

# Swarm Data Protection

---

Databases and big data analytics are extremely I/O-intensive. Currently, these data-heavy applications are dependent on the performance and failure risks of dedicated physical disks.

With Kodiak Data's Virtual Cluster Infrastructure (VCI), all compute, RAM, networking and disk resources are shared across multiple clusters. HDDs and SSDs are shared and virtualized as virtual disks (vDisks), which allows for full I/O capacity to applications, as well as increased CPU efficiency.

## Kodiak Data Fabric and Tile System

Kodiak Data's patent-pending IP-based fabric technology revolutionizes the efficiency of resource sharing by replacing complex file systems and RAID controllers with a highly scalable Intelligent Tile system. Each tile is located on a Fabric server and represents 1GB of a vDisk. Tiles can use a mix of underlying media, and "migrate" without impact to the application. Multiple tiles can form Tile Swarms for a particular virtual cluster.

## Data Protection via N+1 Tile Sets

The most common cause of application downtime and data loss is HDD or SSD failure. Erasure codes that protect against disk failure can be enabled for any cluster of vDisks using a feature called "Swarm Data Protection."

### 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 that 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](http://www.kodiakdata.com).

Tile sets can come together, creating a protection group, and ensure that all data is striped and protected using a 4+1 scheme similar to RAID-5. Any disk can fail with zero data loss. Recovery from fault is 10X faster because the data is restored from N tiles to N tiles and all physical disks are used in parallel. The application is unaware of disk failure.

### Benefits

#### Balance Load Across Disks

- No hot spots
- No parity or spare disk

**Failure Without Data Loss**—Can schedule replacement.

**Rapid Rebuild of Kodiak Data Tile Sets/Swarms**—vDisk replacement is fast, simple and app-independent.

**High Performance**—resources of all disks are fully used

---

### KODIAK DATA, INC.

2570 W El Camino Real #500, Mountain View, CA 94040 | 650 383 8374  
info@kodiakdata.com | www.kodiakdata.com

© 2018 Kodiak Data, Inc. All rights reserved. All other trademarks are the property of their respective owners.