Transparent Compression using NoLoad® Computational Storage Technology

NoLoad provides hardware-based Transparent Compression that enables increased capacity (lower $/TB) without sacrificing performance

Executive Summary

- Eideticom's NoLoad Transparent Compression solution requires no application changes, ties directly into any File System and achieves line-rate compression while using 70% less CPU.
- NoLoad Transparent Compression increases data stored on SSDs by up to 10x through industry-leading Compression technology and higher SSD utilization.
- NoLoad provides better storage scalability, longer-lasting NVMe SSDs, and more efficient CPU utilization.
- NoLoad delivers higher performance by transparently offloading the host CPU (70% less CPU utilized) from compute intensive compression operations, providing vastly improved application performance and superior Quality of Service (QoS).
- NoLoad accelerates a range of database applications such as Hadoop, RocksDB, Cassandra and MySQL by leveraging hardware-based computational offloads with unmatched scalability, performance and efficiency.
- NoLoad drives best-in-class power efficiency (60% lower power) as NoLoad Transparent Compression is dramatically more CPU efficient than software alternatives.
- NoLoad's NVMe compliant interface provides seamless integration for all CPU platforms and has been validated on Intel, AMD, ARM and IBM Power8/9 CPUs.
- NoLoad enables Reduced Storage Cost (90% lower total cost) by increasing storage capacity, reducing server count, optimizing I/O and maximizing storage lifetime.

1 Introduction

In this white paper, we discuss how Eideticom's NoLoad Computational Storage solutions can be used to build performant and cost-effective systems. We outline how to tie our NoLoad Computational Storage technology into the software of the Linux filesystem and the benefits of such a tie-in by demonstrating performance gains for various applications.

Database Acceleration, High Performance Computing (HPC) and many other applications rely on an underlying filesystem to store data and to retrieve it for analysis in a performant manner. Since the compute servers in these systems can generate tens of gigabytes of data a second, the filesystems used must be both scalable and performant. Failure to do so can result in bottlenecks that do not allow the system to operate at its full potential.

In this white paper, we will discuss Eideticom's NoLoad® Transparent Compression solution and how it accelerates data center infrastructure, ensuring affordable scaling and dramatically lowering cost.

2 NoLoad Overview

NoLoad® is an NVMe Computational Storage Processor purpose-built for the acceleration of storage and compute-intensive workloads. NoLoad eliminates performance bottlenecks by embedding processing power directly into the storage fabric. NoLoad takes the form of an FPGA accelerator card that performs a variety of storage-centric offload functions while presenting an NVMe compliant PCIe interface to the host.

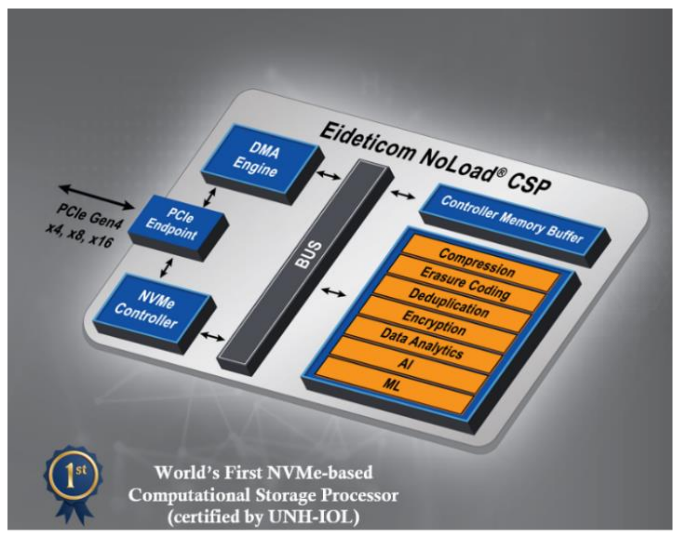

Figure 1: Eideticom's NoLoad

The NoLoad software framework was developed to allow applications such as Databases (Hadoop, RocksDB, Cassandra and MySQL) to offload critical storage tasks to the NoLoad. This offloading leads to improved performance and efficiency and reduced costs for the storage system. Eideticom's NoLoad supports several acceleration functions, ranging from Compression, Encryption, Erasure Coding, Deduplication and Data Analytics.

3 Why NVMe: NoLoad NVMe Computational Storage

Customers are deploying more and more accelerators in data-centers and on the edge and the benefits in terms of performance and efficiency are evident. However accelerators can be hard to consume, especially at scale. Eideticom are making accelerators easier to consume by leading an industry-wide effort around using NVMe Express® for computation as well as storage. NVMe provides a performant data-path and a well understood control and management path. It also aligns with a variety of form-factors that make sense for accelerators.

NoLoad presents acceleration functions to a host via a standard-compliant NVM Express PCIe interface. There are several reasons Eideticom chose an NVMe interface instead of a proprietary PCIe interface, our top six are:

- The NVMe driver is inbox in all major operating systems and the Linux NVMe driver is tuned for latency, performance and efficiency. Competing accelerator solutions utilize custom drivers which require porting or may not even work under all operating systems.
- NoLoad's NVMe compliant interface provides seamless integration for all CPU platforms and has been validated on Intel, AMD, ARM and IBM Power8/9 CPUs and at PCIe Gen 3 and Gen 4 rates.
- By choosing NVMe, we can leverage a rich ecosystem of open-source tools. In addition to drivers, this includes user-space management tools (e.g. nvme-cli) and user-space frameworks like SPDK.
- Aligning with NVMe allows us to implement and deploy on any PCIe enabled FPGA card and use a common software stack across all form factors.
- NVMe has a rich management specification we leverage to manage NoLoad at scale.
- Alignment with the NVMe Computational Storage standardization work being conducted by SNIA and the Technical Working Group of NVM Express.

4 NoLoad Hardware

Eideticom's NoLoad Computational Storage solutions are deployed on a wide varity of PCIe-enabled FPGA cards and form factors.

- The Samsung SmartSSD® (NoLoad SmartSSD) is a PCIe Gen3x4 Computational Storage Drive (CSD) developed by Samsung and integrates Eideticom's NoLoad Computational Storage technology. This solution provides 4TB of persistent data storage and computational services in a 2.5-inch (U.2) form factor. A typical 2 Rack Unit (2RU) server can accommodate 24 NoLoad-SmartSSDs.
- The Xilinx Alveo U50 (NoLoad U50) is a PCIe Gen3x16, or dual Gen4x8 FPGA Accelerator card developed by Xilinx and integrates Eideticom's NoLoad Computational Storage technology. This solution provides computational services in a Add-In-Card (AIC) form factor and is considered a Computational Storage Processor (CSP). A typical 2 Rack Unit (2RU) server accommodates two NoLoad-AICs.
- The BittWare 254-U2 and 250-U2 (NoLoad U2) are PCIe Gen4x4 and Gen3x4 FPGA Accelerator Cards developed by BittWare and integrates Eideticom's NoLoad Computational Storage technology. This solution provides computational services in a 2.5-inch (U.2) form factor and is considered a Computational Storage Processor (CSP). A typical 2 Rack Unit (2RU) server accommodates 24 NoLoad-U2s.

Figure 2: NoLoad enabled Hardware Platforms

5 NoLoad Software

Eideticom's rich suite of software enables the connection between acceleration functions on the NoLoad enabled FPGA cards and applications running on a host CPU. Eideticom provides end-to-end computational storage solutions to unlock the full potential of our customers flash-based storage investment.

5.1 NoLoad Filesystem (NoLoad FS)

NoLoad FS is a Linux-based stacked filesystem that can offload key computation tasks like compression onto Eideticom's NoLoad devices. NoLoad FS resides between the applications and the customer's underlying filesystem (e.g. XFS or Ext4). The NoLoad Software Stack is shown in Figure 3. NoLoad FS provides several advantages:

- **Application Transparent:** NoLoad FS enables storage services like compression to be deployed with no application changes i.e. acceleration without application changes.
- **Lower Filesystem Agnostic:** Customers can deploy NoLoad FS and continue to use their preferred production filesystem. Customers do not have to modify their existing filesystem and can run NoLoad FS over this existing filesystem.

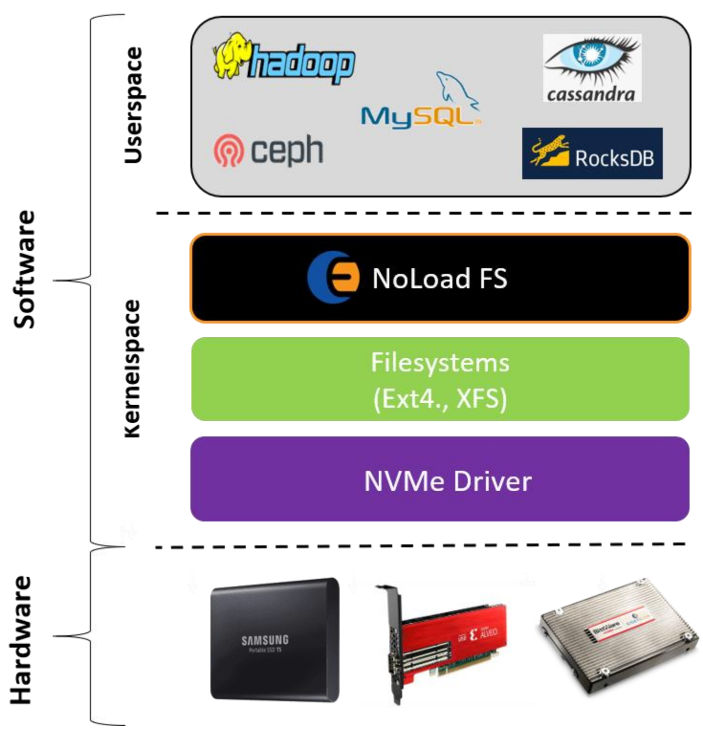

Figure 3: NoLoad Software Stack

6 NoLoad End Solutions

Eideticom provides end-to-end computational storage solutions for a range of applications, such as Database Acceleration and High Performance Computing (HPC). In this white paper, we showcase compression and decompression offload to improve performance and cost-efficiency for these applications. In this section, we provide baseline performance results for NoLoad compression and application performance for Transparent Compression in database applications.

6.1 NoLoad Compression

To establish baseline performance we measured NoLoad compression and decompression performance outside of NoLoad Filesystem. In Figure 4, we present compression and decompression results for several datasets associated with different application spaces. The server configuration was an AMD EPYC 7402 running Ubuntu 18.04.

For each algorithm, we present:

- Compression Ratio (CR)
- Compression Efficiency - compression input throughput per CPU core (MB/s/core)
- Decompression Efficiency - decompression output throughput per CPU core (MB/s/core)

We present results for NoLoad compression, a high throughput software compression algorithm (lz4-1), and a high-compression software algorithm (gzip-9).

The main conclusion from Figure 4 is that NoLoad-based compression **can provide significant Compression Ratio (CR) advantages over lz4-1, while achieving 8-10 times better efficiency for compression and 4-7 times better efficiency for decompression.**

NoLoad's compression ratio is comparable to that of gzip-9 **but achieves CPU efficiencies that are orders of magnitude better than software gzip-9.** It is this combination of improvements in compression ratio and efficiency (MB/s/core) that makes NoLoad attractive to customers for Database Acceleration, High Performance Computing (HPC) and many other applications.

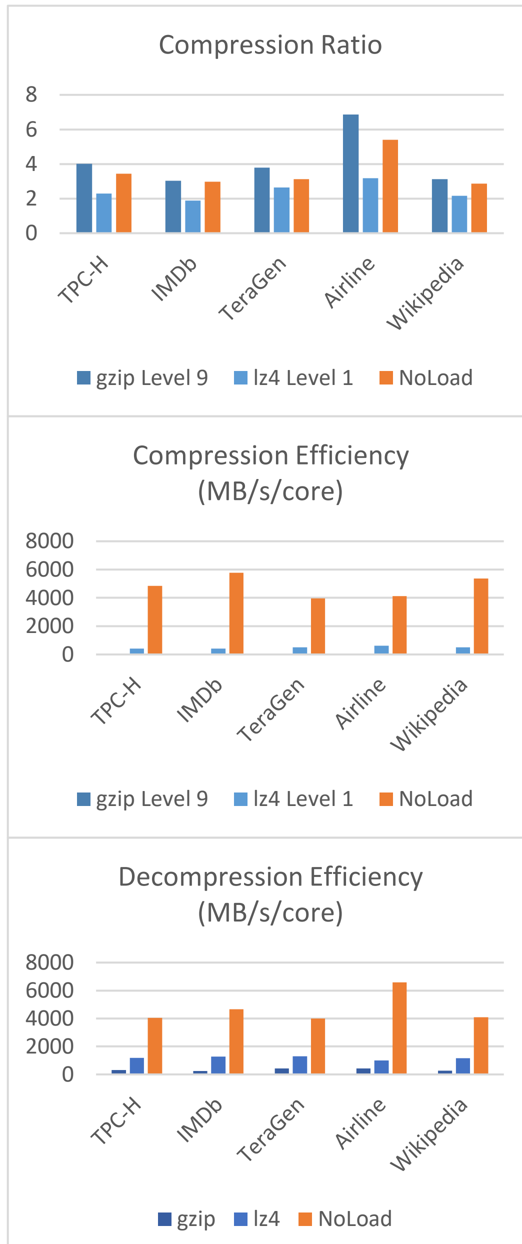

Figure 4: Compression Ratio, Compression and Decompression Efficiency of NoLoad vs. lz4-1 and gzip-9

6.2 NoLoad Transparent Compression using NoLoad File System (NoLoad FS)

To test the performance gains of a NoLoad-based filesystem we collected data on a number of different configurations. The NoLoad Filesystem (NoLoad FS), discussed in Section 5, was installed on these servers, and Big-Data (Hadoop TeraGen) data was written to NVMe SSDs with no compression and with compression, both software and NoLoad, via NoLoad FS. In the sections below we describe NoLoad Transparent Compression peformance on three different hardware platform as follows:

- NoLoad SmartSSD: Compression offload on a Samsung SmartSSD 2.5-inch (U.2)
- NoLoad U50 AIC: Compression offload on a Xilinx U50 FPGA Add-In-Card
- NoLoad U.2: Compression offload on a Bittware 250-U.2 / 254-U.2 FPGA cards

Note that all the above hardware platforms interchangeably use the same Transparent Compression software stack. NoLoad performance scales linearly with the number of NoLoad devices in the system. Additionally NoLoad SmartSSDs, NoLoad-AICs and NoLoad-U2s can be seamlessly mixed in the same system to achieve the expected combined throughput performance. The Eideticom NoLoad software stack discussed in Section 5 is capable of discovering and utilizing all the NoLoad devices in a system even when the devices are deployed on different hardware FPGA cards.

6.2.1 NoLoad SmartSSD Solution

The server configuration was an AMD EPYC 7402 with a single NoLoad SmartSSD. The results linearly scale as multiple NoLoad SmartSSDs are added to the system. The main conclusion from Figure 5 is that NoLoad-based compression can provide **significant disk capacity advantages (lower $/TB) and CPU utilization in Big-Data applications, while also achieving better than line-rate throughput.**

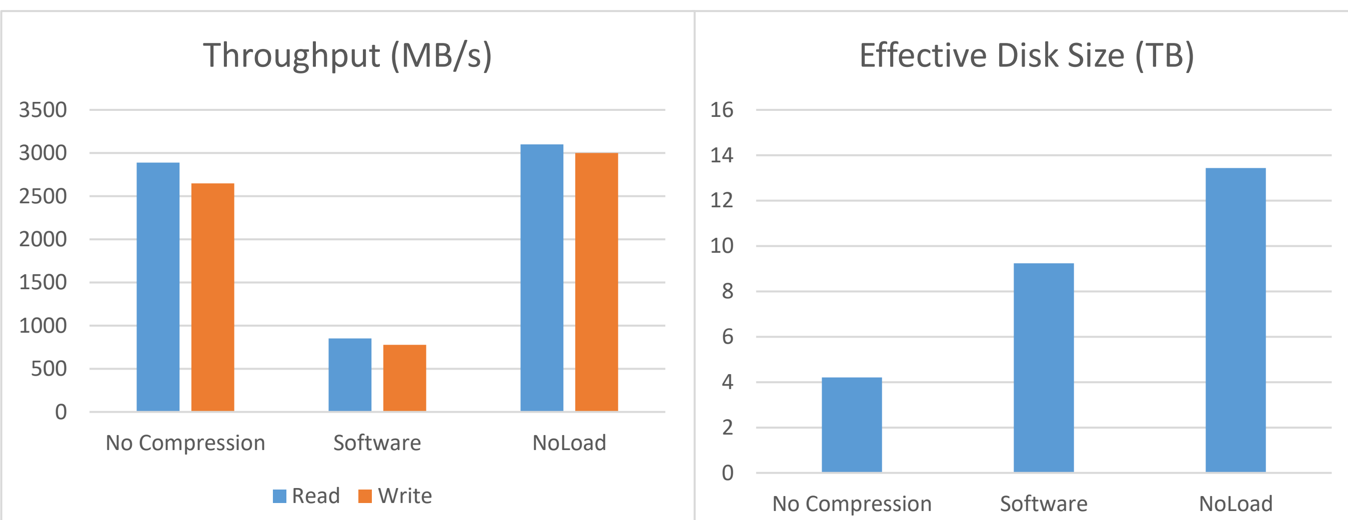

Figure 5: Throughput and Compression Ratio on NoLoad SmartSSD using NoLoad FS

6.2.2 NoLoad U50 AIC Solution

The server configuration was an AMD EPYC 7402 with 10 SSDs. The results achieved use a single NoLoad Alveo U50 Add-In-Card (AIC). The main conclusion from Figure is that NoLoad-based compression can provide **significant Compression Ratio (CR) advantages and CPU utilization in Big-Data applications, while also achieving significally better than line-rate throughput.**

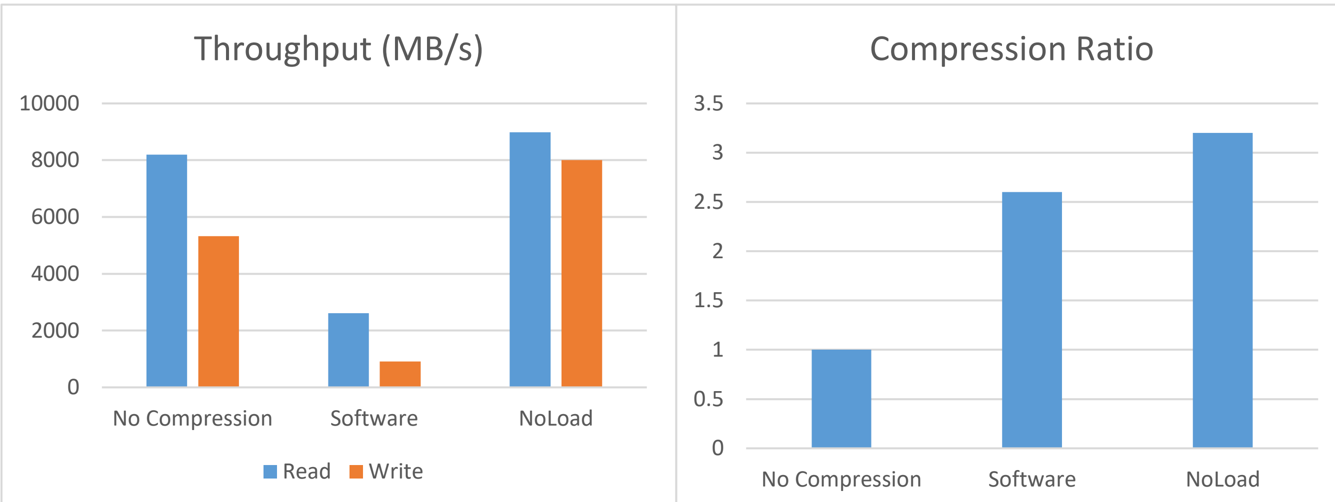

Figure 6: Throughput and Compression Ratio on NoLoad AIC using NoLoad FS

6.2.3 NoLoad U.2 Solution

The server configuration was an AMD EPYC 7402 with 12 SSDs. The results achieved use either four NoLoad 250-U.2 (Gen3x4) or two NoLoad 254-U.2 (Gen4x4). The main conclusion from Figure 7 is that NoLoad-based compression can provide **significant Compression Ratio (CR) advantages and CPU utilization in Big-Data applications, while also achieving significally better than line-rate throughput.**

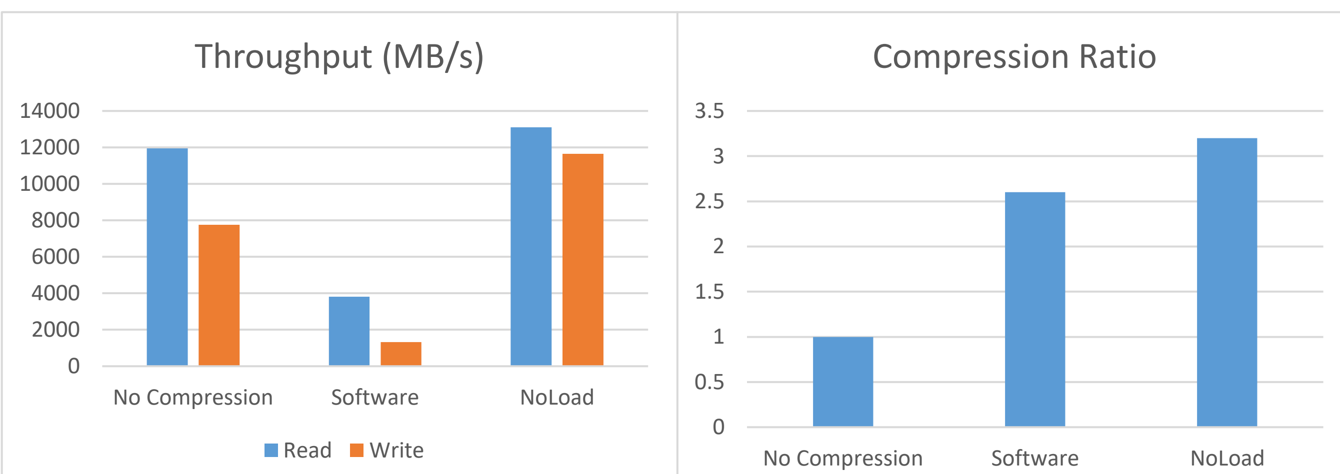

Figure 7: Throughput and Compression Ratio on NoLoad U.2 using NoLoad FS

6.3 NoLoad Transparent Compression Advantages

Hardware based transparent compression can be achieved in different ways, two approaches are:

- File System (Eideticom's NoLoad FS solution)
- SSD Drive Level (Other vendors solution)

In this section we highlight the advantages of Eideticom's File System based approach which is used in the NoLoad Transparent Compression solution. The key advantages are:

- **Improved Compression Ratio:** NoLoad FS approach has the ability to compress larger blocks which leads to better Compression Ratio.
- **Per file compression:** NoLoad FS can selectively compress what is important and avoids compressing small files which is inefficient. Drive-Level approaches have no concept of files and must compress everything.
- **Ability to leverage the page cache:** NoLoad FS utilizes the page cache to optimize performance under certain workloads. Users can choose to leverage the page cache or to bypass it on a per file basis.
- **Minimizes Read/Modify/Write:** NoLoad FS easily handles situations when a file is modified such that its compressed size increases.
- **Ability to mix and match CSDs and CSPs:** Customers have the ability to seamlessly mix and match compression/decompression (or other functions) between Computational Storage Drives (CSD) and Computational Storage Processors (CSP).

7 Cost-Benefit Analysis

Five key factors are considered for a cost-benefit analysis of applications using NoLoad Transparent Compression:

- **Storage Capacity Costs:** NoLoad Compression reduces the effective $/TB of the system by increasing the amount of data stored on SSDs. The performance data demonstrates that NoLoad Compression Ratio (CR) is always higher than software-based lz4-1 compression, therefore increasing the storage capacity of the SSD.
- **Storage Lifetime Costs:** NoLoad's higher Compression Ratio extends the lifetime of an SSD, since for a given compression input throughput it reduces the Drive Writes Per Day (DWPD).
- **Throughput Performance Costs:** The performance data clearly shows that NoLoad compression is 3-6 more CPU efficient than lz4-1 and over 100 times more CPU efficient than gzip-9. NoLoad permits a smaller CPU with less cores to be used to achieve the same throughput. As a result, this leads to fewer servers in the final system.
- **Storage Performance Costs:** As noted in the Storage Lifetime Costs, increased CR means the reduced write throughput to the SSDs. NoLoad allows for the use of cheaper, less performant SSDs and that the SSDs will consume less power.
- **Power Consumption:** The performance data demonstrates that NoLoad compression is dramatically more CPU efficient than both lz4-1 and gzip-9, therefore allowing a given CPU to run cooler and consume less power since less cores are fully loaded. NoLoad results in a power consumption reduction of up to 60% per server.

7.1 Bottomline

The savings from the NoLoad based-system due to Storage Capacity Cost, Storage Lifetime Costs, Throughput Performance Costs, Storage Performance Costs and Power Consumption leads to combined savings of almost **90% per server!**

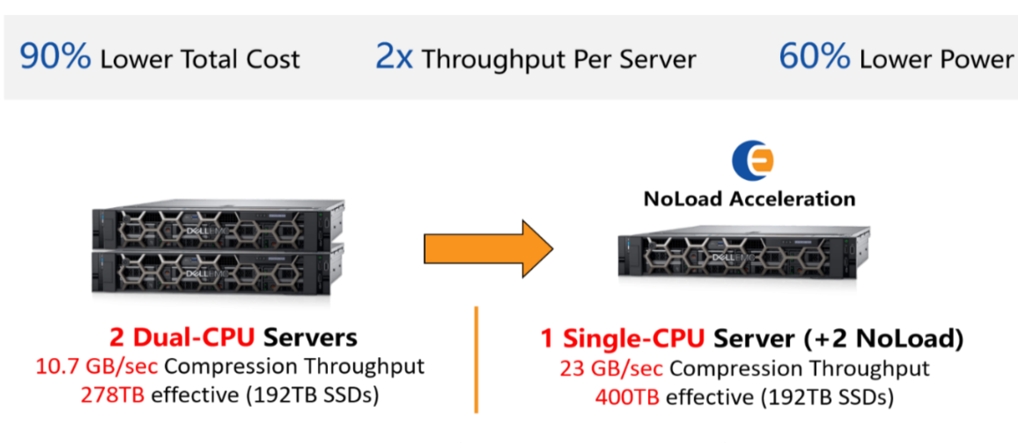

Figure 8: NoLoad Cost-Benefit

8 Customer Success – Los Alamos National Laboratory (LANL)

Eideticom have deployed NoLoad at Los Alamos National Laboratory (LANL). LANL scientists partnered with Eideticom, to integrate NoLoad® into existing parallel file systems thereby creating the first standards compliant, hardware-accelerated parallel file system.

NoLoad® is embedded along the storage system pathway, so compression occurs as the data is in-flight from the processor to the storage device. The performance improvements and data-set reduction are available to users transparently as the file system compresses and decompresses the data automatically. The new data compression capability provided by NoLoad will enable more, and make more effective use of, exascale simulations to support National Nuclear Security Administration (NNSA).

"We are excited to see standards-based computational storage technology being applied to a growing, acute problem in data movement, namely storage server memory bandwidth”

Gary Grider, Deputy Division Leader,
Los Alamos National Laboratory

“The Eideticom NoLoad devices have demonstrated that we can offload storage functions onto accelerators enabling line-rate compression, improving CPU utilization, and reducing memory bandwidth pressure.”

Brad Settlemyer, Senior Scientist,
Los Alamos National Laboratory

Los Alamos
NATIONAL LABORATORY
EST. 1943

Summary

Eideticom's NoLoad® Computational Storage technology utilizes FPGA technology, along with Eideticom's software and industry leading computational accelerator cores, to break through the limitations of CPU-centric computing for storage and compute intensive workloads. As we hit the limits of core and frequency scaling, heterogeneous computing with FPGA technology becomes more attractive for certain workloads.

Computational Storage aims to improve system performance by making compute and storage more efficient and by reducing data movement. In this white paper, we have demonstrated the advantages of Eideticom's NoLoad® Transparent Compression solution and how it results in accelerating data center infrastructure, ensuring affordable scaling and dramatically lowering cost.