Unlocking True Performance and Cost Efficiencies of Storage in Data Centers

This white paper discusses how today’s tiered storage architecture can be supplanted with a simpler, flatter storage architecture leveraging Fungible’s Data Processing Unit (DPU) and in doing so, unlock true performance and cost efficiencies of storage in data centers.

The Nature of Data has Changed

In the past, data as we know it, was mostly structured. Structured data is highly-organized, designed with a fixed schema and typically organized as a table with relations connecting the data points. Relational SQL databases is an example of structured data. Today, data is mostly unstructured. Text, videos, audio, images are all examples of unstructured data. Unstructured data has driven the need for a flexible schema architecture and a storage organization to match. Newer concepts such as data lakes, NoSQL databases with optimized data organization for different use cases e.g. columnar, graph, document etc. are born out of the need to harness insights from unstructured data.

Unsurprisingly, organizations are experiencing various challenges with their current storage solutions. Figure 1 shows how different parameters rank in terms of attributes that need to be improved.

These requirements extend beyond on-premise solutions as organizations look to diversify and future-proof their infrastructure through hybrid and full cloud solutions as well.

On-Premises Infrastructure Attributes to be Improved to Meet the Needs of Customers

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Speed/Performance</td>
<td>46%</td>
</tr>
<tr>
<td>Scalability</td>
<td>44%</td>
</tr>
<tr>
<td>Cost (capex)</td>
<td>42%</td>
</tr>
<tr>
<td>Resiliency</td>
<td>37%</td>
</tr>
<tr>
<td>Self-service infrastructure &amp; application provisioning</td>
<td>34%</td>
</tr>
<tr>
<td>Quality of service</td>
<td>31%</td>
</tr>
<tr>
<td>Pricing variety (e.g. flexible consumption)</td>
<td>20%</td>
</tr>
<tr>
<td>Other</td>
<td>5%</td>
</tr>
<tr>
<td>None</td>
<td>9%</td>
</tr>
</tbody>
</table>

Figure 1. Source: 451 Research’s Voice of the Enterprise: Storage, Workloads and Key Projects 2018

About the Author

Alpika Singh
Technical Marketing Specialist

Alpika Singh is the Technical Marketing Specialist at Fungible. She works with a cross-functional team to ensure that the products that Fungible brings to the market are thriving in a wide variety of customer engagements.

Prior to Fungible, Alpika was the Technical Marketing Manager at DriveScale, HPE, and Broadcom (Emulex Corporation), where she drove the ISV/IHV certifications along with the technical and marketing collaterals for launch, sales, and marketing efforts. She also actively participates in product development and road map activities.

Alpika has a Master of Science degree from Syracuse University and a Bachelor of Engineering degree from the University of Mumbai.
Today's Tiered Storage Architecture

Today, data center operators organize storage into different pools; hot, warm, cold and archive tiers. Each storage tier emphasizes a specific metric - for example, $/GB would be an important consideration for cold data, and latency and throughput figures would be critical considerations for hot data. This tiering strategy is in part due to the nature of different data types having different requirements, but in part also due to technology limitations.

The storage industry, in general, has innovated along the lines of a tiered storage architecture. Data that are mission-critical and workloads that have stringent performance and latency requirements are stored on the more expensive, higher quality media with resiliency options enabled. The media such as Solid State Drives (SSDs) are placed physically "closer" to the compute node in a Direct Attached Storage (DAS) model. On the contrary, data that are less critical and accessed infrequently are stored in the lowest cost, lowest performance media. The media such as Hard Disk Drives (HDDs) could be hosted in a Network Attached Storage (NAS) or Storage Area Network (SAN) topology.

The question begs to be asked - is a tiered architecture flexible enough to meet the demands of today's workloads?

Before we answer that, let's take a look at some of the key challenges of today's storage solutions.

Storage Bottleneck with Limited IOPS and Throughput

Traditional storage solutions consist of either commodity x86-based storage with SAS or SATA HDDs, or SAS, SATA, or NVMe SSDs storage. With slow media like HDDs, CPU performance was not the bottleneck. But, with fast SSD media and protocols such as NVMe, these x86-based implementations cannot fully utilize the performance of the SSDs.

Today's NVMe SSDs can natively deliver about 750K IOPS. However, due to CPU bottlenecks, current solutions are not able to take full advantage of the SSD performance and deliver it to the applications.

Costly High Resilient Storage Solutions

Storage is prone to failures and can cause significant disruptions to applications and their performance. Redundancy schemes are put in place to mitigate faults, prevent data loss and system downtime.

Traditionally, storage solutions have provided data resiliency with techniques such as redundant controller architecture, multipathing and making multiple copies of the data (replication factor). Data replication or mirroring typically requires data center operators to allocate at least 3x storage for each byte of data. The storage requirements for terabytes to petabytes of data add massive pressure on the storage costs.

Modern storage solutions offer more efficient and resilient techniques such as erasure coding. See sidebar on Page 3 for a comparison between erasure coding and replication. While the benefits are significant, erasure coding is both a compute-intensive and network-intensive process.

Limited Scale

Many modern applications such as Big Data Analytics, ML and other workloads require a true scale-out architecture to break free from the capacity or performance constraints so that additional compute or storage can be added on the fly and can span across multiple racks to reduce the blast radius. However, with current software-defined storage solutions, scalability is limited to several racks (or 10s of nodes).

Network congestion is another key contributor to limited scalability. For example, Priority Flow Control and Explicit Congestion Notification enabled in the data center networks still do not prevent widespread congestion failures.
Erasure Coding

Erasure coding creates a mathematical function to break down raw input data into fragments and computes several redundant fragments of parity data.

The original and redundant fragments are always distributed across different locations (e.g., disks, nodes, geographical locations).

Erasure coding is represented by \( (n,k) \) and \( m=n+k \) where "\( n \)" is the original amount of data or symbols, "\( k \)" stands for the additional or redundant amount of parity data or symbols added for redundancy which indicates Failures To Tolerate (FTT). The variable "\( m \)" is the total number of symbols created as part of erasure-coded data.

Let's take a look at a simple example:

Assuming \( n = 6 \) and \( k = 2 \), \( m = 8 \), a 3TB of data would be split into 6 equal parts of 0.5TB with 2 equal parts of parity code amounting to 1TB. The total storage capacity needed to support this scheme is 4TB.

Contrasting this to a triple replication strategy, while the FTT is also 2, the total storage capacity needed to support this scheme is 9TB. This amounts to 6X savings comparing erasure coding and replication.

Deduplication

Deduplication is a technique which removes duplicate or redundant data across all the data in the storage pool. It involves the storage system creating global checksums to identify duplicate or redundant data for matching duplicate data. Once the checksums for new data match an existing checksum, the storage system references the existing data versus storing another copy of the (same) data.

For certain workloads such as virtual desktop infrastructure (VDI) where the volume contains a large chunk of the same data (OS image), deduplication can provide significant savings. For other workloads such as consumer photos or videos, deduplication provides no savings.

High Total Cost of Ownership (TCO)

Data compression and deduplication are techniques commonly used to reduce storage capacity and speed up data transfers, resulting in lower storage costs and network bandwidth utilization. Many of today’s storage solutions offer compression and deduplication as background operations, especially for cold data. When run on software, these techniques can consume many CPU cores. Throughput cannot be sustained at line rate, thus, these techniques are used for cold data. When run on specialized hardware, some of the existing solutions are still bottlenecked by the compute-centric model.

Another challenge of older storage architectures is the need to separate front-end network (north-south traffic) to serve data to the clients and back-end network (east-west traffic) for inter-node communication and data replication. For example, the back-end could be InfiniBand (IB)/Fibre Channel (FC) and the front-end could be NVMeoF/iSCSI, or they could be at different speeds (10/25/100GbE) to serve the data to the clients which requires a mix of IB/FC/Ethernet switches causing a significant increase in CAPEX and OPEX.
A Simpler, Flatter Storage Architecture

The vision for an ideal storage architecture should be one that is cost-effective, highly performant, efficient yet resilient, secure and massively scalable - leading to a simpler, more flexible storage architecture.

To meet these requirements, Fungible has designed a family of purpose-built data center infrastructure processors from the ground up, known as the Fungible Data Processing Unit or DPU™. The DPU is designed to revolutionize the performance, reliability and economics of data centers at all scales.

One of the fundamental feature of the DPU is a programmable software stack, which implements a clean separate of control and data path to drive performance and scale.

Here are some other founding principles of Fungible's DPU:

Efficient, highly durable and available storage

To increase storage utilization and efficiency, Fungible's DPU supports erasure coding and compression in-line and at line rate. Erasure coding eliminates the need for multiple replications of data, thereby providing over 50% reduction in storage overhead and footprint across the data center.

Compression adds to the overall savings of storage and makes data movement faster throughout the data center. The DPU can compress and decompress both "text" data and JPEG images. Standard DEFLATE and LZMA algorithms are used for text compression, allowing the input and output to be industry-standard ZIP files. At a very minimum, compression can provide 3x data reduction, while maintaining line-rate throughput. The DPU also supports deduplication.

The overall efficiency of the data center with erasure coding, compression and deduplication enabled can be higher than 6x with a DPU-based storage architecture.

Massively scalable and secure storage platform

Fungible's DPU enables large-scale high performance disaggregated storage platforms.

The DPU also supports a secure multi-tenant environment through encryption, authentication, overlay and underlay networks. Its AES and SHA cryptographic accelerators enable encryption for data in-transit and at-rest.

See Figure 2 for functions implemented in Fungible's Data Processing Unit. Note that all features are enabled in-line at line-rate.

Unparalleled performance

Fungible's DPU incorporates a full featured, high performant and low latency networking and storage stack, which supports very high throughput and IOPS. The DPU supports industry-standard NVMe and NVMe over Fabrics (NVMe-oF) storage protocols. With the storage and networking stack inbuilt, the DPU can serve as a high-performance controller for vending SSDs to the network. The DPU can deliver line-rate SSD IOPS to workload applications with low end-to-end latencies and tail latencies.

Multi-fabric protocol support

Fungible's DPU includes an embedded networking unit to support standard TCP network along with NVMe and NVMe over Fabrics (NVMe-oF) storage protocols. Instead of creating separate east-west and north-south networking infrastructure, organizations can use a single network for serving data to clients and storage traffic, simplifying topology, and improving performance. The DPU essentially provides an integrated networking model for fast client and data access to the storage nodes.
Fungible’s DPU - Purpose Built to Address Next Generation Storage Requirements

The enormous growth of unstructured and structured data is creating massive business opportunities for organizations. At the same time, it is also creating unprecedented challenges for the underlying storage infrastructure.

Data center architects must look to storage solutions that can deliver extreme performance, efficiency, resiliency, security and true scalability.

Fungible’s Data Processing Unit enables an in-line implementation of data services such as compression, erasure coding, encryption etc. at line-rate and at latencies that are suitable for workloads with the most stringent requirements.

Data center architects now have the option to design a much simpler and flatter storage architecture - and can finally call dealing with data temperature a thing of the past.

---

Fungible, Inc.
3201 Scott Blvd.
Santa Clara, CA 95054
669-292-5522