IBM Spectrum NAS

Easy-to-manage software-defined file storage for the enterprise

Highlights

- Reduce capital expenditures with storage software on commodity servers
- Improve efficiency by consolidating all your file needs on one cluster
- Scale capacity and performance on the fly - just add standard x86 servers
- Support our file-serving needs with native NFS and SMB protocol support
- Tailor application file sets for efficiency, performance and protection
- Monitor and manage the storage cluster with a powerful, intuitive GUI

Faced with a dramatic rise in storage demand, organizations are searching for ways to effectively manage and use data while controlling spiraling expenses. By adopting a software-defined solution, organizations can benefit from increased agility and management simplicity for network-attached storage (NAS) workloads, including general-purpose NAS or home directories, file serving for virtual machines, and NAS for Microsoft applications.

IBM Spectrum NAS software-defined storage helps reduce hardware needs and capital expenditure by working with industry-standard x86 servers deployed as high-performance scale-out storage clusters. The use of a symmetric architecture minimizes the incidence of bottlenecks or hotspots as the cluster scales. With no single point of failure, customers can safely use inexpensive off-the-shelf hardware.

Thanks to its software-defined approach, IBM Spectrum NAS supports organizational evolution from a fixed, rigid, hardware-centric approach to one that is flexible and pragmatic. The solution is hardware-agnostic and readily scalable, helping reduce the cost and worry of future data migrations and hardware replacements.

Overview

IBM Spectrum NAS manages a cluster of nodes, where each node has its own CPU, RAM, storage and cache, with a single file system spread across all nodes. Each node adds access points, cache, storage capacity and performance to the cluster. The IBM Spectrum NAS architecture is balanced by definition, as every node plays an identical role in the storage cluster—there are no dedicated metadata or special-purpose nodes. Users and applications can access the cluster through

any node, spreading the load uniformly and eliminating bottlenecks.

As the needs for storage or bandwidth grow, more nodes can be added to the cluster, to immediately take on their share of the compute and storage load. This helps ensure that what works today will continue to work in the future.

Each node can be either “bare metal” or a virtual machine. IBM Spectrum NAS software comprises a comprehensive software package—including firmware—that is available as a bootable operating system. All nodes in the cluster run the same software stack.

The more nodes in a cluster, the greater that cluster’s capacity and performance. More nodes mean more options for load rebalancing and for accommodating the self-healing process that takes place in the event of a node failure. More nodes also allow more efficient erasure coding, whether for higher data redundancy or smaller footprint.

Applications access the data through the IBM Spectrum NAS virtual file system, which allows multiple clients to interact with the same stored data. IBM Spectrum NAS allows multiple, separate domains and file systems to span a single storage cluster, reducing costs by simplifying administration and allowing more efficient use of storage resources, compared to per-domain storage.

Linear scalability and performance

The IBM Spectrum NAS architecture is designed for linear scalability: adding more storage nodes provides a linear increase in average performance in the cluster. All the resources in the storage cluster are aggregated, including CPU, bandwidth, storage, and cache, giving a persistent high throughput for a high number of simultaneous users. Adding more nodes increases throughput and adding more cache reduces latency. The more nodes in the cluster, the less time it takes to complete tasks.

Cache is used to dramatically reduce latency. All caches throughout the cluster are synchronized to ensure that every node has knowledge about which node owns a copy of any given data, so subsequent reads or writes can be redirected to the appropriate node, which provides faster results than accessing the storage disks.

Reliability and data protection

Reliability is a core design attribute of IBM Spectrum NAS. This is supported by the symmetry of its architecture, with every node being identically configured and running the same small, efficient system core. This architecture, coupled with data redundancy and protection, enhances reliability by avoiding any single point of failure.

Erasure coding provides data redundancy and protection, with data being striped across nodes and locations, not just across disks as with conventional RAID. In the event of hardware failure, the cluster's other nodes are notified, then work in parallel to recreate the missing data. Because all nodes share the work of recreating lost data, recovery time can be dramatically shortened, with minimal impact on system performance.

A virtual IP mechanism is used to ensure that all nodes in a cluster appear available at all times, even when a particular node is taken down for upgrade or has failed. The IP address of an offline node is moved to another node, which becomes responsible for the new IP address in addition to its own. The process is automatic and fast, ensuring that the system continues to service applications uninterrupted.

IBM Spectrum NAS uses a variety of techniques to ensure data integrity even in the face of issues such as heavy multi-concurrent input/output (I/O) activity, sudden power loss or hardware failure. Safeguards include:

- Intelligent locking that allows multiple clients to do concurrent reads and writes by protecting the data at the required level and byte range—ensuring that updates are clean with no intermediate state resulting from power loss or malfunction.
- Snapshots, providing the ability to retrieve earlier versions of files or folders.
- Transaction-based power-loss protection to ensure that changes are stored and confirmed, with no possibility of data left in an invalid intermediate state.
- Hardware failure detection, minimizing the vulnerability window and speeding recovery.
- A self-healing process that constantly scans for consistency issues, recreating any missing data from another copy and scheduling the removal of excess copies. In the event of a node failure, each remaining node will replicate affected files elsewhere on the cluster. The administrator will be alerted so that the broken node can be serviced and reinstated.
- Erasure coding for efficient data protection by means of redundant copies across the nodes of a cluster.

Ease of installation, maintenance and upgrade

IBM Spectrum NAS is designed for ease of deployment and management. A wizard-based guide allows for setup or upgrade in 30 minutes. Identical software per node further simplifies setup and maintenance.

IBM Spectrum NAS supports rolling upgrades over the network, fully transparent to clients and without service interruptions. To deliver uninterrupted service, the system can take down one node at a time for upgrade, passing that node's IP address temporarily to another node, then bringing the node back online before moving to the next.

IBM Spectrum NAS key features

Scaling with business needs	Symmetric architecture, along with intelligent use of flash as non-volatile cache, enables linear scaling of capacity and performance.
Availability and reliability	When a disk or storage node fails, the cluster self-heals by recreating the missing data without interruption to file services. An efficient fault-recovery process using virtual IP address assignment allows for nearly complete redundancy.
Simplified deployment and maintenance	IBM Spectrum NAS provides a single software stack encompassing protocol support, file system and the underlying erasure coding-based disk manager. There is no need for gateways or external storage controllers.
Ability to reduce capital and operational expenses	Hardware-agnostic storage software allows the use of commodity servers or hypervisors, helping reduce capital costs. Consolidating storage for all NFS and SMB applications on a single cluster helps reduce operational expense.
Support for NFS and SMB protocols	You can run all your NAS-based applications with native support for NFS and SMB protocols.

System requirements

	Minimum	Recommended
CPU	x86-64 (4 cores)	x86-64 (4+ cores)
RAM	12 GB	32+ GB
Boot disk	20 GB	60+ GB
Cache disk	20+ GB solid-state disk (SSD)	100+ GB SSD / NVMe
Storage disk	SAS / SATA disks	SAS / SATA disks
Network	1 x Gigabit NIC	2 x 10 Gigabit (for separate back-end)
Number of nodes	4 nodes minimum	Scale out from 4 nodes
Network switch	Gigabit switch or better	2 x 10 Gigabit switches

Why IBM?

Innovative technology, open standards, excellent performance, and a broad portfolio of storage software, hardware and solutions offerings—all backed by IBM with its recognized industry leadership—are just a few of the reasons to consider storage solutions from IBM.

For more information

To learn more about IBM Spectrum NAS, please contact your IBM representative or IBM Business Partner, or visit: https://www.ibm.com/us-en/marketplace/spectrum-nas

This document contains information pertaining to the following IBM products which are trademarks and/or registered trademarks of IBM Corporation:
IBM® Spectrum™

All statements regarding IBM's future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only.