

Paradigm shift in the storage environment

 By [Eran Brown](#)

5 Apr 2019

As Artificial Intelligence (AI) develops, the way we store data also promises to become smarter - and more efficient, flexible, and cost-effective for businesses. Companies in South Africa and across the world that want to reliably store their ever-increasing amount of data (in the growing petabyte range) now have a large - and sometimes confusing - choice. So how can AI disrupt the current storage environment?



Eran Brown, CTO for Infinidat, EMEA

In principle, IT managers want data to be stored on high performing data carriers in order to ensure rapid access to the information at all times. But how sensible is it to store all data on the most powerful media at all times?

For example, Flash is much faster than Near-Line Serial Attached SCSI (SAS) drives, however, it is also much more expensive. It makes little sense to store all your data on Flash, since most data is not used often.

In addition, there is data that must be stored for an extended period of time due to compliance rules, yet there is other data that may need to be accessed on a more regular basis such as for the preparation of long-term analyses. Even backup files do not have to be stored on Flash as they only come into play when restoring a data set.

IT managers must therefore keep a constant eye on their data strategy in order to determine the optimal storage media for each application.

Time consuming manual approaches in use

Up to now, predefined policies have been used to determine exactly what data is stored where. Policies are established at the outset when the corresponding structures are created. They then remain largely the same, even if minor modifications take place during operational activities. Yet the amount and above all, the nature of the data is changing rapidly.

In the past, data was largely standardised due to the limited capacities and capabilities of the IT systems of that time, but today things look different. The constant manual adjustment of policies is becoming more and more complex and increasingly ties up personnel who can no longer perform other important tasks to the full extent. More complex data structures require more regular adjustments as the wrong choice of storage location can either burden the budget by using costly storage for irrelevant data or by disrupting operations through slower access to relevant data.

Groundbreaking solution using AI

So, how can this dilemma be solved? One way out is to use AI. With an automated method, adjustments can be made second-by-second without the need for manual intervention, allowing companies to use more cost-effective storage. Using machine learning, an AI engine can evaluate user behaviour and the nature of access to data and assign the storage location accordingly.

In addition, it can project the patterns according to which accesses need to take place for future usage behaviour. This can also be used to make forecasts of the memory required and the performance required in the future, which can also be reflected in infrastructure and budget planning. An important goal here is to prevent the use of unnecessary resources.

Smart decisions via Neural Cache

For example, AI can be used via a neural cache, a technology that delivers lower latencies than flash by leveraging smart software algorithms. The machine learning algorithms scan the data pool and analyses data patterns to find hidden correlations.

As a result, it decides which data is relevant for immediate access by applications or the user directly. Frequently used data is automatically stored in Random Access Memory (RAM) which is faster than Flash. Next is the “warm” data, which is stored in Flash, and the less frequently used data is stored on Near-Line SAS drives, which are much more cost-effective.

Lower latency and accelerated operations

In a storage array that combines Dynamic RAM (DRAM), Flash media and near-line SAS drives, the neural cache reduces latency and accelerates read/write access. Most applications are transactional. requiring at least two separate Input/Output (I/O) operations.

One operation is always used to integrate the transaction to the logs, the other for the actual write operation of the data. This means that latencies can have an exorbitant effect on performance. Response times of the metadata layer thus affect the maximum performance of the application. Both read and write operations that is, insertions, changes, and deletions from the metadata structure, are processed with the same latency time.

These operations are performed without pre-processing such as pattern removal, compression, or encryption directly in the DRAM of the storage server. Meanwhile, a second copy of the write operation is made in the DRAM of another storage

node with low latency RDMA and only then is a confirmation sent to the host.

Writing directly to the DRAM connected to the server's Central Processing Unit (CPU) results in lower overall latency than directly accessing an external flash device. In addition, the use of a single large memory pool for accepting write access - unlike traditional architectures where the write cache is divided into smaller sections - ensures that larger write bursts can be maintained.

Data that changes frequently can be overwritten with DRAM latency, allowing Neural Cache to intelligently decide which data blocks can be stored on which media. The longer retention of the data in the write cache means that CPU and backends are relieved. The Neural Cache can also accelerate read operations by holding the most active data in the DRAM.

AI builds its experience by analysing large datasets of data and identifying patterns respectively features. It helps IT managers reduce their storage spending - which is already a top line-item in their budgets - and frees money to invest in innovation and transformation.

ABOUT ERAN BROWN

CTO for BMEA at Infinidat

- How agile data infrastructures can save the retail sector - 31 Jul 2020
- How can IT help beat Covid-19? - 4 May 2020
- A look at data over the next decade - 24 Jan 2020
- Elastic infrastructure is key to meeting current and future data storage demand - 12 Dec 2019
- Is AI the key to managing rampant data storage costs? - 23 May 2019

[View my profile and articles...](#)

For more, visit: <https://www.bizcommunity.com>