The Quiet Success of Infinidat

CHRIS EVANS 6 JANUARY 2022 DATA PRACTICE: DATA STORAGE, ENTERPRISE, INFINIDAT, OPINION, STORAGE, STORAGE HARDWARE, SUBSCRIBERS LEAVE A COMMENT

In the modern shared storage market, most of the news and airtime tends to be dominated by the all-flash players. This isn't a surprise, as flash technology has improved performance, reduced latency and, in general, stepped up to the needs of modern applications. Infinidat, however, has bucked that trend and continues to develop systems with most of the data capacity still on hard drives. At the same time, the company has quietly grown into a major enterprise player.

Background

Infinidat was founded in 2011 by Moshe Yanai, arguably the father of all modern shared storage systems. Yanai developed the Symmetrix platform for EMC in the late 1980s and through the 1990s, then founded XIV, which was sold to IBM in 2008. The Symmetrix platform was the first ICDA (Integrated Cached Data Array), where the physical disks were abstracted from the host and used to create logical volumes where the cache optimised read and write I/O. Today we don't think twice about the idea of abstracted LUNs/volumes, but back in the early 1990s, this concept was revolutionary.

Abstraction enabled Yanai and EMC to introduce features like TimeFinder for BCVs (Business Continuity Volumes) (effectively snapshots) and RDF (Remote Data Facility) for site-to-site replication.

The XIV platform was a rethink of the fixed RAID designs seen in the Symmetrix platform. Instead of protecting data with mirroring or RAID-5 groups (typically 4 or 8 disks), the XIV system spreads data across multiple server nodes and media, with a centralised shared network (initially Gigabit Ethernet). This design enabled systems to gain much greater performance from exploiting all the available media while significantly reducing rebuild times on failed devices.

- **Are RAID Rebuild Times Still Relevant?**

Themes

The history of Symmetrix and XIV are essential in understanding the rationale behind Infinidat. Both platforms represented radically different designs from those in the market at the time while addressing end-user concerns about performance and availability.

The technology behind the Infinidat solutions also takes another radical diversion away from perceived wisdom. As we will discuss, the architecture of InfiniBox and associated platform solutions has enabled Infinidat to build storage solutions based chiefly on traditional hard drives, in a market that has moved further towards all-flash. This strategy provides Infinidat with the means to be more financially flexible and operationally efficient than some competitors. In a world moving further towards service-based delivery models, operational and financial efficiency for the vendor will become critical differentiators.

Traditional Solutions

Traditional shared storage solutions have attempted to fix the I/O performance problem of hard drives by caching and tiering. Most applications read and write from the working set, which is a subset of either structured data in a database or unstructured data in a file system (or object store). Few applications (with the notable exception of AI and deep learning algorithms) will generate totally random I/O across an entire dataset. As a result, read caching improves performance for active data, while write

caching improves all write I/O. Tiering places data onto the most appropriate performing media based on a cost/performance model. In the event that data isn't in cache, a fast tier mitigates some of the response time latency.

Both caching and tiering are essentially compromises that deliver improved I/O times with the use of limited and expensive resources. By comparison, the all-flash market takes the view that any I/O deserves to be served from the highest performing media. This approach eliminates "tail latency" as much as possible. The higher cost of flash storage is mitigated through techniques like compression and data deduplication, which are much easier to implement on low latency media (but do require additional compute power).

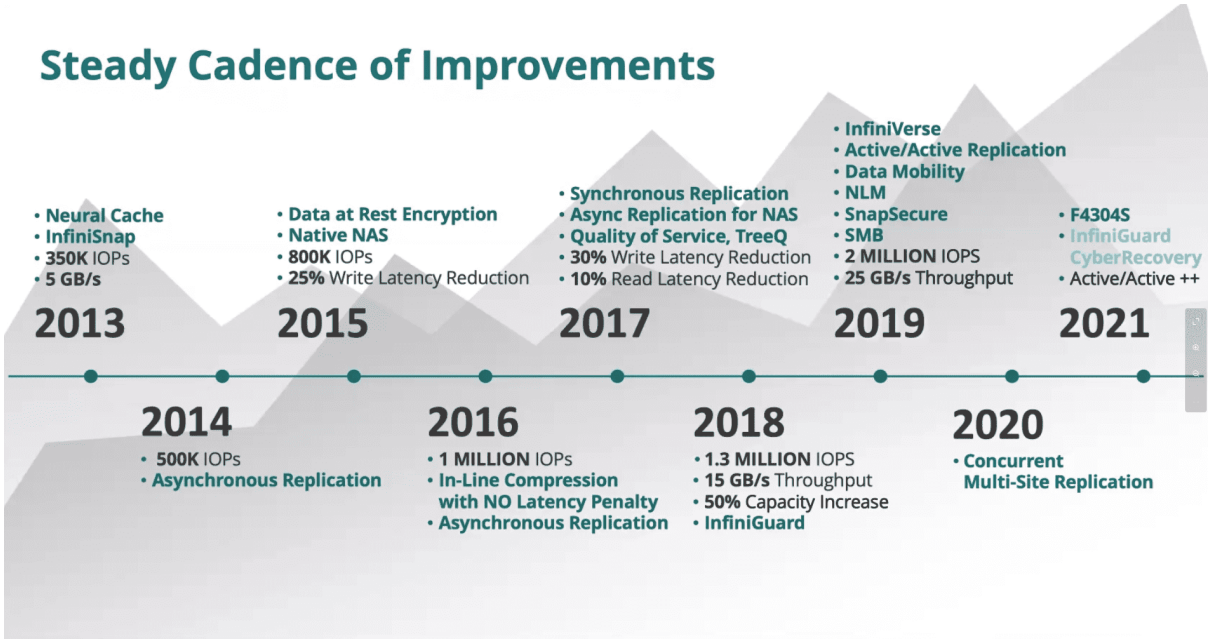
InfiniBox

With the design of InfiniBox, Moshe Yanai and the Infinidat team took a radically different approach compared to the rest of the market. The design of InfiniBox looks at data patterns, placing data on disk with a locality that uses a technique called the Neural Cache.

Hard drives are notoriously bad at random I/O due to the mechanical nature of the read/write heads. However, modern disks have high sequential transfer performance, measured in hundreds of megabytes per second. The Neural Cache architecture aims to group similarly accessed data together, then use two additional layers, one of flash and one of DRAM as caching for the disk I/O, making HDD access as sequential as possible.

The benefit of the Neural Cache design is that around 93% to 99% of read I/Os are typically served from cache. Infinidat claims that real-world field data shows approximately 90% of read I/O comes from DRAM cache, while 100% write I/O is automatically cached through DRAM.

These performance numbers are significant because DRAM operates at nanosecond speeds, whereas flash is orders of magnitude slower (microseconds) and hard drives even slower again (milliseconds).



Triple Redundancy

InfiniBox uses triple redundancy in the form of three active-active controller nodes, each with DRAM and SSDs installed locally. All back-end drives are shared between the controllers, with up to 480 drives deployed in eight disk enclosures. Three controllers provide greater redundancy (seven 9's is claimed) while enabling a high percentage of DRAM per system.

Tail Latency

While 99% cache read success is impressive, any cache miss that results in a read from disk can represent a significant drop in performance for those requests. With many applications, the performance difference between and cache hit and a cache miss may not be significant. However, a small proportion of workloads will find the I/O performance disparity a problem. As a result, Infinidat recently introduced the InfiniBox SSA, a solid-state solution designed specifically to address the tail latency challenges.

Price/Performance

By placing most data on (relatively) cheap media, Infinidat creates some attractive financial and operational benefits.

- **It's Time to Revisit Hybrid Storage**

- **Persistent Memory in the Data Centre**

Looking back to the start of centralised storage, we can see that the $/GB metric has always been used as an indicator of value for money. Tiering alters the $/GB ratio, with more expensive storage offering better I/O throughput and lower latency. In the Infinidat model, most I/O is served from the smallest capacity tiers, namely DRAM and SSD. Infinidat can afford to ship entire systems fully populated and allow customers to grow into the capacity because the incremental growth occurs on the cheapest storage tier, namely the hard drive.

Seeding

Why is the ability to deploy entire fully populated appliances so attractive? In any data centre, change represents risk. Physical changes involve humans entering the data centre and potentially disturbing live running equipment. Mistakes can be made, so many large enterprises operate a policy that ensures hardware is deployed and configured once only. Rack-scale deployments ensure interaction once a system is deployed, with minimal intervention to replace failed components like disk drives.

The ability for Infinidat to deploy entire systems means capacity can be in place for further business expansion, while the marginal cost of leaving unused drives in place is much cheaper than pre-seeding all-flash solutions.

Of course, the cohort of customers that need rack-scale storage is relatively small but arguably increasing every year.

Diversity

As we mentioned earlier, Infinidat has expanded its solutions portfolio with the introduction of InfiniBox SSA in June 2021. These systems offer up to 1.5 million IOPS and 25GB/s throughput. Crucially, the SSA platform addresses the challenge of tail latency and the impact of I/O that isn't cached. Customers now have a choice to deploy one or both solutions to meet the requirements of applications while managing costs.

InfiniBox is also the foundation for InfiniGuard, a petabyte-scale deduplicating backup appliance. InfiniGuard supports all common backup solutions through standard protocols, including NFS, SMB, VTL

(virtual tape) and application-specific protection (e.g. RMAN). The exciting aspect of InfiniGuard versus traditional deduplication appliances is the ability to use Neural Cache in a way that places data in the right location for fast restores. Note: object storage via S3 could be a useful addition for InfiniGuard.

Business Outlook

Infinidat continues to steadily grow the volume of storage deployed in the market. The company also claims to be profitable, which is a remarkable achievement in a competitive market. But what happens next?

Technology

From a technology perspective, hard drive capacity continues to increase, albeit at a slower pace than over the previous decade. At the top end, vendors have introduced 18TB and 20TB models, with some using new techniques like SMR. Dual actuator drives may provide some performance relief with large capacity HDDs, but the promise of 30TB+ drives could be a few years away.

In the flash market, vendors appear to have stalled around the 30TB mark, with only IBM and Pure Storage offering drives with capacities exceeding 40TB (there is the Nimbus Data outlier at 100TB, but this is a SATA/SAS device in a world rapidly standardising on NVMe). The reason for this artificial capacity threshold is probably one of price, with a single drive retailing at thousands of dollars. This represents a costly failure domain as well as an expensive seeding process when many drives are needed to build a redundant solution.

In either case, Infinidat is well placed to use any future new media. We asked the company about the challenges arising from large-capacity drives. The answer is clear, nothing in the InfiniBox architecture precludes the use of any media type. As the SSA demonstrates, future designs could be based on QLC or even PLC technology.

Features/Functionality

The evolution of InfiniBox and InfiniGuard treads a well-worn path of competition in the traditional enterprise market. We believe enterprise sales

(in dollars) has been relatively stagnant for many years, with companies like Infinidat and Pure Storage gaining market share from their laggard competitors. Infinidat continues to build out features and add functionality that addresses the on-premises enterprise market. This strategy has fuelled growth at the expense of others in the market. The big question to ask is where this direction heads next.

The Architect's View™

With masses of data headed to the public cloud, all vendors (Infinidat included) need to find ways to remain relevant. In markets where margins are thin, businesses typically survive by resorting to scaling and price reductions. This direction also requires operational efficiency in solutions operations and the supply chain.

As a consolidation play, Infinidat's solutions offer large enterprises a route to shrink footprint, reduce costs and save on data centre power and cooling. The ability to deploy fully populated systems has a significant business advantage and makes capacity upgrades friction-free (from a data centre deployment perspective).

As an initial strategy, new CEO Phil Bullinger will be focusing on broader worldwide expansion and developing the "as-a-service" offerings, which are biased favourably towards vendors with the ability to seed systems with capacity ahead of time.

In the short term, we see Infinidat doing well in increasing market share across traditional markets. The longer-term strategy is more difficult to predict, as the rise of the public cloud and the future makeup of the split between cloud and on-premises is unclear. Of course, by the time growth becomes an issue for Infinidat, Moshe Yanai may well have developed the next generation of systems to supersede what we have today.