DELL Technologies

Solution brief

The Data Accelerator

One of the world's fastest open-source, NVMe storage solutions solves I/O bottlenecks

Workloads have become increasingly data-centric in recent years, especially as high performance computing (HPC) and artificial intelligence (AI) converge, and IT teams seek to run these data-intensive workloads on the same systems. Growing data sets, coupled with bandwidth- and latency-sensitive workloads are placing simultaneous high demands on parallel file systems. This can create storage input/output (I/O) bottlenecks that increase wait times, and can even render file systems inoperable.

This situation significantly impacts time to results, as well as return on investment for HPC systems. And as workloads become even more data intense, these bottlenecks and service interruptions become more common. Because researchers are often working on critical, time-sensitive work in life sciences, research, engineering and other fields, these slowdowns and interruptions can be particularly devastating.

In recent years, many high performance computing sites have begun using solid-state drives (SSDs) for large-scale HPC systems in an attempt to overcome the growing bottleneck problem between compute and storage resources. Dell Technologies, Intel® and the University of Cambridge have collaborated to enable the next generation of data-intensive workflows with an NVMe-based storage solution: the Dell EMC Data Accelerator. This Data Accelerator (DAC), makes optimal use of modern server NVMe fabric technologies to mitigate I/O-related performance issues.

What's more, the DAC is completely open source and extensible. The configuration scripts and detailed build instructions are freely available to make the technology as accessible as possible to promote further development and testing of SSD I/O accelerators within the HPC community.

Data Accelerator use cases

- On-demand parallel file system for faster processing of data sets
- High-speed low-latency scratch space
- Reducing the impact of noisy neighbors on a capacity-oriented system
- Swap for jobs with non-deterministic memory requirements

Key solution benefits

- Creates an open-source software solution, utilizing infrastructure-as-code and cloud-native technologies built on readily-available server and networking technology.
- Alleviates performance bottlenecks for data-intensive applications on central networked file systems.
- Provides deterministic, high-performance, schedulable I/O resources, providing breakthrough I/O performance.
- Integrates with any traditional HPC storage without redesign to interact with commonly-used scheduling tools.

Get detail on the architecture, including synthetic benchmarking and initial application testing of a performance-optimized reference implementation for the Data Accelerator, which can be deployed as an add-on to Slurm®-enabled HPC systems. The DAC described is currently deployed as part of the 1152-node Dell EMC Cumulus Cloud HPC system at the University of Cambridge and is accessible from conventional HPC systems and Slurm as a Service[1] platform atop of OpenStack®-enabled infrastructure. It comprises 24x DAC nodes providing around 500TB of usable capacity.

Resources

- HPC Engineering: hpcatdell.com
- HPC & AI Innovation Lab
- HPC & AI Centers of Excellence

Learn more

- delltechnologies.com/hpc
- delltechnologies.com/ai

The work started as a co-design project with the University of Cambridge, Dell Technologies, Intel® and StackHPC, and represents the first open source solid-state burst buffer implementation of its kind. This implementation was so successful that it reached #1 in the June 2019 IO500 as the fastest storage system in the world, with almost twice the performance of the second-place entry on the list. With a footprint of just one rack, the solution was able to deliver over 500GB/s bandwidth, 1.3 million file creates per second and 28 million IOPS.[2]

In addition to the hardware layer is a software layer allowing for an on-demand, per-job Lustre parallel file system. The software-defined design approach exploits modern cloud methodologies to promote a reproducible and extensible code base, making heavy use of Ansible for infrastructure as code, together with etcd, a distributed key-value store used, for example, by Kubernetes®. With record-setting performance and software-defined capabilities, the Dell EMC Data Accelerator now forms an integral part of the Cumulus Science Cloud system at the University of Cambridge.

System configuration deployed at Cambridge University

DAC server nodes	Memory	Networking	PCIe storage	NVMe
24x Dell EMC PowerEdge R740xd with 2x Intel Xeon® Gold 6142 16-core processors	24x 16GB 2,666 MT/s DDR4	2x Intel OPA v1 100Gbps (12.5GiB) HFI	2x PLX PCIe switch for NVMe	12x Intel P4600 (1.6TB)

Engineering-validated system configuration choice

DAC server nodes	Memory	Networking	PCIe storage	NVMe
Dell EMC PowerEdge R740xd with 2x Intel Xeon® Platinum 8168 24-core processors	24 x 16GB, 2,666MT/s DDR4	2x Dell EMC HBA330 adapter card with Mellanox® Quantum™ MQM8700-HS2F switch	2x PLX PCIe switch for NVMe	12x Intel P4600 (1.6TB, 3.2TB, 6.4TB, and 7.68TB) or other appropriate drive choices

Dell Technologies, Intel and the University of Cambridge

Dell Technologies, Intel and the University of Cambridge have a long history of collaborating on HPC projects designed to advance the state of the art for researchers. A Dell Technologies HPC & AI Center of Excellence, the Cambridge Dell Intel Centre was founded in 2007 to provide answers to challenges facing the HPC community and share the results back into the wider research community. And in 2020, the University of Cambridge, Dell Technologies and Intel are joining forces yet again to collaborate toward a common goal of making exascale computing a reality.

[1] StackHPC Ltd, "Clusters for Scientific Applications: As-a-Service," August 2017.

[2] Virtual Institute for I/O, IO500 List, June 2019.

DELL Technologies

Copyright © 2020 Dell Inc. or its subsidiaries. All Rights Reserved. Dell, EMC, and other trademarks are trademarks of Dell Inc. or its subsidiaries. Other trademarks may be the property of their respective owners. Published in the USA 05/20 Solution brief DELL-SB-DAC-USLET-101

Intel® and Xeon® are trademarks of Intel Corporation or its subsidiaries in the U.S. and/or other countries. Slurm® is a registered trademark of SchedMD LLC. The OpenStack® word mark and the Square O Design, together or apart, are trademarks or registered trademarks of OpenStack Foundation in the United States and other countries, and are used with the OpenStack Foundation's permission. Lustre® is a registered trademark of Seagate Technology LLC in the United States. Kubernetes® is a registered trademark of The Linux Foundation. Mellanox is a trademark of Mellanox Technologies, Ltd. All other trademarks are proprty of their respective owners.

Dell Technologies believes the information in this document is accurate as of its publication date. The information is subject to change without notice.