

Nimble Storage for HP Vertica

Nimble Storage's Adaptive Flash platform accelerates HP Vertica deployments. Nimble's Adaptive Flash arrays deliver speed and capacity, are easy to deploy and manage, and provide integrated data protection and replication.

Introduction

The IT landscape has changed. The amount of unstructured data being created in the form of blogs, videos, images, texts, and tweets is 10 times greater and growing 10 times faster than structured business data. Within this expanding data repository are important insights for companies that could analyze and act quickly. Unfortunately, most organizations continue to struggle with legacy storage systems that are unable to cope with Big Data's volume and complexity.

Background

Today's businesses must analyze massive amounts of structured, semi-structured, and unstructured data at unprecedented speed, and with a finer level of granularity than ever before. To do that, they need advanced analytics: Products like HP Vertica are being rapidly adopted by organizations seeking a better way to manage, process, and store huge amounts of complex data. Nimble Storage's Adaptive Flash solutions are the ideal platform for sophisticated data analysis products like Vertica, delivering the performance that businesses need at a price they can afford.

HP Vertica

The HP Vertica Analytics Platform was designed with speed, scalability, simplicity, and openness at its core; it was likewise structured to handle analytical workloads via a distributed compressed columnar architecture. HP Vertica delivers ultra fast query response times, along with petabyte-scale out capabilities. HP Vertica can also tightly integrate with a Hadoop framework to ingest and analyze meaningful insights from the ETL (extract, transform, and load) of unstructured data.

Hadoop and ETL

Apache Hadoop is an open-source software framework for highly available and scalable storage, as well as large-scale processing of data sets on clusters of commodity hardware (servers). ETL plays a critical role in gathering and loading data sets from various sources. ETL systems are commonly used to integrate data from multiple applications, which are typically developed and supported by different vendors or hosted on separate computer hardware.

The Problem with Scale-Out Server Clusters

Even as they grow at a rapid pace, organizations need to maintain service level agreements. Most organizations meet growth requirements by adding servers, a practice that poses the following challenges:

- Organizations that process terabytes of data soon realize that their data is growing much faster than their processing requirements. Adding more servers to the cluster becomes a costly way to add storage.

Solution Overview

Nimble Storage's Adaptive Flash is a powerful and flexible platform that delivers actionable intelligence with lightning speed for Vertica environments. Nimble's Adaptive Flash solutions offer:

- **Cost-Efficient Performance** for near real-time insights.
- **Intelligent and Dynamic Scaling** of capacity and/or performance.
- **Efficient and Affordable Capacity** for data growth.
- **Integrated Data Protection and Disaster Recovery** with frequent and fast backups.

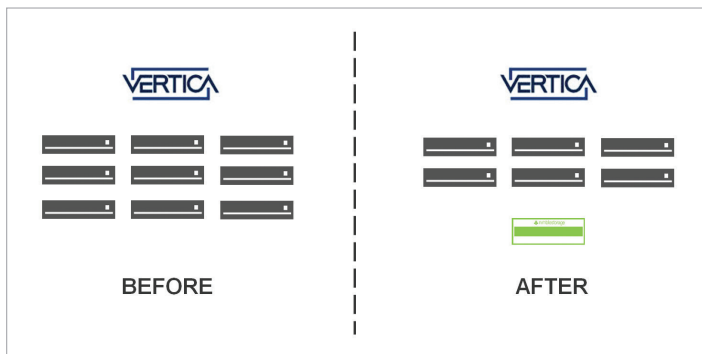
- Hadoop clusters add considerable operational cost and complexity. Hadoop clusters of greater than 6 servers present significant operational challenges.
- Data warehouses are typically assembled from a variety of data sources with different formats and purposes. Vertica's default compression engine cannot make sense of these disparities.
- Additional servers complicate data protection and backup. Basic three-way mirroring is not cost-effective, while site recovery is disruptive and time-consuming.

A Better Approach with Nimble Storage

Nimble Storage's Adaptive Flash solutions have been designed to satisfy the unique demands of business-critical applications, such as Big Data analytics. Nimble solutions deliver cost-effective performance and capacity, are easy to deploy and manage, as well as provide integrated data protection.

Cost-Effective Performance

Nimble Storage arrays deliver adaptive performance for Vertica environments. At the core of every Nimble array is its patented Cache Accelerated Sequential Layout (CASL™) architecture. CASL was designed from the ground up to leverage the random read performance of flash, offering the level of performance demanded by Vertica environments. Vertica handles millions of small interactive queries, a read-only I/O pattern that taxes the resources of traditional storage systems.



The Nimble Storage solution offers greater space reduction by freeing server capacity.

Ease of Management

Nimble Storage arrays are easy to deploy and manage. Nimble arrays can be set up in less than half a day by an IT generalist without prior training. In addition, Nimble solutions free server capacity by moving data to disk, using a fraction of the server capacity required by server-only, scale-out solutions.

InfoSight™ is Nimble's innovative approach to management using cloud-based analytics to deliver true operational efficiency across all storage activities. InfoSight monitors all Nimble arrays, collectively and individually, from the cloud, automatically gathering millions of data points per array each day and making sense of them in real-time. Using that data, InfoSight identifies potential problems — and offers solutions — long before they can bring systems down.

Capacity Efficiency

Nimble Storage's CASL architecture incorporates innovative efficiency features such as inline variable-block compression, offering 2x improvement compared with traditional storage solutions (without Vertica's native compression). The range of data types (random vs. sequential) and data sizes (8k-32K block) is significant in this environment, meaning that post compression is an inefficient way to reduce capacity. Nimble's variable-length compression offers greater space reduction, and improves performance.

Improved Scalability

Nimble Storage solutions allow for independent and non-disruptive scaling of performance and capacity, eliminating the need to add server nodes (compute). With Nimble, enhancing performance is as easy as upgrading controllers. Simply adding SSDs can accommodate more active data; storage capacity can be scaled to hundreds of terabytes merely by adding disk shelves.

Data Protection

Nimble Storage offers maximum data protection for Vertica environments with frequent snapshots, consistent backups, fast restores, and efficient replication for disaster recovery. Nimble solutions deliver simple and affordable data recovery, efficiently protecting months of stored data as well as easing the burden on both server and network. In addition, Nimble solutions efficiently replicate data onto another array by transferring compressed, block-level changes only. These remote copies can be activated even when the primary array is unavailable, so that data recovery is both simple and affordable.

Conclusion

Deployed as part of a comprehensive analytics solution, Nimble Storage delivers actionable intelligence with lightning speed. Nimble solutions are easy to deploy and use, reducing total cost of ownership (TCO) and providing an exceptional return on investment (ROI) for HP Vertica environments.



211 River Oaks Parkway, San Jose, CA 95134
 Phone: 408-432-9600; 877-364-6253
 Email: info@nimblestorage.com
 www.nimblestorage.com



© 2014 Nimble Storage, Inc. Nimble Storage, the Nimble Storage logo, InfoSight, CASL, SmartStack, and NimbleConnect are trademarks or registered trademarks of Nimble Storage, Inc. All other trademarks are the property of their respective owners. SB-HPV-0914