ENTERPRISE STORAGE STACK

ESS Can Make Cloud Clusters Faster, More Robust, and Less Expensive

ESS's linear writing method changes the rules of mass storage. How this impacts cloud and cluster storage architecture depends on the specific goals and usage patterns. ESS benefits for cluster deployment come at several levels:

ESS generates less wear amplification: This lets you use less expensive, lower grade media while still meeting design-life goals. Typically, this can halve OSD node costs.

ESS Random Writes Faster: Due to ESS's linear design, a node with 24 SSDs will typically random write at 6GB/sec – 1.6 million IOPS even with Raid-5 parity.

Use of parity Raid. Environments such as CEPH recommend against parity raid due to slowness and the ability of CEPH to tolerate individual disk failures. Parity and hot spares can improve cluster manageability and virtually eliminate rebalance operations associated with single disk failures.

ESS Compression is amazingly logical on OSD nodes. Even with marginally compressible data, ESS will give you free bandwidth, and free space.

ESS Deduplication verifies and virtualizes duplicate data real time. In some environments, this can radically reduce physical storage needs.

Taking full benefit of ESS with CEPH requires changes to the way you visualize OSD devices.

An ESS "device" should be stable, but does not need to be HA. The cluster will handle a server failure. ESS should be a single volume with data and log on the same device.

In order to spread out access to the, potentially large, ESS volume, it should then be cut up with LVM into smaller "virtual OSD devices". So a 20TB ESS "server" might get 20 1TB OSD daemons. This will help OSD performance.

With ESS data reduction, techniques must be employed to deliver the synthesized storage to the cluster, and to monitor and tune the amount of over provisioned space available. Fortunately, several options are available ranging from deploying extra OSD nodes to varying the fill level of nodes that are backed by ESS with data reduction. In either case, ESS data reduction can be directly delivered to the cluster as usable storage.

Likewise, CEPH load balance ratios might be used for similar reasons. An ESS node might be setup starting with 30% utilization and 3:1 over-commit. As data is added the utilization can be increased as the data reduction ratios are better evaluated.

Much of this talk about CEPH is "conceptual". EasyCo is only just starting active work on clusters. We believe that ESS, when applied to the edge of a CEPH cluster at the storage device "level" can provide a great deal of value, particularly where SSD level performance is desirable, and where SSD costs need to be mitigated. EasyCo is actively looking for cloud partners to see how and where ESS can be exploited to maximize cluster performance, capacity, manageability, and cost.

EasyCo LLC
220 Stanford Drive
Wallingford
PA 19086 USA
Tel: (+1) 610-237-2000
888-473-7866
Email: sales@EasyCo.com
Web: http://WildFire-Storage.com