CHAPTER 4

OBSERVATION AND DESIGN OF DATA BACKUP STRATEGY IN A BUSINESS ORGANIZATION

CHAPTER 4

OBSERVATION AND DESIGN OF DATA BACKUP STRATEGY IN A BUSINESS ORGANIZATION

4.1 INTRODUCTION

The main focus of this chapter is to perform a case study on the disaster recovery planning in a selected business organization. Risk analysis is implemented on the existing backup strategy in the organization, which include identification of threats and vulnerabilities. Sufficient solutions for vulnerabilities found are also provided.

Based on the observation, a local and remote data backup system is designed. The designed strategy has solved the existing vulnerabilities in the business organization. Besides, issues raised in system designing are also discussed at the end of this chapter.

4.2 OBSERVATION ON A BUSINESS ORGANIZATION

On the 8th April 2000, 9 morning, an interview was held with the director of SEA Automation Sdn. Bhd. (189347-T), Mr. Lim Sea Ping, about the existing backup strategy in their LAN. At the early session, the network architecture, backup strategies, backup storage devices were analyzed.

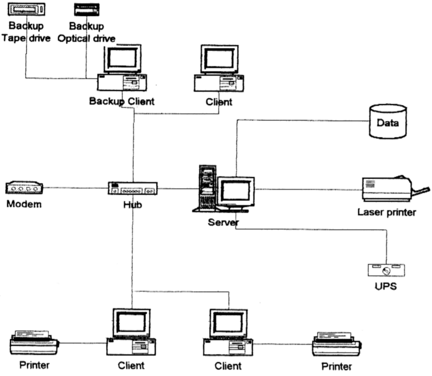

Figure 4.1: Existing Network Architecture and Backup Devices that are utilizing

4.2.1 EXISTING NETWORK ARCHITECTURE

Figure 4.1 illustrates a schematic view of the network architecture and backup strategy used in the target organization. It shows a radical view of a client-server environment. There are three basic components in this environment: client machines, server machines, and a network that connects all of the machines and organizes them into a working system.

The network is originally developed as a means for sharing data and expensive hardware resources such as printers, and for communicating with remote sites. With the evolvement of complex and varied workloads, client and server machines are divided accordingly to their functions. As the data used on these isolated machines grew larger and more important, simple storage devices were used to backup data. Normally, these were standalone tape cassette or optical disks devices.

4.2.2 EXISTING BACKUP STRATEGIES

In essence, data of the business organization will be stored in the server database system. Weekly backup procedure will be done at the backup client by copying the data to a tape or an optical storage device. In addition, a data-mirroring strategy is used in the server by duplicating another hard drive storage. The backup tapes and optical devices are then kept inside a fire proof safe.

In a sound business case, the automation of centralizing server backup may be cost savings. It funnels backups through the database server to backup client, where tape or optical disk backups are performed. Foremost, centralization in administration of server and single location of backup station would definitely minimize personnel force. The server backups would be near-line, so access would be fast and automated, providing quick and easy file recovery. Finally, the backup disk storage silo is a fireproof safe and may protect data integrity.

4.3 IMPLEMENTATION OF RISK ANALYSIS ON EXISTING NETWORK ENVIRONMENT

After observing the existing network architecture and backup strategies in the business organization, a risk analysis was implemented. The objectives of risk analysis are to identify and prioritize organizatíon's data assets, to identify vulnerabilities in network environment and to recommend the solutions for vulnerabilities found.

4.3.1 IDENTIFICATION AND PRIORITIZATION OF DATA ASSETS

In the client and server environment, data are created daily. As more and more applications are utilized, sensitive data may reside on the network. During the second session of the interview, identification of data in various forms is examined. Table 4.1 summarized the various forms of data together with its utilized application.

Table 4.1 Identification and Prioritization of Data Assets

Data Assets	Utilized Application	Prioritization
Account Payable	Accpac Plus – AP	Vital
Account Receivable	Accpac Plus – AR	Vital
General Ledger	Accpac Plus – GL	Vital
Inventory Control	Accpac Plus – IC	Vital
Sales Order	Accpac Plus – SO	Vital
Purchase Order	Accpac Plus – PO	Vital
Bank Records	Microsoft Excel	Vital
Payroll	Microsoft Excel	Vital
Customer lists	Microsoft Excel	Vital
Employee records	Microsoft Excel	Vital
Production data	Microsoft Excel	Critical
Financial data and budgets	Microsoft Excel	Vital
R&D data	Microsoft Word	Time Sensitive
Product and marketing plans	Microsoft Word	Time Sensitive
Executive correspondence	Microsoft Word	Time Sensitive
Quotation	Microsoft Word	Critical
Database System	Microsoft Access	Critical
Electronic Mail	Microsoft Outlook Express	Non-Critical

4.3.2 RISK IDENTIFICATION

Risk identification is one of the procedures that must be performed in risk analysis. It would identify vulnerabilities in the present network architecture. These vulnerabilities involve deficiency of network management, backup strategies and automation. Hence, possible solutions are suggested in favor of compensating the effects of vulnerabilities found.

In summary, the vulnerabilities found during the risk analysis are listed as below:

1. Centralization of server backup
2. Inadequate physical separation between backups and data center
3. Lack of automation control
4. Bandwidth limitation
5. Weak security

4.3.2.1 RISK: CENTRALIZATION OF SERVER BACKUP

As analyzed by Kevin Koski in *Centralizing Server Backups*:
"...Many large companies today are centralizing corporate server backups using a combination of software and hardware solutions. Centralization per se is worth considering; done wisely, it can provide cost and productivity benefits. But we've also seen that carelessly approached, centralization can jeopardize a company's survival, by crippling its ability to recover from a real disaster..."[35]

In this case, a single point of failure exists in the network, hence, it is vulnerable to a failure of server. Besides, a single point of failure also exists at the backup facility, hence, it is vulnerable to a failure of backup machine. This key mistakes which leads to network crashes are common to organizations of all sizes and it happened in AT&T's InterSpan frame relay network on April 13, 1998, as a result of a single event. [37]

4.3.2.2 SOLUTION: CENTRALIZATION OF SERVER BACKUP

The "Single point of failure" vulnerability may be easily solved by managing the network in a distributed manner. The proposed solution is a distributed backup strategy where the storage devices are distributed over the network, but with centralized control. Dr Peter Liu brought up this idea in his paper [26].

"...Since small capacity is no longer a limitation and the availability of carousels and jukeboxes allow for unattended backup, it seems logical to combine the strengths of the isolated backup strategy and the centralized storage strategy into a solution that we propose to call a "distributed backup"..."

4.3.2.3 RISK: INADEQUATE PHYSICAL SEPARATION BETWEEN BACKUP STORAGE AND DATA CENTER

Regional natural disasters like earthquakes, floods or hurricanes can wreak devastation across miles. Storage of backup tapes and disks in local fireproof safe are not recommended for disaster recovery planning. Minimum distance requirements of the off-site storage facilities must anywhere from five to 300 miles away.

4.3.2.4 SOLUTION: INADEQUATE PHYSICAL SEPARATION

Vulnerability of inadequate separation between backups and data center cab be solved by transferring data files to the off-site secured remote facility over the WAN. For data recovery purpose, clients can achieve off-site storage of these resources and restoring data back to the organization.

4.3.2.5 RISK: LACK IN AUTOMATION CONTROL

As is normal with human nature, the responsibilities of backing up data may be neglected. These responsibilities include administrating entire backup storage process, backing up data in scheduled manner and cataloguing and labeling of tapes and optical disks. Therefore, the existing backup strategy is deficiency in automation. Besides, services supported by the existing backup system are also inadequate, including end-user restores, hierarchical storage management, and data archiving.

4.3.2.6 SOLUTION: LACK IN AUTOMATION CONTROL

By using local and remote data backup system, cataloguing and labeling of tapes and optical disks may be eliminated. These have been replaced by the log files that will automatically document the backup progress showing time, date and file transfer. In addition, with the designed system, the local and remote backup process could be performed in an automated and scheduled manner. As a result, with the user-friendly backup application in local and remote data backup system, automation of backup system has been well refined.

4.3.2.7 RISK: BANDWIDTH LIMITATION

As indicated in Figure 4.1, all of the storage devices are attached to a backup client machine. Thus, most of the data to be backed up must travel across the network, excluding the data that held on the backup client machine. Severe traffic problems can result on the network especially when large amounts of data are involved. Bandwidth limitations become a severe handicap, which data compression can only partially relieve.[38]

4.3.2.8 SOLUTION: BANDWIDTH LIMITATION

If several machines are directly connected to storage devices, network traffic will be reduced considerably. Bandwidth problems will be relieved because data can travel directly to these devices rather than traveling over the local network to reach a storage device. Figure 4.2 illustrates the solution of network architecture.

4.3.2.9 RISK: WEAK SECURITY

Since most of the data must travel across the network, it is vulnerable to interception by unauthorized persons because data encryption is not performed in the local area network of the organization. However, these vulnerabilities occur only during a backup operation if an intruder intercepts it while it is traveling on the network to the backup client machine.

4.3.2.10 SOLUTION: WEAK SECURITY

Distributed backup strategy may be the solution for weak security. Data is also more secure when it can be backed up directly because it need not travel on the local network where it might be intercepted. Moreover, the time needed to encrypt and decrypt sensitive data is eliminated when data is backed up to a device directly. Hence, the security services can be emphasized during data transfer in WAN.

4.4 PROPOSED LOCAL AND REMOTE DATA BACKUP SYSTEM

Hence, after the disaster recovery planning case study on the business organization, a backup system will be designed. The main concern of the design is to overcome full range of possible disaster events and the extent to which each event can interrupt business operations and damage corporate data. In fact, requirements of local and remote backup system are to design:

1. a local backup module that suits well in local area network, implementing distributed backup strategy.
2. a remote backup module that backup data to the remote server over the Internet. This ensures that there is sufficient physical separation between the data centers and backup storage.
3. a scheduled backup module that activates backup process in an automated manner at specified time and date.
4. a documentation module that would record progress of backup regarding its time, date and data files. Both the schedule and documentation modules are designed to improve automation of backup

Moreover, a network architecture for local and remote backup system is also designed. In essence, it implements the local distributed backup strategy and remote data backup strategy using FTP. Therefore, centralization of server backup has been enhanced by distributed backup, which distributes the backup devices over the local area network. While FTP server is added for remote backup storage. Figure 4.2 illustrates the proposed network architecture for local and remote data backup system.

In conclusion, the proposed system is best suited for an environment with low communications costs, high labors overhead and increased data volumes. The client machines can easily, and cost-effectively, transfer their data files to the off-site secured remote facilities by simply connecting a corporate LAN to a communication line. In addition, a fault tolerance backup system has been established by implementing both the local media backup and remote backup approaches. Besides, with appropriate management in the network environment, this strategy will compensate the observed vulnerabilities.

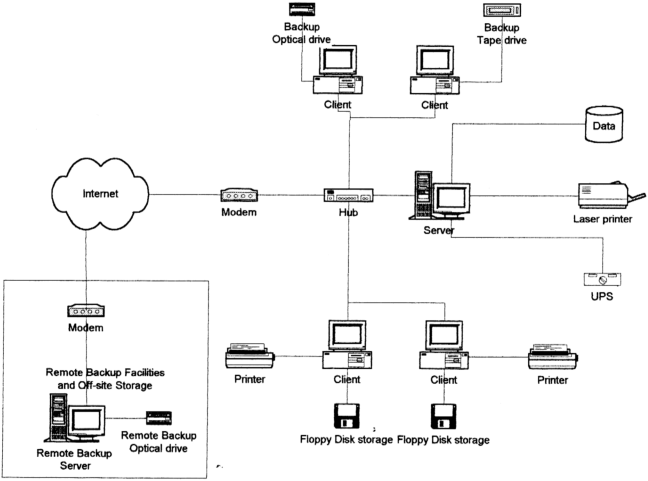

Figure 4.2 Proposed Network Architecture for Local and Remote Data Backup System

4.5 ISSUES RAISED IN SYSTEM DESIGNING

This section discusses the issues, which arises while designing the local and remote data backup system. These issues involve the benefits of distributed local backup strategy and remote data backup system.

4.5.1 BENEFITS OF DISTRIBUTED LOCAL BACKUP STATIONS

As illustrated in figure 4.2, it is called a "distributed backup" because it literally allows storage devices to be distributed throughout a network, while retaining the advantages of a centralized control. The backup's view is that any machine that is connected to a storage device is considered as a backup station rather than using a single backup station. Moreover, implementing a distributed backup results in the following benefits:

1. Better Network Performance:

   As the large number of devices connected to one station has been eliminated, network performance is improved.

2. Increased Flexibility:

   Storage devices are distributed over the backup stations. The placement of storage device are based on the necessity and suitability, where they are needed most or where they fit best, in view of their capacity and whether or not they can be used for unattended backup.

3. Time Efficiency:

   Data backup procedure is shorter when several distributed storage devices perform backup simultaneously. In addition, the time that data would normally take to travel to a storage device over a network is reduced.

4.5.2 REMOTE DATA BACKUP SYSTEM

Remote data backup complements the local data backup methods by protecting organizations from events that can destroy fault-tolerant storage devices. By compensating for the deficiencies of tape backup and fault-tolerant storage devices, remote backup system rounds out a complete solution for organizations requiring robust data protection capabilities as part of their disaster recovery planning.

As in the case of data exchange, remote data backup duplicates the full or partial contents of a disk from one system to a disk on another system by sending data over a local- or wide-area network. All databases and files that have been updated are copied from the source system disk.

4.5.2.1 BENEFITS OF REMOTE DATA BACKUP SYSTEM

1. Protect Against Region-Wide Disasters:

   Earthquakes, floods, hurricanes, extended power outages are the few examples of disasters that will affect entire sites or regions. Consequently, local data backup technologies can be rendered useless, even if fault-tolerant devices are employed. This is due by the destruction of primary operational system, redundant hardware and its backup data. Therefore, a copy of mission-critical data on a target server geographically separated from the source server should be maintained by remote data backup system in order to eliminate the risk that a site- or region-wide disaster will damage both original and backup data and systems.

2. Recover Data Quickly:

When a disaster strikes in an entire organization, both the data backup client or server machines are subject to failure. As a result, recovery data process may be restrained although data storage devices are available. On the contrary, remote backup may recover data in a faster pace because data is mirrored to an active system and critical information can be recovered rapidly from the on-line source. This requirement is particularly important to companies that need to restore operations within mandatory windows of time, sometimes as short as minutes or hours.