



in

# What to do with all that data? Answer: SingleStore on PowerFlex

f

🐦

Wed, 10 May 2023 22:55:28 -0000 | Read Time: 0 minutes

Tony Foster

Every organization has data, every organization has databases, every organization must figure out what to do with all that data from those databases. According to research by the University of Tennessee, Knoxville's Haslam College of Business (<https://haslam.utk.edu/news/happy-global-business-analysis-day-do-you-know-what-your-business-data-is-doing/>) there were 44 zettabytes of data in 2020, and by 2025 it is estimated that 463 exabytes of data will be created daily.

us/dt/storage/powerflex.htm).

Recently Dell and SingleStore released a joint validation white paper (<https://infohub.delltechnologies.com/t/deploying-singlestore-database-on-dell-powerflex-4-0-1/>) on a virtualized SingleStore environment running on PowerFlex. The paper provides an overview of the technologies used and then looks at an architecture that can be used to run SingleStore on PowerFlex. After that, the paper looks at how the environment was validated.

## SingleStore

Before I get into the details of the paper, I suspect there might be a few readers who have yet to hear about SingleStore or know about some of its great features, so let's start there. Built for developers and architects, SingleStoreDB is based on a distributed SQL architecture, delivering 10–100 millisecond performance on complex queries—all while ensuring that your organization can effortlessly scale. Now let's go a bit deeper....

The SingleStoreDB :

- Scales horizontally providing high throughput across a wide range of platforms.
- Maintains a broad compatibility with common technologies in the modern data processing ecosystem (for example, orchestration platforms, developer IDEs, and BI tools), so you can easily integrate it in your existing environment.
- Features an in-memory rowstore and an on-disk columnstore to handle both highly concurrent operational and analytical workloads.
- Features the SingleStore Pipelines data ingestion technology that streams large amounts of data at high throughput into the database with exactly once semantics.



This means that you can continue to run your traditional SQL queries against your every growing data, which all resides on a distributed system, and you can do it fast. This is a big win for organizations who have active data growth in their environment.

What makes this even better is the ability of PowerFlex to scale from a few nodes to thousands. This provides a few different options to match your growing needs. You can start with just your SingleStore system deployed on PowerFlex and migrate other workloads on to the PowerFlex environment as time permits. This allows you to focus on just your database environment to start and then, as infrastructure comes up for renewal, you migrate those workloads and scale up your environment with more compute and storage capacity.

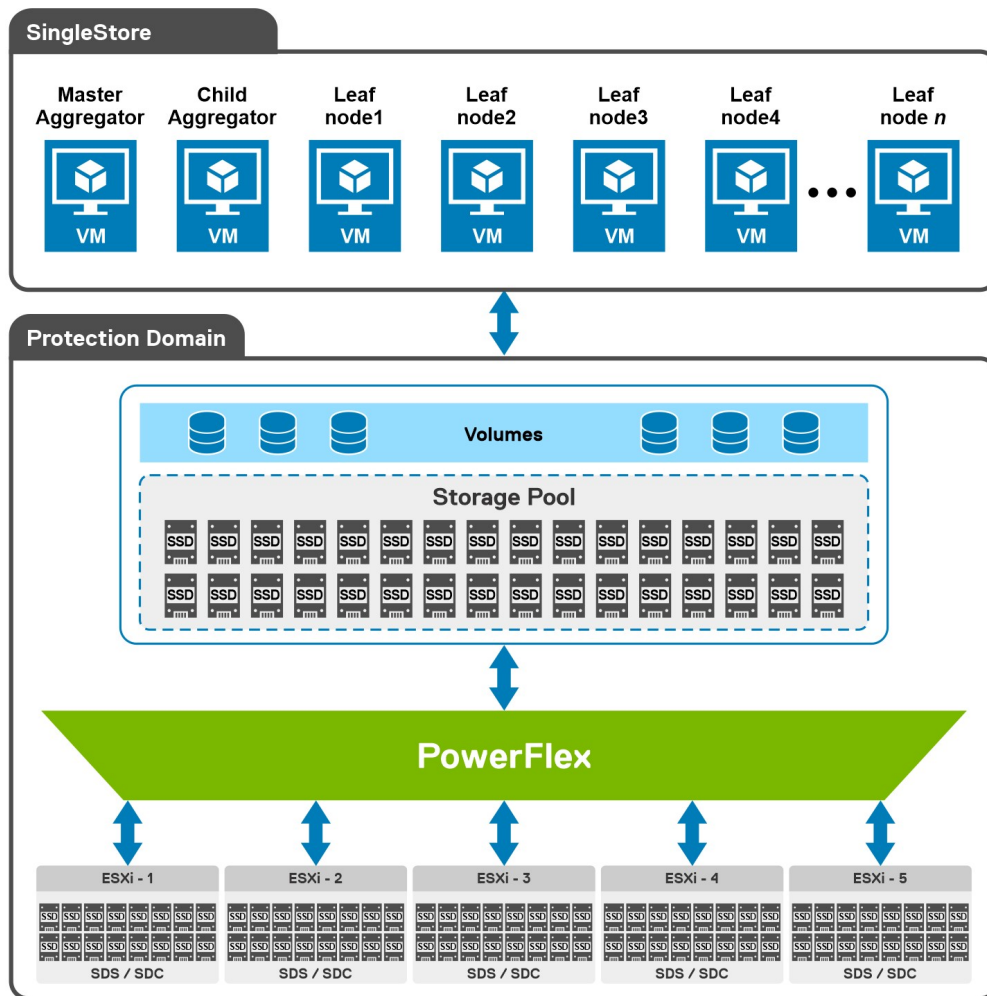
Or maybe you are making a bigger contribution to that 463 exabytes of data per day I mentioned earlier, and you need to scale out your environment to handle your data's growth. You can do that too!

That's the great thing about PowerFlex, you can consume resources independently of each other. You can add more storage or compute as you need them.

Additionally, with PowerFlex, you can deliver bare-metal and virtualized environments without having to choose only one. That's right—you can run bare-metal servers right next to virtualized workloads.

## Architecture

As shown in the following diagram, our engineering team used five Dell PowerEdge R640 servers with dual CPUs, 384 GB of RAM, and eight SSDs per node. These five nodes were configured as HCI nodes and connected with a 25 Gbps network. The storage from across the nodes is aggregated to create a large software-defined pool of storage as a single protection domain that provides volumes to the SingleStore VMs. This is ideal for even the most demanding databases due to its high I/O capability.



For this validation, the SingleStore Cluster VMs consist of two aggregator VMs and multiple leaf VMs. The white paper details the configuration of these VMs.

Additionally, the white paper provides an overview of the steps used to deploy SingleStore on VMware vSphere in a PowerFlex environment. For this validation,

With the environment configured, the white paper then discusses how to validate the environment using TPC-DS (<https://github.com/memsql/benchmarks-tpc/blob/master/tpcds/queries.sql>). This tool provides 99 different queries that can be used to test a database. For this validation, only 95 of the 99 were used. The paper then describes both how the sample data set was created and how the tests were run.

The validation tests were run on 4, 6, and 8 leaf node configurations. This was done to understand the variation in performance as the environment scales. The testing showed that having more SingleStore leaf nodes results in better performance outcomes.

The testing also showed that there were no storage bottlenecks for the TPC-DS like workload and that using more powerful CPUs could further enhance the environment.

The white paper shows how SingleStore and PowerFlex can be used to create a dynamic and robust environment for your growing data needs as you do your part to contribute to the 463 exabytes of data that is expected to be created daily by 2025. To find out more about this design, contact your Dell representative.

## Resources

- <https://www.singlestore.com> (<https://www.singlestore.com>)

**Author:** Tony Foster

Twitter: [@wonder\\_nerd](https://twitter.com/wonder_nerd) ([https://twitter.com/wonder\\_nerd](https://twitter.com/wonder_nerd))

LinkedIn (<https://linkedin.com/in/wondernerd/>)

Tags:    PowerFlex    SingleStore Database

## Related Blog Posts



(/p/new-file-services-capabilities-of-powerflex-4-0/)

PowerFlex File services File storage

## New File Services Capabilities of PowerFlex 4.0

(/p/new-file-services-capabilities-of-powerflex-4-0/)

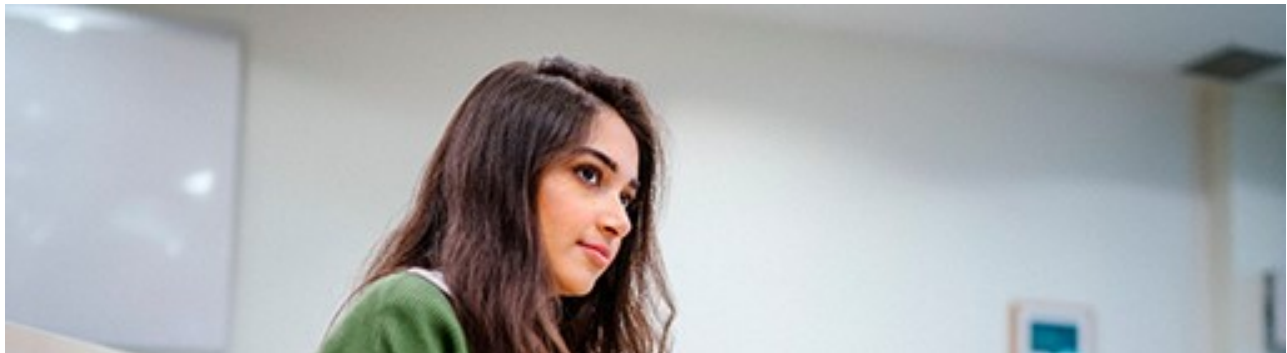


Alan Ng, Tony Foster

Tue, 16 Aug 2022 14:56:28 -0000 | Read Time: 0 minutes

(/p/new-file-services-capabilities-of-powerflex-4-0/)

“Just file it,” they say, and your obvious question is “where?” One of the new features introduced in PowerFlex 4.0 is file services. Which means that you can file it in PowerFlex. In this blog we’ll dig into the new file service capabilities offered with 4.0 and how they can benefit



(/p/introducing-nvme-over-tcp-nvme-tcp-in-powerflex-4-0/)

VMware PowerFlex NVMe/TCP

## Introducing NVMe over TCP (NVMe/TCP) in PowerFlex 4.0

(/p/introducing-nvme-over-tcp-nvme-tcp-in-powerflex-4-0/)



Kevin M. Jones, Tony Foster

Fri, 26 Aug 2022 18:59:38 -0000 | Read Time: 0 minutes

(/p/introducing-nvme-over-tcp-nvme-tcp-in-powerflex-4-0/)

Anyone who has used or managed PowerFlex knows that an environment is built from three lightweight software components: the MDM, the SDS, and the SDC. To deploy a PowerFlex environment, the typical steps are:

logo

© 2023 Dell Inc. (<https://www.dell.com/learn/us/en/uscorp1/site-terms-of-use-copyright>)

Privacy (<https://www.dell.com/learn/us/en/uscorp1/policies-privacy>)

Terms Of Use (<https://www.dell.com/learn/us/en/uscorp1/site-terms-of-use>)

Legal (<https://www.dellemc.com/en-us/customer-services/product-warranty-and-service-descriptions.htm>)

Anti-Slavery and Human Trafficking (<https://i.dell.com/sites/doccontent/corporate/corp-comm/en/Documents/dell-california-trafficcking.pdf>)