UNIVERSITY OF OREGON
APPLIED INFORMATION MANAGEMENT

Presented to the Interdisciplinary
Studies Program:
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

Selecting WAN Capable Replication Technologies for Disaster Recovery Planning in Mid-Sized Organizations: A Guide for IS Managers

CAPSTONE REPORT

Jacob Klearman
Network Administrator
WestStar Bank

University of Oregon
Applied Information
Management
Program

March 2006

722 SW Second Avenue
Suite 230
Portland, OR 97204
(800) 824-2714

Approved by

____

Dr. Linda F. Ettinger
Academic Director, AIM Program

Abstract

for

Selecting WAN Capable Replication Technologies for
Disaster Recovery Planning in Mid-Sized Organizations:
A Guide for IS Managers

Large scale disasters highlight the need to protect and quickly restore mission critical data. A conceptual analysis of articles published from 1999 to 2006 forms the basis for a decision support tool, designed for IS Managers of medium sized enterprises, who need to understand asynchronous WAN replication options. The tool, to be used during business continuity planning, explores the pros, cons, functionality and cost of replication technologies including: file, snapshot, CDP, Block and Byte replication.

Table of Contents

List of Figures and Tables

CHAPTER I – PURPOSE OF STUDY

Brief Purpose

The purpose of this study is to examine selected types of disaster recovery data replication technology suitable for medium sized enterprise (Ayyagari, Beck, Demirgüç-Kunt, 2003). A medium sized enterprise is defined in this study as one with more than 50 but fewer than 250 employees utilizing between five and twenty (Marks, 9/2005) Microsoft Windows based servers in a Wide Area Network (Stalling, 2001) distributed architecture environment (Toigo, 2003).

This study fits in the larger field of business continuity planning (McManus & Car, 2001). The goal of business continuity planning is to maintain the delivery of core business processes after a disaster (Doughty, 2001) through activities designed to prevent and minimize the negative impact of unplanned interruption events (Toigo, 2003).

The intended audience is information system (IS) managers of medium sized enterprises (Ayyagari, Beck, Demirgüç-Kunt, 2003 & Howard, 2005) in the process of updating or creating a business continuity plan (ED. Hiles & Barnes, 1999) or a disaster recovery plan (ED. Hiles & Barnes, 1999). The study is designed for IS managers who already have a working automated tape back up solution in place (Baltazar, 2005) and are exploring options to minimize the cost of down time (Krischer, 2005) due to the unavailability or failure (Toigo, 2003) of mission critical (Toigo, 2003) file servers through a remote site recovery process (ED. Hiles & Barnes, 1999).

A cross section of types of disaster recovery technologies that are capable of WAN replication and are marketed to medium sized enterprises is pre-selected for content analysis

(CSU Writing Center, 2005). The overarching features of the different types of asynchronous replication technology are examined, specifically, the pros, cons, functionality, and cost (Toigo, 2003). The selected types of replication technology reviewed include: snapshot, file replication, block replication, and continuous data protection (Marks, 5/1/2005).

The method of study is a literature review (Leedy & Ormrod, 2001). Selected literature published between January 1999 and September 2005 in the larger context of disaster recovery, business continuity planning, and network and server architecture are analyzed using the conceptual analysis method (Leedy & Ormrod, 2001; CSU Writing Center, 2005)

The outcome of the study is presented in two forms. First, a summary table (see Table 1: Comparative View of WAN Capable Replication Technologies) provides a report of feature sets and attributes of selected types of data replication technologies. The table is designed to help the IS manager weigh the pros, cons, functionality, and costs of different replication technologies in order to choose the one that will best meet their organization’s unique business needs. Then a discussion of this table is provided as an example of how an IS manager might use the table, when compared to a general set of business continuity planning goals, with a focus on activities designed to prevent and minimize the negative impact of unplanned interruption events (FCA, 2005, Toigo, 2003, Daughty, 2001).

Full Purpose

There are many potential disasters that can threaten or cause catastrophic damage to an enterprise. Examples are system failure, incorrect software operation, human error (Chevance, 2005), external man-made events, and internal intentional events (Hiles & Barnes, 1999). Business continuity planning (BCP) and disaster recovery planning address strategies and actions to ensure operation of all critical business processes and activities after a disaster, such as human resources, external dependencies, merchandise replenishment, and key infrastructure (Vancoppenolle, 1999). Assuring data protection and its availability for access following a disaster is of central concern in business continuity planning and disaster recovery planning because without access to critical business data, recovery is not possible (Toigo, 2003).

There is no one business continuity or disaster recovery strategy that would work for every enterprise (Krisher, 2005). The level of data replication necessary for any specific enterprise depends heavily on their recovery point objective, their recovery time objective (Chevance, 2005), and the estimated cost of down-time (Krischer, 2005).

Backup and restore operations are a subset of the larger business continuity and disaster recovery planning (Chevance, 2005). One way to ensure data protection is to replicate mission critical data to an alternate site (Drew, 2005). According to Chevance (2005), the Gartner Group distinguishes five levels of data replication:

- Level 0. No recovery
- Level 1. Backups without verification procedures
- Level 2. Backup of the complete contents of storage at arbitrary times, with a check of the backup contents
- Level 3. Systematic backups of the complete contents of storage, with a check of the backup contents

- Level 4. Systematic backups of only the data managed by the critical applications, backup up only the modified data, with a check of backup contents
- Level 5. Systematic updates of data modified by the application on the backup system as the application executes, with a check of the backup contents.

Tape backup has been the standard for backup since tape was replaced as primary storage (Toigo, 2003). According to Toigo, tape backup solutions are relatively inexpensive, mature, and a robust technology (2003). Yet there are many problems with this standard. Problems include: how to guarantee off site rotation (Connor, 02/2005), a long verification process, and a slow restoration process (Macvittie, 2005).

Since 2004, many data dependant enterprises are moving away from the typical disk-to-tape rotation to a disk-to-disk-to-tape rotation (Kovar, 2005). This move allows for faster backups (Hamblen, 2003), faster restores, multiple simultaneous backups (Marks, 9/2005), and the centralization of backups from multiple servers to one location.

Disk-based backups are, in fact, becoming the new standard for backup (Connor, 6/2005). The Gartner Group estimates that by 2008, 80% of recoveries will be from disk-based backups (Connor, 6/2005). Yet, disk-based backups still have many limitations. Most notably, they do not address geographical redundancy (Chevance, 2005). The replication of mission critical data to a off site location in a close to real-time manner, in addition to a tape backup, addresses many of the limitations associated with disk-to-tape or disk-to-disk-to-tape backup and restore operations (Connor, 02/2005).

Long distance replication technologies are not new, and have been tested and used by large institutions since 2001 (Hamblin, 2003). But until recently, the products capable of off site backups over a WAN were priced out of reach of most medium sized enterprises. There are now many replication products and options available (Connor, 2004). While these products and options address close to real-time mirroring of critical data over a WAN, each available solution has different strengths and weaknesses (Krischer, 2005).

This study is conducted as a literature review. Literature in the general areas of business continuity planning, disaster recovery planning, and network architecture is collected. Literature review, according to Leedy and Ormrod (2001), describes previous research findings and can be used to identify general themes that run throughout the selected literature. It is most useful as a way to both analyze and synthesize existing information. In this case, selected literature is examined that addresses types of replication-based disaster recovery technologies suitable for medium sized enterprises exploring options on how to automatically replicate data from their Microsoft Windows based file servers to an alternate site.

To assist IS managers in finding replication technology that is a “best-fit” for their enterprise, this study uses a conceptual analysis process (CSU Writing Lab, 2005). The goal of this content analysis is to add clarity to the complex topic of BCP/disaster recovery planning by examining how the use of specific data replication technologies may lessen the negative impact of unplanned outage events by having a readily accessible, close to real-time copy of the data at a secondary location. Conceptual analysis allows the researcher to identify patterns and themes in the selected text (Leedy & Ormrod, 2001) and clarify a concept by describing the essential or generic meaning of a concept (McMillen & Schumacker, 1993).

The content analysis process involves searching for instances of a pre-determined set of concepts; including snapshot, file replication, block replication, and CDP within selected texts. The results of the content analysis are developed into a final outcome of this study and summarized into a single table (see Table 1: Comparative View of WAN Capable Replication Technologies). Table 1 serves as a decision support tool that enables comparison of the pros, cons, costs and functionality of selected WAN based data replication technology appropriate for medium sized enterprises. This tool can be used to help guide the IS managers in their decision process while choosing the type of data replication technology best suited for their unique business needs when exploring options to enhance their disaster recovery preparedness. Table 1 is designed to provide a comprehensive overview of each category presented in the results, seen from the perspective of business continuity planning (Hiles & Barnes, 1999, Doughty, 2001, Toigo, 2003,). Then an example of how this tool might be used by IS managers is provided as the Conclusions section of this paper, in the form of a discussion of each of the key concepts. The discussion is framed in terms of three key business continuity planning goals including to (1) minimize disruption of services to the enterprise and its customers, (2) ensure continuity of critical operations and (3) limit financial loss (FCA, 2005, Toigo, 2003, Daughty, 2001). BCP/disaster recovery provides the framing for this discussion because the main goal of BCP is to limit exposure to yourself and your customers from loss due to a disaster. Currently, data replication provides the fastest path to a copy of the data that was lost.

The intended audience for this study is IS managers who understand that the effective storage and accessibility of mission critical data is of major concern. These are IS managers who have already implemented steps to protect the data storage environment including but not limited to: environmental controls, physical access security, quality server hardware, fault tolerant disk arrays, antivirus protection, and scheduled operating system (OS) updates (Toigo, 2003). They know what data is stored, where it is stored, and the relative importance of different data to support their enterprise’s business goals and processes. They have a working tape backup procedure in place but are experiencing issues with getting the tape backup offsite to a secure location, or want to decrease the time it take to restore normal operations after an unplanned event or disaster. Most importantly, they realize that simply mirroring data to another location does not constitute a disaster recovery or business continuity plan but rather that mirroring data over a WAN to a second location is just one part of a much larger project to ensure business continuity following a disaster.

Significance of the Study

It is estimated that 50 percent of businesses that experience major data failure lasting for more than 24 hours go out of business within 2 years (Klein, 2005). As businesses recognize their dependency on their storage media, the market for backup solutions is continuing to grow (Connor, 06/2005). The market for data protection software including backup, archive, hierarchical storage management, snapshot backup, and replication is expected to be $8.1 billon by 2006 (Connor, 2004). An estimated 35 million terabytes of data will need to be restored by 2009 (Connor, 06/2005).

The large scale disasters of the past years, including hurricanes Katrina and Rita and terrorist attacks in the US and abroad, have highlighted the need for enterprises to take appropriate steps to protect and be able to quickly restore mission critical data. An estimated 40% of Fortune 1000 companies are not prepared for a regional disaster. Small and medium sized businesses are even less prepared (Mearian, 2005). Many enterprises were caught off guard by the scale of the recent events, and months after are still struggling to resume operations (Britt, 2005).

The main goals of business continuity planning are to (1) minimize disruption of services to the enterprise and its customers, (2) ensure continuity of critical operations and (3) limit financial loss (FCA, 2005, Toigo, 2003, Daughty, 2001). A business continuity plan attempts to addresses all aspects of returning an enterprise to an operational point within a timely manner after a crises event. Critical components of a BCP include: personnel, technology, data center alternatives, recovery facilities, geographic diversity, file backup, software backup, off-site storage, communication, prevention, and testing (FFEIC, 2005). Disaster recovery planning is one part of business continuity planning that focuses on the technological aspect of enterprise operations (Daughty, 2001). From an IS prospective shortening the time to data after an unplanned crises event is the primary mission of disaster recovery planning (Toigo, 2003).

Currently, tape backup is the primary mechanism for data protection (Toigo, 2003). Tape backup is prone to hardware complications and human error and provide no real time protection (Klein, 2005). Even when backups are done, they are often not taken off site (Hiles & Barnes, 1999). If tapes are successfully stored off site, there can be problems accessing the tapes in an emergency (Toigo, 2003). Many companies were caught off guard after the 9/11 attacks when their backup media could not be retrieved because of air travel restrictions (Toigo, 2003).

Tape backup continues to be the primary means of data protection for all but the most sophisticated or highly regulated enterprises because of the substantial costs associated with mirroring data (Toigo, 2003). But with the recent advances in technology, it is no longer a question of if the technology exists that can replicate data to a secondary location (Pratt, 2005) but rather what type of technology is appropriate for an enterprises.

This study is not intended to create new knowledge, but rather to present existing information in a format which allows IS managers to compare the pros, cons, functionality, and cost of selected data replication technologies enabling them to match their enterprises recovery point objective and recovery time objective and budget (Krischer, 2005) to specific replication technology.

Limitations to the Research

The resources collected to support this study are primarily published between 1999 and 2005, i.e., the most recent five years. This period covers the rapid changes in disaster recovery technologies that have lead to the mirroring of mission critical data over a WAN in close to real-time manner to an off-site location. This study reviews only types of replication technology that moves data over a WAN link from one location to another, concentrating specifically on technology capable of mirroring centralized file data to a secondary location over a shared WAN connection. It is assumed that medium sized enterprises have WAN connections near, at, or above T1 speeds. The relatively high speed WAN connection allows this researcher to include replication technologies that offer a close to real time asynchronous backup. WAN speeds must be sufficient to allow for the extra data traffic over the WAN without interfering with normal business activity.

While there are many data mirroring technologies available today, only a few target medium sized enterprises and address mirroring data from Microsoft Windows based file servers to an alternate site over a WAN. This study focuses on data mirroring products marketed to medium sized enterprises, excluding those intended to address the needs of large organizations utilizing storage area networks, or multiple locations connected by high speed fiber, or small businesses that do not need, or can not afford, this level of data protection. Small businesses are excluded because of the relative simplicity in which a replication solution can be implemented. For small businesses, the question is not what types of technology are available, but rather can a small business afford to implement the solution. Medium sized enterprises with mission critical data running on multiple servers have reached a level of complexity requiring extra care. The replication solutions targeted for small businesses can not support the complexity or become cost prohibitive as the amount of data to be backed up grows. The customized solutions targeted to large enterprises, though meeting all of their business requirements, is priced out of reach.

The focus of this study is on types of replication technology appropriate for medium sized enterprises working in a centralized data processing architecture that are exploring options to limit down time and get a readily accessible copy of file data to an alternate location. The off site data is created in order to allow an enterprise to recover mission critical data in a time period measured in hours rather than days

Some existing replication technologies appropriate for medium sized businesses are not explored in this paper. Examples include virtual tape libraries and devices capable of replicating data over a LAN. Though virtual tape libraries meet the requirement of creating an off site copy of the data, they do not meet the requirement of being centralized. That is, with virtual tape libraries, technical staff would need to be on hand at the secondary location to swap tapes and troubleshoot issues. Additionally, virtual tape libraries are still vulnerable to the same problems as the tape devices located at the main site. Other replication technologies, including LAN only backup options, are also not discussed. Though a readily accessible replicated set of data is created and could be used for some types of disasters, the secondary set of data would be destroyed along with the primary data set in the event of a crisis event that destroyed the main processing center.

Additionally, best practices for on-site hardware redundancy, environmental concerns, or the backup of data distributed on workstations and remote servers, or high-availability options such as server clustering or database replication are outside of the scope and are not discussed. Environmental precautions and database replication is an important part of a complete backup solution and are explored in depth in a business continuity planning process. Some of the technologies explored in this paper are capable of replicating data stored in databases, but because of the need for special considerations for different types of databases, this topic is not a focus of this study.

This study is directed at IS managers of medium sized enterprises, who already have a working tape backup solution in place and are searching for ways to minimize the impact of a disaster by implementing close to real-time (Toigo, 2003) mirroring (FFEIC handbook, 2003) of mission critical data over a WAN (Stallings, 2001) to an alternate site (Laufman, Neaga, Winters, 1997) through the use of currently available software or application based products that cost less than ten thousand dollars per server to implement.

Due to the recent rapid technological changes and nature of this topic, some of the literature is retrieved from professional, regulatory, association and web sites. Full text database searches, available through the University of Oregon WORLDCAT system and Eagle Public Libraries, are also utilized as well as recently published books and e-books.

Databases for searching are selected based on the inclusion of full-text articles and their timeliness and inclusion of relevant data. Databases accessed include: Academic Search Premier, Expanded Academic ASAP, Lexus Nexus, Business Source Premier, General BusinessFile ASAP, and MasterFILE Premier. One problem encountered is that much of the academic research available concentrates on highly technical data inappropriate for inclusion in this study such as the algorithms used to replicate data over a WAN. Collected data include text from regulatory agencies, articles from trade and business magazines, reviews of specific replication software, and white papers as well as business continuity planning and disaster recovery planning books.

The method of study is a literature review (Leedy & Ormrod, 2001). Selected in the larger context of disaster recovery, materials addressing business continuity planning, and network and server architecture are analyzed using a qualitative design with a conceptual analysis methodology (Leedy & Ormrod, 2001; CSU Writing Center, 2004). A literature review is the most suitable method of this study given the limited time frame in which the research must be completed. Given the existence of a pre-defined set of terminology to use during coding, and the goal of clarifying a concept, conceptual analysis is a suitable strategy for data analysis.

Problem Statement

The increasing importance of mission critical data to enterprises, the changing definition of acceptable down time, and the technological enhancements over recent years can be demonstrated through a series of specific quotes. Authors of the book titled Fire in the Computer Room, What Now?, explain the importance of disaster recovery (Nega, N., Winters, B., Luafman, P., 1997). They write:

> As organizations have become more and more dependent on data processing (DP) to conduct their business and to stay competitive, the availability of the processing facilities has become crucial. Today, most businesses require a high, if not continuous, level of DP availability.
>
> As a consequence, most businesses would find it extremely difficult to function without data processing. Manual procedures, if they exist at all, would only be practical for a short period of time. A lengthy outage in a computing environment can result in significant financial losses, especially if management liability is involved. More importantly, one can lose customer credibility and subsequent market share. In some cases, such losses could lead to the total failure of the business (p.4).

In an article entitled “Tragedy Will Sell Companies on Disaster Recovery” (2001**),** published shortly after the World Trade Center was demolished in a terrorist attack, Alterio adds additional insight about the changing expectations of disaster recovery:

> Today there's a PC on every desk that's linked to a server through a network. The focus has shifted from recovering a single data center to resurrecting an entire business. Industry experts call this "business continuity" because nearly every worker needs access to technology to continue to function (http://www.thejournalnews.com/newsroom/091601/16disasterside.html)

Toigo, in his book Disaster Recovery Planning (2003), continues the discussion of the evolving data restoration techniques with focus on the issue of restoration time. He provides three degrees of sensitivity, including: The (1) traditional tape backup, (2) tape vaulting, (3) mirroring (Toigo, 2003). Traditional tape backup and restore software along with the removal of backups to off-site storage is a highly manual process (Toigo, 2003). Restoration of data requires the retrieval of tapes, transporting the tapes to the recovery facility and restoration of data through software (Toigo, 2003). Tape vaulting copies data to a remote tape drive over a WAN connection (Toigo, 2003). This process is less manual than the traditional backup to tape method and there is a shorter time frame to restore data but large scale restores may be difficult or slow (Toigo, 2003). Mirroring further reduces time to data because a close to real-time copy of the data exists and is readily accessible over a WAN without restoration from backup media (Toigo, 2003).

The preliminary research into business continuity planning (BCI, 2005, Doughty, 2001, Toigo, 2003,) shows the direct link between data availability and the ability of an enterprise to survive an unplanned outage event of any severity. Further exploration into data recovery shows fully developed, yet very expensive, options for real-time data mirroring for large enterprises (Toigo, 2003, Baltazar, 2005). Small data dependent businesses can rely on readily available vaulting services to protect relatively small amounts of data in a real time or close to real time manner. However, the market for medium sized enterprises is still emerging. The technologies supporting this market, as previously noted, have not yet been widely accepted but forecasted trends show increasing pressure on businesses to enhance their existing backup and restore capabilities. Information Systems managers exploring options to decrease down time after a disaster need to understand what types of data replication options are available so they can make an informed decision on what technology or technologies combined best suits their enterprises recovery time and point objectives.

CHAPTER II – Review of References

The following is an annotated bibliography of the key references used to frame and build this research paper. They are listed in alphabetical order. Each entry begins with a formal bibliographic citation and is followed by a brief summary of the content used, how it relates to the study and the criteria used to select the reference.

Chevance, R., (2005). *Server architectures: Multiprocessors, clusters, parallel systems, web servers, storage solutions.* Massachusetts: Elsevier Digital Press

This highly technical book explores server architecture including database design and high-availability systems. Sections 6.23 to 6.26 provide valuable information on data backup, restore, optimization of backup resources, and supporting technologies. Chapter 10 reviews hardware and software solutions for high availability including information on mirroring data. This reference supports the purpose of this study. Specifically, Chevance explores the critical aspects and options that need to be reviewed by management in order to have a successful backup and restore process by reviewing and agreeing upon the restore point and restore time objectives.

René J. Chevance has a Doctorat d'Etat Es-Sciences from université Paris VI and an Engineering degree from Conservatoire National des Arts et Métiers. Chevance has been a Professor at Ecole Centrale and ENSTA in Paris and is currently an independent consultant.

CSU Writing Center, 2005. Writing guides: Conducting content analysis. Accessed online on Oct 1, 2005 at http://writing.colostate.edu/guides/research/content/contrib.cfm

This text provides an understanding of the conceptual analysis process and goals. This writing guide is relied upon for the content analysis and research methodology sections of this paper. The writing guide is part of the Department of English at Colorado Statue University, a 4 year nationally-accredited comprehensive research university. This guide is written by graduate students as part of their research methods and theory course. Each year since 1993 CSU graduate students have expanded and revised this guide under the guidance of Professor Mike Palmquist. Professor Palmquist is the director of the composition program and co-director of the Center for Research on Writing and Communication Technologies at CSU.

FCA, (2005). Business continuity: Essential practices for information technology exam manual IT section. Retried on November 15, 2005 at http://www.fca.gov/Download/itbusinesscontinuity.pdf

The Farm Credit Administration (FCA) examination manual provides a comprehensive overview of the goals and the IS department's responsibilities within the larger focus of business continuity planning and disaster recovery planning. Content used from this manual includes the description of the main goals of BCP, presented in the full purpose and data analysis sections of this study.

The Farm Credit Administration (FCA) is an agency within the executive branch of the U.S. government. It is responsible for examination and regulation of the banks and agencies that comprise the farm credit system. Though this examination manual is intended for industries regulated by the FCA the examination is based on best practices as defined by the Federal Financial Institutions Examination Council (FFIEC).

FFEIC, (2005). Federal Financial Institutions Examination Council, *Business continuity planning March 2003 IT examination handbook.* (n.d) Retrieved September 10, 2005 from http://www.ffiec.gov/ffiecinfobase/booklets/bcp/bus_continuity_plan.pdf

This Federal Financial Institutions Examination Council (FFIEC) handbook provides an extensive glossary containing current terminology used in business continuity and disaster recovery planning. Additionally, this handbook provides a well developed list of internal and external threats to an enterprise and a list of critical BCP components. Though this handbook is intended to provide guidance and examination procedures to assist examiners in evaluating financial institution risk management processes, the focus is on industry best practices that can be applied outside of the financial industry. This document is referenced not only in the definition section but in the purpose portion of this paper. The Federal Financial Institutions Examination Council was established in 1979 and is a formal interagency body empowered to prescribe uniform principles and standards for the federal examination of financial institutions.

Hiles, A. & Barnes, P. (Eds.). (1999). The definitive handbook of business continuity management. [Electronic version] Chichester: John Wiley & Sons.

This book provides background on the terminology used in the business continuity and the disaster recovery process. This text includes a detailed review of the process and importance of business continuity planning and practical suggestions on how an enterprise can implement strategies to protect themselves from unplanned interruption events. The information in this book is key in framing the larger issue of BCP, in the brief and full purpose sections of this paper and is relied upon heavily in the definition section.

Peter Barnes is the general manager of Survive a business continuity group a respected business continuity association established in 1989. Andrew Hiles is a published author and international speaker on service management issues.

Krischer, J. (2005). Business continuity there is no one-size-fits-all business continuity strategy, so think of disaster recovery scenarios as modules that can be invoked depending on the situation; think modular for effective recovery plans. *Computer Weekly, Sept 20,* p32. Retrieved October 3, 2005 from Expanded Academic ASAP database.

In this article Josh Krischer explores business impact analysis, one aspect of business continuity planning. Specifically Krisher reviews how recovery time objectives, recovery point objectives and cost of downtime should determine the technologies and backup and restore methods that an enterprise implements. Information from this article was used mainly in the significance section.

Josh Krischer has 36 years IT industry experience. He is a vice president and research director in Gartner Research. His research focuses on high-end computing, storage and central operations. This articles was published in Computer Weekly, a web portal designed for IT professionals that provides business and technical articles and independent analysis on new technologies.

Leedy. P.D. & Ormrod. J.E. (2001). *Practical research: Planning and design* (7th ed.). Upper Saddle River: Merril Prentice Hall.

This text offers a simplified overview of qualitative and quantitative research methods including literature review. This book provides the framework for the understanding of the process of a literature review used in this paper. This is the 7th edition of this widely respected book, first published in 1974.

Marks, H., (05/2005). Storage pipeline: Data-replication software. Retrieved September 18, 2005 from http://www.networkcomputing.com/story/singlePageFormat.jhtml?articleID=161600248

This article is a review of specific replication software that represents most types of mirroring products currently available for Windows servers including: NSI Software's Double-Take, Software Pursuits' SureSync, Veritas Software's Replication Exec, XOsoft's WANSyncHA Server, Softek Storage Solutions's Softek Replicator, Veritas' Volume Replicator, and LeftHand Networks' LeftHand SAN NSM 150. This text presents an informative overview of product options that are available and rates each product on ease of use, features and failover time versus cost, bandwidth consumption, and CPU usage. This article provides the inspiration for this paper and is cited in the purpose and data analysis sections of this paper. Marks' article was downloaded from Networkcomputing.com, an online IS trade magazine that provides product reviews, comparisons, analysis and advice for IT professionals. Alexa.com, a web search engine calculates that Networkcomputing.com has a reach of 49.5%. That is, it is estimated that out of every million web users 495,000 visit Networkcomputing.com. Alexa web traffic details can be viewed at http://www.alexa.com/data/details/traffic_details?q=&url=networkcomputing.com. The author, Howard Marks, is the chief scientist at Networks Are Our Lives; a network design and consulting firm in New Jersey.

Toigo, J.W. (2001). *Disaster recovery planning: Preparing for the unthinkable (3rd ed.).* New Jersey: Prentice Hall, PTR.

Throughout this book Toigo focuses on disaster recovery planning and the importance of protecting, backing up and restoring critical data after a disaster. Chapter 4, Data Recovery Planning (p120 – 180), is most relevant to this research. In chapter 4 Toigo explains that without the ability to restore from backups efficiently many companies could never recover from a major disaster. Toigo explores backup technologies, vendors products and provides a functional overview backup strategies. Toigo's writing is central to framing the purpose and significance sections of this paper. Jon William Toigo is a well known consultant, speaker and author. He has written over 1000 articles for the technology trade press and is the author of 15 books, 3 on the subject of contingency planning.

CHAPTER III – Method

Literature review (Leedy & Ormrod, 1993) is selected as the overarching research methodology for this study. Events such as terrorism and natural disasters that have destroyed data centers, along with lives and property, coupled with businesses increased reliability on data stored on servers, have in part kept disaster recovery on the proverbial “front burner”. Many books, articles, and white papers have been written on business continuity and disaster recovery planning and data protection techniques. The literature review methodology supports the kind of data collection necessary to explore the pros, cons, functionality, and costs of selected types of disaster recovery technologies suitable for medium sized businesses.

Content analysis, as defined by the Colorado State University writing guide, is “a research tool used to determine the presence of certain words or concepts within a set of texts” (http://writing.colostate.edu/guides/research/content/contrib.cfm, 2005). Conceptual analysis, a form of content analysis, is used as the strategy for data analysis for this study. According to McMillen & Schumacker (1993), conceptual analysis can be used to clarify a concept by describing the essential or generic meaning of the concept. The goal of this paper is to clarify the concept of selected WAN capable data mirroring disaster recovery technologies.

Data Collection

The initial literature collection process is cyclical in nature and includes the following iterative steps: identify search terminology, World Wide Web searches, selected online business magazine article searches including CIO.com, Gardnergroup.com and Businessweek.com, selected full text database searches by key words, library catalog keyword searches, and full text database searches by subject. Only texts directly pertaining to the purpose, as described in the full purpose section, are included for potential review in this study.

Literature collection is designed to gather recently published text documents that would help explain selected types of technologies available for medium size enterprises to replicate the file data stored on their Microsoft Windows servers to an offsite location in a close to real-time manner. Criteria for the selection of documents included in this study are:

1. Is the document published within the last 5 years?
2. Are the products or replication solutions appropriate for use by and marketed to medium sized enterprises?
3. Is the document retrieved from a reputable source?
4. Is there a group of articles supporting each key perspective?

If all of these criteria are met, the document is included in the study. Documents retrieved directly from vendor web sites are not included in this study because of the obvious bias, but articles written by professionals associated with a vendor but retrieved from a reputable source are not automatically excluded.

The initial search is conducted through Google and Teoma search engines. Google.com maintains a prominent market position and Teoma.com offers a search page that suggests ways to narrow the search terminology for more targeted search results. Keyword searches include the following terminology: business continuity, data replication, server replication, server mirroring, data protection, backup process alternatives, data recovery, multi-site data replication, disaster recovery, WAN replication, and business continuity planning. Searches are conducted with individual words and word combinations. The search term "system recovery" yielded too many and spurious results. Additionally, some combinations of keywords yielded no results at all, for instance when combining the terms wide area networks, replication, and business continuity.

The initial web search is conducted to identify some potential text to be used in this study but is generally more useful in providing terminology and information on existing backup solutions and technology. The appendices of the Federal Financial Institutions Examination Council's (FFEIC) business continuity handbook located on line at http://www.ffiec.gov/ffiecinfobase/booklets/bcp/bus_continuity_plan.pdf has a well developed glossary, a detailed explanation of the BCP and a short section on replicating data over a WAN. Another valuable source is located online at Networkcomputing.com which has storage and server technology reviews. An article by Howard Marks provides a technical review and overview of specific data replication products.

Selected databases available from the University of Oregon WORLDCAT system (http://libweb.uoregon.edu) and the Eagle Public Library are searched based on the inclusion of full-text articles, their timeliness, and inclusion of relevant data. Databases accessed include: Academic Search Premier, Expanded Academic ASAP, Lexus Nexus, Business Source Premier, General BusinessFile ASAP, and MasterFILE Premier. The Expanded Academic ASAP, Academic Search Premier, MasterFILE Premier and Business Source Premier provide the most relevant and timely full text articles. The same search terms used on the web are used for the database searches both individually and in conjunction. It is necessary to use the advanced search functions on database searches in order to limit the number of spurious articles. For instance, articles with terms such as “SQL” and “Oracle” are specifically rejected with the “not” command in order to perform a directed search to remove articles about technology that is designed to replicate databases over a WAN.

The Eagle and Denver Public Library catalog search by keyword is performed with the same keywords as previously mentioned. This process yields the fewest, but most valuable, resources. The chapter named “Data Recovery Planning” in John Toigo’s book Disaster Recovery Planning (2003) explores the entire process for planning for data recovery including a review of some vendor products. In section 6.23 of Server Architecture (2005), Chevance provides a highly technical overview of data backup and restore options.

Data Analysis

Data analysis is performed following the eight steps presented in the conceptual analysis process as outlined on the CSU Writing Lab website (http://writing.colostate.edu/guides/research/content/contrib.cfm). The approach in each step is outlined below. The coding process is designed to explore the following question, “What types of asynchronous data replication technologies, capable of copying file data from Microsoft Windows servers to a off-site location over a WAN link, are available and appropriate for medium sized enterprises”?

Step One: Level of Analysis: Phrases are chosen as the level of analysis for concept coding because concept coding for a common word such as “file” could skew the results.

Step Two: Number of Concepts to Code: Limiting the number of concepts allows for the coding of very specific concepts. Coding for concepts begins with a pre-defined set of phrases that describe prominent data replication technologies, as defined during preliminary review of the literature, including: snapshot, file replication, block replication, continuous data protection (CDP), and tape backup. During the coding process, if additional relevant concepts are found beyond this pre-determined coding set, they are added as emergent concepts.

Step Three: Code for Existence or Frequency of Concept: Coding is done based on the existence of the concept in a text, not its frequency.

Step Four: How to Distinguish Among Concepts: Concepts are coded loosely in order to provide maximum inclusion. That is, concepts are recorded as the same even when they appear in different forms, allowing for the coding words that imply the same meaning as the selected coding phrases. Additionally, acronyms, such as CDP, are also included in coding. Different spellings or hyphenations are counted as the same word. The hyphenated terms, such as snap-shot, are counted. Figure 1 presents the translation rules utilized in this step.

Figure 1: Translation Rules

CDP	Snapshot	Block replication	File replication	Byte level replication	Tape Backup
Continuous data protection	Snapshots	Block based replication	Replicate files	Portion of file that changes	Tape
Continuous protection	Snapshot type	Replicate changed blocks	Synchronize files	Byte-level	Tape-backup
	Shadow copy	Block asynchronous	File-based replication		Tape solutions
	Snapshot technology	Send changes to the data not the entire file	File-filter asynchronous replication		
		Data blocks	File level replication		
		Changed blocks	File synchronization		
		Block-level	File-level solution		
		Changed data blocks			

Step Five: Rules for Coding the Text: Translation rules are created to give the coding process consistency and coherence. For example, context is reviewed to verify if the selected phrase is used in a negative manner.

Step Six: What To Do with Irrelevant Information: All irrelevant information, i.e., any information not directly relevant to data recovery technologies appropriate for medium sized enterprises and capable of WAN replication, is ignored, because it will not impact the outcome of the coding.

Step Seven: Code the Texts: The reviewed text is coded manually by writing down concept occurrences in a spread sheet. This allows for contemplation of context, recognition of errors and allows for the inclusion of overlooked important codes. Data related to the initial set of concepts (i.e., data recovery technologies including snapshot, file replication, block replication, continuous data protection (CDP), and tape backup), as well as emergent concepts, is coded in this way.

Data Presentation

Step Eight: Analyze Results: The raw tally of occurrences of coding terms identified during the conceptual analysis process is organized into Figure 3: Coding Concept Occurrences. Data identified during conceptual analysis is then sorted into categories by the type of data replication technology, in order to prepare for secondary analysis (see Appendix A). Unwanted material is skipped over. Next, each of these new organizational categories is amplified along four features, including the pros (advantages), cons (limitations and drawbacks), functionality, and cost. The results of this amplification process are presented in the final outcome of the study, in two ways: (1) as a summary table that presents the four features examined in each identified replication technology (see Table 1: Comparative View of WAN Capable Replication Technologies); and (2) in a discussion of the summary table, framed in terms of key business continuity planning goals, which include the need to: (a) minimize disruption of services to the enterprise and its customers, (b) ensure continuity of critical operations and (c) limit financial loss (FCA, 2005, Toigo, 2003, Daughty, 2001).

The summary table is designed as an easy-to-use decision support tool for use by IS managers as a way to compare the pros, cons, functionality and cost of each type of selected asynchronous data replication technologies. A template for the design of Table 1 is presented below, in Figure 2. This researcher intends that Table 1 will be used by IS managers as a way to initially focus their search for the right replication technology to implement in order to meet their recovery goals, that is their recovery time and recovery point objectives, per their business continuity planning process. Table 1 is a summary of the features of four types of selected replication technologies. In the “pros” section, the advantages over traditional tape backup and the other selected types of replication technologies are summarized. The “cons” row is a summary of the limitations and weak points of the selected type of replication technology. The “functionality” row is a summary of the accepted use of the technology. The functionality row explores how a file or volume is replicated and the scope of restoration that is possible. The “cost” row is an estimated range of the per server monetary cost of software specializing in the specific replication technology marketed to medium sized enterprises. The scale used to report cost is simple, assessing cost against the current market rates, and reported as Low, Medium, or High. Actual prices of vendor applications are included when available as part of the data set.

Then an example of how this tool might be used by IS managers is provided, in the form of a discussion. The discussion is designed to advance understanding of the selected types of replication technology, viewed from the standpoint of Business Continuity Planning goals and framed in terms of recovery point and recovery time objectives, by expanding on the points listed in summary table. Each type of data replication technology explored in this study is designed to minimize disruption of services, ensure continuity of critical operations and limit financial loss to an enterprise by creating a, close-to-real-time, replicated data set at a secondary location that can be utilized in case of data loss or disaster at the primary data processing facility. Though each of the replication technologies explored in this study replicate data, the implementation of a specific or multiple types of replication technologies should be based on the recovery point and recovery time objectives of an enterprise agreed upon during the BCP process. The discussion is designed to allow the IS manager to explore the individual options summarized in the summary table in more detail.

Figure 2: Template Table 1 - Four Features Examined in Replication Technologies

Four Features

		Pros	**Cons**	**Functionality**	**Costs**
Disaster Recovery Data Replication Technologies	**File replication**				
	Block replication				
	Snapshot				
	CDP				

Chapter IV – Analysis of Data

The goal of this content analysis is to clarify a concept by describing the essential or generic meaning (McMillen & Schumacker, 1993). Forty-two articles are accessed per the guidelines in the Method chapter. Articles are read in a preliminary review and matched against the selection criteria. Most articles are removed because they focus on technology inappropriate for medium sized enterprises or tape virtualization. Data analysis is executed on a data set of nineteen articles by following the eight steps as defined in the data analysis and data presentation sections of this document. The data set is reviewed and concept occurrences are highlighted. After an initial review of the concept occurrences the phrase “byte level replication” is added as emergent code. The remaining data set is re-read and concept occurrences are coded. The data set is then reviewed and each concept occurrence is manually added to a spreadsheet of raw data from the conceptual analysis (see Appendix A). Duplicate entries are removed. The spreadsheet includes the author’s last and first name, date of publication, and the concepts occurrences in context. The spreadsheet of raw data was then sorted by author and a count of concepts is performed to verify that there was sufficient data for each concept (see Figure 3).

Figure 3: Coding Concept Occurrences

Concept Coding	Concept Count
File replication	9
Tape Backup	7
CDP	7
Snapshot	6
Byte level replication	4
Block replication	4

Table 1: Comparative View of WAN Capable Replication Technologies is designed as a decision support tool for use by IS managers as a way to compare the pros, cons, functionality and cost of each type of the selected asynchronous data replication technology and initially focus their search for replication technologies to implement in order to meet their recovery goals.

Table 1 is a modified and expanded version of John Toigo’s Data Backup and Restoral Alternatives Table where he briefly compares the strategy, description, pros, cons and cost of traditional tape backup, backup to electronic tape vault and disk mirroring (Toigo, 2003, p. 123). The raw data (see Appendix A) is reviewed and sorted by type. Individual entries are cut and pasted into the summary table. Any mention of an advantage of a product is moved into the “pros” column. Any declaration of a disadvantage or limitation is pasted into the “cons” column. The main purpose and goal of a replication technology is added to the “functionality” column as well as any vendor specific application information. Additionally, all information within the data set pertaining to cost is pasted into the “cost” column, including vendor specific information when available within the data set.

The concept of Continuous Data Protection (CDP) proved most difficult to report because of the major differences in CDP offerings that can bring the price well over the $10,000 imposed limitation. Furthermore, many of the articles in the data set that reference CDP do not contain direct information about the cost of specific products. The byte and block coding also is problematic. During the data analysis process the researcher noted that byte and block replication are not distinct types of replication technologies but rather are distinct ways on how to perform the specific types of CDP, snapshot or file replication. Data gathered about byte and block replication is included in the summary table, in the individual sections and in the type of supported replication technology sections of Table 1. Each concept occurrence is reviewed in context to verify that terminology referring to the concept is coded properly.

The concept of file replication is mentioned most frequently in nine (47%) of the articles. Tape backup and CDP are referred to in seven articles each (37%), and snapshot is found in six (32%) of the articles. Byte and block replication are discussed the least, appearing in four (21%) articles each.

Table 1: Comparative View of WAN Capable Replication Technologies

	Pros	Cons	Functionality	Costs
File replication	• File replication can often be transmitted on a continuous basis or according to parameters such as bandwidth thresholds • Depending on vendor can provide asynchronous replication via queuing in both source and target servers reducing interruptions in processing • Not tied to the storage hardware • Can run in almost any environment • Easy to restore • Relatively inexpensive and easy to implement • Can run in many-to-one or one-to-many and daisy-chained configurations	• Considered a target data backup, offering not a complete disaster recovery tool • If file is overwritten, or virus infects the data center, replication systems will comply the unwanted changes • Performance is not a strength of software-based solutions • Functionality varies widely by vendor • Less expensive offerings generally can not throttle bandwidth utilization and do not journal to disk • Upfront planning needed	• Designed to create a readily accessible copy of files • Double-Take processes only byte-level changes to files or volumes after the initial replication • Fujitsu's Softek Replicator device provides block-level data replication • NSI Double-Take uses a file filter to capture data updates and both in-memory and on-disk • Veritas Storage replicator copies only the blocks of data that have changed • WANsynchHA many options from block checksums to compression and bandwidth throttling • NSI Double-Take uses a modified synchronous scheme to reduce the usual delay associated with host based replication	• Low to Medium • Double-Take starts at $4,495 • Softek Replicator starts at $2,300 per single processor server • Xosoft’s WANsynchHA starts at $3,500 • Veritas Volume Replicator starts at $4,500

	Pros	Cons	Functionality	Costs
Snapshot	• Snap-copy (Volume copy) creates a complete second copy of data • A single disk can hold many snapshots and one can choose the newest, uncorrupted version for the restore • Changes to any files are automatically captured and applied • Most new programs use Pointer-based Snapshots that copy only the changed data requiring less overhead • Snapshots can be taken up to every 15 minutes	• After a restore one looses the changes made since the last snapshot • Does not provide the data-level restoration capabilities of CDP • Changes to any files automatically will be captured and applied including unintentional changes • Disadvantage of snap-copy, there must be enough storage to accommodate the snapshot (100% overhead per snapshot) • Pointer based Snapshots are not exact copies of data • Snapshots are better suited for mirroring than CDP block replication	• Designed to capture data at specific points in time • Snapshots are regular, incremental backups to disk • Snapshots restore at the file or volume level • Backup Exec, combined with Replication Exec, provides off host backup by shifting data snapshot backups to alternate storage systems • EMC's Replistor supports Microsoft's Volume Shadow Copy Service which lets the software take periodic snapshots of data	• Medium (Exact pricing not included in data set) • EMC’s Replistor • Veritas Backup Exec combined with Veritas Replication exec

	Pros	Cons	Functionality	Costs
CDP	• Only records changes to data • Shortens time to data after a failure • Logs all change to the disk image as they occur • Can restore any version of a file • Improves on the data protection and recoverability of tape backup • Can scroll back to any version of a file that has been saved • Can restore part of a file • Ideal for retrieving a single file or piece of data • IT manager can roll back the affected volume to a specific point in time	• CDP augments the existing tape backup with short-term snapshots that eventually get sent to backup • CDP is not inherently a WAN replication technology • Restores are permanent if no hardware in place for staging recovery • Replication and snapshots are better suited for mirroring than CDP • Many CDP products are not “true CDP” but rather they utilize Microsoft’s Shadow copy components • Time to restore is based on the number of modifications - the more changes that have occurred the longer the restore will take	• Designed to instantly copy and backup all changes to data and shorten the time to data after a failure • Offers the ability to set and achieve recovery point objectives • Keeps track of all transactions written to disk and stores them into a log file • Provides data-level restoration capabilities that tape, replication and snapshot technologies lack • Often appliance based • Allows creations of baseline – the process of committing all tracked changes to a file • True CDP creates one snapshot for every instant in time that data modification occurs	• Medium to high • True CDP pricing starts at over $10,000

	Pros	Cons	Functionality	Costs
Tape Backup	• Time tested technology • Low price point • Well suited for archiving • Today's SDLT and LTO tape drives are capable of backup rates faster than servers can supply the data • Can solve even the knottiest of data protection problems when used effectively	• After a restore one looses all changes made since the last backup • Does not afford continuous protection • Provides no real-time duplication or swift data recovery after a disaster • Leaves open a big window of data vulnerability • Requires manual labor, error prone, must be labeled, packed, tracked and sent to storage • Restores from tape are slow • Verifying a tape backup can take hours or days	• Designed to restore at the file or volume level	• Low to High

	Pros	**Cons**	**Functionality**	**Costs**
Byte	• Provides added granularity • Can reduce total time and size of backup • Byte level CDP allow restores to any point in time		• Double-Take processes only byte-level changes to files or volumes after the initial replication • XOsoft's latest version of WANsychHA uses byte level replication	• Low to high • NSI Double-Take $2,495 per server
	Pros	**Cons**	**Functionality**	**Costs**
Block	• Block level CDP is a decent solution for bare metal restores • Heterogeneous solution • Can reduce total time and size of backup	• Block replication CDP is little more than mirroring that can be rolled back • Block-replication CDP knows nothing about files or directory • If file is overwritten, or virus infects the data center, replication systems will apply the unwanted changes. • Not as much granularity as Byte level replication	• Most new programs copy only the changed data • SureSync with SPI Agent capable of replicating only changed blocks in a file • Veritas Volume Replicator is block based logs and requires a dedicated volume for each replication group - upfront planning is needed	(Products are listed medium to high cost) • SureSync with SPI Agent • Fujitsu's SOFTEK Replicator • Veritas Volume Replicator

Chapter V – Conclusions

The goal of WAN capable replication technologies is to limit data loss from unplanned interruption events by creating a copy of critical data to a secondary location. But, each type of replication technology reviewed handles this process differently, and each provides for a different level and granularity of restoral possibilities at different cost points. This makes across-the-board comparison difficult and perhaps not very useful. Each type of reviewed disaster recovery technologies has its own strengths and weaknesses and, depending on the unique recovery point and recovery time objectives of an enterprise, may or may not be appropriate to implement or may need to be implemented in conjunction with other forms of replication. During the business continuity planning (BCP) process, potential disasters should be reviewed and assigned a risk rating. Most enterprises will not be able to mitigate all risks and will concentrate on high risk events. Table 1: Comparative View of WAN Capable Replication Technologies, is designed to be used by IS mangers as part of this risk assessment process. The best time to assess risk and implement disaster recovery technologies is before an unplanned interruption event. The following discussion of Table 1 is presented as a way for IS managers to begin their own search and analysis of replication technologies.

To reiterate, tape backup is still a necessary part of a disaster recovery. Backup to tape, with off site storage, meets the basic requirements for disaster recovery. But, for an enterprise that requires a recovery time measured in hours rather than days or can not afford the gap in data left after a restore from the last nights backup, tape backup alone is not enough. As IS managers know, there are many problems with tape backup.

True Continuous Data Protection (CDP) can provide almost immediate restore to any version of a file. The data gap associated with restores from tape is virtually eliminated. Many less expensive CDP products do not provide true continuous data protection. Rather they utilize Microsoft's Shadow Copy components and take snapshots every 15 minutes. CDP provides good options for enterprises that can not afford loosing even a small amount of work. This type of data protection comes at a very high cost and restores, depending on the configuration, often can not be undone. Additionally, the time to restore may be slower than expected if the file is modified frequently. Bandwidth is also a serious concern with CDP replication and file replication and snapshots are more suited for mirroring across a WAN. CDP appears to be better suited for LAN replication and targeted to companies that can not afford to loose small changes to mission critical data.

Snapshot replication creates a complete second copy, up to every 15 minutes, of data on a server and has the functionality of being able to role back data to a specific point in time before the disaster event, thus shortening the time to recovery and meeting all but the most stringent recovery point objectives. The main limitation is that there is no support for data-level restoration. That is, a single file can not be restored from the snapshot. A restoration is an all or nothing event. This technology is designed to remediate for major disasters. If a single file is lost or all image files are destroyed by a virus and restoration from tape was not available a company would have to decide quickly if it was worth destroying all the changes on a server, both the intentional and unintentional, to roll back to a specific point in time. The ability to role back a server might be more useful for transaction-based applications rather than file replication.

The concept of file replication appeared most frequently in the data set. File replication, if implemented properly, provides a readily accessible second copy of data at a remote location. Some vendor solutions even allow for close to synchronous synchronization, meeting the most stringent recovery point objectives given available bandwidth. And, because the data is readily accessible and easily restored, file replication can meet strict recovery time objectives. The main drawbacks to file replication are the performance hit on servers, potential bandwidth issues and that all changes to files are copied to the remote location. The wide variety of vendor offerings and price points will require extra time for review by IS managers and extra upfront planning.

Once the replication technology is chosen, the IS manager must decide on a specific vendor offering. Depending on the vendor, the level of replication can range from file to block or byte. The main goals of block and byte replication are to limit the amount of data that must be pushed over a WAN by replicating only the changed blocks or bytes of data rather than the entire file.

On a final note, the replication technologies reviewed do not take the place of traditional tape backup but rather use different strategies to limit exposure from the restore gap associated with tape backup. Each type of replication technology offers different restore points, ranging from restoring an entire volume to a specific point in time, to being able to restore part of a single file. Each type of replication technology has inherent limitations. CDP, for example, is not appropriate for restoring entire volumes and the benefits of file replication have to be measured against performance issues and potential bandwidth issues. This study revealed that each type of technology reviewed will decrease the time to data restoration after a disaster and the recovery point objectives, ranging from continuous to a scheduled delay, can be met.

Appendix A: Raw Data

Author (Last, First)	Date	CDP	Snapshot	File replication	Block Replication	Byte Level	Tape
Baltazar, Henry	10/20/2003						IT managers are finding that traditional tape backup with Off site storage is not enough
Baltazar, Henry	1/19/2004			Fujitsu's SOFTEK Replicator provides block-level data replication - heterogeneous solution	Fujitsu's SOFTEK Replicator provides block-level data replication - heterogeneous solution		
Baltazar, Henry	1/19/2004			SOFTEK Replicator provides block level replication of volumes	SOFTEK Replicator provides block level replication of volumes		
Baltazar, Henry	1/19/2004			Double-Take is a file level solution			
Baltazar, Henry	1/19/2004			Not recommended for high transaction volume			

Baltazar, Henry	1/19/2004			Softek Replication is a better choice for servers that create a large number of small files because an IT manager just has to setup replication for a data volume			
Baltazar, Henry	1/19/2004			Softek Replicator starts at $2,300 per single processor server			
Baltazar, Henry	4/11/2005	CDP should be use in conjunction with tape					CDP should be use in conjunction with tape
Baltazar, Henry	4/11/2005	Improves on the data protection and recoverability of tape backup					Improves on the data protection and recoverability of tape backup
Baltazar, Henry	4/11/2005	New CDP Products are rushing in to fill gaps in tape backup					New CDP Products are rushing in to fill gaps in tape backup

Baltazar, Henry	4/11/2005						Problem: leaves open a big window of data vulnerability. Can only restore up to last backup point
Baltazar, Henry	4/11/2005	CDP keeps track of all transactions written to disk and stores them into a log file. IT manager can roll back the affected volume.					
Baltazar, Henry	4/11/2005	Problem: data rewinds (restores) are permanent because no hardware in place for staging recovery					
Baltazar, Henry	4/11/2005	XOsoft software based Enterprise Rewinder starts at less than 1,000					
Baltazar, Henry	9/5/2005						Tape is far from dead but should not be the only backup technology in your environment

Baltazar, Henry	9/5/2005			Data Domain Inc's DD Replicator			
Baltazar, Henry	11/7/2005			Available from Symantec, XOsoft Inc. and NSI Software Inc. Work at the host level (not block)			
Baltazar, Henry	11/7/2005			Not tied to the storage hardware - they can run in almost any environment			
Baltazar, Henry	11/7/2005			Performance is not a strength of software-based solutions			
Baltazar, Henry	11/7/2005			Relatively inexpensive and easy to implement			
Clark, Elizabeth	3/1/2003			Double-Take - after the initial replication only byte-level changes to files or volumes are processed		Double-Take - after the initial replication only byte-level changes to files or volumes are processed	
Clark, Elizabeth	3/1/2003			Double-Take also uses sequential transfer to help ensure data integrity			

Clark, Elizabeth	3/1/2003			Double-Take monitors and replicates change store open files as they occur			
Clark, Elizabeth	3/1/2003			Double-Take provides asynchronous replication via queuing in both source and target servers reducing interruptions in processing.			
Connor, Deni	9/1/2005		Microsoft's VSS provides the mechanism for creating point-in-time copies of data				
Connor, Deni	9/1/2005		EMC's Replistor supports Microsoft's Volume Shadow Copy Service which lets the software take periodic snapshots of data				
Connor, Deni	9/1/2005	Replistor saves changes to files on Windows file servers and NAS devices as they are made.	In addition a number of data snapshots are made on the host replistor server.				

Connor, Deni	9/1/2005		Records every change to a file made on the network and will recover only files from known snapshots				
Connor, Deni	9/1/2005		Replistor starts at $1,650 per server				
Connor, Deni	9/1/2005		Snapshot taken up to every 15 minutes				
Don MacVittie	6/23/2005						
Elizabeth Clark	3/1/2003			Changes can be transmitted on a continuous basis or according to parameters such as bandwidth thresholds			
Fonseca, Brian	1/5/2006		Backup Exec and Replication Exec - provide off host backup by shifting data snapshot backups to alternate storage systems				
Fonseca, Brian	1/5/2006		Changes to any files automatically will be captured and applied				

Klein, Karen	4/5/2005						Manual tape backup systems provide no real-time duplication or swift data recovery after a disaster.
Klein, Karen	4/5/2005					NSI Double-Take $2,495 per server	
Lipschutz, Robert	fall 2004			Can backup all types of data files including database and e-mail.			
Lipschutz, Robert	fall 2004			Considered a target data backup offering not a complete disaster recovery tool. Does not backup the operating system or applications and lacks transport encryption			
Lipschutz, Robert	fall 2004			Once VSR agent is installed tracks witch data blocks have changed.			

Lipschutz, Robert	fall 2004			Veritas Storage replicator copies only the blocks of data that have changed.- so far less data crosses a WAN link.			
Lunt, Penny	11/1/2005	Creates an electronic journal of complete storage snapshots with one snapshot for every instant in time that data modification occurs	True CDP creates an electronic journal of complete storage snapshots with one snapshot for every instant in time that data modification occurs				
Lunt, Penny	11/1/2005	Legitimate CDP products add an expensive step to backup. They don't replace your existing backup infrastructure they merely augment it with short-term snapshots that eventually get sent to backup	Legitimate CDP products add an expensive step to backup. They don't replace your existing backup infrastructure they merely augment it with short-term snapshots that eventually get sent to backup				

Lunt, Penny	11/1/2005						requires manual labor, error prone, must be labeled, packed, tracked and sent to storage
Lunt, Penny	11/1/2005			Image Systems provides block and file-level replication software suitable for database and e-mail servers			
Lunt, Penny	11/1/2005			NSI Software pursuits, Veritas software and XOsoft			
Lunt, Penny	11/1/2005	Live vault offers true CDP for backing up databases and files from/to remote sites starting at $25,000 - appliance based					
Lunt, Penny	11/1/2005	Offers the ability to set and achieve recovery point objectives.					

Lunt, Penny	11/1/2005	Pure CDP backs up all protected data whenever a change is made, capturing and time-stamping every transaction.					
MacVittie, Don	6/23/2005		a single disk can hold many snapshots and one can choose the newest, uncorrupted version for the restore				
MacVittie, Don	6/23/2005		after a restore one looses the changes made since the last snapshot				after a restore one looses the changes made since the last snapshot
MacVittie, Don	6/23/2005	CDP	CDP provides the data-level restoration capabilities that tape, replication and snapshot technologies lack				CDP provides the data-level restoration capabilities that tape, replication and snapshot technologies lack
MacVittie, Don	6/23/2005		Most new programs copy only the changed data		Most new programs copy only the changed data		

MacVittie, Don	6/23/2005	CDP Provides the data-level restoration capabilities that tape, replication and snapshot lack	Providing the data-level restoration capabilities that tape, replication and snapshot lack	Providing the data-level restoration capabilities that tape, replication and snapshot lack			Providing the data-level restoration capabilities that tape, replication and snapshot lack
MacVittie, Don	6/23/2005		Regular, incremental backups to disk				
MacVittie, Don	6/23/2005	Replication and snapshots are better suited for mirroring than CDP block replication	Replication and snapshots are better suited for mirroring than CDP block replication				
MacVittie, Don	6/23/2005		Snapshot restore at the file or volume level				Tape restores at the file or volume level
MacVittie, Don	6/23/2005		Some copy the entire area being backed up each time				
MacVittie, Don	6/23/2005						Restores from tape are slow
MacVittie, Don	6/23/2005						Tape has largely become an archiving rather than backup technology
MacVittie, Don	6/23/2005						Tape is the old standby

MacVittie, Don	6/23/2005						Verifying a tape backup can take hours or days.
MacVittie, Don	6/23/2005	Block replication CDP decent solution for bare metal restores			Block replication CDP decent solution for bare metal restores		
MacVittie, Don	6/23/2005	Block replication CDP little more than mirroring that can be rolled back			Block replication CDP little more than mirroring that can be rolled back		
MacVittie, Don	6/23/2005	block-replication CDP knows nothing about files or directory structures - just logs all change to the disk image as they occur			block-replication CDP knows nothing about files or directory structures - just logs all change to the disk image as they occur		
MacVittie, Don	6/23/2005			If file is overwritten, or virus infects the data center replication systems will comply the unwanted changes too.	If file is overwritten, or virus infects the data center replication systems will comply the unwanted changes too.		
MacVittie, Don	6/23/2005	Application aware CDP - allows restores down to the data level - allow the restore of a part of a file					
MacVittie, Don	6/23/2005	Application aware CDP - high performance hit					

MacVittie, Don	6/23/2005	Best question to ask - Will I ever need to restore a single piece of this data?					
MacVittie, Don	6/23/2005	Can restore any version of a file					
MacVittie, Don	6/23/2005	CDP is becoming a critical part of a complete backup strategy					
MacVittie, Don	6/23/2005	Falconstor & Revivio and Microsoft (no inherent remote capabilities)					
MacVittie, Don	6/23/2005	File replication not good for database or transactional applications					
MacVittie, Don	6/23/2005	Good for protecting a limited subset of transactional applications					
MacVittie, Don	6/23/2005	Ideal for retrieving a file or piece of data					
MacVittie, Don	6/23/2005	Most CDP products are sold as appliances - the repository and supporting software can be sold as a package					
MacVittie, Don	6/23/2005	Most CDP systems also creations of baseline - committing all tracked changes to a file					

MacVittie, Don	6/23/2005	Most common - file-replication CDP - watches the dire and logs all changes to files as they occur- requires no agents - single files are easily distinguished and restored					
MacVittie, Don	6/23/2005	Products let you scroll back any version of a file that been saved					
MacVittie, Don	6/23/2005	Retains information about every change to a file over its live					
MacVittie, Don	6/23/2005	Time to restore is based on the number of modifications - the more changes that have occurred the longer the restore will take					
MacVittie, Don	6/23/2005	Would Time to restore is based on the number of modifications - the more changes that have occurred the longer the restore will take					
MacVittie, Don	6/23/2005	XOsoft enterprise rewinders are software packages that use a database a the repository					

Marks, Howard	9/14/2005						falling cost of disk storage and annoyances of tape are driving the interest in disk-to-disk-to tape backup
Marks, Howard	9/14/2005						Tapes are fine if restoring an entire drive but are slow when restoring smaller groups of files
Marks, Howard	9/14/2005						Today's SDLT and LTO tape drives are capable of backup rates faster than servers can supply the data.
Marks, Howard	5/1/2005			SureSync with SPI Agent capable of replicating only changed blocks in a file.	SureSync with SPI Agent capable of replicating only changed blocks in a file.		

Marks, Howard	5/1/2005			VV Replicator Block based logs require a dedicated volume for each replication group - upfront planning needed.	VV Replicator Block based logs require a dedicated volume for each replication group - upfront planning needed.		
Marks, Howard	5/1/2005			Can run in many-to-one or one-to many and daisy-chained configurations			
Marks, Howard	5/1/2005			Double-Take is the current market leader			
Marks, Howard	5/1/2005			Double-Take starts at 4,495 per server			
Marks, Howard	5/1/2005			Integrated with the Backup exec console			
Marks, Howard	5/1/2005			Modified synchronous scheme reduced delay typically found with synchronous mirroring.			
Marks, Howard	5/1/2005			No journaling or maintaining the write order - Not acceptable for database or e-mail - good simple file replication			

Marks, Howard	5/1/2005			NSI Double-Take uses a file filter to capture data updates and both in-memory and on-disk journals to hold updates at the source until the target can receive them			
Marks, Howard	5/1/2005			Replication exec - no application-specific features to support exchange or SQL			
Marks, Howard	5/1/2005			Replication exec - one-to-many, many-to-one and mirror			
Marks, Howard	5/1/2005			Softeck Replicator - block asynchronous approach to replication			
Marks, Howard	5/1/2005			Softek Replicator - can not throttle bandwidth utilization and windows version will not journal to disk			
Marks, Howard	5/1/2005			Softek Replication - needs sufficient bandwidth to be successful			

Marks, Howard	5/1/2005			Softek Replicator - if traffic level exceeds available bandwidth must resynchronize the volumes			
Marks, Howard	5/1/2005			Softek Replicator - journals disk writes to special memory structure called BAB and updates the modified blocks on target system			
Marks, Howard	5/1/2005			Softek Replicator starts at 3495			
Marks, Howard	5/1/2005			Starts at $3995			
Marks, Howard	5/1/2005			SureSync scans directories on source servers and copies changed files without requiring an agent on the target server			
Marks, Howard	5/1/2005			SureSync SPI Agent provides replication-stream encryption and compression			
Marks, Howard	5/1/2005			Veritas Replication exec - file-based replication			

Marks, Howard	5/1/2005			Veritas Volume Replicator - Block Level			
Marks, Howard	5/1/2005			Volume replicator not a true synch transfer - does not wait for write complete message from host.			
Marks, Howard	5/1/2005			Volume replicator supports asynchronous and synch transfer			
Marks, Howard	5/1/2005	XOsoft enterprise rewinder - CDP roll back allowed		WANsynchHA many options from block checksums to compression and bandwidth throttling			
Marks, Howard	5/1/2005			WANsynchHA starts at 3,500			
Marks, Howard	5/1/2005			XOsoft WANsychHA - application specific version available			
Marks, Howard	5/1/2005			XOsoft WANsychHA - File-filter asynchronous replication			

Marlin, Steven	10/31/2005	Because continuous data protection only records changes to data there is less data to be copied. And shortens the time it takes to recover from failure					
Marlin, Steven	10/31/2005	designed to instantly copy and backup all changes to data although it is probably isn't needed by most companies					
Marlin, Steven	10/31/2005	For most companies daily or hourly backups are sufficient					
Marlin, Steven	10/31/2005	Symantec Corp CDP Can replicate the snapshots to iSCSI storage arrays					
Marlin, Steven	10/31/2005	Symantec Corp CDP products that can backup incremental copies of files to local servers and restore quickly					
McCown, Sean	9/1/2005			XOsoft's latest version of WANSyncHA uses byte level replication		XOsoft's latest version of WANsychHA uses byte level replication	

McCown, Sean	9/1/2005			WANSync Creates time-stamped checkpoints for each I/O event and rewinds to any checkpoint			
McKinstry, Jim	2/1/2005		Advantages of Snap-copy is that there is a full copy of the original data				
McKinstry, Jim	2/1/2005		A disadvantage of Pointer based is if source file is write intensive, maintaining the copy on write can affect the performance.				
McKinstry, Jim	2/1/2005		Disadvantages of snap-copy there must be enough storage to accommodate the snapshot (100% overhead per snapshot)				
McKinstry, Jim	2/1/2005		pointer based - (copy on write) requires fraction of original disk space.				
McKinstry, Jim	2/1/2005		pointer based called (copy on write)				

McKinstry, Jim	2/1/2005		Pointer based Snapshots are not exact copies of data				
McKinstry, Jim	2/1/2005		Simplest form of replication is the Snapshot				
McKinstry, Jim	2/1/2005		Snap-copy (Volume copy) creates a complete second copy of data				
McKinstry, Jim	2/1/2005		Snapshots generally controlled by a single device.				
Mearian, Lucas	10/3/2005	CDP come in two iterations - those that record every change to data at the byte level - allow restores to any point in time or Periodic snapshots of data that allow recovery back to specific points in time	CDP come in two iterations - those that record every change to data at the byte level -allow restores to any point in time or Periodic snapshots of data that allow recovery back to specific points in time			CDP come in two iterations - those that record every change to data at the byte level -allow restores to any point in time or Periodic snapshots of data that allow recovery back to specific points in time	

Penny Lunt	11/1/2005	CDP isn't widely available yet. Many vendors actually just supporting Microsoft's Volume Shadow copy service that can take one snapshot an hour	CDP isn't widely available yet. Many vendors actually just supporting Microsoft's Volume Shadow copy service that can take one snapshot an hour				
Toigo, Jon	12/1/2003						20 plus years of service
Toigo, Jon	12/1/2003						Can solve even the knottiest of data protection problems when used effectively
Toigo, Jon	12/1/2003						Does not afford continuous protection
Toigo, Jon	12/1/2003						Low price point
Toigo, Jon	12/1/2003				Revivo block-level data changes (hardware appliance)		
Toigo, Jon	12/1/2003	Connotes an on-going protection method close to synchronized					

Toigo, Jon	12/1/2003	Ideally would be handled at layer 7 of ISO model but strategies have bombed badly					
Toigo, Jon	12/1/2003	Journaling and forwarding I/O activity to an alternate storage repository					
Toigo, Jon	12/1/2003	Layer 7 CDP model introduce unacceptable latency					
Toigo, Jon	12/1/2003	Vendors: Veritas, Datacore, GlaconStore, Fujitsu Softeck, Symantec					
Toigo, Jon	12/1/2003	Work best lower down in ISO stack					

Appendix B: Definition of Terms

Agent: A software-driven process running on a communication or networking device that allows that device to participate in a management system (Toigo, 2003).

Alternate routing: Safety technique enabling communications to continue in the event of nod failure by allowing for data to travel on alternate paths through the network to arrive at the same location (Toigo, 2003).

Alternate site: Another facility, such as a commercial hot site or customer-owned second site, that will become the recovery site in the event of a disaster (Neaga, Winters, Laufman, 1997).

Archival backup: A data backup that will be stored for a long time (Toigo, 2003).

Back up: To make copies of important files in case the originals are damaged (Toigo, 2003).

Bandwidth: A characteristic of a communication channel that is the amount of information that can be passed through it in a given amount of time, usually expressed in bits per second (Toigo, 2003).

Business continuity management: A holistic management process that identifies potential impacts that threaten an organization and provides a framework for building resilience and the capability for an effective response that safeguards the interest of its key stakeholders, reputation, brand and value creating activities (BCI, 2005)

Business continuity planning (BCP): The advance planning and preparations which are necessary to identify the impact of potential losses; to formulate and implement viable recovery strategies; to develop recovery plan(s) which ensure continuity of organization services in the event of an emergency or disaster; and to administer a comprehensive training, testing and maintenance program (Doughty, 2000).

Business process: A group of related activities that support the successful operation of the business or its services (Neaga, Winters, Laufman, 1997).

Byte: The amount of memory space used to store one character, usually 8 bits.

Cold site: A recovery site equipped with sufficient environmental conditioning for a data processing infrastructure (Neaga, Winters, Laufman, 1997) such as air conditioning, electric connectivity, communication access, network connections, configurable space to accommodate the installation and operation of equipment by critical staff required to resume business operations (ED. Hiles & Barnes, 1999).

Continuous availability: The elimination or masking of both planned and unplanned outages so that no system outages are apparent to the end user (Neaga, Winters, Laufman, 1997).

Cost of downtime: The potential losses incurred as the result of the disaster and in recreating lost data (Krischer, 2005).

Crisis: An abnormal situation which threatens the operations, staff, customers or reputation of an enterprise (ED. Hiles & Barnes, 1999).

Critical data point: The point which data must be restored in order to achieve the recovery objective (ED. Hiles & Barnes, 1999).

Data currency: A measure of how close the restored data level matches the data level at the time of the disaster (Neaga, Winters, Laufman, 1997).

Data mirroring: A method of storage in which data from one disk is duplicated on another disk so that both drives contain the same information (Toigo, 2003).

Data processing: Computers and their supporting environment including hardware, software, data, networks and operators (Neaga, Winters, Laufman, 1997).

Data replication: A process that duplicates data to another location over a computer network in real time or close to real time (FFEIC handbook, 2003).

Data synchronization: The comparison and reconciliation of interdependent data files at the same time so that they contain the same information (FFEIC handbook, 2003).

Disaster: Any accidental, natural, or malicious event which threatens or disrupts normal operations, or services, for sufficient time to affect significantly, or to cause failure of the enterprise (ED. Hiles & Barnes, 1999).

Disaster recovery: Process of recovering from a major IT interruption (FFEIC handbook, 2003).

Disaster recovery plan: A plan to resume specific essential operations, functions or processes of an enterprise (ED. Hiles & Barnes, 1999).

Disk array: A group of disk drives that have been combined to appear as a single logical storage unit (Toigo, 2003).

Disk mirroring: A method of storage in which data from one disk is duplicate don another disk sot that both drives contain the same information, thus providing redundancy (Toigo, 2003).

Distributed architecture: A set of interacting computer systems situated in different locations (Toigo, 2003).

Downtime: The time during which a computer is nonfunctional because of problems with hardware, software or environmental failure (Toigo, 2003).

Enterprise: An organization, firm, business entity or establishment, government body department or agency, or business (ED. Hiles & Barnes, 1999).

Failure: The malfunction of a system or component; the inability of a system or component to perform its intended function (Toigo, 2003).

Hot site: A facility with sufficient hardware, communications interfaces and environmentally controlled space capable of providing close to immediate data processing support (ED. Hiles & Barnes, 1999).

Incident: Any event which may lead to a disaster (ED. Hiles & Barnes, 1999).

Incremental backup: A routine that makes it possible to backup only the files that have changed since the last backup, instead of backing up every file (Toigo, 2003).

Local area network (LAN): A communications system that links computers into a network, usually via a wiring based cabling scheme. LANs connect PCs, workstations and servers together to allow users to communicate and share resources and does not run on leased lines (Toigo, 2003).

Mirroring: A process that duplicates data to another location over a computer network in real time or close to real time (FFEIC handbook, 2003).

Mission critical: Any computer process that cannot fail during normal business hours; some computer processes require 100 percent uptime (Toigo, 2003).

Mission critical data: Data or information considered to be so important that its loss would cause grave difficulty to all or part of a business (Toigo, 2003).

Multi-site data replication: Saving changes on a regular or near-constant basis to files on multiple separate locals or remote arrays.

Network attached storage (NAS): This is the provision of storage in a form that is readily accessible on a network. A disk array storage system that is attached directly to a LAN. A typical NAS has a processor, an operating system and processes file I/O protocols (Toigo, 2003).

Off-site location: A storage facility located a safe distance from the primary facility which is used for housing recovery supplies, equipment and vital records (ED. Hiles & Barnes, 1999).

Orphan data: Data that has been entered at the primary data center but had not been transferred to the alternate data center at the time of the disaster (Neaga, Winters, Laufman, 1997).

Outage: Any period of time, within defined service hours, when the data processing service is not available to the users (Neaga, Winters, Laufman, 1997) which may result in the organization’s inability to provide services for some period of time (ED. Hiles & Barnes, 1999).

RAID (redundant array of inexpensive disks): A method of combining hard disks into one logical storage unit which offers disk-fault tolerance and can operate at higher throughput levels than a single hard disk (Toigo, 2003).

Recovery: The process of using alternate resources to restore data processing to an operable state after a failure (Neaga, Winters, Laufman, 1997).

Recovery point objective: The point in the business process to which data must be recovered after a disaster occurs (Krischer, 2005).

Recovery time objective: The length of time between when a disaster occurs and when the business process must be back in a production mode (Krischer, 2005).

Recovery site: A designated site for the recovery of computer or other operations which are critical to the enterprise (ED. Hiles & Barnes, 1999).

Recovery vendors: Organizations that provide recovery sites and support services for a fee (FFEIC handbook, 2003).

Remote mirroring: Mirroring disk writes to a duplicate disk array located at an off-site facility using WAN or Internet link (Toigo, 2003).

Remote site recovery: The ability to continue or resume processing of the critical workload at a remote site in the event of a primary outage (ED. Hiles & Barnes, 1999).

Remote tape vaulting: Writing data backups to a tape library or loader located at an off-site facility via a WAN or Internet connection (Toigo, 2003).

Replication: Saving changes to data on a regular or near-constant basis to files on a separate local or remote array (Toigo, 2003).

Resilience: The ability of a system or process to absorb the impact of component failure and continue to provide an acceptable level of service (ED. Hiles & Barnes, 1999).

Response: The reaction to an incident or emergency in order to asses the level of containment and control activity required (ED. Hiles & Barnes, 1999).

Resumption: The process of planning for and/or implementing the recovery of critical business operations immediately following an interruption or disaster (ED. Hiles & Barnes, 1999).

Server: A computer or other device that manages a network service (FFEIC handbook, 2003).

Snap-shot backup: Regular or incremental backups to disk that copy that take a point in time copy of a backup area (MacVittie, D. 2005 – from network computing)

Storage area network: A network comprising multiple hosts and storage peripherals, currently conceived as Fiber Channel/SCSI command Set based (Toigo, 2003).

System recovery: The procedures for rebuilding a computer system to the condition where it is ready to accept data and applications. System recovery depends on having access to suitable hardware.

T1: A dedicated phone connection consisting of 24 individual channels each supporting 64Kbits per second supporting a total of 1.544Mbits per second data transfer (Toigo, 2003)

Tape backup: the process of streaming your data to tape based storage (MacVittie, 2005)

Vaulting: A process that periodically writes backup information over a computer network directly to the recovery site (FFEIC handbook, 2003).

Wide area network: (WAN): A network which covers a larger geographical area than a LAN and where leased telecommunications links are implemented (Toigo, 2003).

Wide area network replication: Saving changes to data on a regular or near-constant basis to files on a separate remote array via a WAN interconnect (Toigo, 2003).

Bibliography

Alterio, J., (2001).Tragedy will sell companies on disaster recovery. Retrieved on November 7, 2005, from http://www.thejournalnews.com/newsroom/091601/16disasterside.html

Ayyagari, M., Beck, T., Demirgüç-Kunt, A., (2003). Small and medium enterprises across the globe: A new database. Retrieved October 11, 2005, from http://www.worldbank.org/research/projects/sme/abd.pdf

Baltazar, H., (2003). XOsoft gives tape backup a boost. *Eweek 20* (38) p60. Retrieved January 16, 2006, from Academic Search Premier database.

Baltazar, H., (2003). Keeping data safe and sound. *Eweek 20* (42) p43. Retrieved January 11, 2006, from Academic Search Premier database.

Baltazar, H., (2004).Softek aims for best replication on block. *Eweek 21* (3) p53. Retrieved January 16, 2006, from Academic Search Premier database.

Baltazar, H., (2005). CDP wares aid tape backup. *Eweek 22* pS1-S4. Retrieved September 29, 2005, from Business Source Premier database.

Britt, P., (2005) Taking steps for disaster recovery. *Information Today 22* (9) p1-21. Retrieved October 18th, 2005, from Academic Search Premier database.

Chevance, R., (2005). *Server architectures: Multiprocessors, clusters, parallel systems, web servers, storage solutions.* Massachusetts: Elsevier Digital Press.

Clark, E., (2003). Sundt construction nails down disaster recovery. *Network Magazine 18* (3) p54-58. Retrieved January 16, 2006, from Academic Search Premier database.

Connor, D., (2004). Backup-up protection on tap from storage vendors. *Network World 21* (41) p18. Retrieved September 10, 2005, from MasterFILE Premier database.

Connor, D., (02/2005). HP, others tackle data replication recovery. *Network World* p16. Retrieved October 3, 2005, from Expanded Academic ASAP database.

Connor, D., (06/2005). Storage conference focuses on recovery. *Network World 22* (24) p11. Retrieved October 2, 2005, from MasterFILE Premier database.

Connor, D., (09/2005). EMC upgrades windows data replication software. *Network World 22* (38) p14. Retrieved January 16, 2006, from Computer Source database.

CSU Writing Center, 2005. Writing guides: Conducting content analysis. Accessed online on Oct 1, 2005, at http://writing.colostate.edu/guides/research/content/contrib.cfm

Doughty, K. (2001). Introduction. In K. Doughty (Ed.), *Business continuity planning: Protecting your organization's life* (pp xi-xix) Florida: Auerbach.

Drew, R., (2005). Ready for trouble? Faced with potential catastrophe caused by anything from the weather to a malicious attack, companies' need to make sure their disaster recovery plans match best practices. *Computerworld 39* (17) p25-27. Retrieved October 3, 2005, from Expanded Academic ASAP database.

FCA, (2005). Business continuity: Essential practices for information technology exam manual IT section. Retried on November 15, 2005, at http://www.fca.gov/Download/itbusinesscontinuity.pdf

FFEIC, (2005). Federal Financial Institutions Examination Council, *Business continuity planning March 2003 IT examination handbook.* (n.d) Retrieved September 10, 2005, from http://www.ffiec.gov/ffiecinfobase/booklets/bcp/bus_continuity_plan.pdf

Fonseca, B., (2005). Veritas ties backup, storage. *eWeek 22* (4) p19. Retrieved January 16, 2006, from Academic Search Premier Database.

Hamblen, M., (2003). Keeping a safe distance: IP storage allows long-distance data replication and recovery, but the technology is emerging slowly. *Computerworld 37* (46) p42. Retrieved October 3, 2005, from Expanded Academic ASAP database.

Hiles, A. & Barnes, P. (Eds.). (1999). The definitive handbook of business continuity management. [Electronic version] Chichester: John Wiley & Sons.

Klein, K., (2005). Have you backed up your data today? *Business Week Online* 04/05/2004. Retrieved September 10, 2005, from MasterFILE Premier database.

Kova, J., (2005). A backup with integrity. *CRN* 1142, p3a. Retrieved September 29, 2005, from Business Source Premier database.

Krischer, J., (2005). Business continuity there is no one-size-fits-all business continuity strategy, so think of disaster recovery scenarios as modules that can be invoked depending on the situation; think modular for effective recovery plans. *Computer Weekly, Sept 20,* p32. Retrieved October 3, 2005 from Expanded Academic ASAP database.

Laufman, P., Neaga, G., & Winters, B. (1997). *Fire in the computer room, what now? Disaster recovery: Planning for business survival.* New Jersey: Prentice Hall.

Leedy. P.D. & Ormrod. J.E. (2001). *Practical research: Planning and design* (7th ed.). Upper Saddle River: Merril Prentice Hall.

Lipschutz, R., (2004). Protecting remote-office data conveniently. *PC Magazine* 23(19) p55. Retrieved January 16, 2006, from Academic Search Premier database.

Lunt, P., (2005). Better backup strategies. *IT Architect 20* (11) p4. Retrieved January 16, 2006, from Academic Search Premier database.

MacVittie, D., (2005). Continuous data protection: File-level restoration gets real. *Network Computing 16* (12) p10-13. Retrieved September 29, 2005, from Business Source Premier database.

MacVittie, D., (2005). Replication products ease backup pain. *Network Computing 16* (13) p79-80. Retrieved September 29, 2005, from Business Source Premier database.

Marks, H., (05/2005). Storage pipeline: Data-replication software. Retrieved September 18, 2005, from http://www.networkcomputing.com/story/singlePageFormat.jhtml?articleID=161600248

Marks, H., (9/2005). Escape the tape. *Network Computing 16* (18) p67-72. Retrieved September 29, 2005, from Business Source Premier database.

Marlin, S., (2005). Backup your work, one keystroke at a time. *Information Week* 1062 p38. Retrieved January 16, 2006, from Academic Search Premier database.

McCown, S., (2005). WANSynchHA protects databases, servers. *InfoWorld 27* (37) p54-55. Retrieved January 16, 2006, from Academic Search Premier database.

McKinstry, J., (2005). Data replication. *Computer Technology Review 25* (2) p10-12. Retrieved January 1, 2006, from Computer Source database.

McManus, D.J., & Carr H.H. (2001). Risk and the need for business continuity planning. In K. Doughty (Ed.), *Business continuity planning: Protecting your organization's life* (pp 3-10). Florida: Auerbach.

McMillen, J.H. & Schumacker, S. (1993). *Research education: A conceptual understanding.* New York: Harper Collins.

Mearian, L., (2005). Microsoft, Symantec ready CDP products. *Computer World 39* (40) p10. Retrieved January 16, 2006, from Academic Search Premier database.

Mearian, L., (2005). Disaster recovery works for some, but scope of calamity was difficult to prepare for. *Computer World 39* (36) p4-5. Retrieved September 18, 2005, from Business Source Premier database.

Pratt, M., (2005). Redefining disaster: Many CIOs have changed the way they think about disaster recovery. How about you?. *Computerworld* 39 (25) p39-40. Retrieved on October 3, 2005, from Expanded Academic ASAP database.

Stallings, W. (2001). *Business data communications* (4th ed.). New Jersey: Prentice-Hall, Inc.

Toigo, J.W. (2001). *Disaster recovery planning: Preparing for the unthinkable (3rd ed.)*. New Jersey: Prentice Hall, PTR.

Toigo, J.W., (2003). The joy of data replication in 2004. Retrieved on December 24, 2005, at http://searchstorage.techtarget.com/tip/1,289483,sid5_gci942767,00.html

Vancoppenolle, G. What are we planning for?. (1999). In A. Hiles & P. Barnes (Eds.), The definitive handbook of business continuity management. [Electronic version] Chichester: John Wiley & Sons.