ECM and LIS; and
- **Administration systems**, such as word processing tools, spreadsheets, presentations and even accounting software.

Data generated by these applications can also be categorized by whether they are:

- **Structured**, such as dynamic data that is accessed and changed regularly [e.g. database files from MUMPS, Cache, MAGIC, Oracle or SQL];
- **Unstructured**, such as static data that is unlikely to be accessed or changed [e.g. documents, pictures, audio-visual files]; or
- **Semi-structured**, e.g. DICOM files from PACS systems that contain the medical image [that is static] plus additional meta-data [which can be dynamic].

In understanding the context and characteristics of the applications and data, it becomes easier to determine the appropriate Recovery Point Objectives (RPO) – the acceptable amount of data loss measured in time; and Recovery Time Objectives (RTO) – the duration of time in which systems must be restored to an acceptable level of service after an outage has occurred. There is no single RPO and RTO rule that can be applied across healthcare applications – these must be determined by the hospital based on the criticality of the system. For example, an HIS would be considered by most as top of the list for recovery in order to minimize disruptions to patient care.

Best practice suggests that in the case of HIS and EHR systems, which use a manageable amount of structured data, the RTOs and RPOs should be short. Backing up such data is relatively straightforward given SAN technologies and data “snapshotting” and replication techniques and tools that enable the speedy creation of disk-based replicas for quick restore.

More challenging is the case of PACS and ECM platforms that contain many TBs of digital images. While desirable, it is practically impossible [economically and practically] to create a speedy RPO or RTO. However, as the image data is largely static [medical images, in themselves, do not change, though the accompanying meta-data does and, therefore, can be managed in the same way as the HIS or EHR], it is possible to overcome the backup and DR issue utilizing archiving techniques and tools. By creating a few geographically dispersed copies of the data, on multiple media types, you can redirect your applications to the secondary data sources to gain the RPO and RTO requirements demanded by the organization.

Ultimately, if you understand your healthcare IT environment in relation to applications, data and storage, using the appropriate combination of backup and archiving processes and tools, you can navigate your way through the minefield that is Healthcare Disaster Recovery.

---

\(^1\) “Survey on the Government of Unstructured Data”, 30 June 2008, Ponemon Institute
Introduction

Hospitals worldwide are faced with growing volumes of digital data. The adoption and expanded use of Electronic Health Records (EHR) (also known as the Electronic Patient Records (EPR) is being spurred by forces such as the HITECH provision of the American Recovery and Reinvestment Act (ARRA) in the United States, Canada Health Infoway, and the European Institute for Health Records (EuroRec). And while the amount of this digital data (and the storage needed to maintain it) is growing at an unprecedented rate, it pales in comparison to the new data created every day by Picture Archiving and Communication Systems (PACS). Not only are traditional Radiology PACS creating larger digital files as imaging modalities improve, but as other clinical disciplines, such as Mammography, Cardiology, Pathology, Endoscopy and Ophthalmology adopt digital imaging technology, data stores are growing exponentially.

As healthcare providers increasingly come to rely on computerized systems in the course of caring for patients, healthcare IT is under more pressure to protect, retain, and make available 24/7 this vast amount of medical data. In the event of a system outage, their response must be swift, enacted confidently and without hesitation so that the business of patient care can continue with minimal disruption.

This paper will examine the unique challenges faced by healthcare IT professionals as they plan and prepare for the disaster recovery of systems critical to the business of healthcare. We will examine the data profile of a typical provider organization and illustrate how the type of data goes hand-in-hand with the quantity of data being recovered if one is to enact an efficient and effective Healthcare Disaster Recovery strategy. Only by understanding the nature of the data in play can healthcare IT professionals devise a truly comprehensive approach – call it a holistic approach – to Healthcare Disaster Recovery.
Disaster Recovery is Still Job One for Healthcare IT

A recent survey of healthcare professionals revealed that among all the initiatives vying for the attention and budget of healthcare IT, good old backup and disaster recovery remained the top priority by a wide margin.\(^2\) How can that be? Surely, healthcare IT solved this problem long ago? In the United States, for example, the Health Information Portability and Accessibility Act (HIPAA) of 1996 included provisions for data protection and security which required, among other things, a disaster recovery plan. And similar initiatives have been enacted in countries throughout the world to ensure healthcare data remains protected and recoverable in the event of a disaster. Why, after all these years, would healthcare IT still view backup and disaster recovery as a top priority? The answer lies in a growing problem, or rather, a problem of data growth.

PACS and EHRs: An Explosion of Digital Data

In the mid-1980s the first digital X-ray systems began to appear in hospitals. Dubbed Picture Archiving and Communications Systems, or PACS, these systems were to change the face of medical imaging forever. The old way of taking diagnostic pictures of the human body on film, developing that film, sending it to a radiologist for diagnosis, and storing that film for future reference was entering a new era. Now, powerful imaging modalities could transmit their data to PACS via the Digital Imaging and Communications in Medicine, or DICOM, standard. The power of computers allowed these images to be stored in electronic form, along with the relevant patient demographics, study notes, and reports vital to presenting a comprehensive study of a particular patient. While this technology was still mostly science fiction for the vast majority of hospitals back in the mid-’80s, fast-forward nearly three decades and science fiction has become science fact. Now, even a modest community hospital can own and maintain its own PACS.

Digital imaging goes beyond radiology

It’s no longer just Radiology that benefits from digital images. Other disciplines have rapidly adopted the technology over the years. From Cardiology and Mammography to Pathology and Ophthalmology, new digital imaging systems are continually coming online to change the way myriad medical disciplines diagnose patients. Meanwhile, existing imaging technology is not standing still. Technological advances enable more traditional modalities to create images with greater and greater fidelity, in the process consuming greater and greater amounts of disk storage. The combination of new digital imaging systems coming online, along with the increasing resolution of the images produced by current systems, is creating an exponential problem in data growth the likes of which healthcare IT has never seen before.

EHRs result in significant data growth

While PACS may be the largest consumer of data storage in today’s healthcare organization, coming in at an increasingly close second is the Electronic Health Record System, or EHR. Like PACS, EHRs in one form or another have been in use for a number of years in hospitals worldwide, but recent events have put this technology into greater focus. In 2009, the American Recovery and Reinvestment Act (ARRA) was passed in the United States to help combat what was at that time the worst economic recession on record since the Great Depression of the 1930s. As part of ARRA, the HITECH Act was passed, requiring for the first time that physicians and hospitals demonstrate “meaningful use” of EHR technology. With its goal being to improve the quality of patient care, HITECH includes both a “carrot” and a “stick”, providing financial incentives for hospitals to quickly adopt EHR technology as well as financial penalties for those who lag behind. With such a government mandate, systems that were a “nice to have” at many healthcare facilities have quickly become a “need to have” in the current regulatory environment.\(^3\)
Recognizing the prolification of medical data worldwide and its astounding rate of growth, it is no wonder that healthcare IT sees backup and disaster recovery as its top priority. As providers become ever more reliant on their systems, in the course of diagnosing and treating patients, it becomes critical that data is highly available. Even one hour of downtime for a hospital’s EHR could negatively impact patient care. Extended outages could have a more severe effect. In such cases, what are healthcare providers to do? The old recourse was to “go back to paper”, but as hospitals rapidly adopt all-digital models, increasingly there is no paper to “go back” to.

Now put this in the context of today’s reality. A typical community hospital in the United States at any given time may have dozens of TBs of data, depending upon when and how aggressively they adopted digital systems. As the healthcare IT professional responsible for disaster recovery, how do you protect 10TB of data? 50TB? 100TB? How do you back it up every night? The answer is, you don’t. It becomes an economic and, eventually, a physical impossibility to back up that much data on a daily basis. So how do you protect this asset vital to the business of healthcare? Before you can protect the data, first you must understand it.

International focus on digitizing records

The United States is not alone in its promotion of EHR technology. In Canada, the Canada Health Infoway project has dedicated CAD$2.1 billion³ to the adoption of electronic medical records while, in Europe, the independent EuroRec organization has made its mission the promotion of this technology throughout the continent. And many individual countries worldwide have funded their own initiatives to move growing amounts of patient data into digital systems in an effort to increase the speed, efficiency and quality of patient care.

As the deployment of EHRs accelerates, so do efforts to digitize all patient data, making it part of the patient’s permanent legal medical record. Documents such as driver’s licenses, insurance cards, healthcare proxies, and other paper records are being scanned and attached to the EHR. More ambitious facilities are even engaging in retroactive scanning, converting years of historical paper records into digital ones. While having all of this data in digital form creates many new opportunities for treatment and research, it also creates a new problem for healthcare IT: the exponential increase in demand for disk storage to house all of this new digital data. Indeed, recent studies show that going paperless in a hospital requires 60 gigabytes (GB) of data storage per bed per year. For a 100 bed community hospital this equates to 6 terabytes (TB) of storage per year, and this figure does not include retroactive scanning. Taken together, PACS and EHR data require TBs of disk storage on an annual basis for even modest-sized hospitals. With many larger facilities having already passed the petabyte (PB) threshold for total data, could the first healthcare organization with an exabyte (EB) (1,000 PBs) of data be far behind? If that seems unlikely, one need only refer to a recent study that reveals 30% of the world’s storage belongs to healthcare⁴ to realize that patient and administrative data may be growing beyond the limits of control.

³ As reported in “Survey on the Government of Unstructured Data” 30 June 2008
The Healthcare Data Profile

Like any other business, a healthcare organization employs many different software applications to run its day-to-day activities. However, unlike other businesses, most of these systems directly or indirectly impact patients, i.e., when these applications (and the data behind them) are unavailable, patient care may be compromised. Healthcare IT is charged with keeping these systems available virtually 24/7 and, in the event of an outage, is tasked with restoring operational capacity as quickly as possible. In the event of a disaster, the entire hospital must execute its business continuity plan to ensure minimal disruptions to patient care. Healthcare IT plays its part by carrying out the disaster recovery plan ensuring that these systems continue to function during this critical time.

While the stakes may be higher for healthcare IT, are the data protection and disaster recovery requirements fundamentally different than those of any other business? In many ways they are not, but in one key way they are. Healthcare organizations generate a large percentage of static data, the nature of which requires a particular approach to Healthcare Disaster Recovery. Before we examine that approach, let us first examine in more detail the systems in question and the types of data they contain.

Clinical systems

Clinical systems are the lifeblood of healthcare IT. These are the applications that aid care providers in diagnosing and treating patients each day. When these systems are unavailable, the speed of patient care slows, and quality may be compromised. As such, it is critical that these clinical applications are restored first in the event of a disaster. There are a number of systems and sub-systems that fall into this category.

The Hospital Information System (HIS) stores, gathers and displays information critical to caring for patients in a hospital. Depending on the particular implementation, the HIS may encompass or work alongside the Laboratory Information System (LIS), the Radiology Information System (RIS) and the Electronic Health Record (EHR). Complementing the EHR is often an Enterprise Content Management (ECM) system, which aggregates information (both paper-based and electronic) that normally sits outside the EHR/HIS and makes it accessible to these systems to help establish a comprehensive patient record. Finally, many hospitals have a PACS for the capture of digital images used in patient diagnosis. Radiology PACS has become commonplace in hospitals in developed countries, and other medical disciplines are rapidly bringing online their own digital imaging systems, such as Cardio PACS or Digital Mammography.

Administrative systems

While clinical systems may be what set hospitals apart from other businesses, a hospital is still a business – whether it is a for-profit or not-for-profit organization – and as such requires business systems to operate. These include applications such as email, office productivity (e.g. word processing, spreadsheets, presentations), and accounting systems. Although less critical than the clinical systems, in the event of a disaster all of these will also need to be recovered.
The Data: Where, What and How Much?

All of the aforementioned systems create data, but the type and quantity of that data varies from system to system. It is useful to separate this data into three categories: structured, unstructured and semi-structured data.

Structured data

Applications that generate structured data are database driven. The most common example is the HIS, which may store its data in databases such as MUMPS, Cache, MAGIC, Oracle, or SQL. The RIS, LIS, EHR and accounting systems typically fall into this category, as well. The amount of data these systems generate typically falls in the hundreds of GB range, maybe approaching a few TB in larger facilities.

Unstructured data

Applications that generate unstructured data produce discrete files that are not associated with a database. Office productivity suites are a good example of this. Word processing documents and spreadsheets are routinely created by hospital administrative staff, and these files are typically stored on file servers. And often, different users may store individual copies of identical files. Other file types such as presentations, audio and video files, and pictures are also stored in this manner. The file systems storing this data may be of modest size in a small facility but are often significantly greater than larger healthcare IT organizations, where many TBs of unstructured file data can be a challenge to back up and recover.

Semi-structured data

There is a third category of data common to hospitals that we shall call semi-structured. The two best examples of semi-structured data come from PACS and ECM systems. Both maintain a database of information [structured data] that often reference large quantities of discrete files [unstructured data]. A PACS database may run on Oracle or SQL, and its size may be relatively small in relation to the many TB of DICOM images that the database references. You may also find it useful to think of an email system, and attachments associated to individual email messages, in the same manner.

Best Practices

While the methodology of drafting a complete disaster recovery strategy is beyond the scope of this paper, we can specify some data protection best practices applicable to the data types mentioned herein. Before we do so, it may be helpful to define a couple of key terms.

RPO and RTO

RPO, or Recovery Point Objective, describes the acceptable amount of data loss measured in time as defined by the organization. For example, if a hospital defines a one hour RPO, it has determined that in the event of a disaster, data must be restored to within one hour of the time the disaster occurred. Another way to state this is “How much data can we afford to lose?” In this example, the answer is one hour’s worth of data.

RTO, or Recovery Time Objective, describes the duration of time in which systems must be restored to acceptable service levels after a disaster. For example, if an organization defines a three hour RTO, it has determined that in the event of a disaster, systems must be back up and running, operational, and available to end users within three hours of the time the disaster occurred. Another way to state this is “How long can we afford to be without the system?” In this example, the answer is three hours.

Typically, a hospital will not have a single RPO and RTO for all of its systems. Rather, it will prioritize the criticality of systems to determine in which order they must be restored. This prioritization is a key component of the disaster recovery plan as well as the overarching business continuity plan. For example, an organization may deem the HIS its most critical application and assign it the shortest RPO and RTO. Email may also be deemed a critical application necessary to communication in the event of a disaster, so it may also have a short RTO but a longer RPO. Again, these decisions will vary from hospital to hospital, and a discussion of the factors behind such decisions are beyond the scope of this paper. Instead, we shall focus on how the amount and type of data impact both RPO and RTO and, thus, influence the data protection strategies employed to assure the effective disaster recovery of these systems. Our emphasis will be on a few key clinical applications common to the majority of healthcare organizations.
**HIS and EHR**

Ask any doctor or nurse to specify the RPO and RTO for their HIS or EHR, and they are likely to return your question with a blank stare. Instead ask them, in the event of a disaster, how much data can they afford to lose, and how long can they be without the system, they are likely to answer “none” and “not very long”, respectively. While an immediate RPO and instant RTO may well be disaster recovery nirvana, the economic, if not technological, reality requires a compromise to something less than the always-available system.

HIS and EHR systems are dynamic, their databases being frequently updated during the course of a day as patients are cared for. As such, the expectation is that their RPOs will be short, as will their RTOs. Given that the size of these systems is relatively manageable, it is not unrealistic to aim for an RPO of one to four hours, the higher end being easier and more economical to achieve. With a four hour RPO, a full-system point-in-time copy of the RIS and EHR must be created six times each day. Where this once may have meant backing up these systems to tape every four hours, modern SAN infrastructure and data “snapshotting” and replication technology make it easier to achieve frequent disk-based system replicas. This also allows for quicker RTOs, as disk-based system copies do not have to be restored from tape media. The transportable nature of these copies, enabled by current SAN technology, means a secondary data center can contain complete standby systems, which can be made quickly operational from the most recent recovery point.

**PACS and ECM**

While you would be correct to assume that it is also desirable to have short RPOs and RTOs for the PACS and ECM systems, their data profiles present some unique challenges. For example, these systems may contain many TBs of digital images. One cannot quickly create a point-in-time copy of a system this large. As a result, the RPOs will be much longer as will the RTOs because restoring many TBs of data can be a time-consuming process.

Keep in mind, however, that the databases that drive PACS and ECM systems, along with their short-term data caches containing recently created data, are not very large and, in fact, share similar RPO/RTO characteristics as the HIS and RIS. The large quantity of data files these systems reference – the medical and document images – are what present a disaster recovery challenge. However, one interesting characteristic of these files is that they are static in nature. Unlike the associated databases that are frequently being updated, a scanned copy of a patient’s driver’s license or a patient’s chest X-ray do not change: they are created and stored on disk for future reference. While the meta-data associated with the image may change, the image itself remains constant. As such, making frequent point-in-time copies of this data constitutes a poor data protection strategy. Why back up unchanging data night after night, never mind every few hours, when all you are doing is making copy upon copy of the same thing? There is a better way.
Enter Archiving

Point-in-time copies are useful to recover entire systems of structured data. For large amounts of unchanging, unstructured or semi-structured data, archiving is a better approach. By creating a few geographically dispersed copies of the data on multiple media types, you are best positioned to meet your RPO and RTO for these systems. For example, in the event of a disaster, your PACS database can be recovered as quickly as your HIS and RIS, and then pointed at a secondary copy of the image archive, preferably extant on storage in a secondary data center.

A Holistic Approach

This combination of backup and archiving provides an optimal strategy for treating each data type with the right method. By understanding the nature of the data in the critical clinical systems, healthcare IT can deliver both realistic and acceptable RPOs and RTOs to the business. In the event of a disaster, the organization can rest assured that the systems they need will be available when they need them. And doctors and nurses can get on with the business of patient care rather than concern themselves with the business of healthcare IT.

Conclusion

By now it should be clear why disaster recovery continues to be a top concern for healthcare IT professionals. The proliferation of computerized clinical systems and the rapid adoption and reliance upon those systems by healthcare professionals to aid them in the course of patient care mandates that healthcare IT keep these systems highly available. In the event of a system outage, a plan must be in place to recover quickly and efficiently, prioritized by the criticality of each system as it impacts the business of the hospital.

The protection and recovery of these systems is somewhat complicated by the static nature of the majority of healthcare data – clinical images and scanned documents that rarely change once created and that consume vast quantities of disk storage. In order to optimize efficiencies, healthcare IT must protect this data in a way that meets the availability requirements of the organization while containing costs.

A strategic combination of backup and archiving technologies is the best method to ensure that all data types within a healthcare organization are adequately and cost effectively protected.
With 20 years’ experience in data and storage management, and 10 years in healthcare, BridgeHead Software is trusted by over 1,000 hospitals worldwide.

Today, BridgeHead Software helps healthcare facilities overcome challenges stemming from rising data volumes and increasing storage costs while delivering peace of mind around how to **STORE, PROTECT** and **SHARE** clinical and administrative information.

BridgeHead’s Healthcare Data Management (HDM) solutions are designed to work with any hospital’s chosen applications and storage hardware, regardless of vendor, providing greater choice, flexibility and control over the way data is managed, now and in the future.

Learn more about BridgeHead Software at [www.bridgeheadsoftware.com](http://www.bridgeheadsoftware.com)

Follow BridgeHead Software on Twitter at [www.twitter.com/bridgeheadhdm](http://www.twitter.com/bridgeheadhdm)