

ENTERPRISE STORAGE RESOURCES OPTIMIZATION

Ahmed Saad Alsalam

Published on April 30, 2020

Abstract: Big data idea is common in most fields. Analyzing, managing, and storing large amounts of data has become a competitive advantage for most enterprises due to the criticality of data as a constituent of an enterprise. However, this competitive advantage also places the data at risk of loss. The need to analyze and store vast amounts of data increases the vulnerability of the data. A deeper understanding, therefore, on how to prevent loss of stored data improves the efficacy and promotes customer preference for an enterprise.

Keywords: storage capacity, capacity optimization methods, cloud, optimization.

1. INTRODUCTION

Data storage optimization methods primarily involve a reduction in the amount of storage space consumed by data. These methods can affect overall storage and system performance. It is, therefore, crucial to understand the enterprise environment before deploying a data optimization method. Techniques involved include thin provisioning, delta snapshots, data deduplication, and compression.

2. BODY

Compression is the encoding of data to reduce its size. There are two types of compressions. Lossy and lossless compression. In lossy compression, loss of a portion of stored data that is considered less critical from the original data. However, the remnant data should be able to retain its integrity. On the other hand, lossless compression ensures that there is no content loss from the original content. The original content can be easily reconstructed from its compressed form in lossless compression. Compression enables faster transmission of data over networks through the consumption of less bandwidth. Delta snapshot, on the other hand, is a data recovery technique applied in storage. It enables a computer to retain a set of data at a particular instant in time. It only stores the blocks of data that are different from an existing complete copy of the data. Most snapshots comprise of a base file and a "delta" file where the delta file enumerates the alterations of the base file until another instant in time.

Thin provisioning involves the flexible allocation of storage space among multiple users based on the least amount of space that is required by each user at a given time. It allocates storage space on demand rather than the reservation of large amounts of storage space per user. Data deduplication can be divided into two: subfile deduplication and single-instance storage. Subfile deduplication functions at a finer granularity than an entire data set while separate instance storage involves the elimination of redundant data copies, thus reducing storage overhead.

Deduplication and compression have several similarities and differences. For instance, both methods optimize storage aptitude, consume system resources, and depend on data patterns. However, some sets of data can be effectively optimized by one over the other. Despite their difference, deduplication and compression can be used together in cases where data pattern exists. Sub-LUN-Auto-Tiering is a method that involves the storage of data sets in storage mediums, depending on their frequency of access.

COM Method	Physical Capacity Saving (%)	Idle/Active Power Saving	IOPs 70-30 R-W Workload
Compression	50%	38%/35%	0.25X
Thin Provision	50%	38%/35%	0.50X
Thin Provision with De-duplication	75%	57%/68%	0.20X

The table above outlines the benefits of using Capacity Optimization Methods (COM).

The cloud has become the primary way of optimizing storage capacities. Uploading data to the cloud has several advantages such as long term retention, disaster recovery, and consumption of less physical disk space. Optimization through the cloud has increased storage and backup efficacy. Furthermore, it significantly reduces administrative costs, operational costs, and the overall cost of storage. Cloud will be advantageous even to buyers as they will not need to purchase hard disks and other data storage disks to store data when they can upload to the cloud for free.

3. CONCLUSION

In conclusion, optimization has a positive impact on both storage operations and customer budget. The ability of large amounts of data to occupy less space saves on operational, administrative, and purchasing costs. Cloud computing has made data retrieval easier, faster, and convenient. Enterprises have used capacity optimization methods to become more effective in their operations.

REFERENCE

- [1] The Green Grid, White Paper #58- The Green Grid Data Center Storage Productivity Metrics (DCsP)
[https:// www.thegreengrid.org/en/Global/Content/white-papers/WP58-DCsP](https://www.thegreengrid.org/en/Global/Content/white-papers/WP58-DCsP)