NoSQL
Case Solution

PERFORMANCE ISSUES OF NOSQL DATABASES

As more and more deployments move to the cloud and with an increase in size of data sets (Big Data), Non-Relational Databases or NoSQL are being increasingly adopted. NoSQL databases are particularly prone to be affected by performance cliff, which happens when the working set of the application exceeds the system RAM. Due to the inherent way in which NoSQL databases and the applications IO performance issues crop up. Solid State Devices are gaining popularity too, given the low latency high throughput option they present. All cloud providers now have SSD as part of their offerings, as direct attached devices and also as network attached devices. Larger direct attached SSDs are only available with larger and more expensive instance types. The problem with instance store disks are that they can lose data on a hard reset or if the base machine were to restart. Since the data cannot survive past hard resets and restarts, the data on these devices need to be backed up at all times on a more reliable device type.

DEEP INSIGHT INTO IO BEHAVIOR THROUGH ANALYTICS

PerfAccel provides deep analytics into IO behavior, through:

- Tracking I/O, read hits and read misses
- Tracking Latency and Throughput measurement at application level
- Identifying Hot files

Such an IO tracking helps administrators understand dynamics of active data better and use this intelligence to better data placement and improve performance of NoSQL database deployments.

USING CACHE INTELLIGENTLY

Using the data intelligence, administrators can configure their caching policies and rules to ensure that the right working set resides in the cache. This would include:

- Caching only hot areas of the file or frequently accessed files
- Controlling which data to cache and which to not
- Sticky flagging for performance critical data
- Pre-fetching files for large jobs

RESOLVE THE INSTANCE STORE SSD BOTTLENECK

Using the insight from data analytics, PerfAccel can use the instance store devices more efficiently by using the faster device available as a cache and through the optimal placement of frequently used hot data. The application directly benefits since all the reads coming from this device are much faster increasing performance and reducing latency. In addition, since these read operations are offloaded by the cache, the backend storage device which holds the entire dataset is more responsive as it has to serve fewer IOPS. Thus PerfAccel cache not only improves read performance, it also implicitly improves the write performance of the application.

PerfAccel provides:

- Storage visibility through deep file-level analytics

- Intelligent caching & deterministic placement of hot files

- Higher performance using fewer SSDs used optimally

- Cost reduction

SINGLE WINDOW MANAGEMENT

A single GUI pane combined analytics and insight through performance dashboards, as well as a simple command line, for managing across large grid deployments with a centralized data repository for analytics. PerfAccel's flexible interface let users configure their own policies of persistent cache, pre-fetching, predictive cache, real-time cache size configuration and auto-caching hundreds of NFS mount points.

About Datagres

Datagres provides software that helps companies visualize, control and accelerate their application performance using deep storage intelligence. Datagres' flagship productPerfAccel is a very powerful analytics driven software solution that operates at a file level and can show the exact IO pattern of an application data access especially in a scale-out grid environment. As a result, it can provide an effective way of controlling IOs and also accelerate for higher throughput and lower latencies using high-performance SSD devices.

The company is headquartered in Palo Alto, California and is venture-backed by Nexus Venture Partners

For more information, visit www.datagres.com

Datagres Technologies Inc

2600 EL CAMINO REAL, Palo Alto, CA 94306 Phone: 510-402-4365
www.datagres.com

DATAGRES
Application I/O Intelligence