Excelero

NVMesh® SUPERCHARGES BeeGFS

Up to 3x performance with Elastic NVMe

JOINT SOLUTION

The BeeGFS® parallel file system was developed at the Fraunhofer Center for High Performance Computing. It was designed for I/O intensive workloads that require massive performance and capacity scalability. BeeGFS has been deployed by hundreds of scientific and research organizations as it is ideal for demanding, high-performance workloads found in HPC and life sciences. Think of Artificial Intelligence, deep learning, cognitive computing or other data-intensive analytics.

The joint NVMesh-BeeGFS solution creates a high-performance file system that is perfect for all kinds of access patterns: lots of small files, small random access, high streaming throughput. NVMe flash gives BeeGFS customers great performance benefits but when you deploy NVMe in large storage clusters under a shared file system such as BeeGFS, you risk having stranded capacity or unrealized performance. By running BeeGFS on NVMesh, you can leverage the full potential of your NVMe devices: increasing performance and scalability while utilizing a shared namespace across your entire application cluster.

SOLUTION BENEFITS

NVMesh enables you to deploy BeeGFS at massive scale on high-performance NVMe while meeting all your storage requirements

FULL NVMe PERFORMANCE UTILIZATION
Ultra-low access latency
Highest IOPS & metadata performance

MAXIMUM HARDWARE EFFICIENCY
Pool NVMe to maximize utilization
No NVMe target CPU overhead, no noisy neighbors

COMPLETE DATA PROTECTION
Flexible from unprotected scratch to dual parity
Automated drive or host recovery

MINIMIZE HARDWARE INVESTMENT
No Proprietary storage hardware
No need for a separate block storage tier – run block and BeeGFS file services on the same nodes

INTUITIVE DATA MANAGEMENT
Tiered storage solution spanning flash tiers and spinning
Unified monitoring and statistics under Grafana
On-the-fly, automated provisioning for dynamic cluster workloads and schedulers
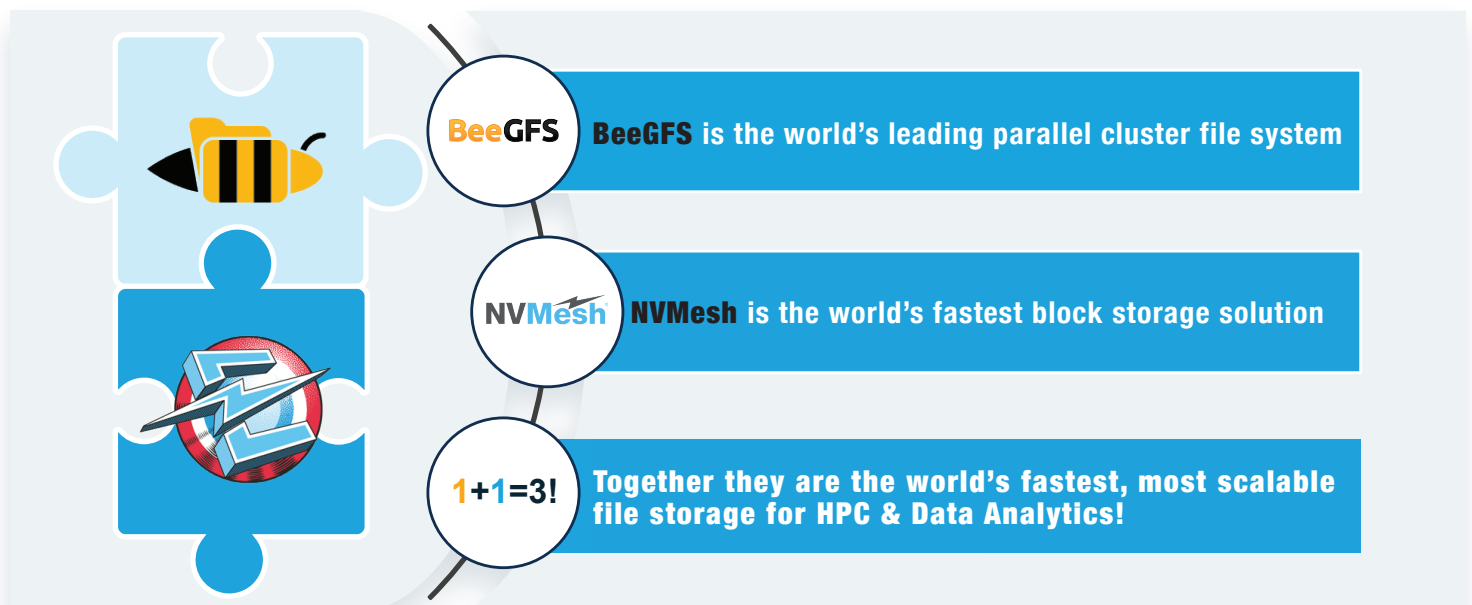
SCALE AS YOU NEED

BeeGFS was designed to scale simply and granularly. By increasing the number of servers and/or drives in the system, performance and capacity of the file system scales out to the desired level. This applies seamlessly to small systems or to massive clusters with thousands of nodes. NVMesh scales the same way: you simply add capacity or performance by adding more NVMe devices to your servers or adding more servers to your environment.

MORE PERFORMANCE
BETTER INSIGHTS
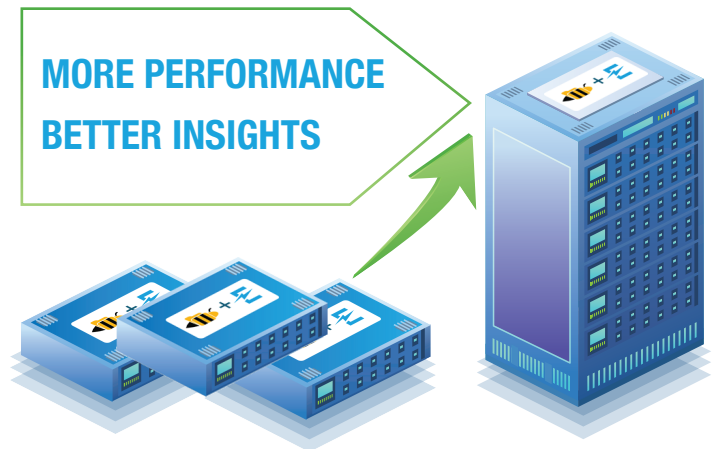
ZERO DATA CENTER FOOTPRINT
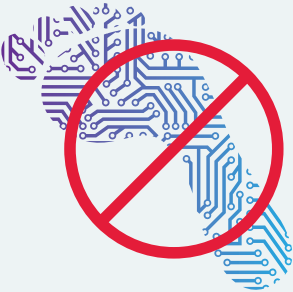
With NVMesh for BeeGFS, you can source standard NVMe drives and can completely obviate the need for proprietary hardware, and even dedicated storage appliances. NVMesh and the patented Remote Direct Drive Access (RDDA) technology allow you to logically disaggregate NVMe drives in the servers away from CPU resources. That is, though the local NVMe drive may be used by remote nodes, that usage does not consume local CPU. Every compute node can have a local NVMe SSD (or multiple drives) and all the drives are pooled for use by the cluster. Alternately, the BeeGFS file server nodes can also serve NVMesh block to each other, with protection, while improving the performance of those nodes. Thus, even if you desire to have dedicated storage nodes, you only need one layer for block and file service.

BeeGFS **BeeGFS** is the world's leading parallel cluster file system

NVMesh **NVMesh** is the world's fastest block storage solution

1+1=3! **Together they are the world's fastest, most scalable file storage for HPC & Data Analytics!**

DATA PROTECTION

NVMesh features a flexible, distributed data protection architecture offering various protection levels, matching resiliency and performance to application needs. Options range from no redundancy, mirroring (1+1) to parity-based (N+M). The latter provides over 90% storage efficiency, yet delivers ultra low-latency performance on large-scale configurations. By using BeeGFS on NVMesh, you automatically add enterprise-grade data availability: BeeGFS pools can have dual-failure protection from drive and host failures.

"We've been impressed with NVMesh's ability to deliver the high IOPS and ultra-low latency of NVMe drives over the network with highly available volumes – as well as options for distributed erasure coding, resulting in a BeeGFS solution with unmatched ability to efficiently handle all kinds of access patterns and file sizes."

Frank Herold, CEO of ThinkParQ

PERFORMANCE

To demonstrate the possibilities of this new scale-out infrastructure, we used the industry standard mdtest and IOR benchmarks. The test system was a compact 2U 4-server chassis with a total of 24 NVMe drives, connected via a 100Gbit RDMA network to 8 BeeGFS client compute nodes. Tests were run on the exact same hardware configuration with BeeGFS utilizing the direct-attached NVMe vs. BeeGFS utilizing NVMesh logical volumes. Taking advantage of NVMesh to offload mirroring operations, BeeGFS file create operations were boosted 3x, while metadata read operations were boosted 2.5x. For small random file access, which is often considered to be especially critical for application efficiency, NVMesh's low latency technology boosted BeeGFS 4K write IOPS to 1.25 million per second, a 2.5x improvement. By leveraging NVMesh distributed erasure coding for BeeGFS, customers can get up to 90% usable capacity while still tolerating drive failures and 75GB/s streaming throughput already out of such an entry-level system.
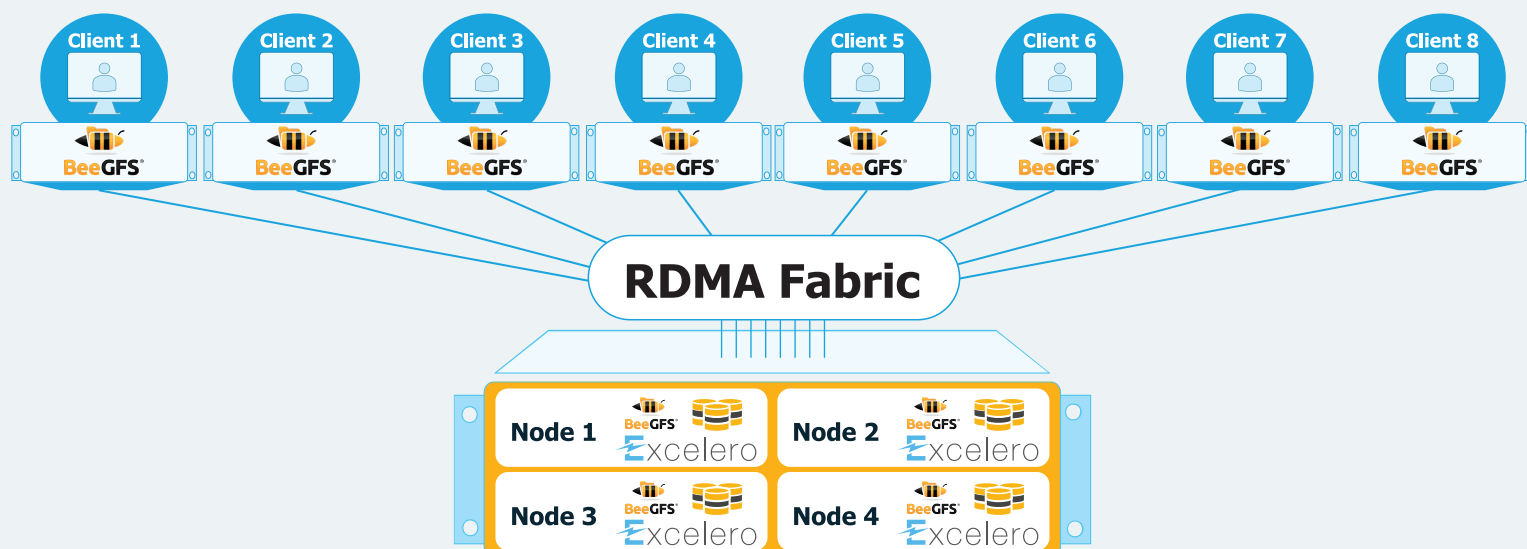
NVMesh for BeeGFS

NVMesh is a Software-Defined Block Storage solution that features Elastic NVMe, a distributed block layer that allows high-performance applications to utilize pooled NVMe storage devices across a network at local speeds and latencies. NVMe storage resources are pooled with the ability to create arbitrary, dynamic block volumes that can be utilized by any host running the NVMesh block client. These virtual volumes can be striped, mirrored, or both while enjoying centralized management, monitoring, and administration. In short, applications can enjoy the latency, throughput, and IOPS of a local NVMe device while at the same time getting the benefits of centralized, redundant storage.

NVMesh provides the ability to attach logical volumes to any host on the network, making all data appear to be local. NVMesh is deployed as a virtual, distributed non-volatile array and supports both converged and disaggregated architectures, giving customers full freedom in their architectural data warehouse design.
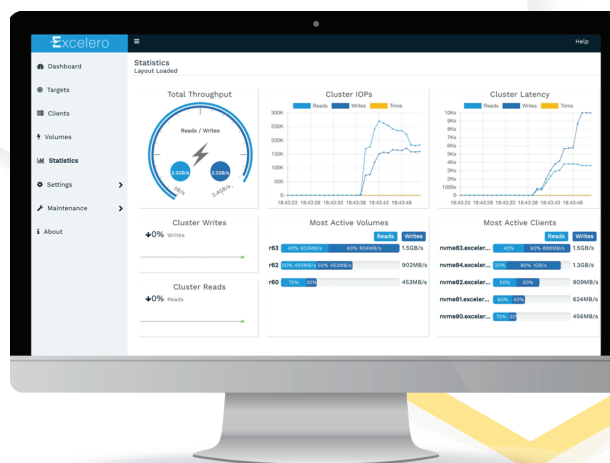
ABOUT EXCELERO

Excelero delivers low-latency distributed block storage for hyperscale applications such as AI, machine learning and GPU computing, in the Cloud and on the Edge. Founded in 2014 by a team of storage veterans and inspired by the Tech Giants' shared-nothing architectures for web-scale applications, the company has designed a software-defined block storage solution that meets the low-latency performance and scalability requirements of the largest web-scale and enterprise applications.

Excelero's distributed NVMe technology enables customers to benefit from the performance of local flash with the convenience of centralized storage while avoiding proprietary hardware lock-in and reducing the overall storage TCO. Excelero's storage software is deployed by major web-scale customers, for data analytics and machine learning applications and in Media & Entertainment post-production and HPC environments.