WHITE PAPER

NEC Sets High Bar in Scale-Out Secondary Storage Market

Sponsored by: NEC

Eric Burgener
August 2015

IDC OPINION

Managing the explosive data growth in 3rd Platform computing environments is driving significant change in how datacenters evaluate storage infrastructures. 3rd Platform computing is predominantly built around virtual infrastructure, heavily leverages flash and cloud technologies, and is used to host legacy applications like relational databases, messaging systems, and file shares as well as next-generation applications built around mobile computing, social media, and big data and analytics. Primary data sets are large, but secondary storage requirements are generally three to four times as large for most enterprises. Because of the significant benefits that disk offers, many customers are deploying secondary storage platforms built around scale-out disk-based architectures. Soon even small and medium-sized enterprises will need to accommodate petabyte (PB)-scale environments, and they need to begin to plan for that in their next storage refresh cycle. Key considerations in this market include scalability, reliability, and overall cost of solution.

Scale-out disk-based secondary storage architectures are available from a number of vendors, including many of the larger players, and a clear set of baseline requirements has emerged for these platforms. They need to deliver high performance in terms of both data ingest and access, support scalability into the petabyte range and beyond, provide high data resiliency that is optimized for today's large-scale environments, meet five-nines availability requirements, leverage inline data reduction technologies to lower effective dollar-per-gigabyte costs, and offer the data services necessary to effectively manage these large-scale data sets. Features that make storage provisioning and management transparent for already overworked administrators help differentiate the better solution offerings in this area. These platforms are generally designed to work with third-party software products that provide application-specific functionality in the areas of backup, disaster recovery (DR), archiving, and fixed content storage in business settings that require long-term retention to meet compliance, regulatory, and other requirements.

NEC was an early entrant into this market (back in 2007) with HYDRAstor, and today it has a very mature and feature-rich offering that is being used in a variety of verticals worldwide to manage some of the industry's largest data sets. While HYDRAstor meets all the baseline requirements for scale-out disk-based secondary storage platforms noted previously, it stands out as the industry leader in terms of data movement (up to 4PB per hour) while delivering proven scalability, resiliency, and availability at very low cost points.

IN THIS WHITE PAPER

The market for scale-out disk-based secondary storage platforms is growing as customers of all sizes continue to look for more cost-effective ways to manage massive data sets in 3rd Platform computing environments. This IDC white paper takes a look at that market and identifies key customer concerns that are not well met by legacy storage architectures. Then it reviews the NEC HYDRAstor scale-out secondary storage platform, discussing its ability to meet customer requirements for long-term storage of various data types.

SITUATION OVERVIEW

Data is expected to continue to grow at explosive rates, and managing data growth is consistently ranked in IDC surveys as a top challenge among administrators. Even in midsize organizations, storage managers are routinely managing hundreds of terabytes (TBs) to petabytes of data, and compliance and regulatory requirements are forcing longer retention terms. Scale of this magnitude introduces its own set of issues as administrators look to optimize single-platform scalability, performance, data resilience, and recovery while continuing to work within their budgets.

Although primary storage environments are growing rapidly, secondary storage environments will always remain much larger. IDC surveys indicate that secondary data sets for backup, DR, archive, fixed content, and other areas generally make up three to four times the capacity of primary data sets for most enterprises. Each of these secondary application environments presents its own challenges. Backup window concerns and meeting recovery point objectives (RPOs) and recovery time objectives (RTOs) are specific to backup and DR environments. RPO defines the amount of data loss acceptable in the event of a failure, while RTO defines the time it takes to bring a failed application service back into normal operation. eDiscovery costs and features to manage the data life cycle are concerns in archive environments, while access performance across massive data sets is key for fixed content platforms. Concerns about data integrity, effective dollar-per-gigabyte costs, administrative costs, and moving massive amounts of data in a timely manner span all secondary storage environments.

There are a number of issues with tape, the historical medium used for secondary storage platforms, that make it difficult to meet this array of requirements. Tape is a relatively slow medium that does not allow random access (just serial access), and such issues impact data movement performance in data protection (backup windows, recovery times), archive (search costs), and fixed content storage (access performance). Tapes must be physically transferred to offsite locations for DR purposes and can be lost or stolen in transit or misplaced during storage. Tapes wear out and need to be replaced, and they also cannot be searched online, resulting in extremely high ediscovery costs when used in archive environments. Most organizations have moved away from tape because of these issues and have looked to disk-based platforms to take their place.

Evolving business models are driving requirements for ever higher levels of availability as well as increasingly stringent service-level agreements (SLAs) for data access–related issues. Backup windows are shrinking and, in many cases, are already nonexistent. SLAs that define RPOs and RTOs are widely used. Just a few years ago, many organizations did not specifically know their cost of downtime or their RPO/RTO requirements. Today, only 8% of organizations do not know their cost of downtime, and 65% of the organizations that do know their cost of downtime place it at $20,000–100,000 per hour for their most critical applications. Nearly 85% of the organizations have RPOs of

less than an hour, and 78% of them have RTOs of less than four hours. As data sets get larger, the problem of how to manage them while still meeting stringent RPOs/RTOs becomes a concern.

How systems handle failures plays a large part in determining overall availability. Failures will clearly occur, and with availability requirements in the five-nines range (roughly 5 minutes per year) for many environments, systems must provide redundancies to address failures, transparently rebalance workloads in the event of component failures, and support nondisruptive expansion, reconfiguration, and firmware upgrades. Scale-out storage solutions that leverage commodity components — and the inherent redundancies in that design — provide a much simpler and more cost-effective alternative to scale-up architectures, particularly if they are built around proprietary hardware.

Today's Best Practice: Scale-Out Disk-Based Secondary Storage

To address backup window, data resiliency, and RPO/RTO concerns, the industry has moved to disk-based data protection approaches. Disk is a much better fit to meet these requirements than tape, and storage efficiency technologies like compression and data deduplication have helped bring the cost down to manageable levels. Disk enables random access to data and much higher data transfer rates, both during backups and during recoveries. Disk allows technologies like replication to be used to build much more reliable DR strategies that have much better RPO and RTO performance.

While disk has addressed certain data protection issues, it has created other challenges. Early disk targets were based around monolithic scale-up storage architectures designed to address storage requirements in the tens of terabytes range. Today's massive data sets have demanded different architectures, and scale-out platforms scale across a far wider range of both performance and capacity. This extensive scalability does not require expensive forklift upgrades or downtime for expansion or technology refresh. Today, there are scale-out disk-based platforms specifically targeted at secondary storage environments like backup, archive, and fixed content retention that scale easily into the petabyte range and beyond. They also offer a much more granular expansion path that gives customers the option of adding performance and capacity independently, supporting linear scalability and much more balanced performance as configurations scale.

Traditionally, data resilience concerns in disk-based environments have been addressed by RAID, but given its performance and capacity overhead, it does not scale well to efficiently provide protection in today's large environments. The use of large-capacity (>1TB) disks in these massive configurations has made drive rebuild times long enough that there is a relatively high probability of a second error within the rebuild window. Performance in degraded modes is problematic as well, and other concerns with capacity overhead (RAID parity, drive-level sparing) impose costs that are quite high at massive scale. For this reason, many of the scale-out disk-based secondary storage platforms have implemented newer data resilience schemes like erasure coding that actually offer better performance (in both regular mode and degraded mode), require less capacity overhead, and deliver higher levels of resiliency for large-scale environments.

On a dollar-per-gigabyte basis for raw capacity, disk is clearly more costly than tape. New disk-based data reduction technologies, such as deduplication, can be brought to bear to lower costs in these environments. Secondary data sets for backup and DR are highly reducible (i.e., they exhibit high rates of data redundancy), and when compression and data deduplication are used together, data reduction ratios can easily be in the 10:1–20:1 range for many secondary data sets. The data reduction ratio increases the amount of data a given raw physical terabyte can store, and these types of ratios drop the effective dollar-per-gigabyte costs of the most efficient disk-based systems today into the range of $0.05–0.10 per gigabyte at scale.

There are three concerns with the use of data reduction technologies, however, that must be addressed. The first concern has to do with performance. Data reduction processing can add latency, and the processing load increases as the data set grows. It is important to evaluate performance-sensitive metrics like backup window, RPO and RTO, access performance, and data movement times in light of any latencies that data reduction processing may add. The second concern has to do with the impact of data loss. Data deduplication operates by identifying redundant blocks, storing a single copy of that block, and replacing all the other blocks with pointers. If the source block were to be lost or become corrupted, a large amount of data could potentially be affected. For this reason, data resiliency needs to be far stronger than it has been with traditional RAID. The third concern has to do with the scope of deduplication: Does system expansion produce deduplication silos that result in lower deduplication ratios, or does the system employ a global deduplication approach that will result in the highest ratios systemwide? Global deduplication almost always results in higher overall data reduction ratios.

When dealing with massive, rapidly growing data sets, storage provisioning tasks associated with infrastructure expansion and reconfiguration can become onerous. With storage management tasks increasingly migrating to less storage-savvy IT generalists, this is even more of a concern. Managing LUNs and volumes to maintain balanced performance as configurations grow is not intuitive, and particularly in midrange markets, storage management is increasingly performed by administrators with server and/or application, not storage, expertise. The best scale-out secondary storage platforms deploy automation that handles these mundane tasks quickly, reliably, and transparently, not only addressing the "skills" issue but also increasing the span of administrative control by improving operator productivity.

The NEC HYDRAstor Platform

NEC Corp. is a Fortune Global 500 company that was originally founded in 1899 and has annual revenue of $30 billion. The company entered the server and storage markets in 1958 and today offers a wide range of storage solutions that include SAN and scale-out grid platforms targeted for use in both primary and secondary storage application environments. IDC ranks NEC, the North American subsidiary of NEC Corp., as a top 10 vendor in terms of worldwide market share (based on revenue) in the external storage market.

NEC initially introduced the HYDRAstor platform in 2007. HYDRAstor, based around a scale-out grid storage architecture, is a NAS target that offers an optimized on-disk data protection scheme based on erasure coding, an object storage–based back end, inline global data reduction, transparent failover in multinode configurations, and replication. HYDRAstor is available in several configurations, including the entry-level HS3 single-node offering and two larger platform offerings, the HS6 and the HS8, that leverage a scale-out grid architecture. All HYDRAstor software offerings are compatible with the HS3 and HS8 Series. The HS3 and HS8 Series systems are for general-purpose use, while the HS6 option provides a lower price point for platforms built around the same hardware architecture and is targeted for use in archiving environments. The HS6 also offers some unique functionality, such as a single global namespace and optional high-speed image compression for bitmap files, that is targeted specifically at archiving requirements.

Targeted at secondary storage environments, HYDRAstor is capable of handling up to 4PB per hour of data movement and is based on cost-effective enterprise SATA hard disk drives (HDDs) to help keep costs low. Its scale-out architecture allows customers to mix and match processing and storage resources as needed to meet performance and capacity requirements and allows customers to simultaneously support up to three generations of product within a single system to ease platform upgrade and expansion, maximizing the longevity of its life cycle. DynamicStor, the HYDRAstor

operating environment, automatically rebalances data as resources are added (or subtracted) to make it easy to incorporate newer CPU, memory, and storage technologies as they become available without causing any application disruption.

HYDRAstor customers choose from two types of nodes when building configurations: Hybrid Nodes and Storage Nodes. Hybrid Nodes include both processing power and storage capacity, while Storage Nodes include just storage capacity, and both node types are based on industry-standard x86 servers and SATA HDDs. HYDRAstor systems scale from single-node configurations (a Hybrid Node) all the way up to 165-node configurations that can ingest up to 4PB of data per hour and support more than 100PB of effective storage capacity. Customers can mix and match Hybrid Nodes and Storage Nodes as needed to meet performance and capacity requirements, and HYDRAstor provides that range of scalability without requiring any forklift or disruptive upgrades.

At the heart of HYDRAstor is an object-oriented data store that supports NFS, CIFS, and SMB access. Part of the DynamicStor operating environment, this data store can be configured into a single large or multiple smaller namespaces and uses an optimized erasure coding approach to protect data that NEC calls Distributed Resilient Data (DRD). DRD uses a wide striping algorithm to provide very high read and write performance that scales linearly as resources are added.

Under DRD, when data is written to a file share, it is broken into chunks. Each chunk is then broken into fragments, which are written across 12 different physical disks (and up to 12 different physical nodes if there are at least that many in the HYDRAstor configuration). A chunk can be reconstructed by accessing a subset of the 12 fragments, with the processing needed to do so spread across up to 12 nodes. In the default setting, DRD writes 9 data fragments and 3 parity fragments, providing the ability to transparently recover from up to 3 simultaneous physical failures — disks and/or nodes — with only a 25% capacity overhead. The level of data protection is configurable, though — if 8 data fragments and 4 parity fragments are written, then the system can transparently recover from up to 4 simultaneous failures. All parity fragment calculations are based solely on the data fragments and are performed in main memory — no data must ever be read from disk a second time for DRD to complete any operations. At the default protection level 3 setting (9 data fragments, 3 parity fragments), DRD is 50% more resilient than RAID 6 with 25% less overhead while supporting extremely high I/O performance. DRD protection levels are configurable at the file system level so administrators can select their desired level of resiliency.

DRD resolves the rebuild risk at scale. When a disk fails, only the data, not the entire disk, is rebuilt. HYDRAstor does not maintain physical disk spares — all sparing is managed at the system level, promoting more efficient capacity utilization. The processing for the rebuilds is spread across up to 12 nodes, allowing the rebuilds to occur very rapidly even when 4TB disks are in use. The wide striping offered by DRD, combined with its multiple levels of protection, helps maintain high performance even in degraded mode while its configurable data protection removes any data loss risk during the rebuild process.

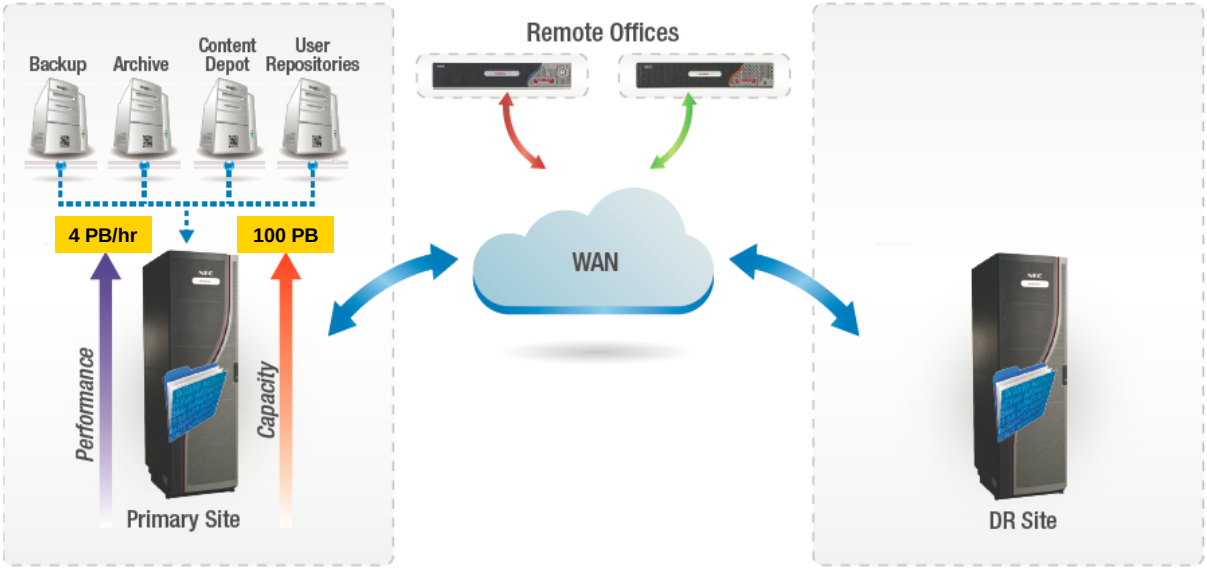
DynamicStor is an auto-provisioning, self-healing storage operating environment with resource auto-discovery that virtualizes all storage resources into a common shared pool (see Figure 1). Storage capacity is dynamically allocated as needed, data is automatically balanced across available resources, and nodes can be added or removed without impacting application availability. Individual nodes are very resilient with active/active controllers and fully redundant power and cooling, all nodes are interconnected over dual-switch interconnects, and DynamicStor works with the inherent redundancies in the grid architecture to support front-end automatic failover. RepliGrid, HYDRAstor's WAN-optimized asynchronous replication facility, enables a variety of replication topologies (one to one, one to many,

many to one) to enable DR that flexibly fits a variety of different requirements. All in all, HYDRAstor includes the features needed to meet enterprise-class availability requirements in the five-nines range.

DynamicStor includes a number of features that simplify data management at scale. Systems employ auto-discovery and auto-configuration capabilities that enable simple, fast deployment – initial system installation can be completed in less than 45 minutes from uncrating to running the first production I/Os. LUN and volume provisioning are performed automatically by the system, and the auto-balancing feature makes time-consuming manual performance optimization unnecessary while delivering optimal performance as configurations scale. The self-healing capability ensures rapid and transparent recovery without data loss, rebalancing data across the remaining resources automatically after a failure.

FIGURE 1

NEC's HYDRAstor Storage System



Source: NEC, 2015

The DynamicStor operating environment also includes DataRedux, HYDRAstor's patented inline data reduction capability. After data is chunked, DataRedux checks a globally distributed hash table to see if the chunk is already stored elsewhere in the HYDRAstor system. Chunking is performed at the subfile level using variable-length windows. If the chunk is unique, then it compresses and writes the fragments representing data and resilience out to disk using the DRD algorithm. If the chunk is already stored elsewhere, it merely updates the relevant pointers in the hash table. The globally distributed hash table ensures that I/O performance stays high even as configurations are scaled because the overall load stays evenly spread across resources as they are added. Individual nodes perform the DRD and DataRedux work for only the data they directly manage, contributing to the system's well-balanced performance.

In addition to the deduplication and compression, DataRedux also leverages some backup application-specific data reduction algorithms to further increase data reduction ratios. Application-aware data deduplication leverages format awareness to filter metadata inserted by the application and deduplicate the data payload separately. Specifically designed for use with data protection solutions like Veritas NetBackup, CommVault Simpana, IBM Tivoli Storage Manager, EMC NetWorker, HP Data Protector, Oracle Recovery Manager, and Windows system recovery, this facility can increase data reduction ratios by up to 130% over a single weekly cycle (full, five daily incrementals, and then another full) and by as much as another three to four times over time.

A full suite of enterprise-class data services that improves administrative options, security, and ongoing management is available for the HYDRAstor platform. In addition to DRD and DataRedux, DynamicStor includes space-efficient snapshots and clones, dynamic data shredding for secure data disposal, and automatic and transparent node failover. Additional options include RepliGrid, HYDRAstor's WAN-optimized replication with in-flight encryption; HYDRAlock for WORM file system functionality; encryption at rest; and the HYDRAstor OpenStorage (OST) Suite. The OST Suite includes features specific to Symantec NetBackup (formerly Veritas NetBackup), such as Express I/O, Dynamic I/O, and Deduped Transfer for higher data movement performance, optimized synthetics, accelerator, optimized copy, and support for NetBackup's Auto Image Replication feature.

NEC has adapted both the Express I/O and the Deduped Transfer features for use with any third-party data protection software that runs on Linux. These versions provide the same high performance and very efficient data movement functionality that the OST Suite-based versions do for NetBackup but make them available for use with other data protection solutions.

NEC has been successful selling HYDRAstor, with well over 1,000 systems across more than 500 customers worldwide. Although the product is applicable across many verticals, NEC has achieved the most success in the financial services, broadcasting/entertainment, medical/healthcare, government, retail, and manufacturing verticals. There are some very large HYDRAstor systems in production, and NEC can point to referenceable customers that have well more than 50 nodes and tens of petabytes of effective capacity under management. Several HYDRAstor capabilities consistently drive purchase decisions across verticals, including the massive scalability of both capacity (supported by the inline data reduction capabilities of DataRedux) and performance, the data integrity of HYDRAstor's DRD erasure-coded data protection, and the flexibility that a scale-out architecture provides for investment protection and the ability to accommodate new technologies nondisruptively.

FUTURE OUTLOOK

3rd Platform computing has driven transformations all across the datacenter, and the rise of disk-based secondary storage platforms is one of them. By leveraging a variety of new disk-based technologies, these platforms have become compelling infrastructure choices not just because of the management advantages that their functionality brings to the table in the areas of data protection, DR, and archive but also because of their scalable performance against massive data sets and the cost advantages disk-based technologies like thin provisioning; space-efficient implementations of snapshots, clones, and replication; and inline data reduction bring to the table. Data growth is not expected to abate, and in a few short years, even small to medium-sized enterprises will be struggling with efficiently managing petabytes of data. There is no doubt that scale-out disk-based storage architectures with self-managing capabilities, such as NEC HYDRAstor, make the Herculean task of managing these large data sets much easier.

These platforms are expected to evolve in two directions: toward broader support of secondary storage applications and toward primary storage. Data protection has become increasingly snapshot based, and these platforms will need to continue to evolve so that they can work better and more broadly with developing technologies in this area. As the percentage of data that is unstructured continues to dominate new data growth, support for new access protocols in these platforms will proliferate so that the storage consolidation message can include all data types that datacenters routinely manage — structured, unstructured, and semistructured. Newer technologies like flash can be intelligently integrated into these environments so that lower latency requirements can be met, allowing them to expand more into primary storage environments. By the nature of the work they do, datacenters will always require different types of storage platforms, but the ability to consolidate more of the workload onto a single, massively scalable platform that can flexibly meet all requirements provides for broader spans of administrative control, making management operations easier and more efficient. These developments will also drive the requirement for better multitenant management capabilities on these platforms.

CHALLENGES/OPPORTUNITIES

When HYDRAstor was introduced in 2007, it was one of the first disk-based secondary storage platforms built around a scale-out architecture — certainly the first one available from a Fortune Global 500 company. Today, HYDRAstor is a very mature product that provides high performance, broad scalability, and excellent availability and reliability and offers a rich feature set for its target markets. HYDRAstor is deployed today at over 500 customers, some of which originally purchased the product back in 2007 and have three generations of HYDRAstor nodes in production. As more customers turn their attention to new storage architectures to help them better address their massive data management requirements, NEC needs to ensure that it makes the appropriate marketing investments to generate awareness about this aggressively competitive platform.

With new flash technologies becoming available, there is a distinct opportunity for scale-out platforms, which to date have primarily targeted secondary storage environments, to expand into primary storage environments. It is not necessary for these platforms to deliver the same type of microsecond latencies that today's leading all-flash arrays (AFAs) provide for them to offer significant value in many primary storage environments. Many software-defined storage (SDS) products designed to be run on scale-out architectures based around commodity x86 servers with internal storage include storage tiering capabilities that give them an ability to incorporate a flash tier and address not just secondary storage applications but also primary storage applications. These vendors are actively pursuing these markets, leveraging the benefits of scale-out architectures against their monolithic scale-up competitors. If NEC were to consider expanding HYDRAstor into these areas, the product would offer a much more mature set of capabilities in the data protection and data reduction areas than many of those SDS platforms can offer, and it has a strong bench of reference customers that can attest to its proven scalability and reliability.

Other opportunities that are a good fit for HYDRAstor would include providing data protection suites similar to its current OST Suite for specific use with other key backup platforms and offering a unified storage platform. HYDRAstor already offers optimized extensions for Symantec NetBackup, which commands the number 1 market share position in software-based backup platforms, but there could be opportunities with other market-leading backup solutions that don't have a strong hardware platform tie-in, such as CommVault and a newly invigorated (and spun off) Arcserve (which focuses on the midmarket). Note that HYDRAstor already supports Universal Express I/O and Universal Deduped Transfer, two features that support very efficient data movement and can be used with any Linux-based third-party backup products. Customers looking to consolidate storage onto a single platform

benefit considerably from offerings that support block-, object-, and file-based access. This may be another area into which NEC could consider expanding with a platform (HYDRAstor) that can already handle many storage needs beyond just data protection requirements.

CONCLUSION

Scale-out disk-based platforms offer a very attractive choice for secondary storage environments today. These will quickly become a mainstay in most enterprise environments that need to manage data sets that are in the 100TB range or beyond — the advantages they bring to the table over scale-up architectures and tape-based alternatives are just too compelling. Scale-out architectures can provide scalability into the 100PB range and support the well-balanced linear scalability that dynamic businesses need to manage 3rd Platform computing data sets. This market has developed quickly, and IDC can clearly identify the baseline functionality requirements for these environments: high performance in terms of both data ingest and access, scalability into the petabyte range and beyond, high data resiliency that is optimized for today's large-scale environments, five-nines availability, inline data reduction technologies to lower effective dollar-per-gigabyte costs, and the data services necessary to effectively manage these large-scale data sets.

NEC HYDRAstor meets these requirements and brings a strong long-term track record in the industry for performance, data resiliency, nondisruptive technology refreshes, and low cost. Initial shipments occurred in 2007, and NEC has a number of customers that have continued to expand their systems across three generations of technology refreshes without any disruption. Effective dollar-per-gigabyte costs are under $0.10 in most configurations. Customers looking for a secondary storage consolidation platform that delivers all the advantages of disk-based data management in a very scalable, cost-effective package should consider HYDRAstor.

About IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications and consumer technology markets. IDC helps IT professionals, business executives, and the investment community make fact-based decisions on technology purchases and business strategy. More than 1,100 IDC analysts provide global, regional, and local expertise on technology and industry opportunities and trends in over 110 countries worldwide. For 50 years, IDC has provided strategic insights to help our clients achieve their key business objectives. IDC is a subsidiary of IDG, the world's leading technology media, research, and events company.

Global Headquarters

5 Speen Street
Framingham, MA 01701
USA
508.872.8200
Twitter: @IDC
idc-insights-community.com
www.idc.com