

# DATA BACKUP AND RESTORE IN THE NETWORK ENVIRONMENT

Ondrej Latta – Jozef Galanda

The article discusses the issues with data backup and recovery in computer network environment, with the focus on data protection against loss or damage due to minor failures in IT systems or other substantial catastrophes. The designed model fits in the aviation enterprise environment and meets the requirements to avoid any malfunctions, possible threats and consequences, in case of the system outage.

**K e y w o r d s:** Backup, data, recovery, disk subsystems

## 1 INTRODUCTION

Nowadays, every person uses IT technologies to make the work easier. Enterprise employees use computers in order to communicate with customers, to communicate between different departments within the same company or with external resources. They deal with various data like emails, text documents, photos, etc on daily basis. In most cases, data is the most valuable asset of the company. The data is nothing but zeros and ones. These zeros and ones carry the encapsulated information of financial documentation, building plans, business plans, business contacts, personal information of customers, who bought the flight ticket, or other confidential information. Certainly, each computer user experienced the situation of data loss, when mistakenly deleted the data or storage medium was seriously damaged, and all the necessary paperwork is lost. It's more than necessary to emphasize the necessity to backup your data and save them against damage or loss. In the area of information technologies, data are being saved on the storage medium and they can be recovered in case of any issues. The main function of storage system is to backup, save, protect, and restore data after a disaster occurrence (natural, operating).

## 2 DATA BACKUP AND RESTORE

From the strategic point of view data represent the most valuable asset and base of the business in company. The data can include business contacts, contracts with customers, accounting records, emails, etc. Therefore the loss of such kind of data or even temporary unavailability can mean huge financial loss, loss of competitiveness, in the worst scenario, suspend activities. Data backup can prevent the worst case scenarios of data loss, hardware and software failure, or other natural disasters. The data is being kept on a safe place and can be recovered anytime.

Data backup is process of copying and saving data from one medium to storage medium. It can be performed in two different ways; there can be backed up either chosen files or the disk image. There are following ways of data backup:

- Full backup – A complete backup of everything you want to backup
- Incremental – The backup software looks at which files have changed since you last did a full

backup. Then creates copies of all the files that are different from the ones in the full backup

- Differential - The backup software creates copies of all the files, or parts of files that have changed since previous backups of any type (full, differential or incremental).

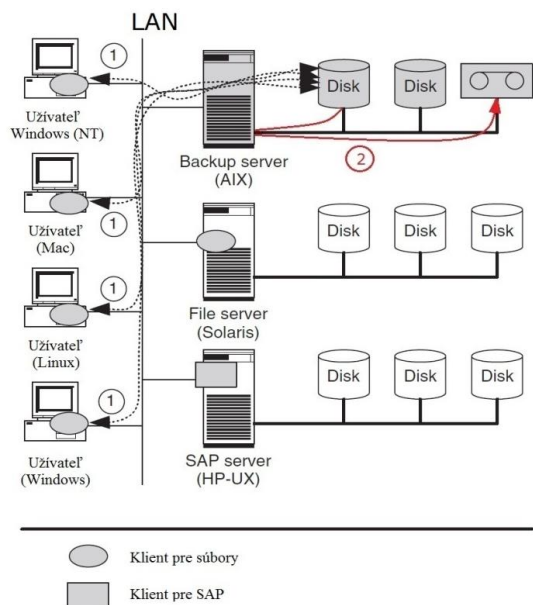
Data recovery is the process of restoring data that has been lost, accidentally deleted, corrupted or made inaccessible for any reason. In enterprise information technology (IT), data recovery typically refers to the restoration of data to a desktop, laptop, server, or external storage system from a backup.

The data recovery process may vary, depending on the circumstances of the data loss, the data recovery software used to create the backup, and the backup target media. For example, many desktop and laptop backup software platforms allow end users to restore lost files themselves, while restoration of a corrupted database from a tape backup is a more complicated process that requires IT intervention. Data recovery can also be provided as service. Such services are typically used to retrieve important files that were not backed up and accidentally deleted from a computer's file system but still remain on disk in fragments.

The current backup technologies can back up data from workstations and servers connected to the network, which operate on different operating systems and hardware environments of different manufacturers. Manual backup is becoming increasingly complicated for system administrators with the rising volume of data to be backed up in the shorter time period. In enterprise environment is important to perform routine operations on daily basis and automatically, with centralized management options.

## 3 CURRENT BACKUP TECHNOLOGIES AND TRENDS

Network backup is any backup system where the data to be backed up; it traverses the network to reach the backup media. Network backup typically requires client-server software architecture. The backup server resides on a centralized server and the backup clients reside on every system to be backed up. Advanced network backup systems can manage backup media which are also connected to the backup server via a network.

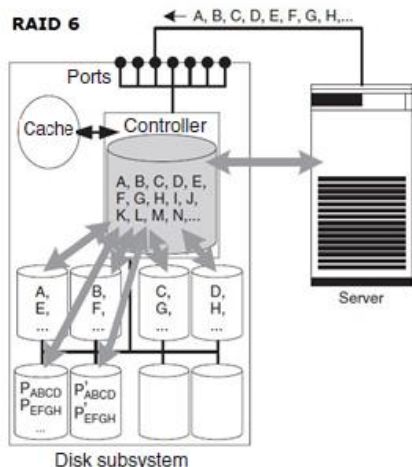


**Figure 1. LAN backup model**

Server stations use disk subsystems as a storage medium. Disks can be set up in different configuration; the most modern is RAID 6 configuration.

### 3.1 Hardware components and technologies

RAID 6 provides fault tolerance up to two failed drives. This makes larger RAID groups more practical, especially for high-availability systems. This becomes increasingly important as large-capacity drives lengthen the time needed to recover from the failure of a single drive.



**Figure 2. RAID 6**

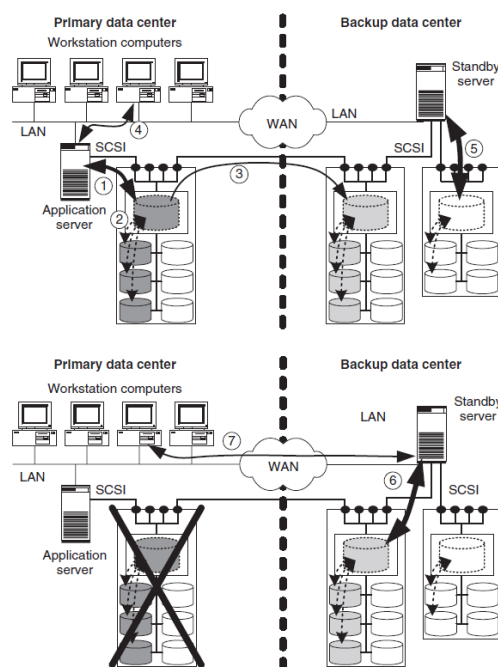
As a medium, SCSI (Small Computer System Interface) defines a parallel bus for the transmission of data with additional lines for the control of communication. The bus can be realised in the form of printed conductors on the circuit board or as a cable. The SCSI protocol defines how the devices communicate with each other via the SCSI bus. It specifies how the devices

reserve the SCSI bus and in which format data is transferred.

iSCSI is Internet SCSI (Small Computer System Interface), an Internet Protocol (IP)-based storage networking standard for linking data storage facilities. iSCSI is one of two main approaches to storage data transmission over IP networks; the other method, Fibre Channel over IP (FCIP), translates Fibre Channel control codes and data into IP packets for transmission between geographically distant Fibre Channel network.

Fibre Channel is currently the technique most frequently used for implementing storage networks. Interestingly, Fibre Channel was originally developed as a backbone technology for the connection of LANs. The original development objective for Fibre Channel was to supersede Fast-Ethernet (100 Mbit/s). Meanwhile Gigabit Ethernet and 10-Gigabit Ethernet have become prevalent in this market segment.

Serial-attached SCSI (SAS) is a method used in accessing computer peripheral devices that employs a serial (one bit at a time) means of digital data transfer over thin cables. Serial-attached SCSI offers advantages over older parallel technologies. The cables are thinner, and the connectors are less bulky. SAS offers data transfer rates in excess of 3 gigabits per second (Gbps) with potential rates of 10 Gbps or more. Devices that employ SAS are compatible with serial ATA (Advanced Technology Attachment) devices. In addition, SAS products are compatible with devices that employ earlier SCSI technologies.



**Figure 3. Remote mirroring**

Remote mirroring offers protection against such catastrophes. Modern disk subsystems can now mirror their data, or part of their data, independently to a second disk subsystem, which is a long way away. The entire

remote mirroring operation is handled by the two participating disk subsystems. Remote mirroring is invisible to application servers and does not consume their resources.

### 3.2 Software

Backup software in general is one integrated product that protects virtual and physical environments, simplifies both backup and disaster recovery, and offers unmatched recovery capabilities. Administration console makes it easier to set up backups, manage backup policies, perform disaster recoveries, and convert backups of servers to virtual machines for instant disaster recovery. Restore Wizards guide the administrator through a process that is designed specifically for the type of data that you want to restore.

Enterprise backup software solutions are realized by proprietary software; its license is not free. It works on specific standards and thus the compatibility can be limited. Top 5 enterprise backup software:

- Symantec Backup Exec 2012 [6]
- EMC NetWorker [7]
- IBM Tivoli Storage Manager [8]
- CA technologies ARCserve [9]
- HP data protector [10]

The source code of the open source software solutions stays available and it is licensed with an open-source license in which the copyright holder provides the rights to study change and distribute the software for free to anyone and for any purpose. List of TOP 5 open source backup software:

- Amanda Enterprise [11]
- BackupPc [12]
- EaseUS Todo Backup [13]
- Bacula Enterprise Edition [14]
- Cobian Backup [15]

Enterprise software as well as open source software provide with key features as centralized management, supporting various operating systems, backup and restore functions, supported media (hard disks, tape devices, cloud, network backup, data encryption, virtualized environment).

### 3.3 Cloud, SAN, NAS

Cloud backup service is a service that provides users with a system for the backup, storage, and recovery of computer files. Online backup providers are companies that provide this type of service to end users (or clients). Such backup services are considered a form of cloud computing. There are many products on the market – all offering different feature sets, service levels, and types of

encryption. Providers of this type of service frequently target specific market segments.

A storage area network (SAN) is a high-speed special-purpose network (or subnetwork) that interconnects different kinds of data storage devices with associated data servers on behalf of a larger network of users. Typically, a storage area network is part of the overall network of computing resources for an enterprise. SANs support disk mirroring, backup and restore, archival and retrieval of archived data, data migration from one storage device to another and the sharing of data among different servers in a network. SANs can incorporate subnetworks with network-attached storage (NAS) systems.

Network-attached storage (NAS) is a dedicated hard disk storage device that is set up with its own network address and provides file-based data storage services to other devices on the network. A network-attached storage device is attached to a local area network and assigned an IP address, allowing both application programming and files to be served faster because they are not competing for processor resources. NAS devices are usually configured with a web browser and do not have a keyboard or display. Network-attached storage consists of hard disk storage, including multi-disk RAID systems. NAS software can usually handle a number of network protocols. Many of NAS devices run their own proprietary operating system.

## 4 DESIGNING AND IMPLEMENTING THE SOLUTION IN AVIATION ENTERPRISE

Our main goal was to design an optimal solution of data backup for any airline. The solution should contain both software and hardware infrastructure. The infrastructure needs to be centralized in order to meet company requirements of easy-to-manage and 24/7 available solution.

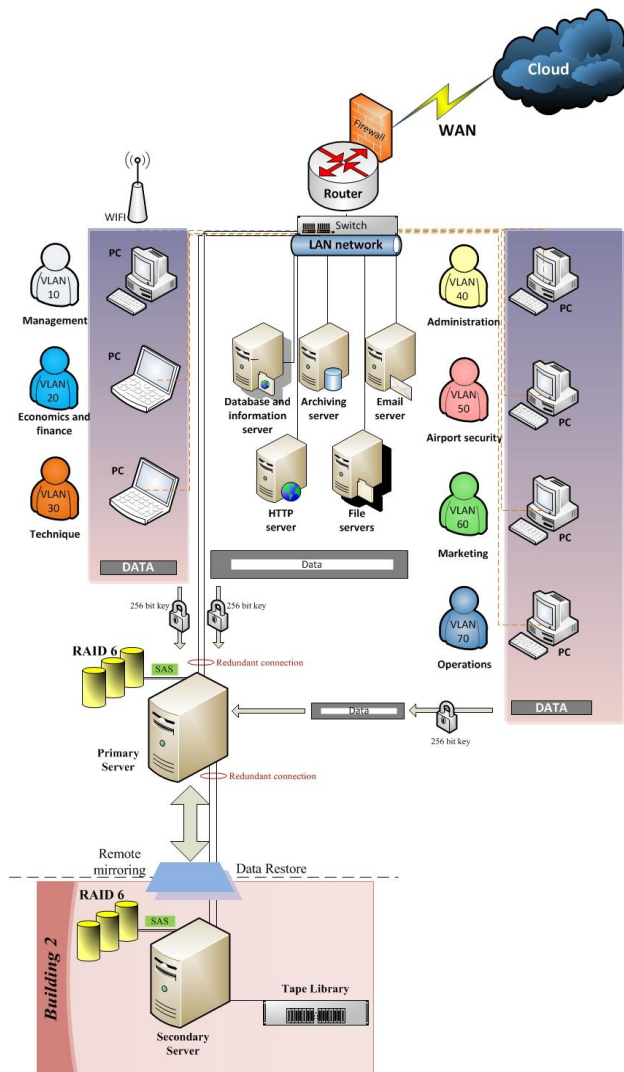
The proposal will be implemented in the existing aviation enterprise network infrastructure containing file, archiving, database, email, and HTTP servers. Employees manually save data on file servers at the end of their work shift. Before designing the tailored solution, we have to figure out few basic questions:

- What data and how much have to be stored?
- How often it's necessary to store the data?
- Is the proposed solution compatible with the existing network infrastructure?
- Who will be responsible for backing up and restoring the data?
- Is the solution centralized?
- Does the solution provide redundancy?

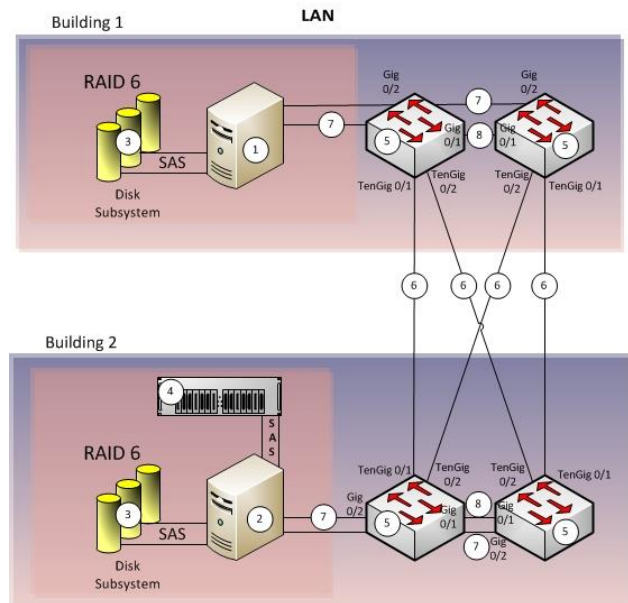
Answering these questions eases to pick up the optimal solution, which meets company requirements to fit in the existing infrastructure.

Users manually store data on the file server after finishing their work. Primary backup server copies data

on local disks. Data is transferred by using 256 bit AES encryption technology to secure data. Secondary backup server replicates data from the primary server. This is being performed by remote mirroring method. Server disks are connected via SAS interface, and are set up to create RAID 6 configuration. Tape library is connected to secondary server via SAS interface. Secondary server and tape library are located in different building, on remote site. Management is centralized and is being performed from backup servers. The data restore will be performed in the same way from centrally managed device, in case of any system failure.



**Figure 4. Proposal for data backup and restore in aviation enterprise**



**Figure 5. Detailed network equipment diagram**

The primary backup server (1) is located in building #1; secondary backup server (2) is located in building #2. Disks (3) are connected to both primary and secondary server via SAS interface; disks are deployed in RAID 6 configuration. Tape library (4) is attached to secondary backup server via SAS interface. Switches (5) are cross-connected with fibre cables (6), fibre cables are compatible with Gbic modules connected in switches; the optical connection provides the speed of 10 Gbit/s. Servers are attached via standard Ethernet interface; data transfer reaches the speed of 1 Gbit/s through network interface card installed in each server (7). There is no single point of failure; redundancy is designed at each point. Server is connected to switches via two separate links. Each switch is connected to another in Full Mesh topology; this topology avoids any case of data unavailability.

The proposal includes the software solution as well. Since the market offers proprietary and open source solutions, we decided to suggest both of them. Decision will be up to leadership. Anyway, proprietary and open source solutions dispose of the same features; they only differentiate in license (proprietary vs. open source).

We chose two software solutions: Amanda Enterprise and Symantec Backup Exec. Amanda Server version is compatible only with Linux Server version; client version can be installed in diverse operating system environment. What comes to Symantec software, server version can be deployed in Windows operating system; client version provides with the same resiliency as Amanda and can be used in various operating system environments.

#### 4.1 Pros and cons

The advantage of the proposed solution is in the high data availability of 24/7, because the data are backed up on primary server from existing production servers; all data from primary server is mirrored to secondary server and then archived to tape library what protects data. Redundant links offer even better and more stable connectivity. It avoids the single point of failure, for example when one cable from primary server to switch is broken there is still another one remaining. Another example is damage of primary server due to disaster like blaze; data is stored on a save place in building #2. Disks of both primary and secondary server store data based on RAID 6 configuration and thus data is protected against any case of loss. Software solutions Amanda and Symantec are compatible with the existing software platforms and ease monitoring and performing backup tasks in meaning of centralized management. All operations are triggered from one place – backup server. Again, data restore can be performed from the same place. The proposed solution provides us with opportunities for business growth and enhances resiliency of the network.

The disadvantage consists in high costs and additional expenses related to training..

#### 5 CONCLUSION

I hope that the information provided as above will help to introduce current technologies used in enterprise network environment to protect and store the data. It is obvious that such network topology as proposed can increase data availability, store it on a save place and ensure the business continuity in aviation enterprise.

#### BIBLIOGRAPHY

- [1] TROPPENS, Ulf - MÜLLER-FRIEDT, Wolfgang - WOLAFKA, Rainer - ERKENS, Rainer - HAUSTEIN, Nils: Storage Networks Explained, Basics and Application of Fibre Channel SAN,NAS, iSCSI, InfiniBand and FCoE, Second Edition. Chichester: John Wiley & Sons Ltd., 2009
- [2] SPALDING, Robert: Storage Networks: The Complete Reference. Berkeley: McGraw-Hill/Osborne, 2003. 529 s.
- [3] DURKIN, Jim - GOODMAN, John - HARRIS, Ron - FERNANDEZ-POSSE, Frank - REZEK, Michael - WALLACE, Mike: Building Multiservice Transport Networks. Indianapolis: Cisco Press , 2006. 539 s.
- [4] LONG, James: Storage Networking Protocol Fundamentals. Indianapolis, Cisco Press, 2006. 516 s.
- [5] FROOM, Richard – SIVASUBRAMANIAN, Balaji – FRAHIM, Erum: Implementing Cisco IP Switched Networks (Switch), Foundation Learning Guide. Indianapolis (IN): Cisco Press, 2010. 526 s.
- [6] Symantec Backup Exec: symantec.com. [online]. Mountain View (CA): Symantec Corporation, 2012. [cit. 2013-04-05]. Available on internet: <<https://www.symantec.com/backup-exec?fid=backup-exec>>
- [7] EMC NETWORKER: emc.com. [online]. Hopkinton (MA): EMC Corporation, 2013. [cit. 2013-04-05]. Available on internet: <<http://www.emc.com/backup-and-recovery/networker/networker.htm>>
- [8] Tivoli Storage Manager: ibm.com. [online]. Armonk (NY): IBM Corporation, 2013. [cit. 2013-04-05]. Available on internet: <<http://www-03.ibm.com/software/products/us/en/tivostormana/>>
- [9] CA ARCserve Backup: arcserve.com. [online]. Islandia (NY): CA, 2013. [cit. 2013-04-05]. Available on internet: <<http://www.arcserve.com/us/Products/CA-ARCserve-Backup.aspx>>
- [10] HP Data Protector software: hp.com. [online]. Palo Alto (CA): Hewlett-Packard Development Company, L.P, 2013. [cit. 2013-04-05]. Available on internet: <<http://www8.hp.com/us/en/software-solutions/software.html?compURI=1175640#.UWYNLnKnfvk>>
- [11] Amanda Enterprise: zmanda.com. [online]. Sunnyvale (CA): Zmanda - A Carbonite Company, 2013. [cit. 2013-04-05]. Available on internet: <<http://www.zmanda.com/amanda-enterprise-edition.html>>
- [12] BackupPc: backuppc.com. [online]. Sunnyvale (CA): Zmanda - A Carbonite Company, 2013. [cit. 2013-04-05]. Available on internet: <<http://www.backuppc.com/>>
- [13] EaseUS: todo-backup.com. [online]. Chengdu: CHENGDU Yiwo Tech Development Co., Ltd., 2013. [cit. 2013-04-05]. Available on internet: <<http://www.todo-backup.com/>>
- [14] Bacula Enterprise Edition v6: baculasystems.com. [online]. Yverdon-les-Bains: Bacula Systems, 2013. [cit. 2013-04-05]. Available on internet: <<http://www.baculasystems.com/products/bacula-enterprise>>
- [15] Cobian Backup: cobiansoft.com. [online]. Fishers (IN): Luis Cobian, 2012. [cit. 2013-04-05]. Available on internet: <<http://www.cobiansoft.com/cobianbackup.htm>>

#### AUTHORS' ADDRESSES

Latta Ondrej, Ing.  
ondrej.latta@gmail.com

Galanda Jozef, Ing., PhD.  
jozef.galanda@tuke.sk

Department of Aerodynamics and Simulations  
Faculty of Aeronautics, Technical University of Košice  
Rampová 7, 041 21 Košice