NEC HYDRAstor: A Fifth-Generation Purpose-Built Backup Appliance

Phil Goodwin

IDC OPINION

Purpose-built backup appliances (PBBAs) have evolved into a cornerstone technology for many organizations' data protection strategy. Data protection has evolved from daily backup and recovery to tape to a multilayered strategy of snapshots, mirrors, and remote replication. Even so, daily backup is still required, involving age-old issues around backup window, manual handling of tapes, and backup failures due to tape infrastructure failures. Daily backup operations are inherently labor intensive, requiring daily loading and unloading of tapes and tracking and transferring tapes to an outside vault. In addition, storage administrators must triage failed jobs and determine the root cause, correct the error if necessary, and restart critical jobs. Organizations may suffer as much as 15% backup job failure on a nightly basis, resulting in dozens of failed jobs to correct. Organizations that aspire to best practices will achieve <5% backup job failure. Backup operations do not directly contribute to a market competitive advantage for organizations (except, perhaps, those that purvey backup services). In other words, organizations that make widgets do not sell more widgets because customers perceive that they have awesome backup. IDC research indicates that the average cost of downtime exceeds $100,000 per hour. Lost revenue can be especially acute for Web-based transactional businesses; thus any reduction of downtime delivered by faster recovery mechanisms offsets the cost of the infrastructure needed to achieve that outcome. IDC believes that organizations looking to improve their backup strategy and improve recovery times will need to consider the following guidelines:

- Look for technologies that reduce the possible points of backup/recovery (B/R) failure, both mechanical failure and human error.
- Minimize labor-intensive operations, to both reduce errors and free up staff to take on more strategic tasks.
- Invest sufficiently in technology that will meet business-driven SLAs, balanced against a calculated cost of downtime, such that the investment meets business needs but no more.

IN THIS STUDY

This document describes the benefits that PBBAs can deliver to organizations in terms of improved backup and recovery operations. More specifically, it examines the fifth-generation release of NEC's HYDRAstor PBBA.

SITUATION OVERVIEW

Increasingly stringent business demands put constant pressure on IT organizations to continually improve system availability and reduce data loss. IDC research indicates that 61% of organizations surveyed have a recovery time expectation of less than four hours and that 61.7% of organizations expect data loss to be less than one hour's worth of data (i.e., RPO <1 hour). State-of-the-art systems installed five years ago are inadequate for current business requirements, and it can be further said with certainty that today's systems will be inadequate five years from now. This constant, incremental change challenges IT managers to make strategic data protection deployment decisions that will meet current requirements, yet evolve over time to match business needs.

One of the biggest game changers over the past decade regarding backup/recovery has been the introduction of purpose-built backup appliances. Certainly, snapshots and replication have yielded significant benefits to the overall task of data protection, but no product category has had more positive impact on B/R specifically than the PBBA.

As a backup target, PBBAs can offer significant advantages over traditional tape-centric backup methodologies. First of all, PBBAs can deliver greater throughput than tape, though organizations can always throw more tape at the problem if budget is not much of a consideration. PBBAs can take advantage of the random nature of disk technology, without the limitations of serial read/write of tape. There is no need to consolidate tapes, worry about conflicting retention policies, or periodically recall and rewrite archived tapes.

Another benefit of PBBAs is that they eliminate the mechanical frailties of tape systems, which bedevil any B/R operation. Tapes break, get lost or stolen, and become unreadable over time. IT organizations are forced to retain multiple generations of obsolete software and hardware to ensure that archive tapes in the vault can be read in the rare case when they're needed. In contrast, the disk-based nature of PBBAs ensures that data will always be accessible in a readable format and will not need to be rewritten for integrity. As PBBAs are upgraded over time, the data is brought along with it.

PBBAs come with common feature sets, most notable of which is aggressive deduplication (dedupe) capabilities (i.e., 20:1). Different dedupe techniques are used, and it's important for buyers to know the difference. Some PBBAs dedupe data on the device itself, which has the advantage of offloading any such processing from the sending server (i.e., backup server). Others use client-side dedupe, or dedupe on the sending server. Here, the advantage is that duplicate data is never sent, reducing bandwidth requirements but does require processing on the server. In addition, some systems offer global dedupe, which applies deduplication across multiple devices when more than one PBBA is deployed. This can be highly space efficient but does require processing and network traffic. Deduplication does entail processing overhead to "rehydrate" data when it is recovered, but this overhead is almost certainly less than the overhead of retrieving, loading, and positioning tape.

Other features common to PBBAs include encryption and data replication. Given the data theft threats faced by organizations, simply encrypting backup data by default is the best practice policy. This is true whether the data is inside or outside the datacenter and regardless of the media use for storing it. Data replication allows data to be moved to secondary storage outside the datacenter. This is usually accomplished in the background so as to avoid impact to any backup operations. It also results in data being moved offsite for disaster recovery purposes sooner than tape would be, and especially in large volumes, it can be recalled faster than tape.

This last point addresses the real crux of the matter, which is better data recovery. Because fewer backup jobs are likely to fail due to mechanical issues or drive-busy conditions, PBBAs improve the probability of data recovery in that critical 48-hour window after backup, when the majority of restore requests are made and data loss is most vulnerable due to failure. PBBAs also eliminate the possibility of unrecoverable data due to a lost or broken tape. Upon recovery, tapes do not need to be recalled and data recovered sequentially; data from a PBBA can be recalled as fast as the network and number of spindles allow.

The advent of cloud computing has also impacted PBBA functionality. IDC research shows that 53% of organizations use cloud for some portion of their data protection scheme, and we expect this number to rise to 88% by 2018. To support this dynamic, PBBA vendors are adding Web storage protocols to their products (e.g., S3, OpenStack). There is also a class of product known as cloud gateways, which do protocol conversion for data streams from on-premise to cloud repositories. The key difference between gateways and PBBA is that gateways do not store data and are not used as sources for data recovery, whereas PBBAs are used for both.

FUTURE OUTLOOK

NEC's HYDRAstor was one of the earlier entrants to the PBBA market, nearly a decade ago. The product line has evolved with the market and technology – most recently reflected in the fifth-generation product, the HS8-5000. The HS family of products ranges from the single-node, stand-alone HS3-510, suitable for remote office/back office (ROBO), to HS8-5000 single hybrid node configuration, with up to 165 nodes configuration combining hybrid nodes and storage nodes, depending upon performance and capacity requirements. Hybrid node provides performance and capacity, and storage node provides just capacity.

The HS8-5000 boasts a 54% effective capacity increase and 25% maximum performance increase over the previous-generation HS8-4000. Moreover, the system has a massive scale-out potential from 1 to 165 nodes, yielding 11.88PB raw capacity (158PB effective capacity for a typical backup data set) and up to 5PB per hour throughput with NEC's light-weight protocol, Express I/O, and client-side deduplication, Deduped Transfer. In this configuration, the product is one of the most scalable in the industry.

One of the new elements that helps expand the capacity of the HS8-5000 series systems is the introduction of 6TB SATA drives. However, such large drives in such extensive configurations increase the risk of significant rebuild times for failed drives, even in a RAID 6 configuration. To eliminate this risk, NEC has implemented erasure-coding technology rather than RAID. Moreover, idle spare drives are not required for erasure coding, meaning that every drive can be put to useful purpose storing data and utilizing more spindles for more performance.

Along with erasure coding comes the ability to choose resiliency levels in scale-out configurations. The default mode protects against three concurrent failures (compared with two for RAID 6) but is user configurable from one to six depending upon the SLA and application requirements. With erasure coding, only data in reconstructed, rather than entire drives, and the data is reconstructed across multiple spindles.

Because of the massive scale-out capability, NEC has built in both client-side deduplication and global deduplication. As noted previously, the client-side dedupe minimized bandwidth requirements, making it well suited to serving geographically distributed ROBO locations and/or many server devices. The global deduplication ensures maximum capacity optimization across the multinode configuration.

HYDRAstor also has the ability to manage data between on-premise systems and remote sites. Data on the primary system is deduplicated at the client (backup server) and transferred to HYDRAstor. HYDRAstor WAN-optimized replication, RepliGrid, can minimize data transfer between two locations by transferring only unique data block and compression, and in-flight encryption is also supported. If sufficient infrastructure is available at the remote site, then the data can be restored at that location, such as in a disaster recovery scenario.

ESSENTIAL GUIDANCE

While tape remains a viable technology for long-term data archive due to its low cost per gigabyte of storage, it is now largely relegated to secondary storage uses. Primary backup storage is disk based, and much of that is implemented as a PBBA. Organizations that rely on tape as a primary backup media will want to consider PBBAs for the following reasons:

- Faster backup and restore operations, resulting in better SLA delivery
- Fewer backup failures
- Capacity optimization with dedupe
- Faster offsite data transfer than with tape rotations

NEC has been delivering solid PBBA systems for as long as many better-known competitors and is now marketing its fifth-generation devices such as the HYDRAstor HS8-5000 series. The product has a vast majority of required features, including deduplication, scale-out capacity, and remote replication. Adding high-capacity disk drives with erasure encoding improves resilience and performance relative to a RAID 6 configuration. The systems are capable of AES-256 inline encryption, though this is an optional feature.

NEC's challenges with HYDRAstor will be adding cloud connectivity so that organizations can utilize low-cost cloud storage for off-premise data protection and possibly disaster recovery. Although NEC is an enormous company with a storied and well-respected history and presence in Japan and other Asian countries, it is not as well known in the United States. It will need to continue to focus on brand recognition and a channel push strategy to gain market share in North America.

LEARN MORE

Related Research

- *Worldwide Purpose-Built Backup Appliance (PBBA) Market Revenue Resumes Its Growth in the Third Quarter, According to IDC* (IDC #prUS40855615, December 2015)

Synopsis

This IDC study describes the benefits that purpose-built backup appliances (PBBAs) can deliver to organizations in terms of improved backup and recovery operations. More specifically, it examines the fifth-generation release of NEC's HYDRAstor PBBA.

"Purpose-built backup appliances offer a significant opportunity to improve backup operations. IT organizations wishing for improved data protection service-level delivery, reduced manual labor associated with backup/recovery, and reduced backup failures will want to consider adding PBBA devices as a primary backup target, in many cases replacing tape," said Phil Goodwin, research director, Storage Systems and Software.

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