

BeeGFS Parallel Clustered Storage I/O Optimized Solutions



Performance-Tuned Storage Solution Parallel clustered storage architecture for Enterprise and HPC



The Challenge

Reliably store, manage, and protect unstructured data at cloud scale with lowered total cost of ownership (TCO).

Key Features

- Data and metadata for each file are separated into stripes
- Secure authentication process between client and server to protect sensitive data
- Fair I/O option on user level to prevent a single user with multiple requests to stall other user requests
- Storage high-availability based on replication
- Automatic network failover
- Optional server failover for external shared storage devices
- Runs on non-standard architectures

BeeGFS (formerly FhGFS) is the leading parallel cluster file system, developed with a strong focus on performance and designed for very easy installation and management. If I/O intensive workloads are your problem, BeeGFS is the solution.

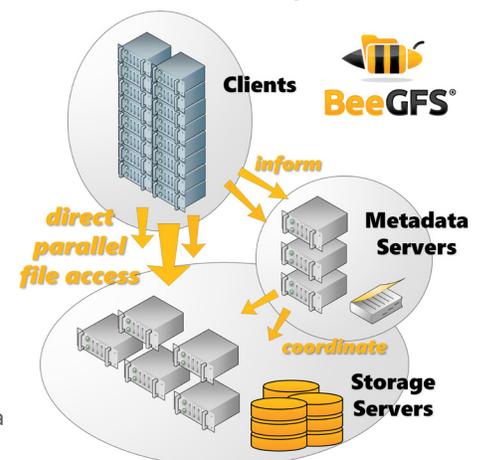
The BeeGFS parallel file system was developed for performance-critical environments, with a focus on easy installation and high flexibility. By increasing the number of servers and disks in the system, the file system's performance and capacity can seamlessly be scaled out from small clusters to enterprise-class systems with thousands of nodes. RAID Inc. has helped to design BeeGFS to be future-proof.

“RAID Inc. is positioned to impact the HPC market by deploying all-flash systems with BeeGFS.”

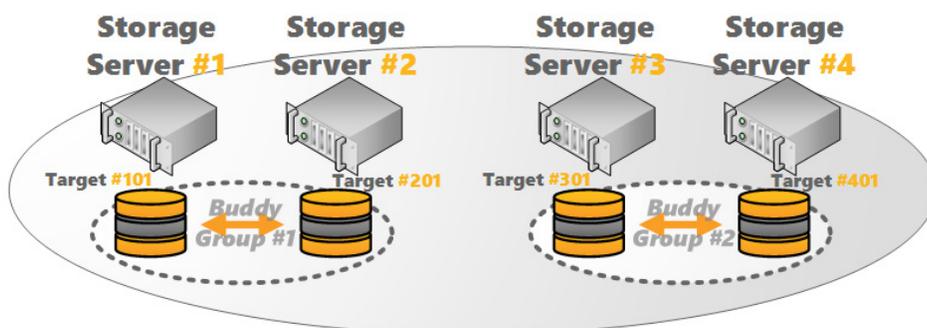
BeeGFS Clustered Storage: A Solution Overview

With performance-centric highly available data increasing in demand, a flexible architecture like BeeGFS eliminates data silos and storage complexity with a solution built specifically for multi-tenancy and cloud. BeeGFS is a parallel file system – by adding more servers, the capacity and performance of them is aggregated in a single namespace. That way the filesystem performance and capacity can be scaled to the level which is required for the specific application. Furthermore, BeeGFS is splitting MetaData from ObjectData. The ObjectData is the data users want to store, whereas the MetaData is the “data about data”, such as access rights and file size – but most important in the MetaData is the information, on which of the numerous storage servers the content of a file can be found.

BeeGFS is based on a lightweight architecture and can be created on a per job basis with the BeeOND (BeeGFS On Demand) application. This BeeOND feature is designed to provide new instances on the fly across all compute nodes being assigned to a particular job. BeeOND aggregates the performance and capacity of internal SSDs or hard disks in compute nodes for the duration of a compute job to increase performance levels with a very elegant method of burst buffering, which can be highly useful in cloud environments and temporary scratch data scenarios.



Buddy Mirroring



- Mirroring capabilities are integrated into the normal BeeGFS services, so that no separate services or third-party tools are needed.
- Mirroring can be enabled on a per-directory basis, some data in the file system can be mirrored while other data might not be mirrored. Data mirroring on the storage targets is based on so-called buddy groups. In general, a buddy group always consists of two targets.

Solution Benefits:

Maximum Flexibility

BeeGFS® supports a wide range of Linux distributions, including RHEL/Fedora, SLES/OpenSuse and Debian/Ubuntu and a wide range of Linux kernels.

- Storage servers run on top of existing local filesystems. Clients and servers can be added to an existing system without any downtime.
- Multiple networks and dynamic failover are supported for when network connections are down.
- Multiple services can run on the same physical machine. This eliminates the need for external storage resources, providing a cost-efficient solution with simplified management.

Maximum Scalability

BeeGFS® offers maximum scalability on various levels. It supports distributed file contents with flexible striping across the storage servers on a by file or by directory base as well as distributed metadata.

BeeGFS® was optimized especially for use in HPC and provides:

- Best in class client throughput (3.5 GB/s write and 4.0 GB/s read with a single I/O stream on FDR Infiniband)
- Best in class metadata performance: linear scalability through dynamic metadata namespace partitioning
- Best in class storage throughput: BeeGFS® servers allow flexible choice of underlying file system to perfectly fit the given storage hardware.

Maximum Usability

The BeeGFS® servers are user space daemons and do not require special administration rights. The client is a kernel module that does not require any patches to the kernel itself.

BeeGFS® also was designed with easy administration in mind. The graphical administration and monitoring system enables users to deal with typical management tasks in a simple and intuitive way: cluster installation, storage service management, live throughput and health monitoring, file browsing, striping configuration, and more.

The BeeGFS parallel file system was developed specifically for performance-critical environments and with a strong focus on easy installation and high flexibility, including converged setups where storage servers are also used for compute jobs. By increasing the number of servers and disks in the system, performance and capacity of the file system can simply be scaled out to the desired level, seamlessly from small clusters up to enterprise-class systems with thousands of nodes. BeeGFS is available free for download from www.beeGFS.com, professional support is available from ThinkParQ.Inc



RAID Inc. was founded in 1994 to deliver end-to-end performance-driven technical computing and storage solutions. The company has earned industry praise for providing platform agnostic technical guidance in high performance computing (HPC), big data, cloud and software-defined data centers—in the most efficient, reliable and cost effective manner. The world's leading research facilities, government, life science, financial, healthcare, energy, and cloud service providers can leverage the our team of engineers' extensive academic, research lab and commercial expertise that makes RAID Inc. a trusted industry leader.