PROVING SECONDARY STORAGE AT CLOUD-SCALE

Cohesity Performance Scales Linearly in Massive 256 Node Test

JULY 2017

Are we doomed to drown in our own data? Enterprise storage is growing fast enough with today's data demands to threaten service levels, challenge IT expertise and often eat up a majority of new IT spending. And the amount of competitively useful data could possibly grow magnitudes faster with new trends in web-scale applications, IoT and big data. On top of that, assuring full enterprise requirements for data protection with traditional fragmented secondary storage designs means that more than a dozen copies of important data are often inefficiently eating up even more capacity at an alarming rate.

Cohesity, a feature-rich secondary storage data management solution based on a core parallel file system, promises to completely break through traditional secondary storage scaling limitations with its inherently scale-out approach. This is a big claim, and so we've executed a validation test of Cohesity under massive scaling – pushing their storage cluster to sizes far past what they've previously publicly tested.

The result is striking (though perhaps not internally surprising given their engineering design goals). We documented linearly accumulating performance across several types of IO, all the way up to our cluster test target of 256 Cohesity storage nodes. Other secondary storage designs can be expected to drop off at a far earlier point, either hitting a hard constraint (e.g. a limited cluster size) or with severely decreasing returns in performance.

We also took the opportunity to validate some important storage requirements at scale. For example, we also tested that Cohesity ensured global file consistency and full cluster resilience even at the largest scale of deployment. Given overall test performance as validated in this report, Cohesity has certainly demonstrated it is inherently a web-scale system that can deliver advanced secondary storage functionality at any practical enterprise scale of deployment.

THE SECONDARY STORAGE SCALE CHALLENGE

We all recognize that traditional approaches to effectively manage secondary storage are crumbling – SLA's are being missed due to shrinking backup windows and data growth, customer expectations for immediate, granular recovery are increasing, and more organizations are leveraging so-called secondary data far more actively for things like big data analytics and highly agile development. Aging secondary storage architectures are rife with complexity, have big gaps in coverage, and present increasing cost and risk with every new web-scale application (e.g. globally deployed and fully containerized mission critical applications and databases).

We'd expect new technology to save us from this data onslaught. However, many of the more modern secondary storage solutions that attempt to unify and simplify secondary storage capabilities, claim optimized secondary data management to reduce capacity, and promise a host of other benefits, are severely challenged to run, much less perform well, at cloud-like scale.

87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557

Cohesity has bravely dropped into this arena with a secondary storage design that will fundamentally scale-out modern storage services (e.g. data management) without hard limits. It's worth mentioning that unlike other secondary storage solutions that are point products or cover only a few bases (requiring IT to assemble and integrate a whole bunch of disparate products), Cohesity was designed to deliver on all common secondary data use cases including primary data protection, long term data integrity/replication, long-tail analytics, agile test/dev, global data deduplication, and global online secondary file access. In a very real way, all this functionality makes delivering large-scale performance much harder. This powerful secondary storage solution would be really valuable to almost all enterprises –if it really performances at scale. To prove exactly that out, we undertook to validate an impressively large-scale performance test.

COHESITY PERFORMANCE TECHNICAL VALIDATION

Just being functional at scale is, of course, important, but really the system has to function and perform well as the system size (and workload) grows. In this validation testing exercise, we've documented a test series in which Cohesity storage nodes are added to the Cohesity cluster to accommodate matching growth in secondary storage IO workload. We validated internal Cohesity test results starting with smaller, reasonable cluster sizes (of the kind their customers already have deployed in production) and then first-hand worked with Cohesity engineers to grow the cluster to a size far beyond the average, just to see what would happen if (or when) future enterprise needs might get to that scale of requirement.

In reality, one would properly plan to deploy storage capacity ahead of expected growth curves, and to that end we did in fact examine Cohesity's built-in call home/support functionality which helps Cohesity help their customers stay ahead of their actual growth curves. For the purposes of this testing, we simply added equivalent storage workload with each node added to the cluster to keep every node adequately utilized.

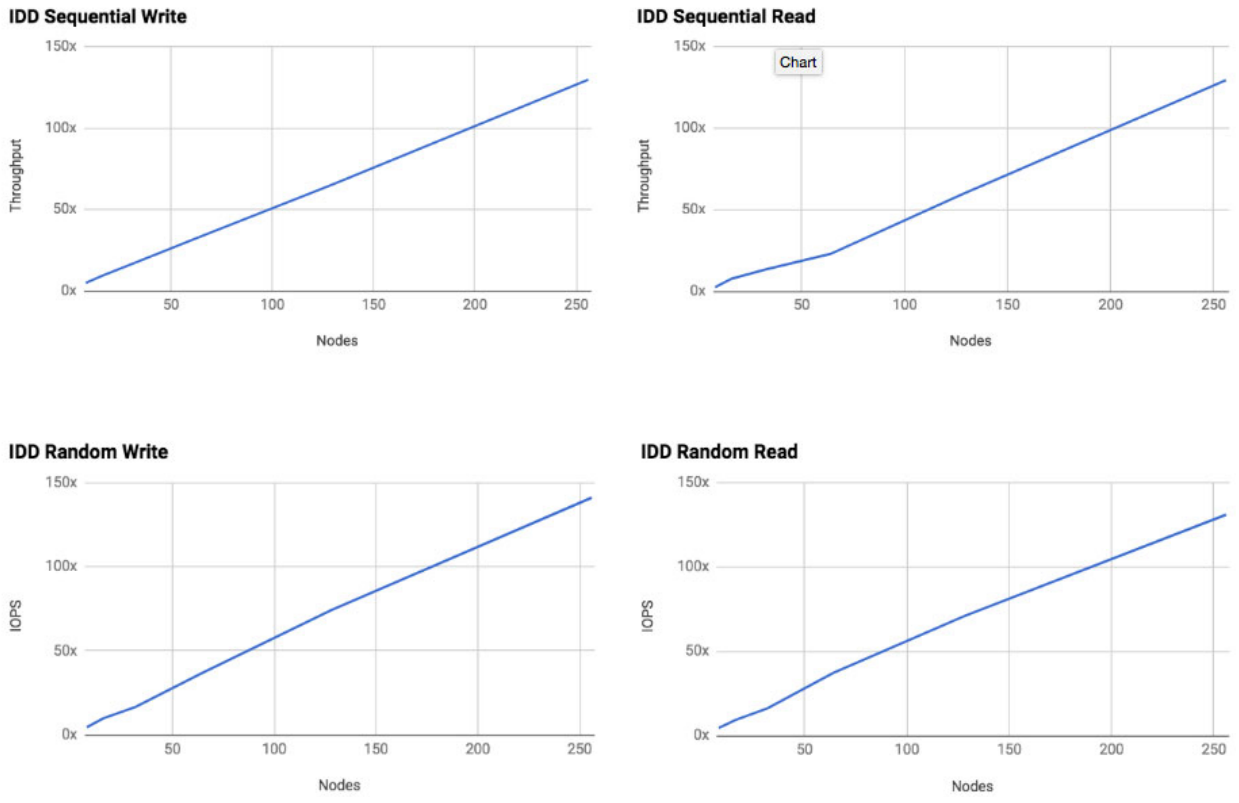


Figure 1 - Large Scale-Out Performance Test Results

87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557

We validated the results for Cohesity clusters ranging from 8 to 256 nodes. Because of the logistical difficulty of assembling such large physical clusters just for testing, the Cohesity engineers implemented their test kit in the Microsoft Azure public cloud environment. Even though it's relatively straightforward to request provisioning, it still took 3-4 hours of time to check out the largest cluster of machines. That and the cost of running such large clusters hard for relatively long periods meant that we judiciously doubled the cluster size between each test. Other than the expected difference between interacting with cloud management to provision machines rather than work with direct on-premise hardware, the resulting Cohesity cluster behavior proved essentially identical whether run in-cloud or on-premises.

The testing suite consisted of four IO test cases–random and sequential reads or writes, powered by the Flexible I/O Tester (popularly known as FIO) load generator. Within each test use case, each node was targeted with reading/writing 2GB files. Inline dedupe was enabled during all testing so that the test results would reflect potential issues with full global dedupe processing at scale.

Each test was repeated three times and averaged. At the largest 256 node test size (3.15 PB), over 6TB of data was read/written for each test case (for more details of the test configuration see Appendix – Notes on the Cloud Test Environment).

During the performance testing observed, actual performance was extremely consistent across all nodes within the cluster. Most nodes were within 1% of the average performance (time, bandwidth, IOPS), and the few outliers we judged likely to have been caused by cloud-processing issues. When running in on-premise physical hardware, Cohesity reports that all storage nodes perform almost identically.

The resulting overall performance over the range of cluster sizes was remarkably linear, showing essentially zero sensitivity to cluster growth. We noted small variations away from absolutely perfect linearity, but that was well within the normal performance variations that could be expected when running hot workloads in a public cloud environment. Essentially, N times more nodes produced N times more performance. There was no cost or overhead in larger clusters, no degradation or flattening off of the performance curves and no point of diminishing returns. This indicates an extremely efficient and well-designed clustering implementation that could be reasonably expected to grow much larger.

KEY FUNCTIONAL VALIDATION TESTING

Global File Consistency

Of course, it's not sufficient just to drive and operate a larger cluster. Cohesity claims the data stored within even the largest cluster is immediately globally consistent and readily available (from any node) within its global namespace.

Using the largest cloud cluster, Cohesity demonstrated to us their ability to write new data to one node and immediately read it from another (we used md5 checksums to verify data integrity). The test was simple but implies a lot of sophistication under the hood. Fundamentally we saw that data can be read from and written to any node in a (even a large) cluster, while remaining fully consistent across the cluster.

Cluster Resiliency And Data Protection

Another aspect we wanted to test was resiliency at scale. We initially wanted to demonstrate failure recovery by killing a storage node (or disk), but as it turns out this wasn't actually easy in the cloud because cloud infrastructure itself is resilient. In the cloud, we couldn't actually force an infrastructure failure, and even if we could the cloud infrastructure (server or disk/volume) would recover transparently (or so we'd expect).

Another complication was that Microsoft had recommended that Cohesity only use RF1 (a replication factor of 1 meaning "no" replication) when running in Azure, as Azure storage is already replicating under the

87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557

hood. Therefore if we tried to simulate a failure by outright deleting an Azure machine (and its internally replicated storage), Cohesity processes wouldn't be able to turn over to a Cohesity replicated data copy (usually set to RF2 or 3 in production).

We decided that we would instead migrate one of the storage nodes to a new Azure machine. This did have the effect of halting new writes to the cluster during migration, as Cohesity manages the whole storage pool as a fully globally "striped" virtual volume. But again, if there was a normal replication factor, IO could turn immediately over to a replicate. Once our node was back up, we were able to verify that the node's data had survived the experience unscathed and that Cohesity had immediately returned to full operation. This is a good sign that even at large scale (and in the face of cloud migrations), Cohesity is robust and can handle node loss with no risk to customer data.

Management at Scale

As a secondary validation, we can document that we productively used the Cohesity management interface during testing. We saw no differences in ease of management as the cluster sizes grew, and found the analytics and management/monitoring to provide exactly the information we needed to validate proper and effective operation.

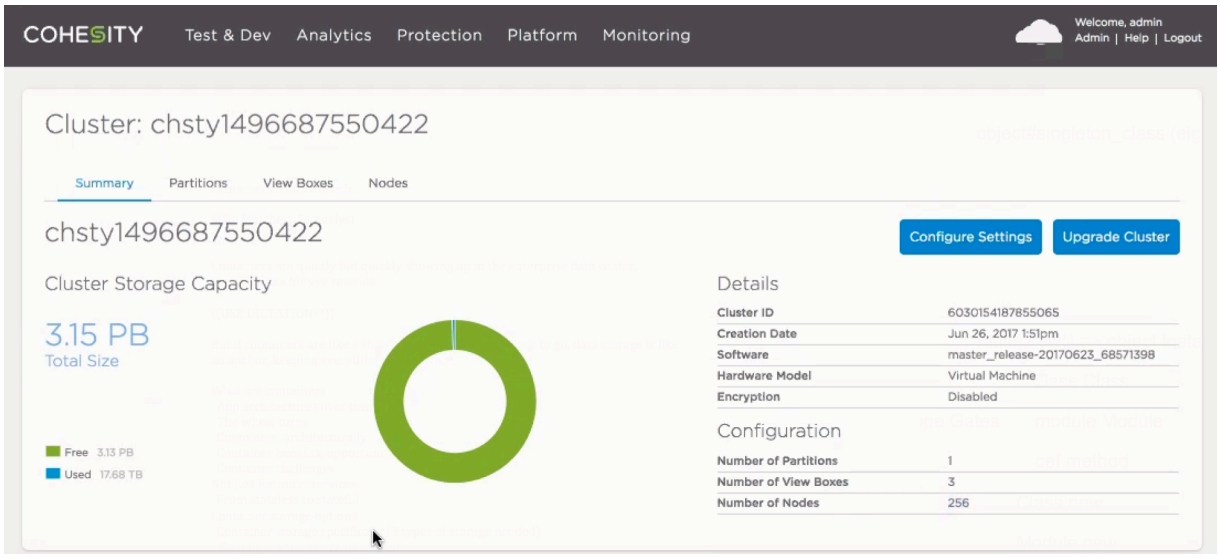


Figure 2 - Cohesity Management

We also reviewed internal call-home Cohesity data collected off the cloud-deployed "client" cluster. This data stream indicated that Cohesity support would be able to effectively monitor, track and proactively help with large Cohesity implementations both on-premise and in-cloud.

TANEJA GROUP OPINION

Linear performance curves with scale-out systems are almost unheard of. At some point even with fully distributed and parallelized functional tasks, increasing system management, overhead processes like scheduling, inline dedupe, East/West network traffic or some other "resource" constraint usually results in decreasing returns as cluster sizes grow. It's clear that the Cohesity architects have applied the lessons of recent large-scale big data and cloud-scale technologies (e.g. Hadoop, Google File System) to enterprise secondary storage, and managed to engineer vast scaling potential into their system.

This testing does beg the question of just how big a cluster Cohesity can grow up to? Honestly, we didn't find a clue in this testing as to any actual design limitations. We can't promise there aren't any, but the absence of observable degradation at the scales tested should be highly reassuring to anyone now trying to

87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557

architect enterprise secondary storage that will be expected to stand the test of time – and incredible data growth.

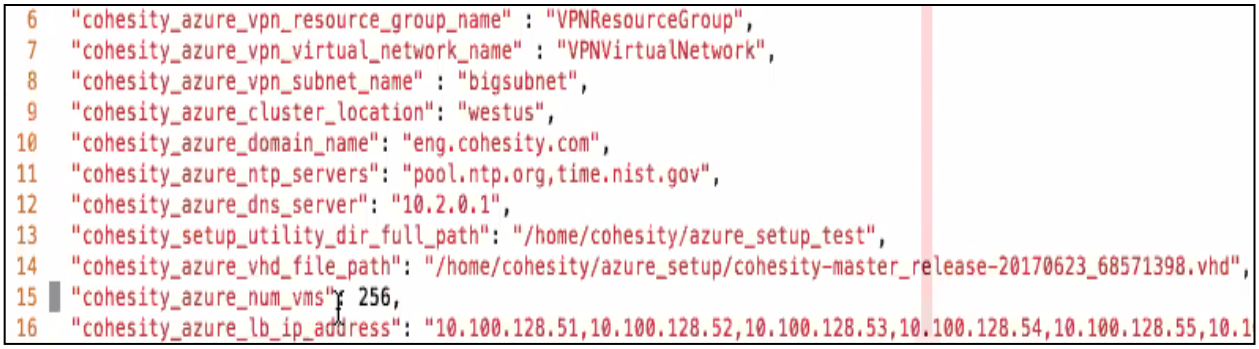
Now we don't mean to suggest with this testing that Cohesity customers need start by thinking of huge initial deployments. The truth, as we saw repeatedly during the testing, is that Cohesity is a very easy solution to deploy at any scale. In fact we'd recommend new customers might think to deploy Cohesity first very productively at small scales on high-value projects. Then, and we think soon after its becomes known just what can now be done with a fully featured and complete secondary storage solution, Cohesity can reliably and without downtime be grown as needs dictate.

We know that Cohesity offers many advantages beyond large scale and performance, like offering an industry leading RTO capability based on a unique "snap tree technology", highly agile copy data management, built-in QoS priority queuing, and automated tiering. Even as initially just an aggregated scalable backup target Cohesity seems unstoppable, but when you want to actively leverage your secondary data, Cohesity might be unbeatable.

APPENDIX - NOTES ON THE CLOUD TEST ENVIRONMENT

Setting up the largest 256-node cluster in a public cloud actually took most of a day, but this was primarily due to cloud provisioning/performance latencies. It would not require much time to set up our testing on an existing virtual or physical cluster (though just racking and cabling a 256 node physical cluster of servers and disks could take weeks).

The actual setup process was very straightforward. Cohesity provides a tool to generate a cluster in Azure, and it is a simple matter to run the tool, driving it with a JSON configuration file. This file specifies details of the test Azure environment, the cluster size (line 15), and the IP address for each node (starting on line 16). Note especially the VM image specified in line 14 -- this is exactly the same system image that would be deployed in an on-premise Cohesity cluster.



```
 6  "cohesity_azure_vpn_resource_group_name" : "VPNResourceGroup",
 7  "cohesity_azure_vpn_virtual_network_name" : "VPNVirtualNetwork",
 8  "cohesity_azure_vpn_subnet_name" : "bigsubnet",
 9  "cohesity_azure_cluster_location": "westus",
10  "cohesity_azure_domain_name": "eng.cohesity.com",
11  "cohesity_azure_ntp_servers": "pool.ntp.org,time.nist.gov",
12  "cohesity_azure_dns_server": "10.2.0.1",
13  "cohesity_setup_utility_dir_full_path": "/home/cohesity/azure_setup_test",
14  "cohesity_azure_vhd_file_path": "/home/cohesity/azure_setup/cohesity-master_release-20170623_68571398.vhd",
15  "cohesity_azure_num_vms": 256,
16  "cohesity_azure_lb_ip_address": "10.100.128.51,10.100.128.52,10.100.128.53,10.100.128.54,10.100.128.55,10.1
```

The setup tool when executed uploads the VM image to each node in the cloud. Since many Cohesity clients run VMware environments, our test also happens to demonstrate the ability to upload the basic VMware VM image and have it convert it to a Hyper-V VM within Azure.

Note the setup configuration for these tests provisioned 1 TB of Premium Azure storage and 12 1TB volumes of Azure standard storage per node. The Cohesity OS then virtually striped all these volumes together into one cohesive(!) storage unit of 3.15PB at the 256 node cluster size.

We've reported the test results in relative performance terms as the important analysis is the linearity of the scaling, not the absolute performance measurement. Obviously, the size, type and capacity of provisioned storage nodes and storage services could vary widely, and any variations in underlying infrastructure configurations would be expected to change the absolute values of observed performance results. We do not believe that additional capacity would have any impact, while "faster" (or slower) storage resources would likely only shift the absolute performance line down (or up). The key observation, linear performance with cluster size, is independent of actual absolute underlying disk/cloud storage performance (assuming it in-turn remains constant at scale, which should be the case with big provider cloud storage services).

87 Elm Street, Suite 900 · Hopkinton, MA 01748 · T: 508.435.2556 · F: 508.435.2557