Chapter 7
High–End Storage

Isabel Schwerdtfeger
IBM, Germany

ABSTRACT

This chapter discusses the challenges high-end storage solutions will have with future demands. Due to heavy end-user demands for real-time processing of data access, this need must be addressed by high-end storage solutions. But what type of high-end storage solutions address this need and are suitable to ensure high performance write and retrieval of data in real-time from high- end storage infrastructures, including read and write access from digital archives? For this reason, this chapter reviews a few disk and tape solutions as well as combined disk- and tape storage solutions. The review on the different storage solutions does not focus on compliance of data storage management, but on available commercial high-end systems, addressing scalability and performance requirements both for online storage and archives. High level requirements aid in identifying high-end storage system features and support Extreme Scale infrastructures for the amount of data that high-end storage systems will need to manage in future.

INTRODUCTION

Today, due to huge data growth, more digitization needs, mobile devices producing more data, faster supercomputer power and applications, end users expect to store and archive their data in real time and have continuous real-time access to that data. According to recent studies by IDC, the worldwide data volume in the year 2020 will be 35 Zettabytes (Erickson, 2010). Data growth is expected to increase by factor 45 compared to the existing volume. One Zettabyte is a one followed by 21 zeros. Keeping up with this fast-growing demand of user behavior of "self-service" real-time data access, common and established IT-infrastructures are put under stress. Storing everything on disks becomes too large to create a backup, which leads to long backup windows and

increasing energy costs for the power consumption of disks systems. This becomes even more critical to evaluate, if some data has to be stored for at least 10 years or longer.

This chapter describes some high-end storage systems for real-time processing and will describe storage solution scenarios. First, there will be an overview of high-end disk-based storage systems and the general characteristics of storage infrastructures. This part includes a short overview over the developments of hard disk drives (HDD) and introduces the alternative of Solid-state disks (SSD). As it is important what type of hardware technology is supporting high-end storage solution, it is also important to ensure data integrity in storage solutions; file system programs play a vital part of a storage solution for data integrity. A file system is necessary to enable access to data by file name or directory and needs to be able to directly access data regions on a storage device (Wikipedia.org, 2012). For the purpose of this analysis two worldwide well-known file systems programs, as characterized as distributed parallel fault-tolerant file systems, IBM General Parallel File Systems (GPFS) and Lustre, an open source file system, will be shortly described.

The second part covers high-end tape-based storage systems, while the third part discusses combined disk- and tape-based solutions, including high-end storage software for archiving and hierarchical storage management (HSM) solutions. As for Future analysis, it will be discussed how the project approach on "Pergamum", an approach for a only Disk-Based Archival Storage, is a viable alternative to Tape-only based for archival storage.

OVERVIEW OF HIGH-END STORAGE AND SOLUTIONS

High-End Disks and Disk Subsystems Overview

High-end disk systems exist in organizations and companies that need high performance and huge storage capacity. These types of complex storage infrastructures need solid management and often include features such as replication, virtualization and support for multiple storage tiers (Perkins, 2007). International Business Machines Corp. (IBM) has a broad range of storage systems available, addressing various needs for storage infrastructures, i.e., for open systems and optimized for z/OS and System i, for block-based or file-based storage infrastructures. The strategy of the IBM System Storage products is aligned to the need of delivering the right system "fit for purpose" for any workload.

High-end storage components are subject to the same characteristics as any storage solution. Therefore, the main characteristics of storage infrastructures are:

- Availability of overall storage infrastructure
- Backup and Recovery policies as well as existing data centre infrastructures to be included into the overall IT environment,
- Systems Management and Service Management tools in use,
- Performance criteria to be met by desired trough-put and end user expectations,
- Capacity planning and sizing to meet expected needs in data growth,

Related Content

Where Do All My Keys Come From?

Andreas Daniel Sinnhofer, Christian Steger, Christian Kreiner, Felix Jonathan Oppermann, Klaus Potzmaderand Clemens Orthacker (2018). *Solutions for Cyber-Physical Systems Ubiquity (pp. 278-300).*

www.irma-international.org/chapter/where-do-all-my-keys-come-from/186911

Using Executable Slicing to Improve Rogue Software Detection Algorithms

Jan Durand, Juan Flores, Travis Atkison, Nicholas Kraftand Randy Smith (2011). *International Journal of Secure Software Engineering (pp. 53-64).*

www.irma-international.org/article/using-executable-slicing-improve-rogue/55269

Object-Role Modeling: Principles and Benefits

Terry Halpin (2010). *International Journal of Information System Modeling and Design (pp. 33-57).*

www.irma-international.org/article/object-role-modeling/40952

Towards Tool-Support for Usable Secure Requirements Engineering with CAIRIS

Shamal Failyand Ivan Fléchais (2012). *Security-Aware Systems Applications and Software Development Methods (pp. 269-284).*

www.irma-international.org/chapter/towards-tool-support-usable-secure/65853

A System for Predictive Data Analytics Using Sequential Rule Mining

Sandipkumar Chandrakant Sagare, Suresh Kallu Shirgaveand Dattatraya Vishnu Kodavade (2020). *International Journal of Software Innovation (pp. 78-94).*

www.irma-international.org/article/a-system-for-predictive-data-analytics-using-sequential-rule-mining/262101