



# DriveScale-PROMISE Reference Architecture

# Table of Contents

## Contents

<b>Introduction</b> .....	<b>3</b>
<b>Audience and Scope</b> .....	<b>3</b>
<b>DriveScale Advantage</b> .....	<b>3</b>
<b>Promise Technology Advantage</b> .....	<b>4</b>
<b>Industry standard x86 Servers</b> .....	<b>4</b>
<b>DriveScale   Promise Solution Overview</b> .....	<b>4</b>
Entry or POC Configuration .....	4
Full Rack/Compute Rack Configuration .....	6
Storage Rack Configuration .....	7
<b>Rack Scalability</b> .....	<b>8</b>
<b>Summary</b> .....	<b>8</b>
<b>Conclusion</b> .....	<b>9</b>

# DRIVESCALE-PROMISE

## Introduction

DriveScale has engineered the next -generation Software Composable Infrastructure (SCI) that is designed to fundamentally change the way data center architects design, deploy, manage and consume the hardware and software resources. DriveScale provides capabilities to IT operators to connect disaggregated resources in an intelligent manner, the ability to manage, modify and scale these components over time. SCI results in a simpler deployment model with a fluid pool of resources that can be used for modern workloads such as Apache Hadoop, Spark, Cassandra etc. with significantly improved agility.

This document is a high-level design reference architecture guide for implementing a DriveScale solution with industry standard x86 servers and Promise Technology JBODs for Big Data workloads.

The reference architecture introduces all the high-level hardware and software component that are included in the stack. Each high-level component is then described individually.

## Audience and Scope

This reference architecture guide is for modern workloads and IT architects who are responsible for the design and deployment modern workload solutions on premises, as well as for Big Data administrators and architects and data center architects/engineers who collaborate with specialists in that space.

## DriveScale Advantage

Modern workloads such as Big Data workloads have become integral part of the traditional data centers. Typical Big Data deployments in any enterprise start with small and isolated deployments. They are designed to scale by adding more hardware resources to the same cluster. This flexibility is built into the Big Data software.

However, typical Big Data deployments have multiple limitations:

- Administrators can't respond quickly to changing application stacks and data velocity.
- Deployments are over-provisioned with resources and under-utilized in order to provide service level guarantees.
- Multiple silos of hardware are created for each application workload.

Once installed, the DriveScale Management System automatically discovers the installed components and their network connectivity and inventories them. It also discovers network bandwidth domains to ensure that drives in the software composed servers are with in the same bandwidth domain for optimal performance. It then uses this inventory to compose logical servers in software using either Restful API that can be interfaced to customer's management system, or use DriveScale GUI.

DriveScale software composed servers deliver nearly the same performance as custom configured hardware servers in all Big Data micro-benchmarks, and match the performance in real Big Data application loads.

DriveScale Management System enables a customer to define alarms on a number of hardware metrics to enables alerts to be generated when these alarms are triggered. The DMS enables a customer to diagnose the system at logical server level or physical component level.

DriveScale’s rack scale architecture provides the solutions to all the modern Big Data workload deployment limitations. **Now administrators can flexibly deploy and manage independent pools of resources at a lower cost without making changes to the application stack.**

### Promise Technology Advantage

The VTrak J5000 Series delivers an affordable high-performance SAS solution with advanced enterprise level reliability and functionality for Big data and modern workloads. The four external 12 Gb/s SAS ports per I/O module provide the connectivity for large capacity solutions requiring multiple cascaded JBODs, and high bandwidth solutions requiring speed. It features fully redundant I/O modules, power supplies and cooling units, combined with a cable-less system design, providing exceptional data availability even during component failures for the enterprise environment. Promise technology’s VTrak J500’s series JBOD is designed to deliver:

- Lower cost high-performance SAS solution with advanced enterprise level reliability and functionality
- Fully redundant JBOD
- Lower IT operation costs

**Table 1: DriveScale certified JBOD**

JBOD Vendor	Model Number
Promise Inc.	J5300s - 3.5", 12 bays, 2U, redundant expanders, 2 x 2 x mini-SAS-HD 12G
	J5320s - 2.5", 24 bays, 2U, redundant expanders, 2 x 2 x mini-SAS-HD 12G
	J5600 - 3.5", 16 bays, 3U, redundant expanders, 2 x 2 x mini-SAS-HD 12G
	J5800 - 3.5", 24 bays, 4U, redundant expanders, 2 x 2 x mini-SAS-HD 12G

### Industry standard x86 Servers

With DriveScale solution, we recommend customers should use industry standard x86 servers. Customers can choose to expand and build the datacenter with their existing server vendor. DriveScale solution integrates with any industry standard x86 server. DriveScale is a certified partner with Dell, Cisco and HPE.

### DriveScale | Promise Solution Overview

#### Entry or POC Configuration

The reference architectures shown below describes a typical single-rack configuration for deploying a Big Data infrastructure utilizing direct-attached storage in the form of Promise JBOD’s, industry standard servers, networking ethernet switches and DriveScale Adapter chassis with four adapters.

The first design shown is a suggested entry configuration that might be deployed for a small workload or for proof-of-concept testing. This design highlights the value proposition of the DriveScale/Promise solution in that it demonstrates how customers can achieve significantly higher levels of **flexibility** in deployment, **reduction in cost** by starting small and **seamless integration** with Big Data workloads such as Hadoop with a highly software composable solution.

The entry or POC configuration is based 4 industry standard x86 disk lite servers serving as node. The storage for the Big Data workloads are provided by the 24 drives hosted in the Promise JBOD. Using the DriveScale Adapter and DMS, customers can mount drives to each of the nodes. They can also expand the capacity of the cluster by adding more nodes if they application requires more compute or add more drives to the existing nodes from the JBOD. Customers have the flexibility to spread the 24 drives across the 4 nodes or add more nodes to be used with the existing 24 drives in the JBOD.

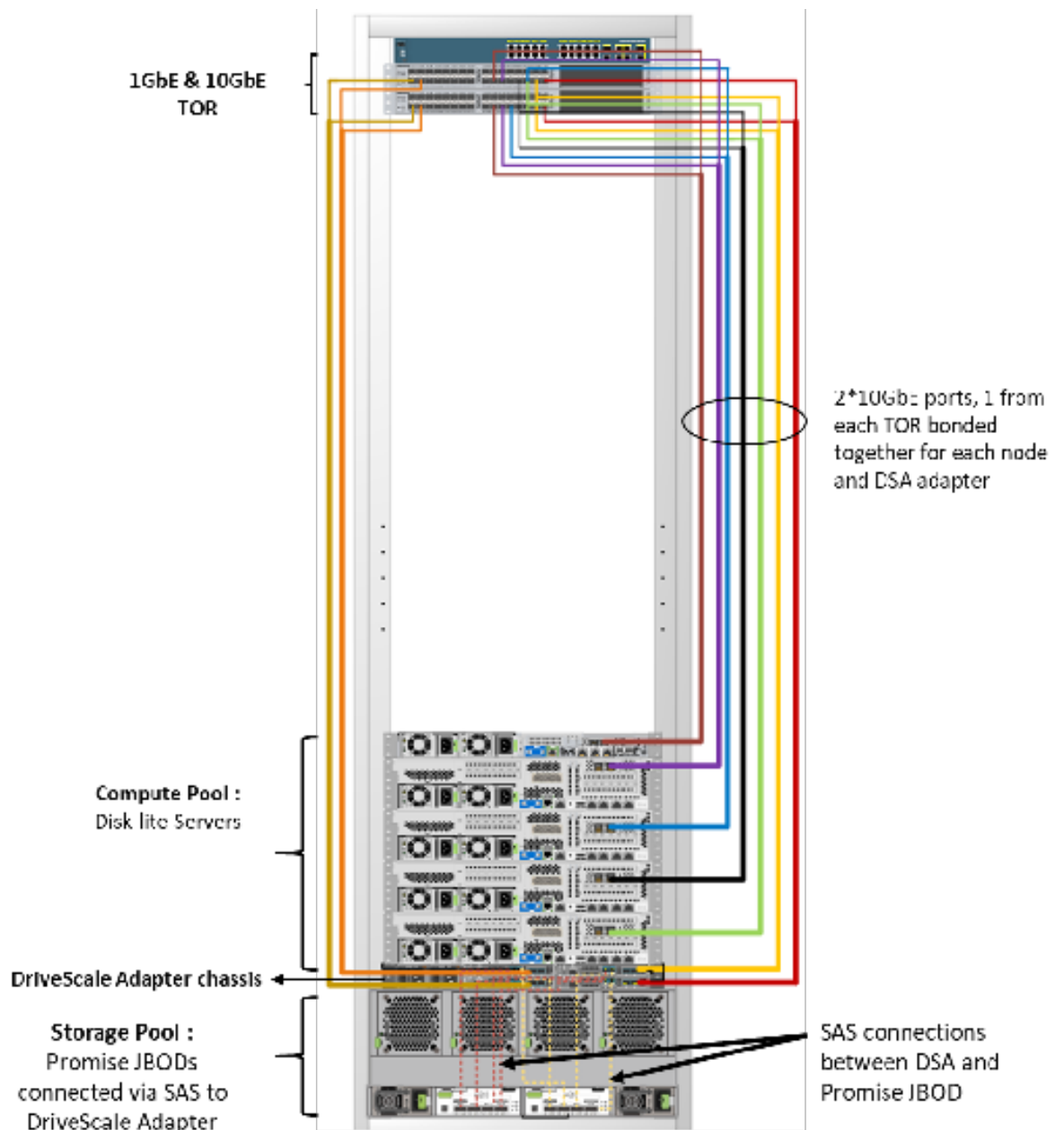


Figure 1: POC/Start up config

### Full Rack/Compute Rack Configuration

The design shown in this section builds on the first reference design shown above, to show what a full rack might look like. In this architecture, the aim is to achieve the highest density for servers by using the industry standard servers, Promise JBOD's with high bandwidth networking and storage. The design also takes into consideration the optimal ratio of drives to servers and the bandwidth to each drive, to build out a rack-level configuration that is optimized for the typical big data workloads. Scaling is achieved by simply replicating this rack-level configuration. Customers have the flexibility to spread the drives across the nodes or add more nodes to be used with the existing 96 drives or have less nodes with 96 drives in the JBOD's. Customers can work with DriveScale Solution Architects to optimally grow the rack and cluster.

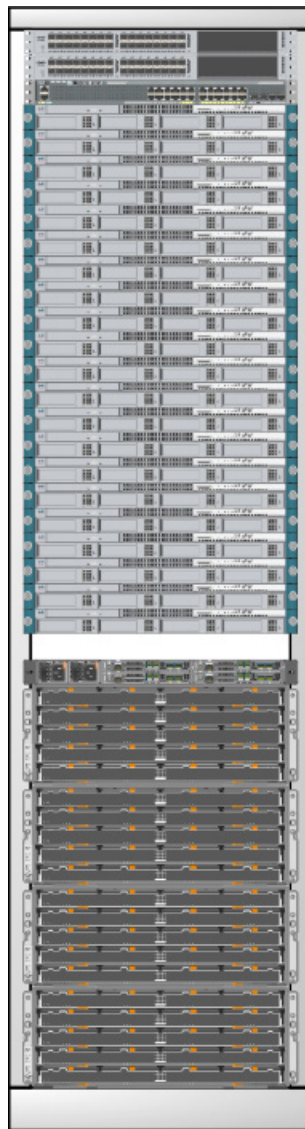


Figure 2: Full rack or Compute rack config

### Storage Rack Configuration

The design shown in this section builds on the first reference design shown above, to show what a full rack might look like. In this architecture, the aim is to achieve the highest density for servers by using the industry standard servers, Promise JBOD with high bandwidth networking and storage. The design also takes into consideration the optimal ratio of drives to servers and the bandwidth to each drive, to build out a rack-level configuration that is optimized for the typical big data workloads. Scaling is achieved by simply replicating this rack-level configuration. Customers have the flexibility to spread the 144 drives across the 13 nodes. Customers can work with DriveScale Solution Architects to optimally grow the rack and cluster.

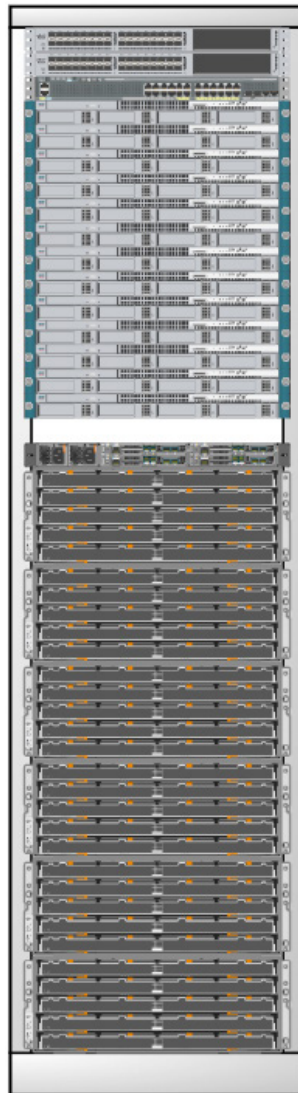


Figure 3: Storage rack config

### Rack Scalability

This design also demonstrates that one can start small and add components as needed, to scale the system. Customers can grow their compute and storage depending on the application needs. If the application demands more storage then they add Promise JBOD's to the cluster and if the application demands more compute, then they can add industry standard servers. They can scale from few nodes and Promise JBOD's to thousands of nodes and Promise JBOD's.

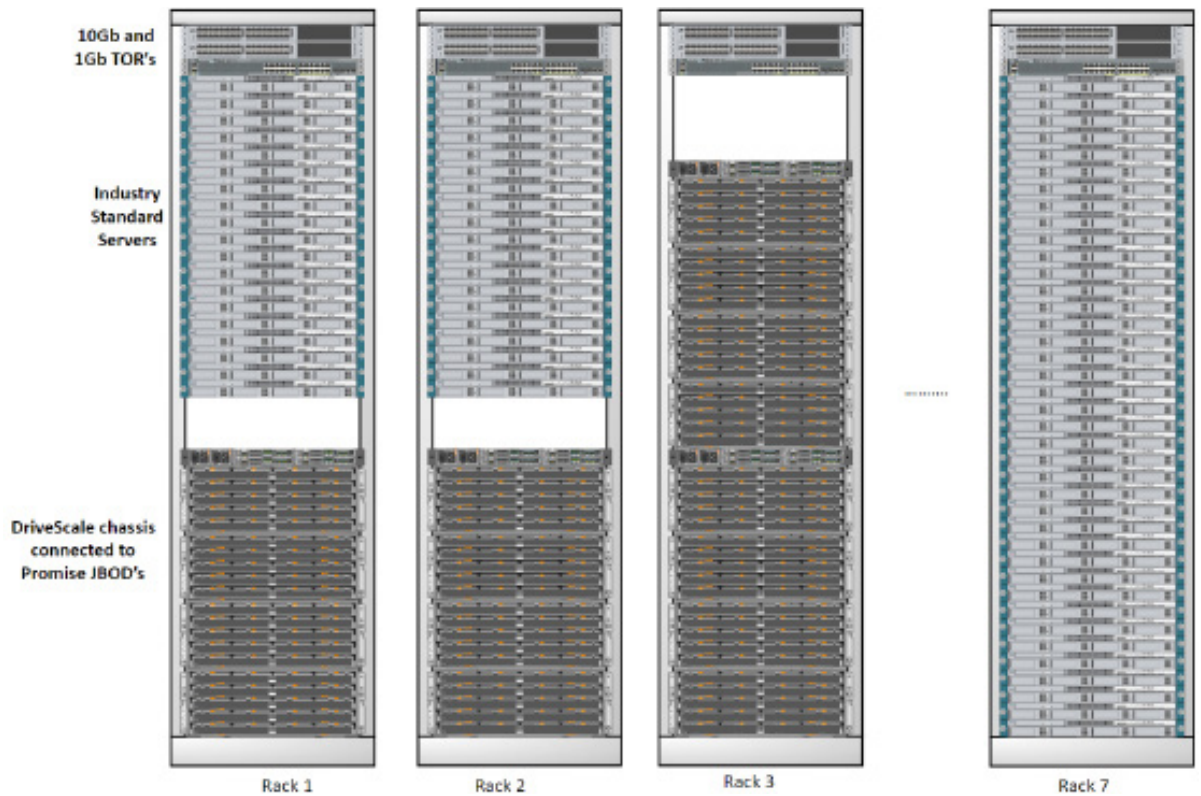


Figure 3: Rack scalability

### Summary

In all the above reference designs, DriveScale Adapter chassis, Promise JBOD's and industry standard servers JBOD's are co-located in the same rack. This ensures that drives are always only one Ethernet switch hop away from the servers that they are attached to, ensuring optimal performance. DriveScale's solution is designed to maintain the data locality that Big Data applications such as Hadoop requires and provide performance to disk that is equivalent to direct-attached drives.

Therefore, so long as we design the systems with no more latency in the network than that produced from a single Ethernet switch, the systems will perform optimally.

Given the software composable nature of the Promise/DriveScale solution, one can scale as needed on either the compute elements or the storage elements very easily, thereby adding capacity as required. This gives the solution a very flexible architecture, not just in terms of composability but also in



terms of capacity scaling. Customers can consult with the DriveScale Solution Architects to optimally design and deploy the modern workloads including Big Data hardware infrastructures.

### Conclusion

The DriveScale/Promise reference architecture described in this document will run any standard modern workload such as Big Data workload deployment including Cloudera, Hortonworks, MAPR and Apache Hadoop. DriveScale solution architects are available to help customers design their infrastructure for optimal cost and performance.

With DriveScale, customers can now operate their on-prem datacenter with the flexibility of a cloud environment while spending less on infrastructure and operations. The architecture also integrates quickly and easily into the existing environment with no additional changes to the application stack.



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

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