

ThinkParQ announces BeeOND - a new parallel storage based on BeeGFS

Kaiserslautern (Germany), February 10th, 2015 - Today ThinkParQ announced a new product based on the widely used scalable parallel filesystem BeeGFS. The new product is named BeeOND, which stands for "BeeGFS on Demand" and allows users with HPC clusters to generate job-dedicated parallel filesystem instances on demand.

Today's HPC workloads are demanding when it comes to storage performance. More and more HPC cluster systems come with a central storage system based on a parallel filesystem. Still, for some workloads with special I/O-patterns, compute nodes are often equipped with local harddisks or SSDs. The issue with that approach is that one doesn't have either the advantage of a single name space nor the flexibility and performance of a number of aggregated drives.

This is where BeeOND comes into play by providing a shared parallel filesystem on a "per job basis" across all compute nodes being assigned to that particular job. It is based on the extremely lightweight architecture of BeeGFS and also includes a fully scalable metadata architecture - therefore BeeOND doesn't interfere with the running application.

BeeOND targets especially to aggregate IOPS, bandwidth and capacity of local SSDs or harddisks in compute nodes for the duration of a compute job. This adds a new way of flexible usage to many compute clusters. Furthermore it integrates with workload managers such as torque or PBS Pro for easy usage. The startup and destruction of the filesystem don't take more than a few seconds due to the efficient architecture of BeeOND.

"With the announcement of BeeOND we take the next step in parallel storage. It can not only be used as a burst buffer for write-intensive workloads, but enables system administrators to free up their central storage from all kinds of performance-critical workloads - and at the same time boost the performance of them. We see that the adoption of SSDs for storing data in compute nodes is not very broad yet, mostly because capacity is still expensive while a single SSD provides more than enough performance for a single compute node. BeeOND will solve this problem by aggregating the capacity of several reserved compute nodes into one shared name space. And with its efficient architecture it will also deliver the aggregated performance. We consider this a game changer for a range of applications."
Jan Heichler - ThinkParQ

Workloads that make use of a dedicated shared storage can be found in a range of fields of scientific computing. Especially life science workloads with frequent access to genome repositories and chemistry codes that store data structures on persistent storage are predestined. But also BigData Analytics matches the profile very well: BeeOND enables general purpose clusters to run map/reduce workloads without big changes on the software side.

BeeOND will be available March 2015 with commercial support through ThinkParQ. It will also come with a special tool for efficient stage-in and stage-out of data from and back to the global cluster storage system. BeeOND can be used independent of whether the global shared cluster file system is based on BeeGFS or on other technology.

About BeeGFS: BeeGFS is the new name for the well known parallel filesystem FhGFS. FhGFS was developed at the Fraunhofer Institute in Kaiserslautern. It provides excellent performance and scalability - but combines it with

ease of use, which is unique for a filesystem targeted at High Performance Computing. Additionally, it is extremely flexible and doesn't lock users in with specific Linux distributions or kernel versions. With these key differentiators, BeeGFS is adopted in a large number of scientific and commercial sites as the default choice for a work-filesystem for HPC.

About ThinkParQ: ThinkParQ was founded as a spin-off from the Fraunhofer Competence Center for HPC in Kaiserslautern to bring BeeGFS to the market. ThinkParQ supports system integrators creating turn-key solutions, is responsible for support, provides individual presentations on demand, organizes events, attends exhibitions and works very closely with system integrators.