InfiniBox: Storage That Learns

Silverton Consulting, Inc. StorInt™ Briefing
Introduction

Founded in 2011 by distinguished engineering talent from other major storage vendors such as EMC and IBM, INFINIDAT is a relative newcomer to enterprise storage. Their approach to storage has been extremely innovative, and their high-availability, multi-parallelism storage architecture has gone well beyond other vendor solutions. Moreover, their state-of-the-art caching algorithms provide memory access speeds using flash and high-capacity disk.

The sections below introduce the INFINIDAT product portfolio, highlight INFINIDAT InfiniBox enterprise-class storage and its performance capabilities and provide a real-world example of these capabilities.

INFINIDAT product portfolio

INFINIDAT's product portfolio includes InfiniBox enterprise-class storage; InfiniGuard data protection/backup storage appliance; InfiniSync survivable, zero-second RPO (Recovery Point Objective) replication solution; and Neutrix Cloud, a multicloud-adjacent, storage-as-a-service/disaster recovery (DR)-as-a-service solution.

InfiniBox enterprise-class storage

INFINIDAT's InfiniBox provides PB-scale storage that meets 21st century enterprise needs. The InfiniBox product family includes the following storage systems series:

<table>
<thead>
<tr>
<th></th>
<th>F2000 Series</th>
<th>F4000 Series</th>
<th>F6000 Series</th>
</tr>
</thead>
<tbody>
<tr>
<td>Base capacity</td>
<td>499TB</td>
<td>1.4PB</td>
<td>4.1PB</td>
</tr>
<tr>
<td>Effective capacity</td>
<td>998TB</td>
<td>2.7PB</td>
<td>8.3PB</td>
</tr>
<tr>
<td>DRAM Cache</td>
<td>768GB</td>
<td>2.3TB</td>
<td>3.1TB</td>
</tr>
<tr>
<td>Flash Cache</td>
<td>103TB</td>
<td>103TB</td>
<td>207TB</td>
</tr>
<tr>
<td>IOPS</td>
<td>500K</td>
<td>750K</td>
<td>1M</td>
</tr>
<tr>
<td>Throughput</td>
<td>7GB/sec</td>
<td>10GB/sec</td>
<td>12.5GB/sec</td>
</tr>
</tbody>
</table>

InfiniBox host connectivity options include 24 x 8 Gbps FC and 12 x 10 Gbps Ethernet ports.

InfiniBox storage also supports seven nines (99.99999%) data availability (<3 seconds of downtime per year), built on top of three-controller storage and patented InfiniRAID™ data protection. To our knowledge, no other storage vendor achieves InfiniBox's data availability levels. In fact, INFINIDAT has just announced their Always On availability guarantee for InfiniBox storage.
The InfiniBox three-way controller architecture, connected over InfiniBand fabric, is unique in the industry. This architecture is used in an **active-active-active** configuration. That is, InfiniBox software has been designed so that all functionality runs on all operating controllers at all times. Other industry architectures run dual controllers in active-passive or at best active-active mode, where each controller services IO for only half the LUNs.

InfiniBox’s three-controller architecture also increases compute power. As a result, InfiniBox can offer more sophisticated caching algorithms that can outperform all-flash array (AFA) storage solutions. Further, the InfiniBox systems listed above provide an average sub-millisecond IO response time at their specified IOP rates.

In addition, InfiniBox storage offers unified FC-iSCSI block and NFSv3 file storage. NFSv3 storage is available in the base product functionality without the need for any additional file protocol hardware or software. As a result, file and block storage is configured, monitored and operated all within the same management framework.

Furthermore, InfiniBox storage provides advanced, enterprise-class storage capabilities, such as **space-efficient, high performing, read-writeable snapshots, synchronous and asynchronous (<4 sec RPO) replication, IO quality of service and data compression-thin provisioning**, as well as **HTML5 GUI, CLI and RESTful API management interfaces**. With InfiniBox snapshots, customers can take nearly infinite point-in-time copies of live data for dev, test, QA and other purposes.

All the above software and hardware functionality integrates well with OpenStack, VMware® vSphere® hypervisor, Linux®, Unix® and Microsoft® Windows Server® operating systems, along with Commvault® and Veritas™ data protection software, SAP® applications, and Oracle® database solutions. InfiniBox’s VMware integration includes a vCenter® plugin, as well as support for VMware Site Recovery Manager™ (SRM) and VASA and VAAI vSphere APIs. Further, InfiniBox storage also supports Host PowerTools with extensive configuration validation to ensure proper deployment and to maximize IO performance in customer data center environments.

Finally, by supplying DRAM IO speed, InfiniBox offers PB-scale storage at a price ($/GB) that few, if any, AFA systems can match. Because InfiniBox uses space-efficient snapshots, thin provisioning, data compression, and large disks, capacity can easily cost less than deduplicating AFA storage and especially so, if application
encryption is active in a customer environment. Moreover, in most cases, such as when operating with 100s of VMs and multiple databases, InfiniBox outperforms AFA storage.

**InfiniGuard data protection/backup storage appliance**

The InfiniGuard is tier 1, InfiniBox storage reconfigured with multiple data deduplication engines and for high throughput IO workloads to supply the best performing, highest available backup storage in the market today. It includes data deduplication, compression and encryption to reduce the storage footprint for data center backups.

INFINIDAT’s **InfiniGuard B4260N** backup appliance supports up to **20+ PB** and the B4212N supports up to **40+ PB of effective backup capacity** in a single 42U rack. Both InfiniGuard systems use inline, variable-length dedupe and data compression. Capable of **up to 74 TB/hour data throughput with NetBoost** acceleration, InfiniGuard systems can support most any enterprise data backup/restore need. Moreover, InfiniGuard supports VTL (FC), NFSv3, CIFS (SMB1), OST and Oracle® RMAN backup protocols, meaning it works with most backup software and most database applications used today.

InfiniGuard also offers bandwidth-optimized, encrypted data replication in 1 to 1 (dual-site redundancy), 1 to 2 (three-site redundancy) or N to 1 (central site repository for remote/branch office) configurations. These replication configurations can be used to support business continuity (BC)/DR for site outages.

**InfiniSync synchronous RPO zero replication over unlimited distance**

Synchronous replication usually incurs write performance penalties when transmitting data to remote sites, while asynchronous replication doesn’t offer zero RPO. To provide synchronous replication with no performance penalty, INFINIDAT offers InfiniSync, a no impact, zero RPO InfiniBox replication solution that uses a nearby synchronous replication InfiniBox storage tailored to support DR for data center catastrophes.
InfiniSync is designed to survive most any disaster – **extreme fires, floods** and **building collapses** – while still protecting customer data without impacting IO performance. InfiniSync’s “black box”, a nearly invulnerable storage replication appliance, is located onsite near primary InfiniBox storage in order to eliminate performance penalties.

After a data center disaster, InfiniSync can **transmit up to 1.6 GB/sec of data** using wired, Wi-Fi wireless or 4G-LTE cellular protocols to any internet-connected location, anywhere in the world. The system comes equipped with an **internal battery for 36 hours of continued operation** after a disaster happens.

InfiniSync is unique in the industry, as nothing similar is offered by any other major storage vendor. Other survivable storage solutions appear to be little more than hardened cases surrounding basic block storage, that’s not fully integrated with primary storage replication services and without any internal, multi-technology data transmission capabilities for data center disasters.

**Neutrix Cloud, cloud-adjacent high availability, high performing storage**

INFINIDAT Neutrix Cloud is InfiniBox storage, that’s located near-to (cloud-adjacent) and with high-speed access to public cloud compute services. It is available on a pay-as-you-go model and used to support multi-cloud and cloud-native application data requirements.

Neutrix Cloud storage is located in INFINIDAT sovereign data centers with storage owned and operated by INFINIDAT and with direct, fast-path connections to AWS, Microsoft Azure and Google Cloud Platform. Applications executing on compute resources in any of those public clouds can access and run with iSCSI block storage residing on InfiniBox, enterprise-class storage. In addition, applications employing filesystems for storage can be deployed to multiple clouds and simultaneously access the Neutrix Cloud filesystem.
Neutrix Cloud storage can also be used as a target for offsite, asynchronous replication to supply DR as a service for other InfiniBox storage located in primary data centers. As with all InfiniBox storage, Neutrix Cloud supplies high-performing, seven nine’s data availability storage but in this case, available to support cloud-native and multi-cloud application data requirements.

**InfiniBox innovative performance capabilities**

INFINIDAT calls their caching management algorithm the **Neural Cache**. Neural Cache takes advantage of innovations in the InfiniBox architecture, including the following:

- **InfiniRAID** – InfiniBox uses a software-defined RAID scheme, that prevents hotspots by using thousands of virtual RAID groups to spread user data across all drives.
- **Log structure file disk layout** – InfiniBox has a backend layout that transforms random writes and sends them to the backend (destaged to disk) as sequential streams of data. Hot data can be overwritten multiple times, but at some point, it is destaged to disk as well. Depending on access frequency, some data is also written to flash cache.
- **Destaged data in performance order** – InfiniBox destages data in performance order so that frequently accessed data is written to disk in continuous segments that can be read back via a single request.
- **Larger cache** – InfiniBox supports up to 3TB of DRAM cache which allows data to be retained in memory much longer (5 minutes) than other vendor systems that use media for performance.

As discussed above, InfiniBox’s three-controller architecture enables all data services to run in all nodes, providing more compute power to speed up caching and more IO parallelism to stage/pre-fetch data that’s typically serialized on other systems.

Data is pre-fetched into DRAM cache, in order to be in memory when an application requests it. The key to better random read cache hit rates is to pre-fetch the **right data** into cache before its needed. Further, with the massive parallelism available from the three-controller system, high speed cluster interconnect from InfiniBand, and wide striping from InfiniRAID, disk reads execute very quickly.

**Read Rates and Response Times**

Another key to InfiniBox’s strong performance lies in how it manages the memory hierarchy, which consists of DRAM, flash and disk storage. For InfiniBox, flash is just another cache used to satisfy IO requests and accommodate workload changes. InfiniBox outperforms AFA storage because more than 90% of its reads typically come out of DRAM rather than flash or disk.
InfiniBox also maintains performance metadata, which tells the system what to pre-fetch into DRAM cache when data is read. All Neural Cache metadata is kept in memory, which allows for quicker access. Neural Cache metadata is constructed as a Trie structure, which uses a LUN block address as an index and supplies performance and other information at the leaves of the Trie. As a result, InfiniBox can maintain information on billions of objects, accessing and updating that metadata in microseconds.

Further, for each data segment or chunk, InfiniBox maintains spatial performance metadata (other data blocks accessed) and temporal performance metadata (times of most recent accesses). Using this information, the Neural Cache can pre-fetch data based on adjacent data accessed when a specific block was last accessed as well as on the time of the block’s last few accesses. As a result, InfiniBox’s Neural Cache can ensure that referenced data is already in cache when needed rather than having to wait for it to be staged from disk.

On the rare occasion when data is not in memory when an IO requests it, there’s a high likelihood that the data is in the process of being pre-fetched from disk or is available directly from flash.

Similar logic applies to flash cache, with the exception that it’s significantly larger and is loaded primarily from memory during destage activity. Neural Cache metadata helps the system decide whether to just destage data to disk or to also copy it to flash. Data read directly out of flash can be used to quickly support IO activity while the rest of the data is pre-fetched into DRAM.

This Neural Cache intelligence enables customers to typically achieve a combined DRAM and flash cache read hit rate of 97-100% and an average response time in the sub-millisecond range.

Constant Learning and Improved Performance

In addition, Neural Cache is constantly learning how to perform better. InfiniBox uses Trie metadata to access and update spatial and temporal data block information as data is being referenced during customer workload activity.

As host IO patterns change, Neural Cache metadata is updated to reflect this new activity. The next time a data block is referenced, pre-fetch activity is based on the spatial and temporal aspects of the most recent access. In this fashion, DRAM cache is always populated with data having the highest probability of being requested by another IO. As a result, InfiniBox Neural Cache is continuously learning the best way to cache customer workloads.

In addition to on-prem Neural Cache learning, InfiniBox provides telemetry data back to INFINIDAT HQ. This data is used to pinpoint where the Neural Cache missed
opportunities to pre-fetch the right data in order to tweak caching algorithms and improve subsequent code updates.

**InfiniBox Neural Cache field performance**

Let’s look at an example of the ability of Neural Cache to learn a customer workload.

Figure 1 shows one large financial services company’s results when running a complex, lengthy RDBMS online transaction processing (OLTP) report query against their database server using InfiniBox storage. [The company supplies payroll processing, human resource services, accounting and financial services, a portfolio of Human Capital Management (HCM) offerings and HCM outsourcing services.]

As seen in these first run results, the point displayed (detailed along the right side and highlighted at the left) shows that InfiniBox is providing 28.2K read IOPS and 13.3K write IOPS for a total of **41.5K IOPS**, with read throughput of 220.4 MB/sec and write throughput of 103.8 MB/sec at an **average response time of 1.79 msec**.
Figure 2 shows the same workload run a second time, 24 hours later with significantly improved IO performance resulting from the Neural Cache learning.

In the second run, InfiniBox delivered 32.7K read IOPS and 15.3K write IOPS for a total of **48.0K IOPS**, with read throughput of 255.7 MB/sec and write throughput of 119.6 MB/sec at an **average response time of 0.90 msec**. The elapsed time to produce the report also decreased from ~58 minutes to ~46 minutes.

On just about every IO metric, InfiniBox IO performance improved by ~15%, the system’s **average response time was cut in half** from the original run (~1.8 msec to 0.9 msec) and application **elapsed time was reduced by over 20%**. These improvements resulted from just two runs of the same application one day apart, suggesting even greater improvements with IO workloads over time.
Summary
INFINIDAT has taken an innovative approach to enterprise-class storage. InfiniBox supplies triple redundancy whereas other products offer dual redundancy, and InfiniBox offers hybrid storage at faster than AFA performance, for a lower $/GB while other products have moved to all flash. In addition, InfiniBox has developed strong cache algorithms to take advantage of a unique architectural design that provides DRAM performance with hybrid storage.

Moreover, this same advanced technology permeates throughout INFINIDAT’s portfolio, improving performance for data protection storage, BC/DR survivable storage, and storage-as-a-service for multi-cloud application environments.

In addition, InfiniBox Neural Cache learns from daily IO activity to optimize cache loading and improve storage performance for real-world customer workloads over time.

While these activities require more intelligence and processing power, InfiniBox’s innovative architecture gives it all the intelligence and power needed to ensure that customer data is in memory when IO requests occur. As such, InfiniBox delivers PB-scale storage at a more economical price than other vendor solutions and outperforms AFA storage solutions.

Silverton Consulting, Inc., is a U.S.-based Storage, Strategy & Systems consulting firm offering products and services to the data storage community.