

Data backup and disaster recovery

Data protection in the enterprise is a vast topic, so much so that vendors who offer a data protection solution offer several products just to cover all the different use cases. There are solutions that address backup, recovery, and disaster recovery, as well as off-site backup and replication. This paper will focus on data protection for the data center, which is the area that houses the largest amount of data and backup infrastructure.

It is a little-known fact that the amount of storage in the data protection tier surpasses “primary” storage totals by as much as a factor of 4. In the past, the raw volume of data alone, combined with yearly, monthly, weekly, daily, and even sometimes hourly retention schedules, could cause clients to end up with as many as 30 copies of their original data, which consumes a vast amount of storage capacity.

This data volume explosion ushered in a host of next-generation data protection solutions, especially purpose-built backup appliances (PBBAs). To help with the speed of moving vast amounts of data, clients moved away from tape and implemented disk-based backup. The next issue that had to be dealt with was the overall volume of data stored on disk for data protection. Something had to be done to shrink this data volume, which opened the door for deduplication technologies. Moving data quickly to an optimized disk platform brought about a new era in which massive amounts of “secondary” storage was deployed to support backup environments in almost all enterprise accounts. Additionally, with deduplication technology, clients could efficiently replicate data after the backup, to another deduplication target or PBBA off site to create a disaster recovery solution.

These solutions enabled the industry to better meet the recovery point objective, and recovery time objective (RPO and RTO) SLAs of the business. However, one thing in this whole scenario didn't change: data growth.

We are now at a crossroads when it comes to the traditional data protection paradigm. Analysts have often quipped, “How do you backup a petabyte (PB) of data?” “You don't,” is the typical response. Disk-based backup appliances help from a performance perspective, but these devices weren't designed to scale to manage PBs of data and hence are not meeting the ever-evolving SLAs of the business.

NEEDS ARE CHANGING

Backup deployments of any significant size involve the deployment of many servers just to manage the ingest speed required to meet backup windows and track the metadata for cataloging the backup. All enterprises have a defined backup window, schedule, and retention period, which should be the foundation of the backup design. This design will then dictate the overall infrastructure needed to support the backup policy. Growth, whether organic or by acquisition, often stresses the design that is in place.

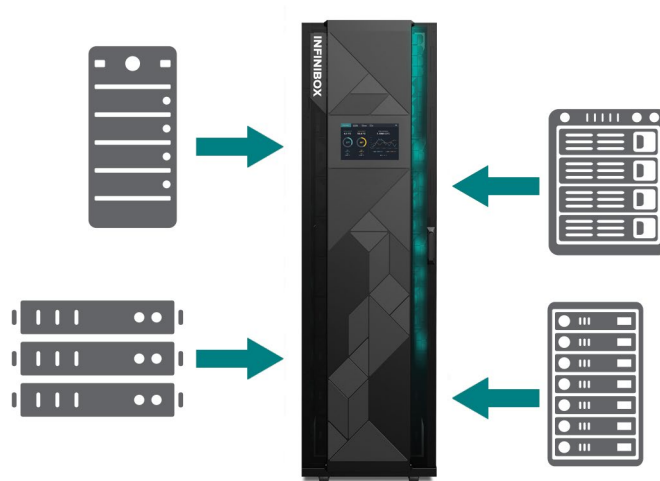
To meet the SLAs of today's business, many of the traditional approaches need to evolve. Standard RPOs and RTOs have changed. For example, businesses have new types of SLAs, version recovery objectives (VROs) and geographic recovery objectives (GROs), that are now vital to meeting businesses' data availability SLAs. The use of snapshots and replicas to meet the new data protection requirements has moved to the forefront as the primary method of operational recovery. There are a lot of benefits that come from this. First and foremost, most enterprise arrays offer snapshots and replication capabilities. In addition, what better way is there to capture and store your data for protection purposes? What backup software provides you with an RPO of near zero? What backup software provides you with an RTO of near zero? How do you achieve these same SLAs for VROs and GROs? The next important question you need to ask is: which array allows you to meet all of your desired SLAs and at the right cost?

INFINIDAT FEDERAL IN THE ENTERPRISE BACKUP ENVIRONMENT

All this now changes with the introduction of InfiniBox®. We recognize that changing processes is one of the hardest things to do in IT, so we looked at solving the data protection challenges of today in two ways: by acting as a target to solve the PBBA box sprawl and also leveraging the snapshot and replication capabilities in the InfiniBox to achieve better data protection SLAs. This combination not only offers better overall data protection, but also does so while lowering costs from both a CapEx and OpEx perspective.

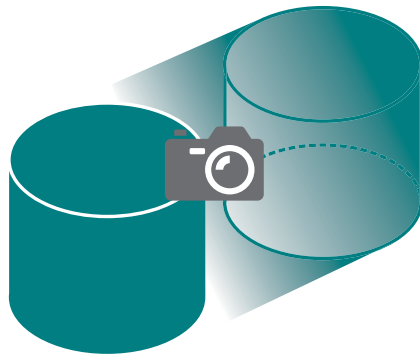
First, combine the capabilities of a large scale storage solution with seven nines of availability with backup vendors providing deduplication inside the backup application package, and you have a powerful data protection solution. Let's take a look at the impact to the environment when replacing deduplication appliances with the backup vendor's native deduplication and an InfiniBox solution.

The first step is to enable deduplication at the backup/media server. This will have the effect of reducing the amount of data being sent across the storage network while limiting the change to the existing environment. It is important to note that the InfiniBox solution can be deployed as a file or block solution. The addition of an InfiniBox, as the backup storage pool, offers the ability to greatly reduce the number of deduplication appliances needed. There are significant features that come with the deployment of the InfiniBox. Performance and scalability are key factors, allowing up to 8:1 consolidation of high-end deduplication appliances.



The InfiniBox provides much better bandwidth than any deduplication appliance while at the same time providing a much denser capacity footprint. The consolidated infrastructure allows for much easier management and a much better environmental footprint. The first significant savings are in power, cooling, and floor space, as multiple deduplication appliances can easily be replaced by a single InfiniBox array that operates at a max of 8kW while providing multiple petabytes of effective storage in a single 42U rack. Finally, a large number of switch ports will be freed up as this consolidation occurs.

The InfiniBox solution is an inherently much more reliable platform than one that was built specifically to be secondary storage. The InfiniBox system is designed with seven nines availability. Adding InfiniBox to the backup environment reduces management complexity and improves reliability, ultimately resulting in improved backup success rates and lower costs.



Snapshots/Replicas for All Operational Recovery

Next, in order to meet the new SLAs of the business, let's look at InfiniBox's efficient snapshot and replication capabilities and the ability to use these for both local and remote operational recovery. The InfiniBox system with its patented InfiniSnap® technology, is designed to provide higher availability, not only for active data, but also for its snapshot data. In addition, InfiniSnap's snapshot and replication capabilities help provide RPOs, RTOs, VROs, and GROs that are unmatched in the industry for uptime and data availability. InfiniSnaps provide IT with the ultimate granularity in setting up recovery point objectives and with the ability to recover any snap from any point in time, providing VROs that were once not possible without time-consuming recoveries. Clients can take as many as 100,000 snapshots (at a pace of 25 per second) with zero impact to the performance of the array. The replicas provide the same granularity in any location, helping IT meet all of its geography-specific recovery needs and have RPOs as low as four seconds. In addition, mounting a snapshot or replica and running off a clone can happen in seconds, creating a near-zero RTO. InfiniBox also has tight integration with VMware and other backup vendors to enable better snapshot management and application consistent snapshots.

Combine all these capabilities with the integration InfiniBox has with VMware's data protection APIs, specifically vStorage API for Array Integration (VAAI), and the vCenter plugin, and you have a simple, self-service, single-instance backup and restore solution (without any dependency on vVol or third-party tools). The entire data protection process can be fully automated via the InfiniSnap tool, which maintains consistent copies of thousands of VMs with relatively small retention schedules. InfiniBox operations are done via VAAI, utilizing only storage resources. The actual snapshot taken per VM using InfiniSnap technology has zero impact on an unlimited number of snapshots. This brings traditional backup and restores of single VM instances from days and hours to minutes or even seconds. Customers struggling to backup and restore their always-growing VM environments can immediately benefit from the powerful InfiniSnap technology.

CONCLUSION

Adding InfiniBox to an existing backup infrastructure can greatly reduce the sprawl of backup appliances that is occurring in most enterprise data centers. By enabling the deduplication feature of the backup software being used today, you can limit the needed network bandwidth and the number of standalone deduplication silos that you currently support.

InfiniBox is the first primary storage array to be cost-effective enough in capacity and performance to satisfy enterprise backup environments at large scale.

| Enterprise Metrics | Deduplication Appliance | InfiniBox |
|-----------------------------|---|---|
| Performance Features | | |
| | Bottlenecks on NFS and CIFS | Consistent and high performance |
| | Limited metadata acceleration options | High IOPS and high read/write throughput |
| | No SAN-based backup supported | Uses flash to boost metadata performance |
| Scalability | | |
| | Limited storage capacities (~590TB) | Multiple petabytes of effective capacity per rack |
| | Limited throughput (~500MB/s) | Average throughput of 4+GB/s |
| Availability | | |
| | Secondary storage array | Seven nines availability, time-based snap, with CDP-like protection |
| | Single controller architecture | |
| Flexibility | | |
| | Purpose built backup storage | Shared storage platform |
| Environment | | |
| | More switch and environmental expense | Dramatic switch port and environmental savings |
| Management | | |
| | More management points to include software options for best results | Seven nines availability, time-based snap |
| | | Simple maintenance |
| | | Built-in RESTful API for easy automation |

By taking the next step in data protection and utilizing the InfiniSnap at the storage layer, data protection is moving closer to the primary storage which allows IT to better meet its data protection SLAs. In addition, driving better data protection lowers CapEx by reducing the amount of data protection software, hardware, and infrastructure needed to support the environment. InfiniSnaps help reduce OpEx, by making the management of the “protected” environment much more a part of the primary storage operation.