

Active Storage

Description: With the rapid pace in which data is proliferating throughout business and society, it is exposing the limitations of current storage solutions available to organizations which are both costly and inefficient storage and retrieval frameworks.

How can organizations actively store data for immediate access and reduce storage costs? Active retrieval of compressed data is a process that is costly in terms of both capability and processing.

Challenge:

Nearly every digital transaction today is captured and retained in some form for some duration by organizations and becomes part of that organization's most valued assets. To analyze that data to understand competitive position, internal efficiency and productivity, and customer behavior, is now more vital than ever. More data equals expanded storage and increasingly sophisticated information technology (IT) to maintain and manage it. The mission today of the Analytics industry is to provide for both timely and comprehensive analysis. Indeed, the frequent analysis of detailed data has totally eclipsed the periodic review of sampled data.

Cloud-Edge Active Storage provides a unique data storage solution with highest level of compression, but at the same time keeping the data proactive for application consumption.

Traditionally storage has been viewed either as block, file or object type storage. Data stored in either of these formats is considered passive from the source system(s) in which it originated and is relatively distant data. Passive data in the block, file, and object storage formats requires complex post storage operations to become consumable again. Due to its passive nature, it is remotely stored. The alternative to passively stored data is an RDBMS system, which become performance and cost prohibitive for increased data volume requirements. Even with proliferation of the recent big data technology solutions, these challenges continue to exist. This is due to limited functionality provided by the big data technology solutions to optimize and process the large volume of disparate data.

Action with Cloud-Edge:

Cloud-Edge provides an industrial grade storage platform for transactional and post transactional data. Cloud-Edge's Active Storage Platform (ASP) provides a new and innovative set of storage and compression features. Cloud-Edge provides innovative features to address the limitations with the current storage solutions available in the market today, which are Compression, Sorting, Scanning and Indexing.

Compression:

Cloud-Edge compression algorithms have the ability to compress and decompress data 'on-the-fly' with ratios that often exceed 20:1 and average in the 9:1 range. Proprietary techniques force compression levels comparable with columnar data stores while still storing data in complete rows. Such compression enables the data to be collected and analyzed at a detail level without losing performance. While the Cloud-Edge Active Storage compression algorithms result in minimal direct performance improvement, they accomplish two important things:

- 1.They improve the overall machine workload output (by better balancing I/O and CPU throughputs) and,
- 2.They provide huge storage savings with a resulting myriad of side benefits.

Sorting:

Database engine sorting has traditionally been a weak link in data processing solutions and has opened the market to third-party solutions offering add-on products to address the sorting challenge with current RDBMS platforms. Cloud-Edge's sorting algorithms eliminates the need for third party solutions by providing a higher performing 'sort solution' built into our relational data engine kernel. With Cloud-Edge heuristic algorithms linear sorting is initiated, providing significant performance gains when compared to other third party solutions.

Scanning:

The Cloud-Edge's Active Storage Platform (ASP) Engine uses an integrated RAID technology to move and filter data, in parallel, to and from disk I/O subsystems. Virtual database volumes are defined, allowing easy tuning of desired performance levels for any set of database tables. The scanning algorithms work closely with other low-level algorithms to keep pipelines and caches loaded efficiently.

Indexing:

The dynamic Cloud-Edge indexing algorithm utilizes the parallel processing of our scanning and sorting algorithms to rapidly create indices, which increase the performance for queries. Multiple indices can be created simultaneously with a single data scan, allowing for rapid data update processes. Raw data used by the indices are heuristically examined to create a dynamic index structure. The system automatically modifies its index structure as needed, to achieve and maintain an optimal memory usage and one-disc-access performance. All data is accessed in the compressed state. Because of our economy of RAM usage, hundreds of indices can be kept open simultaneously.

Engine Speed:

In recent tests, the Cloud-Edge technology performed a full, all-columns scan of 100 million rows (100 bytes wide), executing complex calculations on every single row (all the columns, not just an index or extract), in less than 8 seconds. That rate – roughly 12 ½ million rows per second – scales linearly, even if the scan were increased to 100 billion rows. Using the TPC-H industry benchmark, Cloud-Edge (ASP) can load the entire 100 GB database in less than 14 minutes (on a mid-range server – older, slower, and with less memory than the state-of-the-art hardware used by industry benchmark leaders SQL-Server and Sybase). It should be emphasized that Cloud-Edge (ASP) does not require any of the preprocessing of data during the load phase that other solutions require. For example, there is no need for optimization of a pre-selected set of access paths.

Result and Business Advantage

Traditional and some current storage systems are inadequate for the challenges associated with the current proliferation of data, both from a financial and operational perspective. With this type of growth and associated costs, organizations now have a new option to consider in managing storage and retrieval of their data assets with the Cloud-Edge Active Storage Platform (ASP).

