



Product Spotlight

NEC HYDRAsTOR – Generation 4

NEC has made a significant architectural enhancement to its grid-based storage system. HYDRAsTOR is a scale-out NAS optimized for use as an extremely high-throughput backup target and long-term data archiving platform. With global deduplication, erasure-coded data resiliency that avoids the problems of traditional RAID, replication for disaster recovery and remote office backup, seamless scalability, and non-disruptive technology upgrades, HYDRAsTOR has long been one of the most impressive storage platforms available in the market. Generation 4 enhancements make it even more compelling.

NEC launched its innovative HYDRAsTOR technology in 2007 and has continuously evolved it into a mature full-featured platform, yet too many storage buyers remain unaware of its existence. That's unfortunate, because HYDRAsTOR is an outstanding product, outclassing more well-known competitors in several respects. Many enterprises would benefit from deploying a HYDRAsTOR system, if only they knew about it. This report is Data Mobility Group's attempt to shine a much-needed spotlight on this unique and impressive storage platform.

Let's begin with a quick overview of HYDRAsTOR's most distinguishing characteristics:

- **Scale-out grid architecture:** HYDRAsTOR is composed of independent nodes that work together to comprise a shared pool of storage that is self-organizing and self-healing. Simply by adding nodes, HYDRAsTOR can seamlessly scale to massive capacities and throughputs (we'll provide numbers later in this report).
- **Global deduplication:** HYDRAsTOR deduplicates data across the entire grid and compresses data inline, yielding an effective storage capacity as much as 20x greater than raw capacity. Its inline deduplication engine cleverly utilizes a distributed hash table and harnesses the power of CPUs throughout the grid for unmatched input performance.

Copyright © 2002-2013 Data Mobility Group, LLC. All Rights Reserved. Reproduction of this publication without prior written permission is forbidden. Data Mobility Group believes the statements contained herein are based on accurate and reliable information. However, because information is provided to Data Mobility Group from various sources, we cannot warrant that this publication is complete and error-free. Data Mobility Group disclaims all implied warranties, including warranties of merchantability or fitness for a particular purpose. Data Mobility Group shall have no liability for any direct, incidental, special, or consequential damages or lost profits. The opinions expressed herein are subject to change without notice. Any companies mentioned in this report may be past or present clients of Data Mobility Group.



- **Efficient, flexible data resiliency:** To overcome the performance and reliability problems of traditional RAID—problems that are becoming more pronounced with the adoption of large-capacity drives and arrays—some newer storage systems have turned to erasure-coding as an alternative. HYDRAsstor has employed erasure-coded data resiliency since the beginning, and NEC’s implementation enables the system to reconstruct data with little or no performance degradation. What’s more, HYDRAsstor allows administrators to dynamically configure different resiliency levels for different file systems just by turning a metaphorical dial from 1 to 6. The default level of 3 provides several times more data protection than RAID 6, yet uses less disk space.
- **Non-disruptive technology upgrades:** The HYDRAsstor grid supports mixing nodes of different generations. This enables an organization to retire older nodes in a piecemeal fashion, replacing them with newer nodes that have faster CPUs, more energy efficiency, and larger or faster disks. Because HYDRAsstor is self-organizing, it will automatically incorporate the new nodes and redistribute data with no downtime or performance penalty. This gives HYDRAsstor users a high level of investment protection as technology improves over time.
- **Optimized WAN replication.** Enterprises can replicate data between grids for disaster recovery and business continuity, as well as replicate data from HYDRAsstor’s “mini-HYDRA” (the single-node HS3-410 model often deployed in remote offices and small businesses) to another mini-HYDRA or a larger centralized grid.

HYDRAsstor includes a plethora of other features that help make it a robust storage platform for enterprise backup and long-term data retention: data encryption, both in-flight and at rest; WORM (Write-Once, Read-Many) and data shredding for compliant retention; file system snapshots/clones and instant file copies.

In addition, NEC has an extensive NetBackup/OpenStorage integration with Symantec, has forged technology partnerships with Dell/Quest and Veeam for advanced data protection of virtual machines, and provides application-aware integrations with a variety of other backup software, including IBM Tivoli Storage Manager, EMC NetWorker, CommVault Simpana, and HP DataProtector.

Like we said, HYDRAsstor is a mature and full-featured product. So...

What’s new in Generation 4?

Several things, of course, but the three most important changes NEC has made are these:

1. Introduced a new type of integrated node for greater efficiency.
2. HYDRAsstor now uses 2TB and 4TB drives instead of 1TB.
3. Performance and throughput is improved across the board.



As we'll explain shortly, these changes have boosted HYDRAsstor's already staggering ingest performance into the realm of the truly mind-boggling, and have increased the system's overall effective capacity to somewhere in the neighborhood of 100 petabytes.

In HYDRAsstor's prior generations, the building blocks of a grid consisted of two different types of nodes: Accelerator Nodes, dedicated to intensive data processing tasks, and Storage Nodes, chockful of hard drives. If you needed more throughput and performance you could add more Accelerator Nodes, while if you needed more capacity you'd add more Storage Nodes.

HYDRAsstor Generation 4 introduces a new **Hybrid Node** that combines the functionality of both Accelerator Node and Storage Node, providing added throughput/performance and capacity in a single node. This new Hybrid Node takes the place of the Accelerator Node in HYDRAsstor grids. The second type of node, the pure Storage Node, remains available as a lower-cost alternative to Hybrid Nodes; it has the same capacity as a Hybrid Node, but does not offer any additional performance.

The introduction of Hybrid Nodes simplifies and shrinks HYDRAsstor configurations. For example, in Generation 3, the smallest fully redundant configuration required 12U of space for two accelerator nodes plus four storage nodes, whereas Generation 4 only requires 8U, for two hybrid nodes plus two storage nodes. Factor in the use of 4 TB drives instead of 1 TB drives, and Generation 4 crams 192 TB of raw capacity into 50% less space than Generation 3 used for 48 TB. With deduplication and compression, NEC estimates the effective capacity of this 8U configuration (actually 10U, if you add in 2U of switches) tops out at 2.5 PB—your mileage may vary, of course, but the prospect of that much effective capacity in such a small footprint is impressive.

The gains from using larger hard drives apply from the smallest HYDRAsstor to the largest. The Generation 3 single-node mini-HYDRA was available with raw capacities of 4, 8, and 12 TB. In Generation 4, the mini-HYDRA uses 2TB disks instead of 1TB disks, doubling raw capacities to 8, 16, and 24 TB. In larger multi-node HYDRAsstor grids, maximum capacities have more than quadrupled, because not only does Generation 4 use 4TB drives instead of 1TB drives, it also replaces storage-less Accelerator Nodes with storage-full Hybrid Nodes. Each Generation 4 node, whether a Hybrid Node or a Storage Node, contains 48 TB of raw capacity (624 TB effective). A full HYDRAsstor grid of 165 nodes (in 11 racks) now has 7,920 TB of raw capacity, with effective capacity estimated at an astounding 103 PB.

Even more astounding than the storage capacity of the Generation 4 HYDRAsstor is its performance as a backup target. Performing inline deduplication on the backup stream, the high-end HS8-4000 model can ingest 802 TB/hour (921 TB/hour using OST Express I/O). To give you an idea of how that compares to competitive systems, the maximum throughput of EMC's DataDomain DD990 system is claimed to be 31 TB/hour according to the company's own marketing materials.




A future-proof storage platform

NEC freely admits that none of their customers actually need the maximum performance that HYDRAsstor is capable of delivering. Indeed, none of their customers presently require even 100 TB/hour throughput. Nor do any of their customers presently require dozens of petabytes of capacity. But the key word is presently—because storage needs are constantly growing, and in the future these seemingly extreme performance and capacity figures will likely become common mainstream enterprise requirements. Organizations that deploy less scalable storage systems are volunteering for disruptions, inefficiencies, information silos, and an all-too-familiar litany of other issues that arise when rapid data growth meets inadequate, inflexible infrastructure.

In the short term, organizations that deploy HYDRAsstor get to enjoy the many benefits of highly performant, self-organizing, self-healing grid storage, but they also gain the comfort of knowing they've set themselves up nicely for the longer term.

HYDRAsstor is built for the long haul. It provides plenty of headroom in terms of capacity and performance. Perhaps even more importantly, it is designed to continuously incorporate new and better technology components in a seamless, non-disruptive manner. The roadmap to the future for HYDRAsstor customers is simple and clear: a decade from now they will be running Generation 8, 9, and 10 nodes, having long since retired their obsolete Generation 4 components, and they will have experienced no downtime or data loss in the process.

Summing up

HYDRAsstor stacks up extremely well against the competition. The changes in Generation 4 have made an already strong product even better. The only caveat we'd mention is that HYDRAsstor was designed to meet the requirements of the most demanding enterprises, and therefore it may not be a good fit for smaller environments that project little data growth. (That said, even small businesses and remote offices will likely find the mini-HYDRA, or single-node Hybrid Node starting point, to be a compelling solution.) Any organization in the market for a massively scalable platform for long-term data archiving, or an extremely high-throughput backup target, or both, should absolutely put HYDRAsstor on their shortlist. It's frustrating to us that NEC hasn't been more aggressive in marketing HYDRAsstor. More organizations need to know about it. 

For more information:

NEC's HYDRAsstor landing page can be found at: <http://www.necam.com/HYDRAsstor/>

If you have questions about NEC HYDRAsstor and how it compares to available alternatives, we invite you to contact Data Mobility Group and arrange a FREE teleconference to discuss it. We encourage you to invite members of your technical and business staff to join the conversation.