

# AppNexus Scales with DriveScale

Fast-Growing Digital Advertising Company To Build Out Data Centers with a Cost-Effective, Flexible and Agile Approach



DriveScale's Software Composable Infrastructure (SCI) is a next generation data center architecture that uses software to optimize infrastructure, delivering significant gains in efficiency and agility, while lowering costs.

## Executive Summary

As one of the premier digital advertising technology companies, AppNexus depends on data, time and scale to power the marketplace for digital buyers and sellers. Since its founding in 2007, AppNexus has created a world-class platform that requires an infrastructure which dynamically scales and grows. AppNexus selected DriveScale's unique Software Composable Infrastructure as a critical element of its data center architecture to ensure it can continue to scale with agility and flexibility at the right price point.



## AppNexus Background

AppNexus is the largest independent digital advertising marketplace and is a technology company that harnesses data and machine learning to power its solution. The company began by building a powerful, technology platform in which it embedded core capabilities, tools, and features. On top of that platform, AppNexus built both a suite of cloud-based enterprise products that enable the buying and selling of digital advertising as well as a scaled, transparent, and low-cost marketplace where companies can transact — using either AppNexus's enterprise products or those offered by other companies.

AppNexus executes more than 10 billion transactions per day, each a real-time auction conducted in a fraction of a second. Every transaction is conducted with maximum efficiency and the lowest possible cost.

### Peak daily feeds/speeds

- 10.7 billion impressions transacted
- 123 billion auctions
- 5 million queries per second
- 150TB data generated & processed
- 400gbps in/out of the Internet

### AppNexus Faced an Infrastructure Challenge

AppNexus has approximately 8,000 servers in seven data centers on three continents that run around the clock. Servers are grouped into four general major categories: real-time processing (the actual ad serving), big data pipeline, client facing tools and infrastructure. Tools and applications include Hadoop (HDFS, Hive, Spark) and the Vertica Analytics Platform that comprise the big data pipeline. Both Hadoop and Vertica are disk and I/O intensive, cluster-based applications. In addition, some of the infrastructure applications are also disk-intensive and clustered.

At AppNexus, the big data apps are all clustered, and all clusters are built with commodity servers. The big data app clusters historically have used internal disks while almost all of the other clusters are very light on disk. With the evolution of NVMe and SATADOMs, most of the servers not used for big data are becoming diskless from the application point of view.

Given that the server is the fundamental building block of any cluster, regardless of whether more compute power or more storage is needed for a given workload, adding more servers limits flexibility and creates substantial inefficiencies. At AppNexus, these inefficiencies derived from two major factors. First, not all of the workloads required the same ratio of compute to storage. Some workloads are more compute-intensive and less dependent on storage while others require far more storage. By using standardized server configurations for each workload, the clusters often had underutilized compute or storage, which is inefficient and expensive.

Second, the compute-to-disk ratio is generally fixed at the time of purchase. Of course, adding more disks was a possibility up to a point, but doing so in the field was both labor intensive and complicated from a logistics and asset management standpoint. Additionally, the maximum number of disks was limited by the physical slots in the server, which fell into the “more storage requires more servers regardless of whether the workload needs more compute” conundrum and thus exacerbated the inefficiencies.

“The solution is to have the silos disappear. Silos fundamentally limit flexibility and efficiency.”

Timothy Smith, SVP TechOps  
AppNexus

Each application cluster became a silo due to its configuration of storage and thus its compute-to-storage ratio. This precluded moving resources among the clusters running each type of workload, resulting in significant inefficiencies. These inefficiencies were then compounded because workloads can vary significantly over time, and thus the ratio of compute-to-storage changes as applications evolve. The end result is both unused resources trapped in silos and configurations that are no longer optimized for workloads and thus workloads may not perform optimally.

Another consideration for AppNexus was that servers and storage have significantly different lifecycles. Typically, servers have an effective lifetime of three years before they are obsolete whereas storage generally has an effective lifetime of four to six years, which can sometimes be stretched even longer. With the storage residing inside servers, there are several inefficiencies and “hidden” costs when upgrading the servers. The storage is essentially being replaced long before its effective end of life. Furthermore, when this is done, it necessitates a massive data migration, which is time consuming, logistically difficult, and high risk.

### DriveScale Offers a Different Approach

After extensive research to find a solution to its infrastructure challenges, AppNexus chose DriveScale. With DriveScale’s Software Composable Infrastructure, AppNexus can now disaggregate its compute and storage resources, directly addressing the fundamental problems caused by the storage physically residing inside the server. By moving the storage out of the servers and onto the network, and with the ability for the servers to access storage easily, securely, and with performance equivalent to local attached storage, the DriveScale solution provided a cost effective and flexible solution.

### How DriveScale Works for AppNexus

Instead of buying industry standard servers with internal drives, AppNexus now buys diskless servers along with JBODs (Just a Bunch of Disks). DriveScale easily “composes” servers with whatever amount of compute power and disk storage is needed from the server resource pool and the disk storage resource pool. The DriveScale solution delivers the correct balance of compute and storage for each workload, and can be adjusted easily depending on the workload and balance over time with software. These adjustments are made on-demand via the DriveScale GUI or from scripts that leverage its RESTful API.

## DriveScale Optimizes AppNexus Clusters

DriveScale allows AppNexus to add or remove resources from each cluster easily, allowing it to manage precisely how much compute power and storage are deployed for each workload. Because this is all done with software, it can be performed quickly and easily. No technician is necessary to physically touch the servers in order to make these changes, which is particularly important to AppNexus because it operates its data centers with a minimum of resources. In fact, in some data centers, AppNexus deploys a technician only quarterly.

AppNexus will achieve its goal of eliminating silos entirely once it completely migrates to DriveScale. Once complete, all of the compute and storage resources will be consolidated into just a few common resources pools that can be used to create correctly sized clusters for each workload. In addition, when AppNexus's workloads change, it can easily reallocate resources. Since compute and storage are now disaggregated, or separated into diskless servers and JBODs, AppNexus can upgrade one without touching the other.

When the company wants to upgrade servers to take advantage of new processors or new memory architectures, it can simply replace the servers and recombine the disks with the new servers. This not only saves the cost of buying new disks, but greatly simplifies logistics and reduces risk, because it doesn't require copying data from old to new disk drives. Furthermore, AppNexus no longer has to initiate complex and time-consuming cluster rebalances to bring new nodes into the cluster and rebalance them, nor must it remove old nodes from the cluster and rebalance, which triggers massive data copies.

## No compromises with DriveScale

AppNexus benchmarked local attached storage against DriveScale and JBODs using both synthetic benchmarks and its actual workloads. Performance for AppNexus workloads on servers with internal drives was nearly identical to the results achieved with DriveScale using diskless servers and JBODs. In some cases, DriveScale was faster than local attached storage. Overall, the performance difference was negligible.

In deploying the DriveScale solution, AppNexus worked with the DriveScale founding technologists who have developed some of the most successful data center infrastructure solutions ever available from companies like Cisco and Sun Microsystems.

As Timothy Smith, SVP TechOps at AppNexus relates, "Everyone I have met at DriveScale, from the CEO down, has a singular focus on the customer and the customers' needs. Our success is clearly as critical to them as it is to us. They provided great support from initial proof-of-concept through production deployment. We are very happy with our decision to use DriveScale."



DriveScale, Inc  
1230 Midas Way, Suite 210  
Sunnyvale, CA 94085

Main: +1(408) 849-4651  
[www.drivescale.com](http://www.drivescale.com)