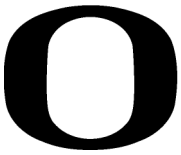
Presented to the Interdisciplinary Studies Program:



UNIVERSITY OF OREGON
APPLIED INFORMATION MANAGEMENT

Applied Information Management
and the Graduate School of the
University of Oregon
in partial fulfillment of the
requirement for the degree of
Master of Science

Optimizing Recovery of Critical Systems Through Replication to the Cloud

CAPSTONE REPORT

Barry J. Gillin
System Application Engineer
Oregon Health & Science University

University of Oregon
Applied Information
Management
Program

Fall 2018

Continuing and Professional
Education
1277 University of Oregon
Eugene, OR 97403-1277
(800) 824-2714

Approved by

_____

Dr. Kara McFall
Director, AIM Program

Optimizing Recovery of Critical Systems Through Replication to the Cloud

Barry J. Gillin

Oregon Health & Science University

Abstract

This annotated bibliography addresses the business requirements for disaster recovery planning in order to avoid the loss of critical data and provide business continuance with a focus on replicating critical data using cloud technology. The references included enable IT managers to make and justify informed decisions prior to implementing a disaster recovery plan. The references related to recovery methods and cloud technology were published between 2010 and 2018.

Keywords: disaster recovery, business continuance, recovery methods, critical systems, data replication, cloud services

Table of Contents

Introduction

Problem

Oregon Health & Science University (OHSU), Oregon's only academic health center, is a public corporation with multiple missions: "excellence in education, research and scholarship, clinical practice and community service" (About OHSU, n.d., para. 3). One of the primary missions of OHSU is to "deliver excellence in healthcare, emphasizing the creation and implementation of new knowledge and cutting-edge technologies" (About OHSU, n.d., para. 4). OHSU has implemented an Electronic Health Record (EHR) system to assist in delivering this excellence. An EHR system is an electronic system which allows for the recording and retrieval of the medical record, the management of the results of laboratory and imaging tests, the tracking of immunizations, and the use of clinical guidelines and protocols (Gans, Kralewski, Hammons, & Dowd, 2005).

A 2011 study of treatment outcomes for diabetes patients in the New England Journal of Medicine documented improved patient care when an electronic health record system was in use. The researchers found "For diabetes outcomes, 43.7% of patients at EHR sites and 15.7% of those at paper-based sites had outcomes that met at least four of the five standards [desired outcomes for blood pressure, weight, blood sugar and cholesterol levels], a difference of 28.0 percentage points" (Cebul, Love, Jain & Hebert, 2011, p. 825-826). The use of EHR systems at hospitals has become ubiquitous; the US Department of Health and Human Services reports that "in 2017, 96 percent of all non-federal acute care hospitals possessed certified health IT" (Percent of Hospitals, n.d., para. 1).

OHSU's conversion to an EHR system has resulted in dependence on the system, and any issues that impact access to the system such as hardware failures, software bugs, electrical

outages and network communication issues all have the potential to have a profound impact on patient care. In a study performed by Kossman and Scheidenhelm (2008), the researchers examined the adoption rate of EHR systems and the effect that the systems had on nursing staff, including the impacts to the nursing staff that resulted from issues that prevented access to the EHR system, such as system downtimes. In discussing the nurses' responses to these access issues, the authors noted "problems encountered with EHR use caused frustration and a sense of decreased effectiveness in job performance and patient care" (p. 72). Severe disruptions in a hospital's EHR systems can even force patients to be diverted to other hospitals, as occurred during a 2016 incident in the United Kingdom due to a ransomware attack that shut down a hospital's EHR system (Vayena, Haeusermann, Adjekum & Blasimme, 2018).

OHSU's EHR system is maintained in the organization's two main data centers, which are "facilities that house high-performance computers, storage servers, computer servers, networking or other information technology (IT) equipment" (Turnbull, Ochieng, Kadlec & Shropshire, 2013, p. 44). Portions of OHSU's EHR system are housed in each data center, which are located approximately 15 miles apart. OHSU has implemented a multi-faceted business continuance and disaster recovery plan to address many of the known risks that have the potential to impact the data centers, and thus OHSU's EHR installation.

Developing and implementing extensive business continuity and disaster recovery plans for continuing patient care during planned and unplanned downtimes of IT is critical for hospitals (Zhong, Clark, Hou, Zang & Fitzgerald, 2013). In this context, a business continuance plan is defined as "the processes and procedures an organization puts in place to ensure that essential functions can continue during and after a disaster" (Mohamed, 2014, p. 57). A disaster recovery plan is defined as the plan "for restoring IT services, but not necessarily restoring

specific hardware and software architectures. Examples of IT services include internet connectivity, telecommunications, and data storage and processing" (Mohamed, 2014, p. 57).

The disaster recovery plan of an IT service, such as an EHR system, is dependent on the extent of the damage and mitigations in place (Turnbull et al., 2013). A 2013 study on disaster recovery plans for hospitals reported, "hospital resilience is the capability to absorb the impact of disasters without loss of functions (termed resistance); maintain its most essential functions (called absorption and responsiveness); and 'bounce back' to the pre-event state (termed recovery)" (Zhong et al., 2013, p. 4).

Some of the major risks that could result in a disaster for any IT organization include human error; power outages; environmental issues, such as air conditioning failures; and network issues (Turnbull et al., 2013). Natural disasters like earthquakes, hurricanes, floods and tornadoes also pose risks to OHSU's data center; while the probability of occurrence of such risks is lower than those previously listed, the potential impact is much higher (Maurer & Lechner, 2014). One study reported that when classifying the cause of a computer downtime, "less than 1% is caused by high impact low probability (HILP) events" (Maurer & Lechner, 2014, p. 46).

Unlike most organizations, the data centers employed by OHSU are all at risk of destruction from a subduction zone earthquake (Toomey et al., 2014). Both data centers are located in the Cascadia subduction zone, which extends over 1,100 km from Cape Mendocino, California, to northern Vancouver Island, British Columbia (Toomey et al., 2014). As Toomey et al. noted in 2014, a major earthquake "has, and will again, hit the Pacific Northwest" (p. 139). The most recent earthquake along this boundary was in 1700; earthquakes in the Cascadia subduction zone occur with a frequency of between a few hundred and 1,000 years (Toomey et

al., 2014). The result of a regional disaster such as a major earthquake could render both of OHSU's data centers inoperable and as a result, make all of the patient data inaccessible.

OHSU's business continuance and disaster recovery plans include the use of mirrored database servers; in this configuration, the source server is referred to as the *principal server* and the destination server is referred to as the *mirror server* (Danel & Neustupa, 2011). As records are changed on the principal server, the principal server simultaneously sends the changes to the mirror instance (Danel & Neustupa, 2011). OHSU's EHR database is mirrored in real time between two data centers located approximately 15 miles apart. In the event of an incident that leaves one data center unable to host the EHR system, operations can be transferred to the other data center. To recover from other issues, such as human error or database corruption caused by software errors, OHSU uses storage system snapshots, or "a point-in-time image of a collection of data" (Chervenak, Vellanki & Kurmas, 1998, p. 2.) and backups where "user data are transferred to a tape, a virtual tape, or a disk" (Chervenak et al., 1998, p. 2.) While OHSU's plan is comparable to plans implemented by other organizations (Mohamed, 2014), the plan does not account for a natural disaster such as an earthquake that renders *both* data centers inoperable.

For critical operations, the goal of disaster recovery is to minimize the effective recovery time objective (RTO), defined as "the length of time a business can be without data processing availability" (Mohamed, 2014, p. 61) and avoid poor recovery point objective (RPO), defined as "how old the data will be once the systems are recovered" (Mohamed, 2014, p. 61). A desirable RTO for a hospital recovering from a disaster is less than one hour and an acceptable RPO is also less than one hour (Peixoto, Duarte, Abelha, Santos, & Machado, 2012). OHSU's current backup processing and offsite replication is a process which takes up to 48 hours to complete. Since "the RPO of a DR system represents the point in time of the most recent backup prior to any failure"

(Wood et al., 2010, p. 2), the resulting RPO for OHSU would be 48 hours or more. The result of a regional disaster impacting both data centers for OHSU is an effective RTO that is very long and a poor RPO.

One potential approach to disaster recovery for organizations that are at risk for natural disasters and cannot afford a lengthy RTO or tolerate the loss of data with a long RPO is using cloud technology to create another replica of the data to a functioning database located in a geographically separated region (Alcântara, Oliveira, & Bessani, 2017). "Cloud computing deals with computation, software, data access and storage services that may not require end-user knowledge of the physical location and the configuration of the system that is delivering the services" (Jadeja & Modi, 2012, p. 877). Infrastructure as a Service (IaaS) provides the required infrastructure components to allow an organization to host a replica of a database at the provider's data center, which can be geographically separate from the organization's data center (Jadeja & Modi, 2012). "Tolerating disasters requires placing backup resources in a geographically separated location so that the same disaster does not affect the primary and the backup infrastructures" (Alcântara et al., 2017, p. 248).

The replication between a primary data center and the cloud backup entails "a continuous transmission of data between the two sites with copies of critical data available at both sites" (Mohamed, 2014, p. 59). For OHSU, this configuration would result in the maintenance of a current copy of the database at both of OHSU's data centers as well as the cloud site. This configuration would result in an RPO that results in almost no lost data; as Alhazmi and Malaiya (2013) note, "If the backup is a synchronous mirrored system, RPO is effectively zero" (p. 3). Since OHSU's database would be in a usable format, no restore would be necessary and thus the RTO would be almost immediate. This technique is recommended by Alcântara, Oliveira, and

Bessani (2017), who note that the system continuously replicates its data to an online remote mirror, which ensures the continuity of the system if a disaster occurs.

Purpose Statement

With the negative impacts that occur with downtimes of EHR systems, information technology support staff members are expected to support the "expectations by health care organizations for highly reliable and nearly always available EHR systems" (Shepard, 2017, p. 187). The purpose of this annotated bibliography is to present literature that addresses the use of cloud technology to establish and maintain an offsite functioning replica of mission critical systems such as EHR systems for disaster recovery purposes and identifies best practices in establishing and maintaining this configuration. While the primary problem of study is related to disaster recovery of EHR systems, this study will also be of use to IT support staff for other mission critical systems.

Research Question

Main Question. What are best practices in establishing and maintaining cost effective data replication to the cloud for mission critical IT systems to minimize downtime and loss of data in the event of a disaster?

Sub-Question. What are the critical concerns with the selection and configuration of a cloud-based data replication solution?

Audience

The stakeholders from OHSU who are impacted by an effort to create a cloud-based offsite storage facility for key patient systems include hospital administrators and the IT managers responsible for providing healthcare IT services. Hospital administrators who fund the

ongoing costs of healthcare data center solutions will be interested in both the efficacy of the model and the initial and ongoing costs.

IT departments are responsible for identifying limitations with their current environments and designing and proposing solutions to address the shortfalls. By performing a study of techniques to help mitigate the risks posed by geographically homogenous data centers, and developing an understanding of the scope of work that would be required to move to a cloud solution, IT employees will gain the knowledge to design solutions and formulate proposals that minimize the impact of regional disasters.

The concepts in this study are applicable to any database system which provides replication capability, and thus are applicable to any organization that considers a low RPO and RTO of critical importance.

Search Report

Search Strategy. The University of Oregon library website provides access to a variety of research guides and I focused my search within the Computer and Information Science Guide. The guide provides access to a variety of search engines; the two most valuable to researching my topic are the ACM (Association of Computing Machinery) Digital Library and IEEE (Institute of Electrical and Electronic Engineers) Xplore. ACM is focused on all aspects of computing, and has an extensive library of peer-reviewed journals that document research in many areas of computer science. IEEE provides a broader range of publications, but also contains research documents that pertain to computer science. The journals published by ACM and IEEE are peer reviewed journals.

To cross reference articles from these sources to other articles that cited them, I used Google Scholar to search for the original article by name, and then used the "cited by" link to locate research related to the original article.

Keywords. I used the following keywords while searching the primary databases. I combined the keywords using Boolean operators to help limit the search results returned.

- database replication,
- database mirroring,
- cloud,
- DR,
- disaster recovery,
- recovery point,
- recovery,
- disaster planning and mitigation,
- impact of disasters on IT services,
- healthcare service dependence on IT,
- disaster recovery in healthcare, and
- business continuance best practices.

Documentation Method

Documentation approach. I downloaded for review articles that looked promising and provided full text access. I reviewed each article starting with the abstract and then performed a quick scan of the article. I saved the full texts of the promising articles for more in-depth review. I made a second pass of promising articles with a more critical eye towards whether the content was pertinent.

I recorded all articles that I kept for future reference in a Microsoft Word document, including the complete American Psychological Association (APA) citation, abstract, local file name and brief notes about the content of the article. I listed each article on a separate page, which allows the pages to be printed and sorted, marked up or collated.

The title, author and brief notes were also copied into a Microsoft Excel spreadsheet, allowing for easy sorting of the articles into the pertinent sections of the annotated bibliography.

Reference Evaluation

Reference evaluation criteria. I evaluated the articles based on the criteria defined in the document "Evaluating Information Sources" published by the University of Florida Center for Public Issues (2014). The criteria are:

- *Authority*. I retrieved most articles from peer reviewed journals. I also performed cross references of whether an article was cited by other articles and whether the author published other articles. For articles that I could not verify were peer reviewed, I examined the author's affiliations, whether the article was cited by authors publishing in other peer reviewed journals, and whether the author had multiple articles published on this or similar topics.
- *Timeliness*. There has been a significant expansion in cloud capacity over recent years, along with increased adoption of cloud usage. Major developments include Amazon's release of AWS in 2006, and Microsoft's Azure release in 2008 (Qian, Luo, Du & Guo, 2009). For this reason, I restricted the searches to 2010 until present to ensure I incorporated the most recent information in the study. I restricted articles that document natural disasters and the impact to computer systems to those published since 2000.

- *Quality*. I evaluated the quality of the writing in areas of grammar, punctuation and spelling, and generally avoided sources that evidenced errors in these areas. Some of the articles were from foreign authors, and I was therefore more lenient of these errors with journals published in other countries.
- *Relevancy*. I evaluated the journal articles to ensure that they were applicable to the main research question and sub-question.
- *Lack of Bias*. I reviewed but eliminated any articles attributed to employees of companies providing cloud services to avoid the risk of bias. I evaluated the remaining articles to determine whether the article was balanced and if there were citations which addressed multiple points of view.

Annotated Bibliography

Introduction to the Annotated Bibliography

This annotated bibliography contains sixteen references that address the concepts of recoverability of mission critical systems such as EHR systems, and the use of replication to the cloud to provide an effective recovery method. The references are categorized as relating to one of three subjects: (a) impact to organizations due to computer downtimes, (b) methods of providing business continuance and disaster recovery services, or (c) best practices for replication to the cloud.

Each annotation contains three elements: (a) the full bibliographic citation in APA format, (b) an abstract, and (c) a summary. The abstracts included are complete as published, or in the absence of a published abstract, written by the author of this annotated bibliography based on the published material. The summary was written by this author and identifies the sections of the areas of the paper relevant to this study.

Impact to Organizations Due to Computer Downtimes

Brotherton, H., & Dietz, J. E. (2014, April). Data center business continuity best practice. *2014 11th International Conference on Information Technology: New Generations (ITNG)* (pp. 496-501). IEEE. doi: 10.1109/ITNG.2014.8

Abstract – This qualitative multiple case study analysis reviews well documented past information technology disasters with a goal of identifying disaster recovery and business continuity best practices. The topic of cyber infrastructure resiliency is explored including barriers to cyber infrastructure resiliency. Factors explored include: adherence to established procedures, staff training in recovery procedures, chain of command structure, recovery time and cost, and mutual aid relationships. Each of these factors is

analyzed in the four cases included in the study and recommendations are presented. Automated fail over and regular disaster recovery drills were found to be key factors for success in actual disaster situations.

Summary – This article is a qualitative case study of high impact IT infrastructure recovery processes. The authors examined four events where incidents occurred and the actions and resolution were documented. The authors analyzed the documentation to identify recurring themes or practices along with the positive and negative results of those practices. The four studies included the failure of Commerzbank's data center due to the September 11 World Trade Center attack; a blackout that effected most of the Northeast U.S. in 2003 due to the failure of a monitoring system at FirstEnergy's data center; Hurricane Katrina's impact to Tulane University in 2005 and a failure at Northrop Grumman which impacted 27 state agencies for the Commonwealth of Virginia. The analysis included examination of the *before incident* planning and mitigation, the performance *during the incident*, and *post incident* plans for improvement.

The findings of the study include issues the authors found with the data center configurations of the organizations included in the research. The authors found that even though Commerzebank had a redundant data center, it took four hours to resume operations, which did not meet the RTO. FirstEnergy had redundant systems within their data center, but this redundancy did not prevent a software-induced failure. Tulane University relied on offsite backup but did not have a recovery site, and some critical systems, such as the student information system, were not included in the DR plan. The Commonwealth of Virginia relied on hardware redundancy, but human error resulted in a

two-day downtime. In all cases, the recovery did not meet the desired recovery time periods.

The authors' evaluation of the before incident planning showed flaws in parts of the plans. Commerzebank employed different configurations at each site, which resulted in incompatibilities and delays. FirstEnergy did not institute sufficient monitoring, did not follow documented procedures and did not have a redundant site. Hurricane Katrina rendered Tulane's backup data center unusable. The Commonwealth of Virginia had redundancy within the data center and a redundant site, but the alternate site was not activated.

All organizations in this study evaluated their existing DR plans and modified them. Commerzebank converted to an active-active configuration and FirstEnergy implemented a new system with better compatibility between two locations. Tulane now has offsite backup servers for systems considered critical and a contract for a mobile data center for recovery of other systems. All of the organizations increased spending on disaster preparation and have included DR training within the plans.

This source is relevant for this study because it clearly demonstrates the impact that unplanned incidents can have on an organization. The four incidents were caused by four distinct issues and impacted four different business sectors. The major conclusions of this study were that redundant systems, if used correctly, can help to achieve the desired RPO and RTO and that the geographic placement of recovery systems is of great importance.

Degaspari, J. (2012). When disaster strikes: How technology drives better preparation. *Healthcare informatics*, *29*(9), 22-23. Retrieved from https://www.healthcare-informatics.com/article/when-disaster-strikes-how-technology-drives-better-preparation

Abstract – Three hospital systems provide details about how technology has influenced the way they prepare for disasters and what they have learned from their experiences.

Summary – This article is a case study of three hospital systems concerning the planning for and the recovery from disasters. The article is based on interviews with senior IT staff from three hospitals that planned for or experienced disasters. The three hospitals and observations were:

- St. John's Regional Medical Center in Joplin, Missouri was hit by an EF-5 tornado. The local servers in Joplin were destroyed, but the data center housing the EHR system had recently been moved to a remote location about 250 miles away. As a result of this move, the patient data remained available, and continued mirroring to a secondary data center.
- Florida Hospital, part of the Adventist Health System located in Orlando, Florida was hit by three hurricanes in a single year. While they did not declare a disaster, and did not need to trigger their recovery processes, the experience demonstrated the shortcomings of the existing recovery processes. The recovery processes relied on restoring backup tapes that would have been several days old, and thus, if they had been forced to trigger the recovery processes, the restore from backup tapes would not have met the RPO. As a result of these disasters, Florida Hospital implemented an asynchronous mirror to a location nearly 1,000 miles away.

- Cooley Dickinson Hospital in Northampton, Massachusetts did not experience any disasters but recognized that the backup environment in place would not be sufficient in case of a disaster. Specific shortcomings they identified included individual tape drives for each system, which resulted in a large number of tape formats and types of tape drives and separate procedures to recover each system, which increases the complexity of the restoration and reduces the efficiency of the staff performing the recovery. To address these shortcomings, Cooley Dickinson implemented an enterprise backup system that backs up complete system images to disk and then replicates the backups to tape for offsite storage. The hospital is also developing a secondary site on their campus for real-time replication.

This source is relevant for this study because of the discussion of the impact to critical EHR systems due to disasters. Two of the hospitals were affected by natural disasters, which forced them review the business continuity and disaster recovery plans that were in place, and to evaluate whether their disaster recovery plans were still valid with the increased dependence on computerized healthcare systems. Key findings include that reliance on tape backup cannot meet the RPO and RTO required of an EHR system, and that the best RPO and RTO is provided by real-time replication between geographically dispersed locations.

Lawson, J. (2005, December 9). A look back at a disaster plan: What went wrong -- and right. *Chronicle of Higher Education*, B20–B22. Retrieved from https://www.chronicle.com/article/A-Look-Back-at-a-Disaster/10664

Abstract - The article explores the strengths and weaknesses of a disaster plan, which was implemented by Tulane University. The disaster plan was based upon time, location,

and intensity of natural disaster. Tulane did not have a formal disaster-recovery plan for replacement of machines with any outside vendor or institution. The communications plan was not as robust as was needed. CIOs were advised to have an old-fashioned radio system for backup communications on campus.

Summary – This article was written by the vice president for information technology and chief information officer at Tulane University after Hurricane Katrina hit New Orleans in August 2005. The author discusses what went right with the disaster plan, which included that everyone was safe, the plan was followed, backups were created of all the systems within the data center and the tapes were safely stored for later recovery. Unfortunately, there was no contract for a recovery site and not all systems were included within the data center DR plan. Two examples mentioned were the accounts receivable system, which was being run out of the controller's office, and the payroll printer.

Major negative impacts to the university following the hurricane included disrupted communications to employees and students, financial systems such as payroll and accounts receivable not functioning, and the inability to open the campus for classes. Observations after the disaster highlighted the critical importance of the university's email and web sites to disaster communication, and the incomplete plans for the restoration of those services. The DR plan called for the transfer of email services to Yahoo if the university's system was not available, but an incomplete user list resulted in many people not having access. The emergency web sites had been moved to a vendor who used virtual IP addresses, but the resulting dynamic nature of the addresses prevented user access until the vendor updated the configuration with static addresses. The email and web difficulties prevented effective communication to staff and students.

The author also notes the ongoing expenses associated with having a robust plan, such as the anticipated $300,000 per year to fund a recovery site, and that not having a recovery site was a conscious decision to save money and accept the risk.

This source is relevant for this study because it demonstrates that disaster planning is extremely important for any organization and that even with a plan, the recovery will take time and might not be complete.

Shepard, A. (2017). Disaster recovery and the electronic health record. *Nursing Administration Quarterly*, *41*(2), 187-189. doi: 10.1097/NAQ.0000000000000213

Abstract – Abstract written by the author of this annotated bibliography in the absence of a published abstract. There have been several federal initiatives that have increased the reliance on the electronic health record (EHR) for recording clinical observations and providing archival storage of patient personal health information. These initiatives include the HITECH (Health Information Technology for Economic and Clinical Health) Act of 2009 and the Centers for Medicare & Medicaid Services program to report electronic Clinical Quality Measures implemented in 2017. These measures result in the expectation of highly reliable and available EHR systems. Health care organizations need to be prepared to deal with disasters such as a tornado that destroyed a hospital in Joplin, Missouri in 2011, or the damage to multiple hospitals in New Orleans caused by Hurricane Katrina in 2005.

This article discusses the impacts to healthcare organizations due to failures of EHR systems and inadequate preparation requirements to be able to provide effective care to patients during crisis situations.

Summary – The author examines the impact that a computer system downtime has on healthcare organizations and the steps required for the organizations to continue to function in the case of a downtime. The article reports that the expectation is for "nearly always available EHR systems" (p. 187) and that the financial cost associated with a downtime that impacts the EHR system can be up to $17,000 per minute. The author asserts that in the case of a natural disaster, like the tornado in 2011 that destroyed a hospital in Joplin, Missouri or Hurricane Katrina hitting New Orleans in 2005, a strong business continuance plan is essential. The plan must include components that mandate who is in charge, what type of backup alternative information must be maintained, standard procedures for patient treatment to be followed, and how the manually transcribed information will be entered into the EHR system once functionality is restored.

The conclusion of this paper states that dependence on the EHR system has become the norm, but that emergencies and disasters will occur and the medical staff must be prepared in order to protect the health of the patients.

While this source does not address the IT techniques to mitigate primary failure points or to recover from failures that occur, it is valuable for this study because it shows the dependence of healthcare organizations on the EHR and the true impact to healthcare organizations whenever the EHR system is not available.

Wang, Y., Coiera, E., Gallego, B., Concha, O. P., Ong, M. S., Tsafnat, G., ... & Magrabi, F. (2016). Measuring the effects of computer downtime on hospital pathology processes. *Journal of biomedical informatics*, *59*, 308-315. doi: 10.1016/j.jbi.2015.12.016

Abstract – *Objective*: To introduce and evaluate a method that uses electronic medical record (EMR) data to measure the effects of computer system downtime on clinical processes associated with pathology testing and results reporting.

Materials and methods: A matched case-control design was used to examine the effects of five downtime events over 11-months, ranging from 5 to 300 min. Four indicator tests representing different laboratory workflows were selected to measure delays and errors: potassium, hemoglobin, troponin and activated partial thromboplastin time. Tests exposed to a downtime were matched to tests during unaffected control periods by test type, time of day and day of week. Measures included clinician read time (CRT), laboratory turnaround time (LTAT), and rates of missed reads, futile searches, duplicate orders, and missing test results.

Results: The effects of downtime varied with the type of IT problem. When clinicians could not logon to a results reporting system for 17-min, the CRT for potassium and hemoglobin tests was five (10.3 vs. 2.0 days) and six times (13.4 vs. 2.1 days) longer than control ($p = 0.01–0.04$; $p = 0.0001–0.003$). Clinician follow-up of tests was also delayed by another downtime involving a power outage with a small effect. In contrast, laboratory processing of troponin tests was unaffected by network services and routing problems. Errors including missed reads, futile searches, duplicate orders and missing test results could not be examined because the sample size of affected tests was not sufficient for statistical testing.

Conclusion: This study demonstrates the feasibility of using routinely collected EMR data with a matched case-control design to measure the effects of downtime on clinical processes. Even brief system downtimes may impact patient care. The methodology has potential to be applied to other clinical processes with established workflows where tasks are pre-defined such as medications management.

Summary – The authors discuss an analysis performed in an attempt to quantify the impact to patient care that results from downtimes that affect EHR system availability. The authors note that there are numerous studies of the types and durations of downtimes that impact computer systems, but their focus is on the frequent issues that result in EHR systems being unavailable for durations of multiple hours. Their analysis is an attempt to quantify the impact of an EHR downtime on patient health by employing information contained within the EHR system.

The authors conducted a review of selected common lab tests that are critical to patient care, including potassium, troponin and hemoglobin levels, in an attempt to determine a correlation between system downtime and the performance of these tests. Key timings recorded within the EHR system included the times that the blood test is ordered, blood is drawn, the sample is run, the results are posted and the clinician reviews the results.

The authors identified five fully documented downtime events that occurred within a one year period at a metropolitan teaching hospital. The duration of the events ranged from five minutes to five hours, with two of the incidents lasting more than four hours. The authors report “the contingency process for significant interruptions to system availability was to hand deliver urgent test requests to the laboratory and results were communicated to clinicians by telephone” (p. 310).

The authors performed a statistical analysis of the target lab tests during the events to analyze the impact of EHR downtimes on the key timings recorded within the EHR systems. The authors determined that the timings were impacted by the downtime, but the limited sample size prevented conclusions from being drawn. The authors also noted that while the EHR downtime delayed the delivery of test results to the clinicians, the impact of the delay on patient outcomes is an area that has not been determined.

This source is relevant for this study because it documents the frequency and duration of system downtimes. The study also describes clinical workflows that are impacted by system downtimes, includes details of contingency processes that hospitals have implemented, and outlines the potential risk that disruptions in the workflows pose to patient care.

Methods of Providing Business Continuance and Disaster Recovery Services

Alhazmi, O. H., & Malaiya, Y. K. (2012, November). Assessing disaster recovery alternatives: On-site, colocation or cloud. *2012 IEEE 23rd International Symposium on Software Reliability Engineering Workshops (ISSREW),* (pp. 19-20). IEEE. doi: 10.1109/ISSREW.2012.20

Abstract – Every organization requires a business continuity plan (BCP) or disaster recovery plan (DRP) which falls within the cost constraints while achieving the target recovery requirement's in terms of recovery time objective (RTO) and recovery point objective (RPO). The organizations must identify the likely events that can cause disasters and evaluate their impact. They need to set the objectives clearly, evaluate feasible /DRPs to choose the one that would be optimal. Here we examine tradeoffs involved in choosing among the disaster recovery options. The optimal disaster recovery

planning should take into consideration the key parameters including the initial cost, the cost of data transfers, and the cost of data storage. To evaluate the risk, the types of disaster (natural or human-caused) need to be identified along with the probability of a disaster occurrence and the costs of corresponding failures needs to be evaluated. An appropriate approach for the cost evaluation needs to be determined to allow a quantitative assessment of currently active disaster recovery plans (DRP) in terms of the time need to restore the service (associated with RTO) and possible loss of data (associated with RPO).

Summary – This article gives an overview of disaster recovery planning for IT systems. After addressing the basic need for a DR plan, the authors discuss the concepts of cold, hot, or warm standby sites to meet specific RPO and RTO objectives. The authors define cold standby sites as requiring hardware, operating system and application installation and configuration prior to use; hot standby sites as sites that can take over processing within seconds, as no operating system or application installation is required; and warm standby sites as a tradeoff between hot and cold sites, where some installation or recovery must be performed prior to use. The authors note that the DR site can be onsite, at an alternate location or in the cloud. The decisions made concerning hot or cold backup sites and the location of the backup site to meet the RPO and RTO have direct correlations to the cost and time to recover. The authors note that hot backup sites are much more expensive but have much shorter RPO and RTO than cold backup sites, and the geographic location results in a statistical correlation between concurrent failures at primary and backup locations due to regional issues. The authors attempt to use the

information to develop analytic formulas to enable organizations to prepare quantitative evaluations of the cost and recovery time of the DR plan options.

This source is relevant for this study because it provides a good overview of the choices that are made when designing and funding a DR plan. The concepts apply to all organizations and can form the starting point for determining the funding required for a DR plan.

Bajpai, A., Rana, P., & Maitrey, S. (2013). Remote mirroring: A disaster recovery technique in cloud computing. *International Journal of Advance Research in Science and Engineering*, *2*(8), 3. Retrieved from https://pdfs.semanticscholar.org/9c87/f581d4e7c39a9db9867ec90aa8fbda43de80.pdf

Abstract – As we know Cloud Computing offer many advantages such as scalability, multi-tenancy of resources and cut hardware costs using virtualization technologies like Xen, Hypervisor software . VMware is also a brand name in virtualization technology. But what happens when a catastrophic failure occurred to the whole cloud storage system Until now there is no such provision to secure our private and confidential data. We have three services models available in the market PaaS (Platform as a service), SaaS (Software as a service) and IaaS (Infrastructure as a service) . In this paper we proposed another cloud service model disaster recovery as a service and a architecture how we can increase the fault tolerant of cloud service model during catastrophic failure. So we use a technique called remote mirroring to protect our data using ISCSI initiator.

Summary – This article details the need for a disaster recovery plan for systems which are hosted in the cloud. The authors state that IT managers assume that since cloud providers maintain multiple data centers, data is therefore replicated between centers, but

note that not all cloud providers provide replication between centers. This fact forms the basis for the article, which is the idea that cloud data must be protected by a valid DR plan. The paper presents multiple options for implementing DR solutions, including backup to tape or disk, and focuses on remote replication of data.

The authors describe the remote replication options of synchronous mirroring, where each write operation is acknowledged once it has been completed at the both sites, and asynchronous mirroring, where the write is acknowledged after the write completes at the local site and the write occurs at the remote site at some point in the future. The author states that synchronous replication provides for no data loss between data centers, but due to long round trip transmission times, is only suitable for distances up to 160 kilometers. The author also states that asynchronous replication can work at distances of up to 1,000 kilometers and requires low network bandwidth, but the RPO can be from minutes to hours behind the primary site. The author does not quantify the required bandwidth for synchronous or asynchronous replication.

This source is relevant for this study because it provides a good overview of alternatives for data restoration as part of a DR plan and the implications to RPO for the various alternatives.

Mohamed, H. A. R. (2014). A proposed model for IT disaster recovery plan. *International Journal of Modern Education and Computer Science*, *6*(4), 57. doi: 10.5815/ijmecs.2014.04.08

Abstract - IT disaster recovery planning is no longer an option. Reliable IT services have become an integral part of most business processes. To ensure the continued provision of information technology, firms must engage in IT disaster recovery planning.

Surprisingly, there is little research on this topic. IT disaster recovery planning has not been fully conceptualized in mainstream IT research. A previously framework for assessing the degree of IT disaster recovery planning. Practitioners can use this study to guide IT disaster recovery planning. Our Disaster Recovery Plan is designed to ensure the continuation of vital business processes in the event that a disaster occurs. This plan will provide an effective solution that can be used to recover all vital business processes within the required time frame using vital records that are stored off-site. This Plan is just one of several plans that will provide procedures to handle emergency situations. These plans can be utilized individually but are designed to support one another. The first phase is a Functional Teams and Responsibilities the Crisis Management Plan. This phase allows the ability to handle high-level coordination activities surrounding any crisis situation. We will also discuss the development, finally maintenance and testing of the Disaster Recovery Plan.

Summary – This article addresses the basic need for organizations to develop, maintain and test IT disaster recovery plans. Reasons that prompt the need for recovery include mistakes that can be made by system users which corrupt data, the prevalence of network attacks such as virus and hacking exploits, hardware failures and natural disasters. The author discusses the benefits of having a DR plan, including minimizing economic loss, disruption of mission critical functions, and potential legal liability and improving the ability to recover after an incident.

A significant amount of the paper is devoted to the various options available for building a DR plan. The first option presented is backup to tapes, which are prepared locally and then transported to an offsite storage facility without a predetermined recovery facility.

The recovery time is noted as being more than a week. The next option expands on the first by including a recovery site that has sufficient hardware and network infrastructure to support the critical systems. The author notes that including the recovery site increases the cost but decreases the recovery time to "more than a day" (p. 59). The third tier of disaster recovery presented by the author is labelled electronic vaulting. The author states that "electronic vaulting consists of electronically transmitting and creating backups at a secure facility, moving business-critical data offsite faster and more frequently than traditional data backup processes allow" (p. 59). The increased frequency of the backups reduces the amount of data loss in the case of disaster. The author states that a recovery facility requires equipment to be running to support the electronic vaulting, which increases costs, but the time to recover is "about one day" (p. 59).

The next tier of disaster recovery also includes a recovery data center, but utilizes "continuous transmission of data between the two sites with copies of critical data available at both sites" (p. 59). Recovery needs to be performed from the data copy, with the recovery time being comparable to the recovery time day in tier three, but with less data loss due to the continuous transmission of data. Tier five as presented by the author is a configuration that includes database mirroring between sites. The author notes that this tier requires "that both the primary and secondary platforms' data be updated before the update request is considered successful" (p. 60), which is frequently described as synchronous replication. The author notes that there is minimal data recovery required, but network reconfiguration is required, and the result is a recovery time of "less than 12 hours" (p. 60).

The author presents the “ultimate level of Disaster Recovery” (p. 60), which includes the database mirroring from tier five, but also includes advanced coupling with the application, allowing an automatic switchover in case of disaster. The author notes that this is the most expensive DR solution, but it provided the fastest recovery, typically a few minutes.

The author then presents an example DR plan based on tape backups being stored offsite with the recovery performed at a backup facility. The author uses this specific DR plan to serve as an example of a plan that contains the critical elements of a DR plan, and not as a desired solution. The critical elements listed in the DR plan presented include managing the recovery operation effectively, restoring critical applications to the most current backup copy, and outlining the duties and responsibilities of data center staff and management to recover from a disaster, with an eye towards limiting the impact while remaining cost effective.

This source is relevant for this study because it provides detailed examples of the conceptual options available for disaster recovery planning. The recovery plan formulated is a simple plan with a long RTO and a RPO of up to one week. While this plan does not meet the RPO and RTO for critical database systems, and it does not address the use of the cloud for recovery purposes, it provides useful conceptual information for the recovery options that are available, and the detailed recovery plan can be modified to fit the solution desired.

Turnbull, L., Ochieng, H., Kadlec, C., & Shropshire, J. (2013). Improving service continuity: IT disaster prevention and mitigation for data centers. *Proceedings of the 2nd Annual Conference on Research in Information Technology* (pp. 43-46). ACM. doi: 10.1145/2512209.2512213

Abstract – Data centers provide highly-scalable and reliable computing for enterprise services such as web hosting, email, applications, and file storage. Because they integrate a range of different systems, data center administration is a complex process. Managing the risk of IT disaster is especially difficult. Layers of interrelated infrastructure multiply the effect of system malfunctions. Seemingly-small problems can turn into major disasters and take entire data centers offline. To cope with the myriad risks, this research develops a matrix of IT disaster prevention and mitigation techniques for data centers. The matrix is organized along two dimensions: attributes of data center infrastructure and elements of the IT disaster recovery process. It includes 134 specific techniques which were clustered into 49 cells within the matrix. An expert panel assessed the validity of the matrix and ranked the techniques within each cell. The result is a comprehensive tool for improving the resilience of data centers.

Summary – The authors performed a content analysis of over 200 peer reviewed articles related to disaster recovery for data centers in order to develop a matrix of data center layers and disaster recovery processes which can be applied to those layers. The researchers considered 114 of the articles to contain information relevant to the study of disaster recovery of data centers. Multiple researchers reviewed each article and classified which layer of the data center environment was impacted by the reported

failure. The layers include the physical structure, utilities, networking, computer hardware, operating system, and application.

The researchers also categorized the recovery processes mentioned in the article and the data center layers to which the processes applied. The recovery processes include detection and notification, preparation of recovery team members, service analysis to determine threats and prioritize for recovery, recovery process, backup procedures, offsite storage, testing and maintenance. The authors' goal was to build a matrix to be used by organizations to help identify areas where data center layers have a risk, and for which no recovery process has been implemented. The authors included a summary of the recovery processes as they applied to the data center layers from the articles included in the review.

This source is relevant for this study because it provides a comprehensive list of the layers within the data center model that can be impacted by risks that can render them unusable. The authors identify recovery methods where risks cannot be eliminated, which include planning and training for recovery, prioritization of recovery, backup and restore procedures, and testing and maintenance of the plans. The combination provides failure scenarios and alternatives for recovery to be considered during disaster recovery planning.

Zhu, T., Xie, Y., Song, Y., Zhang, W., Zhang, K., & Gao, F. (2017). IT disaster tolerance and application classification for data centers. *Procedia Computer Science*, *107*, 341-346. doi: 10.1016/j.procs.2017.03.115

Abstract – The increased demand for continuous operation of data centers in recent years has created great interest in remote backups, which allow systems to continue their

application and operation even in the presence of disasters. Data centers (DCs) provide computing service, storage service and network service. All of these different services mean that DCs management is a complex task. Disaster such as power outage and natural hazard may happen at a certain probability, so DCs management is not easy, and essential measures should be taken. In this paper we discuss the technique of disaster recovery (DR) and classification of disaster tolerance (DT). Then to cope with the various risks as mentioned before, a matrix of IT DT measures and different ranks for DCs applications is created. Experts who are specialized in DT assessed the validation of the matrix. The result is an effective tool to improve DCs high availability, and to accelerate the employment of DR. At last a SDN-like distributed DT architecture is described.

Summary – The authors discuss the basic needs for a disaster tolerance (DT) plan, which is a plan designed to enable a computer system to survive failure scenarios with little or no impact to the user. The failure scenarios presented include simple hardware failures, multiple failures and the failure of a data center. The authors present a brief background of standard disaster recovery terms such as RTO, the time required to restore functionality, and RPO, the amount of data loss tolerated. The authors state that the combination of the RTO and RPO will dictate the recovery facility type required to meet the business need, with choices of cold, warm, and hot standby facilities. The authors describe the implementation of a cold standby facility as an implementation where data is restored from backups to hardware at the time of disaster. This results in a long recovery time and loss of data from the time of last backup until the disaster occurred. Warm and hot standby facilities are implemented by having data continually transferred to the recovery facility and differ in the aspect of whether recovery operations are required prior

to activating the second site (warm standby) or whether the data is up-to-date within the database (hot standby).

The authors note that the replication method employed to maintain the information at the standby facility can be accomplished at different layers within the application stack. These layers include application, database, host, storage controller or storage layers. Application and database replication functionality is provided within the application or database software. Host-based replication occurs with the operating system's volume management layer. Storage controller and storage layer replication is implemented by storage vendors.

The authors use the information presented in the introductory sections to introduce a concept where multiple levels of redundancy are in a design that is referred to as a distributed disaster tolerant environment. The concept employs data replicated between data centers, application servers that have access to storage in both data centers and shared network space so that application failover is transparent to the user. The solution is only briefly presented and only contains the concepts. The authors acknowledge that the solution needs further research.

This source is relevant for this study because it introduces the concept of performing replication at multiple levels within the application stack and introduces the software layers used to support the replication. The ability to perform replication using different methods could be useful if there are multiple types or sources of data requiring replication.

Best Practices for Replication to the Cloud

Alcântara, J., Oliveira, T., & Bessani, A. (2017, December). Ginja: One-dollar cloud-based disaster recovery for databases. *Proceedings of the 18th ACM/IFIP/USENIX Middleware Conference* (pp. 248-260). ACM. doi: 10.1145/3135974.3135985

Abstract – Disaster Recovery (DR) is a crucial feature to ensure availability and data protection in modern information systems. A common DR approach requires the replication of services in a set of virtual machines running in the cloud as backups. This leads to considerable monetary costs and managing efforts to keep such cloud VMs. We present GINJA, a DR solution for transactional database management systems (DBMS) that uses only cloud storage services such as Amazon S3. GINJA works at file-system level to efficiently capture and replicate data updates to a remote cloud storage service, achieving three important goals: (1) reduces the costs for maintaining a cloud-based DR to less than one dollar per month for relevant databases' sizes and workloads (up to 222 times less than the traditional approach of having a DBMS replica in a cloud VM); (2) allows a precise control of the operational costs, durability and performance trade-offs; and (3) introduces a small performance overhead to the DBMS (e.g., less than 5% overhead for the TPC-C workload with ≈ 10 seconds of data loss in case of disasters).

Summary – This article focuses on disaster mitigation where a disaster is described as affecting "the whole (or at least a big part of the) infrastructure where the system is hosted" (p. 248). The authors' goal is to present information on the use of cloud resources to host critical systems in geographically separated locations. The authors present the information through the discussion of the details of a system called Ginja. The authors note that the objectives of Ginja are: "low operational costs, fine-grained control over the

data that can be lost due to a disaster, low performance overhead, ease of use, and portability among different DBMS" (p. 248).

The authors present an overview of a traditional disaster recovery solution based on backup and restore operations and note that this is a low-cost solution but requires a significant time to recover (RTO) and frequently results in the loss of data (RPO). The data loss and recovery time is not acceptable to many organizations and thus these organizations move towards real-time replication between data centers. The replication can be synchronous, where processing is stalled until both replicas are updated, or asynchronous, where the processing can continue once the local copy is updated. The authors note that synchronous replication has the benefit of data consistency between data centers at the cost of performance. The authors further note that replication in this manner can be expensive, as it requires a fully functional database system to be acquired and maintained at both locations.

The authors transition into explaining the details of an application environment that they assert works with several types of databases, and functions by intercepting writes to the local disks and replicating that information to low-cost, cloud-based storage such as Amazon S3 (Simple Storage Service). Simple cloud storage services provide the ability for customers to upload and store information in the cloud, but do not include compute services. The authors quote prices for Amazon S3 of $0.023 per GB per month. They compare this to the lowest cost option for Amazon Elastic Cloud Compute (EC2) service that includes compute and storage resources and is suitable for databases, which is $48.24 per month. The authors state that by using a storage service such as Amazon S3, the users incur significantly lower costs with simplified configuration and maintenance.

The authors present extremely detailed information explaining the disk operations performed within relational databases and how these operations are intercepted and replicated to the cloud storage. The details of the application include the safeguards within the application to ensure valid data and negligible impact to the performance of the database.

A cost comparison between cloud and a full virtual system shows that cloud storage used in this manner is 2% to 10% of the cost of a full virtual system in the cloud. The authors note that in case of a disaster, recovery requires the restoration of the data stored within this storage system to a system capable of hosting the database application, resulting in a long RTO but a small RPO.

This source is relevant for this study because of the extensive background in cloud technologies and techniques the authors discuss, not as a recommendation of the application they described. The technologies the authors discuss include the differences between cloud storage services and cloud infrastructure services, and the advantages and disadvantages of traditional synchronous and asynchronous techniques. This background information provides information that is relevant to any replication to the cloud.

Alhazmi, O. H., & Malaiya, Y. K. (2013, January). Evaluating disaster recovery plans using the cloud. *Reliability and Maintainability Symposium (RAMS), 2013 Proceedings-Annual* (pp. 1-6). IEEE. doi: 10.1109/ISSREW.2012.20

Abstract – Every organization requires a business continuity plan (BCP) or disaster recovery plan (DRP) which falls within cost constraints while achieving the target recovery requirements in terms of recovery time objective (RTO) and recovery point objective (RPO). The organizations must identify the likely events that can cause

disasters and evaluate their impact. They need to set the objectives clearly, evaluate feasible disaster recovery plans to choose the DRP that would be optimal. The paper examines tradeoffs involved and presents guidelines for choosing among the disaster recovery options. The optimal disaster recovery planning should take into consideration the key parameters including the initial cost, the cost of data transfers, and the cost of data storage. The organization data needs and its disaster recovery objectives need to be considered. To evaluate the risk, the types of disaster (natural or human-caused) need to be identified. The probability of a disaster occurrence needs to be assessed along with the costs of corresponding failures. An appropriate approach for the cost evaluation needs to be determined to allow a quantitative assessment of currently active disaster recovery plans (DRP) in terms of the time need to restore the service (associated with RTO) and possible loss of data (associated with RPO). This can guide future development of the plan and maintenance of the DRP. Such a quantitative approach would also allow CIOs to compare applicable DRP solutions.

Summary – This article addresses the general need to have a robust disaster recovery plan, backed up with real world costs incurred by various industries as a result of computer downtime. After giving detailed information concerning the standard DR topics of RPO, RTO, and geographic separation between data centers, the authors discuss the tiers used within DR plans for recovery options. These tiers include backup tape, backup to disk, remote logging of backup information, mirrored data and mirrored data with automatic failover. The primary focus of the paper presented by the authors is the use of cloud services, such as provided by Amazon Web Services (AWS), to host a DR backup site. The authors present key concerns with the use of the cloud, including the fact that

the geographic separation increases the likelihood of surviving a regional disaster, but also increases the transmission delay between sites; security concerns of storing data outside of the organization's control; and the tradeoffs of the savings which are realized by on-demand capacity versus the potential of the capacity not being available in the event of a regional disaster that suddenly increases demand.

The authors note that there are some limitations that may be encountered with the use of the cloud for DR purposes. These limitations include potential oversubscription of resources if there is a large-scale disaster, the potential for (unspecified) cloud-specific vulnerabilities, and outages which affect the cloud itself.

This source is relevant for this study because of the discussion of the concerns with implementing a cloud-based DR plan. The overview of the options available, along with the discussion of the tradeoffs, will help to enable thoughtful strategic decisions.

Bahale, M. S. V., & Gupta, S. (2014). Virtualizing disaster recovery management based on cloud computing. *International Journal of Research in Advent Technology*, *2*(5), 397-401. Retrieved from http://www.ijrat.org/downloads/may-2014/paper%20id-25201464.pdf

Abstract – Almost from the beginning of widespread adoption of computers, organizations realized that disaster recovery was a necessary component of their information technology plans. Business data had to be backed up, and key processes like order entry, billing, payroll and procurement needed to continue even if an organization's data center was disabled due to a disaster. Growing reliance on crucial computer systems means those even short periods of downtime can result in significant financial loss, or in some cases even put human lives at risk. Many business and government services utilize

Disaster Recovery (DR) systems to minimize the downtime incurred by catastrophic system failures. Cloud computing provides the third leg of a disaster recovery plan that is essential for business continuity. Cloud-based storage services take advantage of Internet access to deliver reliable, low-cost online storage, helping you to bounce back from a full-scale data center disaster for less than the cost of a dedicated online storage solution. Virtualization is the means of ushering in a new, productive era of cloud computing, driven by this need for cost management and increased agility. Virtualization can also provide the basic building blocks for your cloud environment to enhance agility and flexibility. This paper delineate how virtualization of cloud computing can be used to address the concerns resulting in improved computer infrastructure that can easily be restored following a natural disaster ,reduced expenses, improved scalability, better performance and is easier to manage.

Summary – This article contains four major sections concerning virtualized cloud platforms that can be used to provide low cost DR solutions. In the first section, the authors present information about data recovery requirements such as the recovery time objective (RTO), which is the time that it takes to recover after a disaster, and the recovery point objective (RPO), which describes the loss of data. This section also includes a section on geographic separation of data centers where the authors state that a geographic separation provides increased protection against a regional disaster but at the cost of increased transmission latency. The authors note that the latency can impact performance, which can be mitigated by the use of asynchronous replication at the expense of loss of data in a disaster.

The authors address traditional approaches to disaster recovery based on backup and restore operations in the second section. Traditional backup and restore operations, where data is backed up to tape or disk, suffer from long times to recover and large data loss due to the time between backups. The authors use the next section to discuss replication of systems for either hot (actively mirrored), warm (replicated but offline) or cold (equipment available to perform restore) site recovery, with an overview of the recovery operations and time required for recovery. The authors note that a hot site is quick to recover with a minimum loss of data.

The authors present the main focus of the article in the final section, addressing the use of the cloud for hosting the data replicas. This section addresses major concerns that organizations will have to address with a cloud-hosted recovery system. These concerns include network configuration issues such as the requirement to override existing DNS entries, which will allow access to the failover site. Organizations that are concerned with data security will need a guarantee from the cloud provider concerning privacy of storage, network and virtual machine resources. The authors note that virtualization systems used to provision and manage cloud-based servers function differently than on-premise virtualization, which can lead to challenges for organizations in maintaining compatible systems within two environments. The authors also note the benefits of capacity on demand pricing and the ability to scale virtual servers as conditions change, which allows organizations to configure DR sites at a price lower than traditional owned or leased recovery facilities.

This source is relevant for this study because of the information provided regarding concerns about networking changes required for failover, the operational issues with

virtualization environments and security concerns that pertain to any organization that is considering using the cloud as a replication target for DR purposes.

Regola, N., & Chawla, N. V. (2013). Storing and using health data in a virtual private cloud. *Journal of Medical Internet Research, 15*(3), 1-12. doi: 10.2196/jmir.2076

Abstract – Electronic health records are being adopted at a rapid rate due to increased funding from the US federal government. Health data provide the opportunity to identify possible improvements in health care delivery by applying data mining and statistical methods to the data and will also enable a wide variety of new applications that will be meaningful to patients and medical professionals. Researchers are often granted access to health care data to assist in the data mining process, but HIPAA regulations mandate comprehensive safeguards to protect the data. Often universities (and presumably other research organizations) have an enterprise information technology infrastructure and a research infrastructure. Unfortunately, both of these infrastructures are generally not appropriate for sensitive research data such as HIPAA, as they require special accommodations on the part of the enterprise information technology (or increased security on the part of the research computing environment). Cloud computing, which is a concept that allows organizations to build complex infrastructures on leased resources, is rapidly evolving to the point that it is possible to build sophisticated network architectures with advanced security capabilities. We present a prototype infrastructure in Amazon's Virtual Private Cloud to allow researchers and practitioners to utilize the data in a HIPAA-compliant environment.

Summary – The authors' goals in this paper are to demonstrate the necessary steps to set up a virtual server that uses the HL7 message passing protocol, used to pass healthcare

related information, in an Amazon virtual private cloud (VPC) environment and address Health Insurance Portability and Accountability Act (HIPAA) compliance. HIPAA requires national standards for health care transactions and code sets, unique health identifiers, privacy of patient health information (PHI) and security of the servers on which the PHI is stored. Since the authors are utilizing existing applications that use defined transactions, code sets, and identifiers, the paper focuses on the privacy and security aspects of HIPAA, as the these are defined within the computational environment created.

The HIPAA privacy rule is designed "to ensure that consumers' PHI is handled appropriately within the health organization and only shared with outside entities according to the uses permitted by law" (p. 3). The authors state that this paper is written with the assumption that the organization has the right to access the patient information and is not exchanging the information with other entities, as the exchange of data with other organizations is subject to legal requirements that are independent of where the data is stored. The security rule covers the security of the server, which depends on "the physical security of the server and network, the operators and users of the server, and the configuration and management of the applications, operating system, and network" (p. 3). The authors focus on infrastructure as a service (IaaS) implementations, which allow "customers to run an operating system of their choice and maintain full control of their operating system and network environment" (p. 5). The full control allows the customer to configure the system to include firewalls, encryption and other security-related features to meet the security requirements.

To meet the requirements of the security rule, the authors identified several risk areas requiring mitigation. The first is administrative safeguards, which sets the organization's security policy. Items in this section include having a documented security policy, defining the roles that are allowed to access the patient data, providing authorization and supervision to anyone who has access to protected data, and ensuring that users are trained on protecting the data. The authors assert that user access concerns remain the same whether the data is on premise or in the cloud.

Cloud implementations do have concerns related to physical access and the controls needed to limit physical access to the facilities while allowing authorized access. The authors report "Amazon EC2's service has completed a SAS 70 Type II audit, obtained ISO 27001 certification, and PCI level DSS validation as a level 1 service provider" (p. 6), indicating Amazon's commitment to adhering to stringent security practices related to their hosting services. Technical safeguards in the HIPAA environment include access control, audit controls, integrity controls, and transmission security. To avoid a known issue with Secure Shell (SSH) password based logins, the authors used SSH2 RSA keypairs, where each user is assigned a passphrase-protected key that is used for authentication. The access control was enhanced by providing three additional layers of protection, including host-based firewall rules, Amazon Security Group firewall rules and network access control rules. A separate audit server was defined within the cloud to store the audit information. The authors note that changes to the configuration have the potential to compromise the integrity of the system. This was prevented by restricting the ability to make changes to the Amazon Web Services (AWS) web client, which was set up with multifactor authentication. To ensure transmission security, all communications

were performed over a virtual private network (VPN) connection using industry standard encryption.

The result of the configuration is a system which the authors state meets HIPAA requirements. The authors present several configuration areas as targets for potential improvement, including issues with the way IP address ranges are assigned and the resulting trial and error required to set up outbound firewall rules. Auditing from the security group firewall and the access control rules are not logged to the auditing server, which prevents the information from being used to spot suspicious activity. The authors include a reminder that decommissioning the cloud environment will require the same data clearing actions as any other system when it is retired.

This source is relevant for this study because any attempt to replicate patient health information to the cloud must be viewed within the bounds of the HIPAA requirements that are clearly presented in the paper. For data not protected under HIPAA, there might be other regulations concerning the protection of the data such as FERPA for student information, and numerous rules concerning the disclosure of financial information. Any applicable regulatory requirements must be investigated prior to storing any protected information in the cloud; for this study, HIPAA provides the governing regulatory requirements.

Wood, T., Cecchet, E., Ramakrishnan, K. K., Shenoy, P. J., van der Merwe, J. E., & Venkataramani, A. (2010). Disaster recovery as a cloud service: Economic benefits & deployment challenges. *HotCloud*, *10*, 8-15. Retrieved from https://www.usenix.org/legacy/events/hotcloud10/tech/full_papers/Wood.pdf

Abstract – Many businesses rely on Disaster Recovery (DR) services to prevent either manmade or natural disasters from causing expensive service disruptions. Unfortunately, current DR services come either at very high cost, or with only weak guarantees about the amount of data lost or time required to restart operation after a failure. In this work, we argue that cloud computing platforms are well suited for offering DR as a service due to their pay-as-you-go pricing model that can lower costs, and their use of automated virtual platforms that can minimize the recovery time after a failure. To this end, we perform a pricing analysis to estimate the cost of running a public cloud based DR service and show significant cost reductions compared to using privately owned resources. Further, we explore what additional functionality must be exposed by current cloud platforms and describe what challenges remain in order to minimize cost, data loss, and recovery time in cloud based DR services.

Summary – The authors discuss the basics of disaster recovery and the key requirements of any DR solution, which includes the RPO, RTO, performance, consistency and geographic separation. The authors note that the geographic separation adds challenges due to latency and state that synchronous replication, where data must be committed to both replicas before continuing, allows the RPO to be zero, but is limited to distances within 10s of kilometers of the primary.

The authors focus the main portion of the article on a comparison of the costs associated with implementing the failover services in the cloud versus the cost of renting resources from a colocation facility. Cloud services support the concept of capacity on demand, where a server can be sized small as a target for replication, and then dynamically expanded when the server is functioning as the production server. The dynamic sizing is a benefit when evaluating the cost of cloud servers. The authors caution that the comparison is meant to be illustrative rather than definitive due to wide variations in the costs associated with cloud and colocation services.

The authors present a price analysis for two case studies: a multi-tier web-based database application and a large data warehouse application. The web-based database application, listed as a small database (30GB) and based on an assumed 99% uptime of the primary data center, results in a cloud cost of $1,562 per year, versus colocation costs of $10,373 per year. In the large (1 TB) data warehouse application, due to daily processing and storage requirements, the prices for cloud ($2,832 per year) and colocation ($3,186 per year) are comparable. The authors suggest alternatives for the data warehouse application, such as making a daily backup to the cloud instead of continuous updates, which has the potential to lower the cost while incurring a weaker RPO.

The authors highlight other challenges that require evaluations prior to finalizing the decision to use the cloud as a DR solution, including the possibility of oversubscription in the case of an outage that affects many organizations, such as a regional disaster. Other concerns include the operations required to clone to the cloud due to potential differences between the virtualization implementations on site and in the cloud, or failover to the cloud copy due to networking configuration changes which might be required. The

authors also identify security concerns with protected data being stored in the cloud due to the shared public nature of the cloud computing practices.

This source is relevant for this study because of the insight provided into the cost structures of using the cloud for different types of applications versus the cost of colocation and how the considerations it provides for the use of the cloud for DR activity.

Wood, T., Lagar-Cavilla, H. A., Ramakrishnan, K. K., Shenoy, P., & Van der Merwe, J. (2011). PipeCloud: Using causality to overcome speed-of-light delays in cloud-based disaster recovery. *Proceedings of the 2nd ACM Symposium on Cloud Computing* (p. 17). ACM. doi: 10.1145/2038916.2038933

Abstract – Disaster Recovery (DR) is a desirable feature for all enterprises, and a crucial one for many. However, adoption of DR remains limited due to the stark tradeoffs it imposes. To recover an application to the point of crash, one is limited by financial considerations, substantial application overhead, or minimal geographical separation between the primary and recovery sites. In this paper, we argue for cloud-based DR and pipelined synchronous replication as an antidote to these problems. Cloud hosting promises economies of scale and on-demand provisioning that are a perfect fit for the in-frequent yet urgent needs of DR. Pipelined synchrony addresses the impact of WAN replication latency on performance, by efficiently overlapping replication with application processing for multi-tier servers. By tracking the consequences of the disk modifications that are persisted to a recovery site all the way to client-directed messages, applications realize forward progress while retaining full consistency guarantees for client-visible state in the event of a disaster. PipeCloud, our prototype, is able to sustain these guarantees for multi-node servers composed of black-box VMs, with no need of

application modification, resulting in a perfect fit for the arbitrary nature of VM-based cloud hosting. We demonstrate disaster failover to the Amazon EC2 platform, and show that PipeCloud can increase throughput by an order of magnitude and reduce response times by more than half compared to synchronous replication, all while providing the same zero data loss consistency guarantees.

Summary – This article is focused on the use of replication of data between data centers as a primary method of providing a DR solution for critical applications. The authors state that cloud data centers provide a good option as targets for replication due to the resource on demand model, where the cost is based on the resources used, and the geographic separation that can be provided by cloud data centers. The authors caution that a major barrier is the fact that the distance between data centers increases the latency between computer systems. The increased latency forces a decision to be made between synchronous and asynchronous replication. Synchronous replication is implemented by ensuring that the write operations are completed at both the local and remote site before acknowledging the operation is complete. This approach increases the time to perform the write and decreases the performance of the application, but ensures that the systems are consistent. The author lists an example round trip time from Massachusetts to Virginia as 16 milliseconds, with the distance being 567 kilometers.

Asynchronous replication is implemented by acknowledging the write to the application when the local copy is committed, and queueing the remote write to be completed later. Asynchronous replication improves performance by eliminating the latency between data centers but leads to inconsistent data between data centers. The result is a RPO which is non-zero, meaning some data could be lost on failover to the remote data center. The

authors cite a test performed using a workload where 8 byte records were added into an existing database, with a 50 millisecond round trip time between data centers. Asynchronous replication was able to process approximately 2,300 requests per second, while synchronous replication was able to process approximately 100 requests per second.

The authors use this detailed analysis in testing a proposed application referred to as pipe-cloud. The proof of concept presented uses a web server as a front end application to access a database with both the primary and backup systems running on virtual servers. The use of virtual servers allows the pipe-cloud application to intercept IO operations to the local disks and replicate them to a remote disk located in the cloud without stalling for IO completion. To ensure the consistency from the user perspective, acknowledgement is not returned to the web server application until all remote IOs are completed. In experimental tests, the authors were able to improve performance while replicating over the longer distance. In repeating the previous test with the 8 byte records being added with the 50 millisecond delay, the use of the pipe-cloud application processed approximately 1,000 requests per second.

This source is relevant for this study because of the analysis the authors present, which demonstrates the performance impact of latency due to the distance between geographically separated data centers. The source also provides a potential application to reduce the effects of latency in geographically dispersed primary and backup data centers.

Conclusion

Organizations have become dependent on computer systems and experience major impacts to operations whenever an issue disrupts normal processing (Brotherton & Dietz, 2014). In the event of a disaster, critical operations must continue (Shepard, 2017). The ability to recover from disruptions in service is disaster recovery (Wood et al., 2010). Disaster recovery plans are blueprints for recovering from disruptions such as natural disasters with the goal of reducing data loss (Turnbull et al., 2013). Disaster recovery plans are of critical importance (Turnbull et al., 2013; Zhu et al., 2017), and typically provide one or more options for the recovery of computer services (Mohamed, 2014).

Some organizations are utilizing cloud services as part of their disaster recovery plans (Alhazmi & Malaiya, 2012; Bajpai, Rana & Maitrey, 2013). Cloud services offer the ability to geographically separate sites, which helps reduce the risk of regional disasters, but at the cost of transmission time between data centers (Alhazmi & Malaiya, 2013; Wood et al., 2011). Cloud services are not a complete solution for disaster recovery; they are a component in the DR plan and must include procedures for activating the cloud copy (Bahale & Gupta, 2014; Wood et al., 2010) and protect the data that is stored in the cloud (Regola & Chawla, 2013).

This annotated bibliography presents scholarly sources that discuss the impacts to organizations due to computer system downtimes, outline the options available for implementing recovery services and detail the implementation concerns when utilizing cloud resources for recovery services. While the primary problem of study is related to disaster recovery of healthcare systems, this study will also be of use to IT support staff for other mission critical systems.

Impact to Organizations Due to Computer Downtimes

Computer system downtimes can occur for a variety of reasons (Brotherton & Dietz, 2014). Brotherton and Dietz (2014) performed a case study of four events that had a large impact on computer systems, analyzing the impacts of the downtimes in terms of financial cost and impact to operations. The cases included Commerzbank activating their DR site due to the September 11 World Trade Center attack and Tulane University suffering an extended downtime due to Hurricane Katrina in 2005. One finding from the study relates to the cost of downtimes; Brotherton and Dietz (2014) note the cost can be high and provide as an example the 2003 blackout that affected the northeast United States, estimated to be $7-10 billion. In other cases the impact can be wide reaching; Brotherton and Dietz (2014) provide the example of a storage failure in the Commonwealth of Virginia in 2005 that took 483 of Virginia's servers offline for a period of more than a week.

The impact of Hurricane Katrina on Tulane University was a long duration recovery lasting multiple weeks (Brotherton & Dietz, 2014; Lawson, 2005). Lawson (2005) reported that while Tulane had a disaster recovery plan and the plan was being followed, the recovery plan was ineffective. Backups were produced and safely stored on the 14th floor and the systems were shut down, so there was no loss of data from within the data center, but the staff was not able to access the tapes and there was no recovery site funded. Tulane's vice president of information technology explained that this was a conscious decision to save money and accept the risk: "I presented the plan for off-site disaster recovery, it was for $300,000 a year or so. We decided that we could not ask the deans to pay for that because they were already upset about how much technology was costing them" (Lawson, 2005, p. 2).

Tulane was able to survive a long recovery period (Lawson, 2005), but computer services are more critical for other organizations. In the case of healthcare organizations, Shepard (2017) notes the criticality of electronic health record systems. “Dependence on EHRs has quickly become the norm for nurses and other clinicians. Emergencies and disaster events will disrupt the norm. It is important that nurse leaders prepare their teams to provide care without this tool, for the good of our patients” (Shepard, 2017, p. 189). The fact that EHR downtimes impact hospitals was further shown by Wang et al. (2016) in a study that attempted to determine the impact to patient care due to EHR downtimes. The study was based on five fully documented downtime events that occurred within a one-year period at a metropolitan teaching hospital. The duration of the events ranged from five minutes to five hours, with two of the incidents lasting more than four hours. While the study was inconclusive as to patient impact, the hospital was forced to trigger the contingency process for significant interruptions to system availability, which was to deliver urgent test requests to the laboratory by hand and communicate results to clinicians by telephone (Wang et al, 2016).

These events demonstrate that there are risks that can impact all types of organizations (Brotherton & Dietz, 2014; Wang et al., 2016). For organizations that have time-critical processes, business continuance is vital (Mohamed, 2014). In some cases, the impact of a disaster can be lessened with planning (Zhu et al., 2017). DeGaspari (2012) reported that an EF-5 tornado hit St. John’s Regional Medical Center in Joplin, Missouri in May 2011. The hospital had recently relocated the EHR system to a remote data center 250 miles away, which was outside the path of destruction. The old computer center did not survive the tornado. If the tornado had hit prior to the move, the hospital “would have lost all of the systems” (DeGaspari, 2012, p. 2). To provide further redundancy, St. John’s Regional Medical Center also has the

clinical data replicated to a second data center located approximately 50 miles from the primary site (DeGaspari, 2012).

Methods of Providing Business Continuance and Disaster Recovery Services

To design business continuance and disaster recovery plans, organizations must analyze the business requirements (Mohamed, 2014). The primary requirements are the recovery time objective (RTO), defined by Mohamed as "length of time a business can be without data processing availability" (2014, p. 61) and the recovery point objective (RPO), defined by Zhu et al. (2017) as the time between the latest backup and the time of the disaster, which defines the quantity of data lost.

The RTO and RPO are the critical metrics when designing the recovery process (Alcântara et al., 2017; Alhazmi & Malayia, 2012; Brotherton & Dietz, 2014; Mohamed, 2014; Wood et al., 2010; Zhu et al., 2017). Multiple authors described tiers of recovery services designed to meet the RTO and RPO requirements (Alcântara et al., 2017; Alhazmi & Malayia, 2012; Bajpai et al., 2013; Mohamed, 2014; Zhu et al., 2017). For an organization that can tolerate long recovery times and the loss of a significant amount of electronic data, Mohamed (2014) reports that an organization can meet the requirements if it "backs up and stores its data at an offsite storage facility" (p. 59). Relying on backup and restore activity has the downside that "the typical length of time for recovery is normally more than a week" (p. 59). To reduce the recovery time and reduce the data loss, Mohamed (2014) describes the use of electronic vaulting. "Electronic vaulting consists of electronically transmitting and creating backups at a secure facility, moving business-critical data offsite faster and more frequently than traditional data backup processes allow" (p. 59). A data center with functioning equipment is required for this

option (Mohamed, 2014). The cost of this option is higher, but with the advantage that "the typical length of time for recovery is normally about one day" (p. 59).

While Mohamed's (2014) option for electronic vaulting results in more frequent backups than traditional data backup processes, disasters will still result in the loss of recent data, as the backups are not continuous. Any data changes that occurred between the last data transfer and the time of the event will be lost (Alcântara et al., 2017; Mohamed, 2014). To eliminate the data loss, data replication between the data centers must be implemented (Alcântara et al., 2017; Alhazmi & Malayia, 2012; Bajpai et al., 2013; Brotherton & Dietz, 2014; Mohamed, 2014; Zhu et al., 2017). Zhu et al. (2017) state that this replication can occur at multiple layers within the application stack, including within the application, at the database layer, at the file system layer and at the storage layer.

Independent of the layer within the application stack, replication can be synchronous or asynchronous (Bajpai et al., 2013; Mohamed, 2014; Zhu et al., 2017). Synchronous replication is a technique where "each local I/O transaction shall be released after receiving the confirmation for completion of remote replication. Synchronous mirror enables remote copy to always match the replicated content requested by the local machine" (Bajpai et al., 2013). Asynchronous replication is implemented by acknowledging the write to the application when the local copy is committed, and queueing the remote write to be completed later (Wood et al., 2011).

Using synchronous replication ensures that the recovery point object will be zero (Alhazmi & Malayia, 2012; Bajpai et al., 2013; Wood et al. 2010). The tradeoff made with synchronous replication is a reduction in performance (Alcântara et al., 2017; Wood et al., 2010; Wood et al., 2011). The time required for the I/O request to be transmitted to the remote site, serviced and acknowledged is referred to as the latency or round-trip time (Wood et al., 2011).

Wood et al. (2011) measured a baseline round-trip time from a data center in Massachusetts to a data center in Virginia as 16 milliseconds, with a distance of 567 kilometers. In further application testing, Wood et al. (2011) measured the impact to throughput over a link between the data centers that experienced a 50 millisecond round-trip time using a test application. The test demonstrated that the application performance degraded from 2,300 requests per second using asynchronous replication to 100 requests per second using synchronous replication (Wood et al., 2011).

Replication that is performed in an asynchronous manner restores the application performance (Bajpai et al., 2013; Mohamed, 2014; Zhu et al., 2017). "Asynchronous replication improves performance since the primary site can proceed without waiting for the replication to complete" (Wood et al., 2011, p. 18), which "minimizes the impact on system performance, [and] ensures long transmission distance" (Bajpai et al., 2013, p. 169). Wood et al. (2011) caution that any disk writes made at the primary site after the last replicated snapshot will be lost in the event of a disaster. The data loss is however limited to the data that is in the process of being replicated, but has not been committed at the remote location (Alcântara et al., 2017; Mohamed, 2014).

The analysis of the RTO and RPO requirements, the selection of the recovery model to meet the RTO and RPO, and the potential decision of synchronous versus asynchronous replication are the major decisions that are used to build the disaster recovery plan (Alcântara et al., 2017; Alhazmi & Malayia, 2012; Bajpai et al., 2013; Mohamed, 2014; Zhu et al., 2017). The formulation of the plan starts with the determination of potential risks (Bajpai et al., 2013; Mohamed, 2014; Turnbull et al., 2013). Based on the risks and the recovery tier selected to meet the requirements, an organization can calculate the costs to determine whether the organization

can afford to implement the solution (Alhazmi & Malayia, 2012). Alhazmi and Malaiya (2012) provide the cost of a DR solution as the sum of the initial cost, ongoing cost, and expected annual cost of potential disasters. They define the ongoing cost as the sum of ongoing storage cost, data transfer cost, and processing cost (Alhazmi & Malaiya, 2012). They calculate the annual disaster cost as the total expected cost of disaster recoveries plus the cost of unrecoverable disasters (Alhazmi & Malaiya, 2012). When faced with these costs, Tulane University decided not to fund a full disaster recovery plan (Lawson, 2005).

A complete recovery plan is based on the design of the recovery system. Turnbull et al. (2013) state that a recovery plan should have the following sections:

- IT disaster identification and notification – Defines what is a disaster and the communication techniques during a disaster.
- Preparing organizational members – Defines the selection and training of disaster recovery team members.
- IT services analysis – Classifies and prioritizes the services requiring recovery.
- Recovery process – Documents procedures for restoring the systems to operational.
- Backup procedures – Prepares techniques that ensure a viable replica of the critical data.
- Offsite storage – Provides for storage of replicas in a location other than the data center.
- Maintenance – Defines the procedures to test and sustain the disaster recovery plan.

Best Practices for Replication to the Cloud

The use of cloud services as a target for replication to meet DR requirements was suggested by multiple authors (Alhazmi & Malaiya, 2012; Bajpai et al., 2013). The primary

benefits presented with the use of cloud services are reduced costs (Alcântara et al., 2017; Alhazmi & Malayia, 2012; Bahale & Gupta, 2014; Wood et al., 2010) and the ability to provide geographic separation from the primary data center (Alcântara et al., 2017; Alhazmi & Malayia, 2013; Bahale & Gupta, 2014; Wood et al., 2010).

Wood et al. (2010) present a cost comparison of resources rented from a colocation facility versus an infrastructure as a service (IaaS) environment set up in the cloud for use as a disaster recovery backup site. An IaaS solution includes storage and compute resources (Wood et al., 2010). They note that a large benefit of using the IaaS server is that the server can be configured small to support replication activity, and resized when responding to a disaster to reduce costs. Wood et al. (2010) used this design in their analysis; the cost comparison demonstrated that for some application environments, the cost of colocation versus cloud might be comparable, but for other applications, the cloud implementation could be 80% less expensive. Factors that caused the cost of a collocated facility to be comparable to a cloud alternative included higher costs for a cloud server that will support the high I/O load and the cloud storage cost associated with the large database size (Wood et al., 2010).

Wood et al. (2011) present considerations with synchronous versus asynchronous replication over long distances and note that there are limited sites offering cloud services. The detailed analysis provided by Wood et al. (2011) demonstrates that there is a large impact to performance when replicating over long distances, and the authors further note that "Even for relatively short distances, e.g., from Massachusetts to the [Amazon] EC2 data center in Virginia, the response time degrades noticeably" (p. 19). Wood et al. (2011) therefore conclude that when geographic separation between primary and backup data centers is necessary to reduce the likelihood of a regional disaster, an organization will probably need to use asynchronous

replication. Wood et al. (2011) present a prototype that overlaps replication with application processing on multi-tier servers, referred to as pipelined synchronicity; the results of their testing provided RPO equivalent to synchronous replication with the performance closer to asynchronous replication.

Alcântara et al. (2017) introduce an alternate technique using asynchronous replication to the cloud, with the target being simple storage service (S3). Simple cloud storage services provide the ability for customers to upload and store information in the cloud, but do not include compute services (Alcântara et al., 2017). The benefits include the ability to locate backup resources in a geographically separated location so that the same disaster does not affect the primary and the backup infrastructures (Alcântara et al., 2017). The analysis indicates that this implementation is a fraction of the cost of a full virtual system in the cloud (Alcântara et al., 2017). The authors do caution that in case of a disaster, recovery requires the restoration of the data stored within the storage system to a system capable of hosting the database application, resulting in a long RTO but a small RPO (Alcântara et al., 2017).

Any use of cloud services, whether IaaS or S3, requires the maintenance of data security (Regola & Chawla, 2013). Regola and Chawla (2013) provide a detailed analysis of security concerns that organizations address when storing protected health information in the cloud. Regola and Chawla (2013) used HIPAA requirements to demonstrate that when properly implemented, patient health information can be stored within an Amazon virtual private cloud environment. They focused on the HIPAA security rule, which covers the security of the server and depends on "the physical security of the server and network, the operators and users of the server, and the configuration and management of the applications, operating system, and network" (Regola & Chawla, 2013, p. 3). Regola and Chawla (2013) used an IaaS

implementation for the study, which allows "customers to run an operating system of their choice and maintain full control of their operating system and network environment" (p. 5).

Regola and Chawla (2013) assert that patient health information can be protected in the cloud when implemented correctly. The first requirement presented is that the organization must implement administrative safeguards, including having a documented security policy, defining the roles that are allowed to access the patient data, providing authorization and supervision to anyone who has access to protected data, and ensuring that users are trained on protecting the data (Regola & Chawla, 2013). Physical access to the servers is covered by audits, certification, and ongoing compliance; "Amazon EC2's service has completed a SAS 70 Type II audit, obtained ISO 27001 certification, and PCI level DSS validation as a level 1 service provider" (p. 6). Other recommendations include correctly using secure shell and virtual private network communication techniques and implementing multiple levels of firewalls to protect the data from outside attack (Regola & Chawla, 2013). Regola and Chawla (2013) conclude "we suspect that the debate on these security issues is just starting and will continue" (p. 10) but also note "we believe that our prototype can be implemented in a HIPAA compliant manner" (p. 10). The analysis can be used as a starting point to analyze other protected information and offerings from other cloud providers.

Despite promising research, there are other concerns that organizations considering implementing a DR solution using cloud services must address. Bahale and Gupta (2014) present a concern over activating the cloud replica for service as the primary replica, asserting that "public Internet facing applications would require additional forms of network reconfiguration through either modifying DNS or updating routes to redirect traffic to the failover site" (p. 399). From an operational perspective, many organizations support server virtualization within their

data centers, and the virtualization within the cloud might introduce a new virtualization stack that will increase operational costs associated with maintaining the on-premise copy and the cloud-based copy (Bahale & Gupta, 2014).

Alhazmi and Malaiya (2013) express concerns that "It is possible that a cloud site may serve as a disaster recovery site for a number of customers from the same region. Thus, it may be overwhelmed if it encounters a sudden high demand from many customers" (p. 5). They also state that there are potential cloud outages and undiscovered security threats that are only applicable to cloud based servers (Alhazmi & Malaiya, 2013).

Final Thoughts

While most organizations depend on IT services, in the case of healthcare organizations, "Dependence on EHRs has quickly become the norm for nurses and other clinicians" (Shepard, 2017, p. 189). Any issues that prevent access to an organization's data have the potential for major impacts (Alhazmi & Malaiya, 2012). Independent of DR planning, some risks cannot be avoided, as demonstrated by the references to hurricanes, tornadoes and terrorist attacks that have struck and negatively impacted organizational data centers (Brotherton & Dietz, 2014).

To be able to recover from disasters, organizations require disaster recovery procedures (Alhazmi & Malaiya, 2012; Brotherton & Dietz, 2014; Mohamed, 2014; Wood et al., 2010) For organizations that can only tolerate short time periods without data access and cannot risk the loss of any data, the recovery solution is a replication between data centers (Alcântara et al., 2017; Alhazmi & Malayia, 2012; Bajpai et al., 2013; Brotherton & Dietz, 2014; Mohamed, 2014; Zhu et al., 2017). One option for a replication site is to utilize cloud services as a secondary site (Alcântara et al., 2017; Alhazmi & Malayia, 2012; Alhazmi & Malayia, 2013; Bahale & Gupta, 2014; Bajpai et al., 2013; Wood et al., 2010; Wood et al. 2011). An

organization considering the use of cloud services as a backup site will have to analyze the performance impact of replicating over a long distance (Bahale & Gupta, 2014; Wood et al., 2010; Wood et al. 2011), security concerns with data being stored outside of their physical control (Alhazmi & Malayia, 2013; Bahale & Gupta, 2014; Bajpai et al., 2013; Regola & Chawla, 2013; Wood et al., 2010; Zhu et al. 2017), and implementation and activation challenges when formulating their plans (Alhazmi & Malayia, 2013; Bahale & Gupta, 2014), but the option of replicating to the cloud offers promise in providing geographic separation between data centers to mitigate the effects of regional disasters (Bahale & Gupta, 2014; Wood et al., 2010; Wood et al. 2011) while providing cost effective computing resources (Alcântara et al., 2017; Alhazmi & Malayia, 2012; Alhazmi & Malayia, 2013; Bahale & Gupta, 2014; Wood et al., 2010; Wood et al. 2011) and is therefore worth investigating.

References

About OHSU. (n.d.). Retrieved October 29, 2018, from https://www.ohsu.edu/about/ohsu-vision-mission-and-values

Alcântara, J., Oliveira, T., & Bessani, A. (2017). Ginja: One-dollar cloud-based disaster recovery for databases. *Proceedings of the 18th ACM/IFIP/USENIX Middleware Conference* (pp. 248-260). ACM. doi: 10.1145/3135974.3135985

Alhazmi, O. H., & Malaiya, Y. K. (2012). Assessing disaster recovery alternatives: On-site, colocation or cloud. *2012 IEEE 23rd International Symposium on Software Reliability Engineering Workshops (ISSREW),* (pp. 19-20). IEEE. doi: 10.1109/ISSREW.2012.20

Alhazmi, O. H., & Malaiya, Y. K. (2013). Evaluating disaster recovery plans using the cloud. *2013 Proceedings-Annual Reliability and Maintainability Symposium (RAMS),* (pp. 1-6). IEEE. doi: 10.1109/ISSREW.2012.20

Alizadeh, M., Greenberg, A., Maltz, D. A., Padhye, J., Patel, P., Prabhakar, B., ... & Sridharan, M. (2011). Data center tcp (dctcp). *ACM SIGCOMM Computer Communication Review, 41*(4), 63-74. doi: 10.1145/1851182.1851192

Bahale, M. S. V., & Gupta, S. (2014). Virtualizing disaster recovery management based on cloud computing. *International Journal of Research in Advent Technology*, *2*(5), 397-401. Retrieved from http://www.ijrat.org/downloads/may-2014/paper%20id-25201464.pdf

Bajpai, A., Rana, P., & Maitrey, S. (2013). Remote mirroring: A disaster recovery technique in cloud computing. *International Journal of Advance Research in Science and Engineering*, *2*(8), 3. Retrieved from https://pdfs.semanticscholar.org/9c87/f581d4e7c39a9db9867ec90aa8fbda43de80.pdf

Brotherton, H., & Dietz, J. E. (2014). Data center business continuity best practice. *2014 11th International Conference on Information Technology: New Generations (ITNG)* (pp. 496-501). IEEE. doi: 10.1109/ITNG.2014.8

Cebul, R. D., Love, T. E., Jain, A. K., & Hebert, C. J. (2011). Electronic health records and quality of diabetes care. *New England Journal of Medicine, 365*(9), 825-833. doi: 10.1056/NEJMsa1102519

Center for Public Issues Education. University of Florida. (n.d.). Evaluating information sources. Retrieved from https://canvas.uoregon.edu/courses/120122/files/5706985/download?wrap=1

Chervenak, A., Vellanki, V., & Kurmas, Z. (1998). Protecting file systems: A survey of backup techniques. *Joint NASA and IEEE Mass Storage Conference* (Vol. 99). Retrieved from http://www.storageconference.us/1998/papers/a1-2-CHERVE.pdf

Danel, R., & Neustupa, Z. (2011). Data continuity solution in fault-tolerant information systems. *2011 12th International Carpathian Control Conference (ICCC)* (pp. 66-69). IEEE. doi: 10.1109/CarpathianCC.2011.5945817

Degaspari, J. (2012). When disaster strikes: How technology drives better preparation. *Healthcare informatics*, *29*(9), 22-23. Retrieved from https://www.healthcare-informatics.com/article/when-disaster-strikes-how-technology-drives-better-preparation

Gans, D., Kralewski, J., Hammons, T., & Dowd, B. (2005). Medical groups' adoption of electronic health records and information systems. *Health Affairs, 24*(5), 1323-1333. doi: 10.1377/hlthaff.24.5.1323

Jadeja, Y., & Modi, K. (2012). Cloud computing-concepts, architecture and challenges. *2012 International Conference on Computing, Electronics and Electrical Technologies (ICCEET)* (pp. 877-880). IEEE. doi: 10.1109/ICCEET.2012.6203873

Kossman, S. P., & Scheidenhelm, S. L. (2008). Nurses' perceptions of the impact of electronic health records on work and patient outcomes. *CIN: Computers, Informatics, Nursing*, *26*(2), 69-77. doi: 10.1097/01.NCN.0000304775.40531.67

Lawson, J. (2005). A look back at a disaster plan: What went wrong -- and right. *Chronicle of Higher Education*, B20–B22. Retrieved from https://www.chronicle.com/article/A-Look-Back-at-a-Disaster/10664

Maurer, F., & Lechner, U. (2014). From disaster response planning to e-Resilience: A literature review. *Bled eConference* (p. 32). Retrieved from https://domino.fov.uni-mb.si/proceedings.nsf/Proceedings/35921A21AFEC35C9C1257CEE004655D9/$File/04_Maurer_Lechner.pdf

Mohamed, H. A. R. (2014). A proposed model for IT disaster recovery plan. *International Journal of Modern Education and Computer Science*, *6*(4), 57. doi: 10.5815/ijmecs.2014.04.08

Peixoto, H., Duarte, J., Abelha, A., Santos, M., & Machado, J. (2012). ScheduleIT – Open-source preventive actions management platform in healthcare information systems. *Procedia Technology, 5*, 734-742. doi:10.1016/j.protcy.2012.09.081

Percent of Hospitals, By Type, that Possess Certified Health IT. (2017). [Graph illustration of percentage of non-federal acute care hospitals, by type, that possess certified electronic health record technology]. *Certified Health IT from the ONC/American Hospital Association (AHA), AHA Annual Survey Information Technology Supplement, 2015.* Retrieved from https://dashboard.healthit.gov/quickstats/pages/certified-electronic-health-record-technology-in-hospitals.php

Qian, L., Luo, Z., Du, Y., & Guo, L. (2009). Cloud computing: An overview. *IEEE International Conference on Cloud Computing* (pp. 626-631). Springer, Berlin, Heidelberg. doi: 10.1007/978-3-642-10665-1_63

Regola, N., & Chawla, N. V. (2013). Storing and using health data in a virtual private cloud. *Journal of Medical Internet Research, 15*(3), 1-12. doi: 10.2196/jmir.2076

Shepard, A. (2017). Disaster recovery and the electronic health record. *Nursing Administration Quarterly, 41*(2), 187-189. doi: 10.1097/NAQ.0000000000000213

Toomey, D. R., Allen, R. M., Barclay, A. H., Bell, S. W., Bromirski, P. D., Carlson, R. L., ... & Forsyth, D. W. (2014). The Cascadia Initiative: A sea change in seismological studies of subduction zones. *Oceanography, 27*(2), 138-150. doi: 10.5670/oceanog.2014.49

Turnbull, L., Ochieng, H., Kadlec, C., & Shropshire, J. (2013). Improving service continuity: IT disaster prevention and mitigation for data centers. *Proceedings of the 2nd Annual Conference on Research in Information Technology* (pp. 43-46). ACM. doi: 10.1145/2512209.2512213

Vayena, E., Haeusermann, T., Adjekum, A., & Blasimme, A. (2018). Digital health: meeting the ethical and policy challenges. *Swiss Medical Weekly, 148*, w14571. doi: 10.3929/ethz-b-000239873

Wang, Y., Coiera, E., Gallego, B., Concha, O. P., Ong, M. S., Tsafnat, G., ... & Magrabi, F. (2016). Measuring the effects of computer downtime on hospital pathology processes. *Journal of biomedical informatics*, *59*, 308-315. doi: 10.1016/j.jbi.2015.12.016

Wood, T., Cecchet, E., Ramakrishnan, K. K., Shenoy, P. J., van der Merwe, J. E., & Venkataramani, A. (2010). Disaster recovery as a cloud service: Economic benefits & deployment challenges. *HotCloud*, *10*, 8-15. Retrieved from https://www.usenix.org/legacy/events/hotcloud10/tech/full_papers/Wood.pdf

Wood, T., Lagar-Cavilla, H. A., Ramakrishnan, K. K., Shenoy, P., & Van der Merwe, J. (2011). PipeCloud: Using causality to overcome speed-of-light delays in cloud-based disaster recovery. *Proceedings of the 2nd ACM Symposium on Cloud Computing* (p. 17). ACM. doi: 10.1145/2038916.2038933

Zhong, S., Clark, M., Hou, X. Y., Zang, Y. L., & Fitzgerald, G. (2014). Development of hospital disaster resilience: conceptual framework and potential measurement. *Emerg Med J, 31*(11), 930-938. doi: 10.1136/emermed-2012-202282

Zhu, T., Xie, Y., Song, Y., Zhang, W., Zhang, K., & Gao, F. (2017). IT disaster tolerance and application classification for data centers. *Procedia Computer Science*, *107*, 341-346. doi: 10.1016/j.procs.2017.03.115