An Architecture Designed for Fast Data Solutions

Kodiak Data Delivers Memory Speeds

Just having the fastest NVMe memory and 100GbE interconnect is not enough: you need an architecture which can take advantage of this cutting edge performance. Kodiak Data's patent-pending technology brings the full capabilities of all levels of hardware to big data in the cloud, for an environment that runs up to 10X faster than anything else! And this technology can run on your existing hardware, achieving higher levels of performance on the systems you already own.

The key is the message-based fabric, which creates a Virtual Cluster Infrastructure (VCI) fabric of processors, memory and I/O devices, giving each and every application the full performance of the entire environment. And performance will grow more than linearly as you non-disruptively add new capabilities.

Run Your Software. No Change. Period. On Your Hardware Or Ours.

You have total flexibility in how you can implement Kodiak Data in your environment. Whether it is Oracle RDBMS, Microsoft SQL Server, Hadoop, Spark, Cassandra or any other application, use your existing servers and storage with Kodiak Data VCI, and get TeraByte per second bandwidth and sub-millisecond latency. Move your apps to a hyper-converged Kodiak Data SaaS Appliance on-prem for even higher performance. Or move the processing to a Kodiak Data MemCloud hosted environment to realize all of the advantages of a public cloud at these same incredible speeds.

Deliver Fast Data Clusters in Minutes

Today's fast data analytical demands require cluster infrastructure to be delivered quickly and cost effectively to analyze the torrents of data flooding enterprises. Analysts and developers can no longer afford to wait weeks or months for a cluster to be online. Kodiak Data VCI gets you going in minutes!

Total Flexibility In Configuration
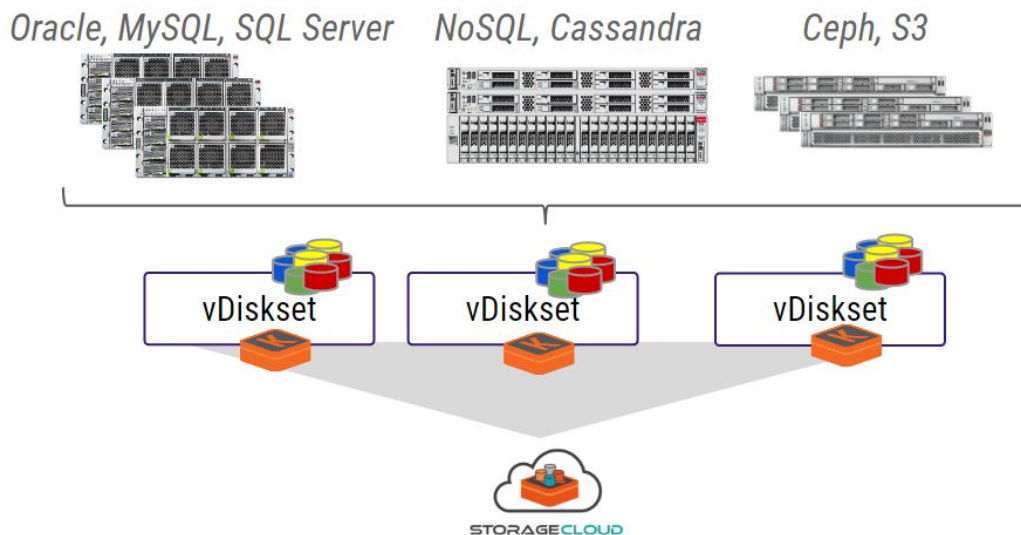
Keep your current database and application servers! Use Kodiak Data VCI StorageCloud™ as embedded storage, a vNAS (NFS or iSCSI), or object-based storage environment, at up to 1TByte/second bandwidth and 125µsecond latency.

How Do We Do That?

Kodiak Data's Virtual Cluster Interconnect technology consists of two core parts:

VCI Fabric Layer	VCI Composer
A pioneering fabric that blends compute, storage and networking into a uniform cloud layer.	A structured and simple way to build a policy driven virtual cluster including compute, storage, network and roles. The compute can either be virtual machines or bare metal. The Composer method is 100x more compact than scripting.

VCI Fabric Layer

Virtual Cluster Interconnect (VCI) creates a distributed data-plane fabric. VCI disaggregates storage into memory based objects, or tiles, and forms a data-swarm of potentially millions of these tiles with internet style routing between them. All I/O transactions are performed by those objects. All of the servers, storage and network controllers, storage and network paths, and physical storage devices, are available in parallel for each and every application instance.
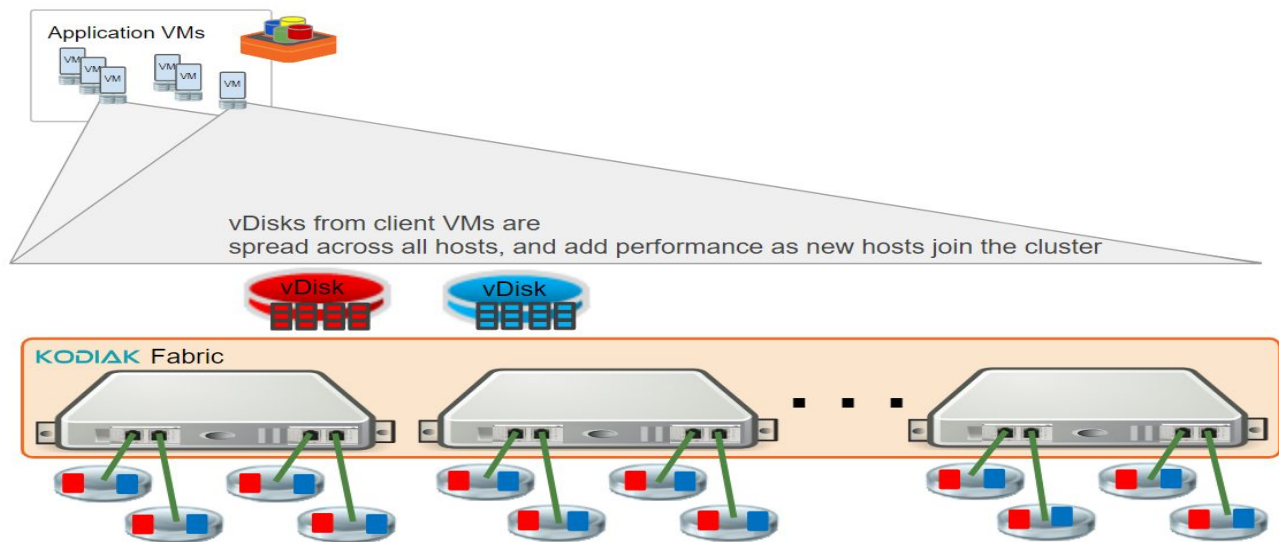
Every application instance distributes its I/Os across the servers, storage controllers, data paths and drives of the entire environment, not just the physical server it happens to be running on. Every application instance has the bandwidth and low latency of the entire fabric. As the environment grows transparently, the speed of every volume grows with it.

VCI gives you extreme availability as well. By using user-selectable erasure coding redundancy, vDisks continue to serve data at full speed even if individual drives or servers fail. And the rebuild time of failed storage is reduced dramatically over RAID designs, by distributing the rebuild across the servers, storage controllers, data paths and physical drives, in the same way that I/O is distributed across all of those components.

VCI Composer

Working with the VCI Fabric, the VCI Composer enables the entire environment to be managed as a single cluster. DataContainer XML manifests specify attributes like the application to be run, the processors, memory and storage capacity required which are drawn from the cluster as a whole, and the policies which control these resources. This enables one-click deployment of very large application sets across very large fabrics, greatly simplifying the management of your environment.

Kodiak Data VCI delivers both massive parallelization of transactions and simplicity in deployment.

TeraByte RAC on Kodiak (TRACK)

Kodiak Data's TeraByte RAC on our hyper-converged reference architecture has achieved simply stunning I/O rates.
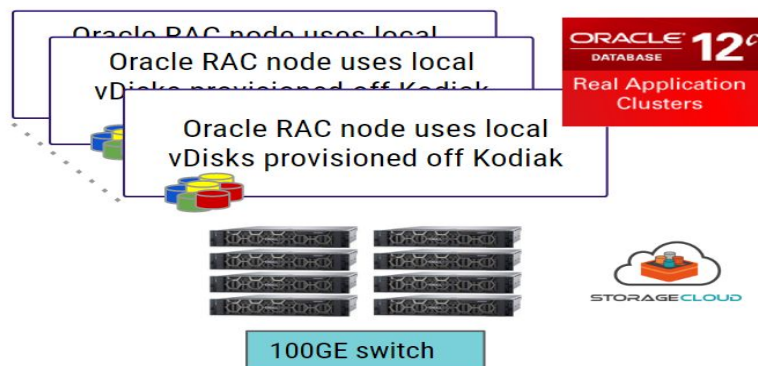
Using dual-socket 18 core Intel servers, 384GBytes of memory and 4x100GbE networking, with ASM Oracle disk groups of 10 NVMe SSDs per server, Kodiak Data has demonstrated:

	Oracle bandwidth	Oracle latency
8-node cluster	¼ TBytes/sec	125μsecond
32-node cluster	1 TByte/sec	125μsecond

Microsoft SQL Server 2017

Microsoft SQL Server 2017 is fully supported on Kodiak Data VCI. Whether you use Red Hat Enterprise Linux 7.3 and above, SUSE Enterprise Linux V12 SP2 and above, or Ubuntu 16.04 and above, you can run Microsoft SQL Server. And it is fully supported with Docker running any of the above Linux distributions.

Now you can use your current Microsoft tools such as Visual Studio, SQL Server Management Studio and SQL Server Data Tools for Visual Studio, in the highest performing cloud environment.

High Ingest DVR

All vendors happily quote read performance, but many customers face limitations in the write performance which are actually achieved by most configurations.

Using dual-socket 8 core Intel servers, 32GBytes of memory and 4x100GbE networking, with 25 NVMe SSDs per server, Kodiak Data has demonstrated 25 GBytes/second sustained 100% write performance per server.

Note that at these levels, all existing SSDs have only a 7- to 9-month expected lifetime due to the continuous writes. But Kodiak Data's high availability storage environment allows the replacement of these SSDs without impacting either performance or availability.

All With Full Access to S3 Cloud Data

And the speed doesn't stop when you need to access public cloud data. Using Kodiak Data EdgeCache™, you can connect your S3 compliant public cloud data to your on-prem environment, at full speed. For example, in an Oracle environment, you can use EdgeCache to store your RMAN backups at the speed and latency of your local datacenter, and have them transferred transparently in the background to a public cloud for long term storage.
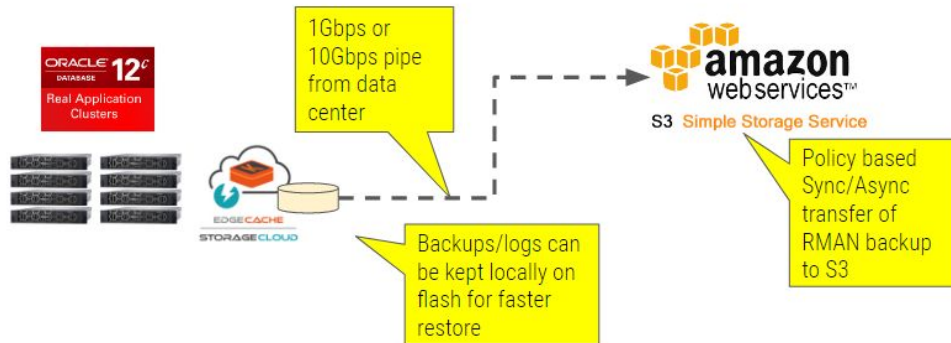
Either way, you can then use EdgeCache™ to dramatically reduce backup and restore times, and provide higher availability by storing your backups in a public cloud.

About Kodiak Data

Kodiak Data is a leading cluster virtualization technology company that allows customers to easily deploy and scale Big Data infrastructure in both public and private clouds. The Kodiak Data Virtual Cluster Infrastructure (VCI) platform is the only solution than can create, within minutes, code-ready virtual clusters that run at memory-speed and scale to the needs of big data applications. For more information about Kodiak Data, visit www.kodiakdata.com.