NEC HYDRAstor and NetBackup Advantage

In Pursuit of Perfection

HYDRAstor

NEC HYDRAstor Advantages

For more than 100 years, NEC has been dedicated to pursuing frontiers and leading technological innovation. With over 67,000 patents worldwide, NEC maintains deep academic affiliations at its 12 R&D facilities in the United States, Japan, Germany & China. In the top rankings in U.S. patents for the last 12 years, HYDRAstor continues the NEC innovation record with nine technology patents filed by NEC as part of the HYDRAstor grid storage platform.

The HYDRAstor architecture announcement in March 2007, from its inception date in 2005, has been specifically optimised, marking the culmination of more than four years of intense research, development, and market validation. Nothing is more satisfying for an industrial research team than to see their innovation solve real customer problems. The HYDRAstor project started with a pair of researchers obsessed with their vision of a whole new way of providing storage. This grew into a global product realisation effort that drew on NEC expert resources worldwide. NEC HYDRAstor is a massively scalable distributed grid storage platform built to support long-term data retention by reducing storage capacity consumption by up to 95% or more using inline global deduplication. The online upgrade/expansion and data in-place technology refresh capability make it unique in the industry. Data is migrated to newer technology seamlessly with no migration planning or effort.

NEC HYDRAstor’s scale-out grid architecture provides performance and capacity scalability, capacity optimisation, advanced erasure-coded data protection, and high availability with node-level resiliency. By incorporating newer generation hardware into the same grid system, HYDRAstor can be further expanded and refreshed online without data migration, maximising investment protection.

NEC HYDRAstor advantages (cntd.)

A HYDRAstor grid used both HNs (hybrid nodes) and SNs (storage nodes). The hybrid nodes are, just like the storage nodes, used to store data but also deliver the 10GbE front-end connectivity. Because of this, the hybrid nodes are configured with better CPUs and more RAM for higher performance than the storage nodes.

In our solution, we make use of the HS8-50 nodes with 8TB disks and 6TB disks to have a balanced solution that can scale in capacity by adding storage nodes only or in capacity and performance & throughput by adding hybrid nodes to the grid.

A hybrid node is configured with the following hardware:

Host Interface (and Replication Port)		1000BASE-T x 6
		1000BASE-T x 4 + (10GBASE-SR x 2 or 10GBASE-T x 2)
		1000BASE-T x 2 + (10GBASE-SR x 4 or 10GBASE-T x 4)
CPU		Xeon E5-2660 (2.60GHz 10 Core) x 2 CPU
Memory		128GB (16GB DIMM x 8, DDR4 2133MHz)
Disk Drive	Capacity	72TB/96TB (SATA 3.5” 7200rpm, 6TB/8TB x12)
	Hot Swap	Supported during HDD replacement
Footprint		2U
Power		Redundant Power Supply (AC100~240V)
FAN		Redundant FAN
Other		Upgrade from the 1GbE model to the 10GbE model is possible.

A storage node is configured with the following hardware:

Host Interface (and Replication Port)		None
CPU		Xeon E5-2620 (2.40GHz, 6 Core) x 2 CPU
Memory		128GB (16GB DIMM x 8, DDR4 1866MHz)
Disk Drive	Capacity	72TB/96TB (SATA 3.5” 7200rpm, 6TB/8TB x12)
	Hot Swap	Supported during HDD replacement
Footprint		2U
Power		Redundant Power Supply (AC100~240V)
FAN		Redundant FAN
Other		None

NEC HYDRAstor advantages (cntd.)

Nota Bene:
A HYDRAstor Grid can contain nodes with different capacities. Each node uses 12 disks, with a total capacity of 72TB (6TB disks), 96TB (8TB disks), or 168TB (14TB disks).

Variant 1 of our solution uses 8TB nodes to meet the Polish Bank's capacity and performance requirements.

We require a minimum of 48 nodes to meet the TB/h performance requirements. By using smaller capacity nodes, we can not only achieve the required performance but also not oversize the solution in terms of RAW Capacity and increased costs. When using the 14TB nodes, we would increase the RAW capacity of the solution by 75% compared to the 8TB nodes.

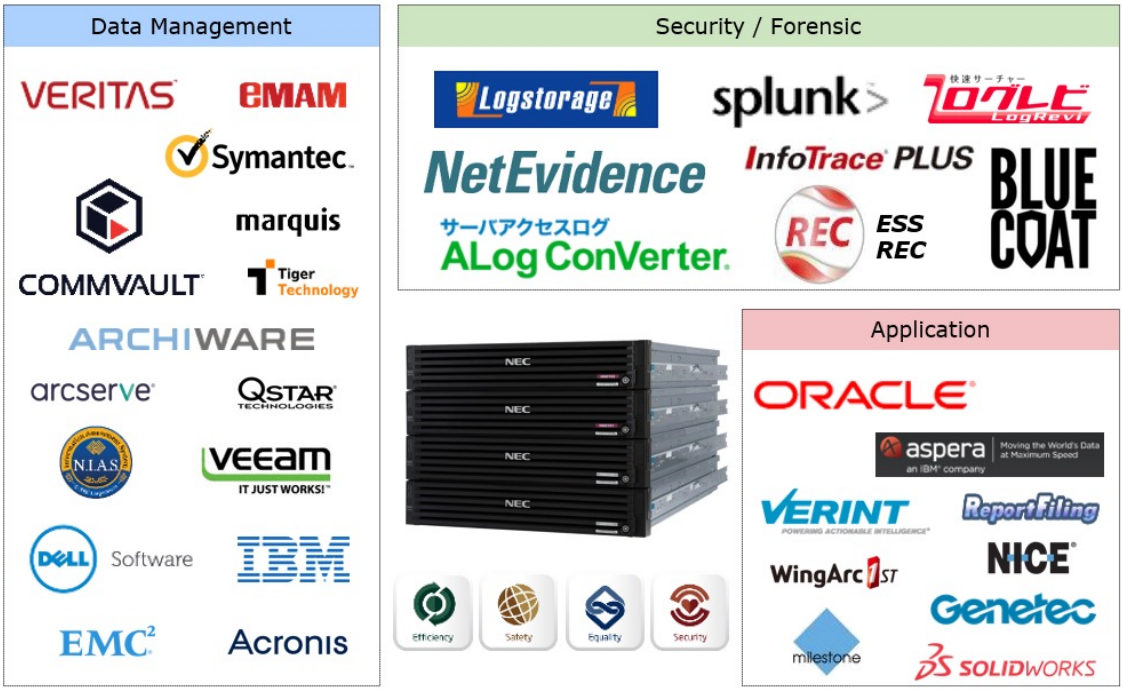
a) Advantage: Unmatched scalability and performance

HYDRAstor is composed of independent nodes that work collectively to provide a shared pool of storage that is self-organising and self-healing. By adding nodes, HYDRAstor can seamlessly scale in performance and capacity. The maximum size of the system is 165 nodes, offering over 16 PB of raw storage space. With deduplication, such a system can keep 160 PB or more of backup data with writing speeds exceeding 4PB/hour for deduplicated data.

There are two types of nodes: a hybrid node (HN) adds both storage capacity and performance on the front end, whereas a storage node (SN) adds storage capacity.

b) Advantage: standard and special interfaces

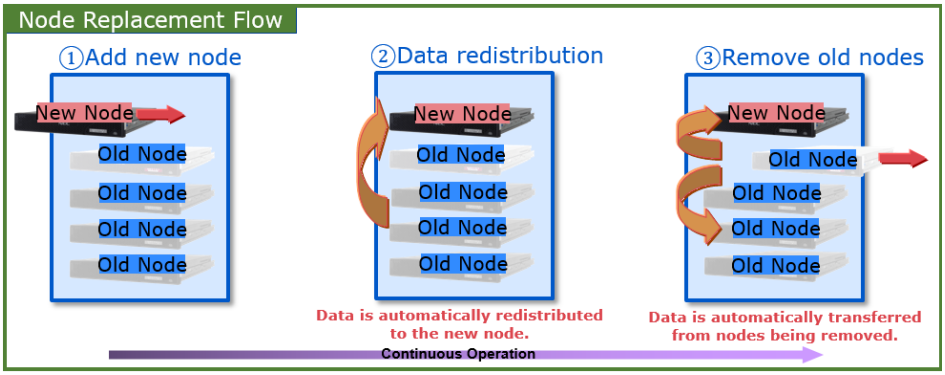
By supporting standard interfaces like NFS and S3, HYDRAstor can work out of the box with many backup and archival applications that utilise these interfaces. At the same time, HYDRAstor implements specialised storage interfaces like Veritas OST, delivering rich functionality and further improvements in deduplication and performance.



c) Advantage: Incremental extensions and non-disruptive technology upgrades

HYDRAstor, through its unique grid-based architecture, allows for incremental upgrades to the existing grid by offering linear performance (with the addition of a single hybrid node) and/or capacity (by adding a storage node). The existing grid can, therefore, be upgraded to meet either a performance or a capacity requirement, online and non-disruptively, with the added benefit that the grid will automatically re-balance data across all the nodes and, therefore, improve overall redundancy. Up to three generations of node architecture are supported within a single HYDRAstor grid.

Enabling three generations of technology to co-exist at any point within the grid allows customers to avoid periodic data migration projects and ensure data availability from inception for many years from creation, if required. This also means that the customer can benefit from new technological advances (performance, capacity, and functionality) within the existing grid infrastructure without overhauling the entire infrastructure.



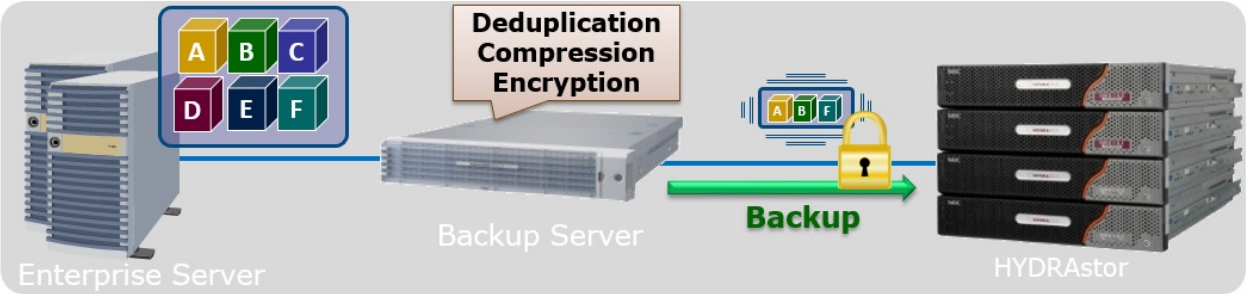
d) Advantage: Global in-line deduplication

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 intelligently utilises a distributed hash table and leverages the power of CPUs throughout the grid for unmatched input and output performance.

Data is deduplicated for all applications writing to a given HYDRAstor system. In-line deduplication enables extremely high throughput for deduplicated data and avoids reserving additional storage in post-process deduplication.

e) Advantage: Source-level deduplication

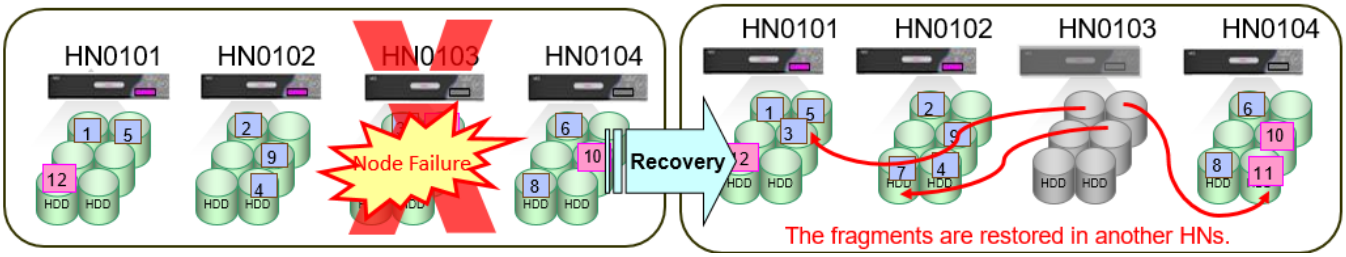
With special plug-ins, HYDRAstor allows for deduplicating data on the source of data, reducing network traffic and improving write performance.



f) Advantage: Efficient, flexible data resiliency

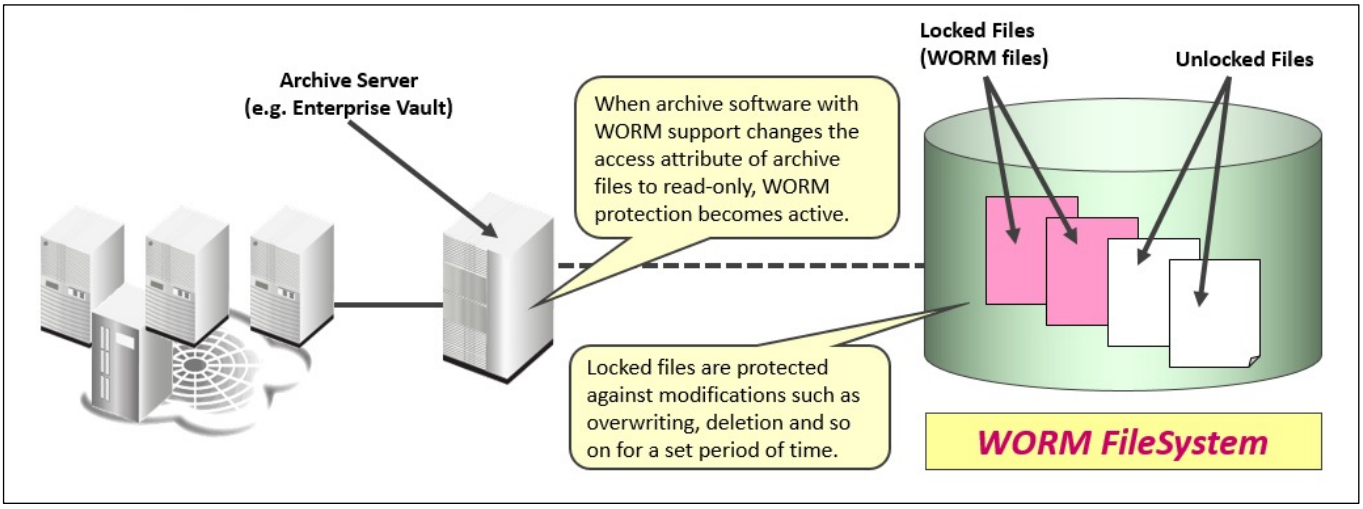
Some newer storage systems have turned to erasure-coding as an alternative to overcome the performance and reliability problems of traditional RAID—problems that are becoming more pronounced with the adoption of large-capacity drives and arrays. HYDRAstor has employed erasure-coded data resiliency since the beginning, and NEC's implementation enables the system to reconstruct data with little or no performance degradation.

What's more, HYDRAstor 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 more data protection than RAID 6 yet uses less disk space. Greater Data Resiliency with Advanced Erasure Coding – HYDRAstor Distributed Resilient Data™ (DRD) delivers greater protection with lower overhead and faster data rebuild than traditional RAID. HYDRAstor DRD protection can tolerate up to six (6) concurrent drive or node failures while maintaining normal I/O. HYDRAstor DRD recovers data faster than traditional RAID by reconstructing only lost data rather than the entire drive and reconstructing the data across multiple drives instead of a single hot spare.



g) Advantage: Worm for ransomware protection and guaranteed retention

Write-Once-Read-Many supports HYDRAstor integrated with selected backup applications, which offers protection against malicious data encryption and ensures guaranteed data retention.

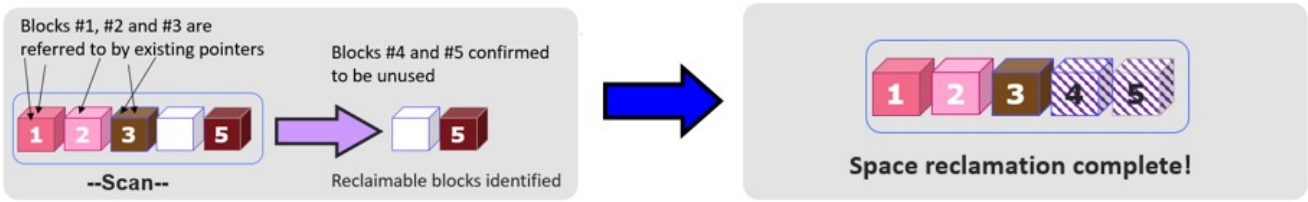


h) Advantage: Predictable space reclamation and garbage collection

One of the key challenges with any storage device is managing deleted data over time – which is exaggerated in disk-based backup targets. The operational process for backup is to store continuously, which causes data access to become increasingly random. This leads to inconsistent and ultimately degraded performance over time. To overcome this, the deduplication devices require an additional process – often called garbage collection – whereby data fragments are re-aligned to provide sequential processing and simultaneously release deleted blocks for re-use. In many cases, this process results in high processor and disk utilisation. As a result, garbage collection may extend over several days, greatly impacting backup and restore times.

As the HYDRAstor architecture distributes the load across all the nodes in the grid, the garbage collection process is managed independently by each node. This means that each node must process a fixed set of usable data in each cycle versus a single processing unit having to sift through all the data. This deletes data as the retention periods are enforced, and the placement and reclamation of deleted data space needs to be managed to avoid fragmentation.

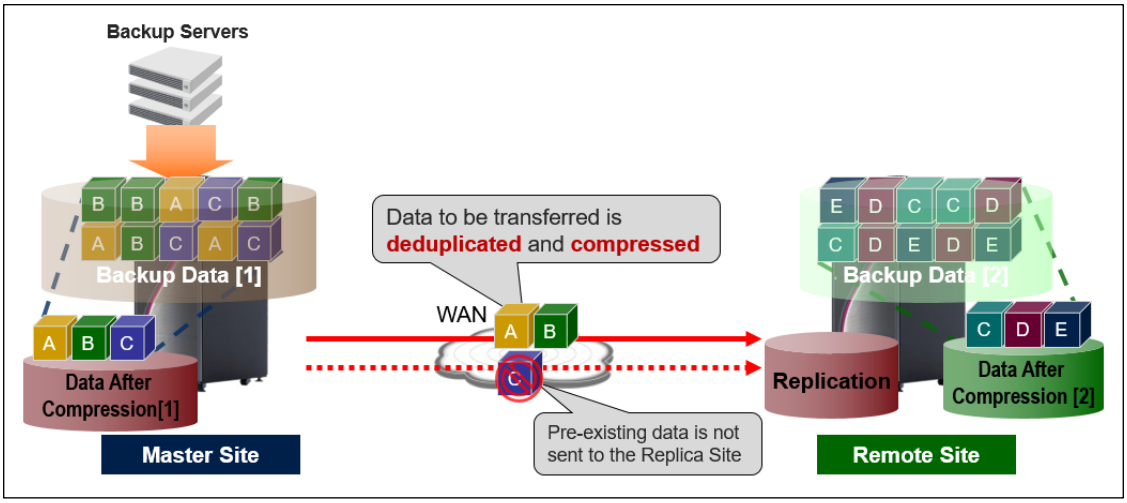
Fragmentation is highly prevalent in disk-based storage targets that do not have deduplication capabilities, resulting in performance degradation. Continuous deletion of backup data results in fragmentation and significantly reduces the time to garbage collect unused space from days to hours, but it has no significant effect on the grid's performance.



i) Advantage: Optimised WAN replication

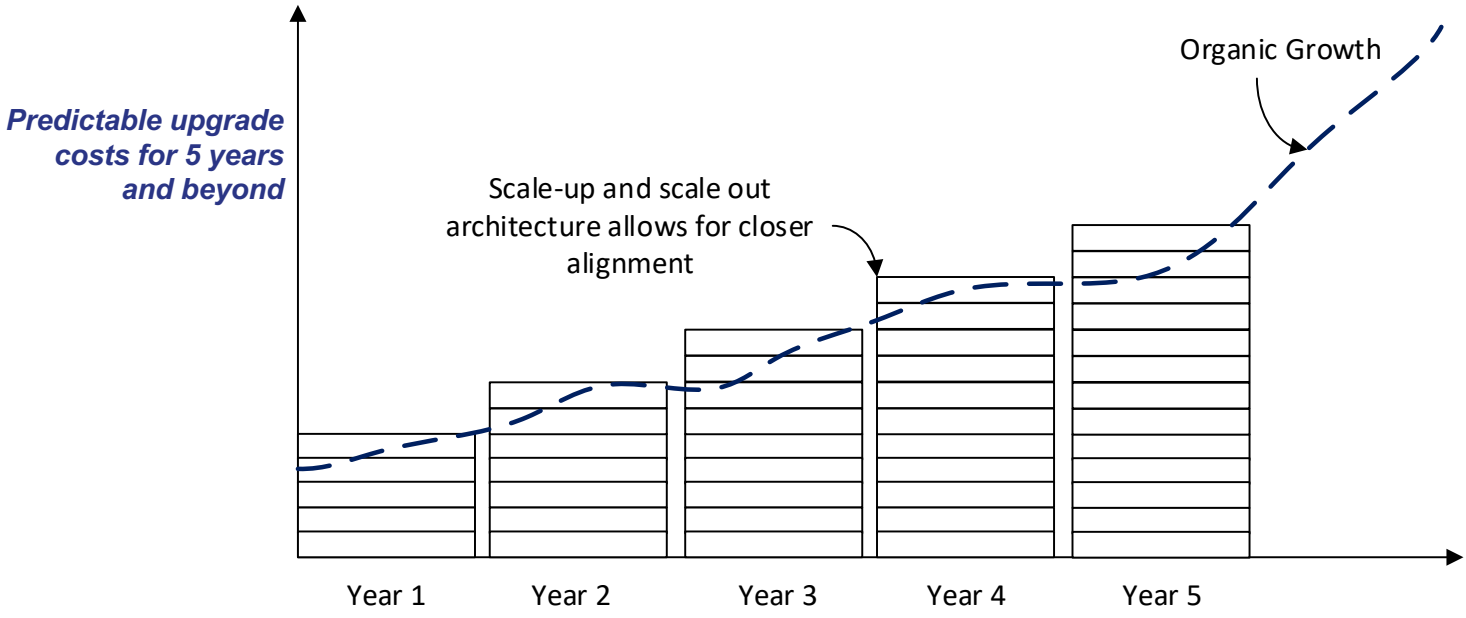
HYDRAstor supports one-to-one and one-to-many asynchronous optimised WAN replication. Deduplication is utilised, so only blocks not present on the target are transmitted, reducing the required network bandwidth.

Replication can be run in parallel to the source backup, so replication usually finishes soon after the source backup finishes.



j) Ultimate Advantage: Significantly lower TCO

- Savings due to lower acquisition cost—not only does HYDRAstor require fewer resources like disk space and network bandwidth due to heavily used deduplication and compression, but the HYDRAstor grid architecture allows for just-in-time purchases of additional nodes dictated by changing business requirements.



- Savings due to operational efficiencies—as one storage grid per data centre, HYDRAstor is easier to deploy, manage, and grow than many storage and performance islands. The utilisation of one large flexible system can be much higher than multiple smaller systems, each imposing its own capacity and bandwidth limit. Higher utilisation of HYDRAstor results in lower operational costs in power, cooling, data centre space, and manpower.

- Savings due to consistent asset life – almost all other technology pricing models are based on co-termination (co-term) , which implies that the solution has a limited lifespan. Once the base requirement is purchased, any further upgrades to the technology will be limited to the lifespan of the original purchase. The upside of the co-term effect is that any new purchases will have a lower maintenance/support cost. However, this has the opposite effect from a financial perspective as the asset life of the upgrade is reduced – effectively making the upgrade substantially more expensive. HYDRAstor, with its modular grid architecture, removes this co-term effect as each node has its maintenance lifespan and, therefore, has a consistent asset life. After purchasing the base requirement, any subsequent purchases will continue as a functional asset even after the base has been evacuated or replaced, resulting in a much higher return on the asset.

- Savings due to reduced cost of migration - migration is becoming a major indirect cost when moving from one technology to another, independent of whether this is within the same vendor or across technology suppliers. As the data capacities increase exponentially, the cost of changing technology platforms has become enormous. Often, migration has a service impact on business and is often a lengthy process that places undue strain on the operational staff. As HYDRAstor can insert and extract nodes from the grid transparently, the migration effort and risk are nullified. The HYDRAstor automatically rebalances data across the available nodes without any effort from the operations team.

- Savings due to improved business continuity—very high data resilience and system availability with nonstop system extensions and upgrades, ransomware protection with WORM, and WAN-optimised replication for DR significantly reduce threats to business continuity.

About Kaizen Technologies Kaizen Technologies, rooted in the Japanese philosophy of continuous improvement, embodies its mission, "***In Pursuit of Perfection***," in every facet of its operations. This guiding principle shapes our approach to distributing the cutting-edge NEC HYDRAstor product range globally and underpins our vision for the future of technology distribution and support.
For more information, visit Kaizen Technologies at: https://www.kaizentechnologies.online/about-us/