SFD15 Hedvig

*Angela Restani, VP of Marketing
https://www.hedvig.io

https://www.hedviginc.com/hubfs/Website_Resources/Hedvig-Brochure.pdf?
t=1508457985427

http://silvertonconsulting.com/blog/2016/07/23/hedvig-storage-system-docker-
support-data-protection-that-spans-data-centers/#sthash.IQirdJeP.dpbs

*Avinash Lakshman, CEO, https://www.linkedin.com/in/avinashlakshman/

Everyone's trying to modernise
Security is becoming more and more important
CapEx is easy, reducing OpEx is hard

Modernisation
Almost everyone wants to future proof - clear path to hybrid cloud
Location transparency (blurring the lines between on-premises and cloud)
Needs to be software based

Data Sovereignty
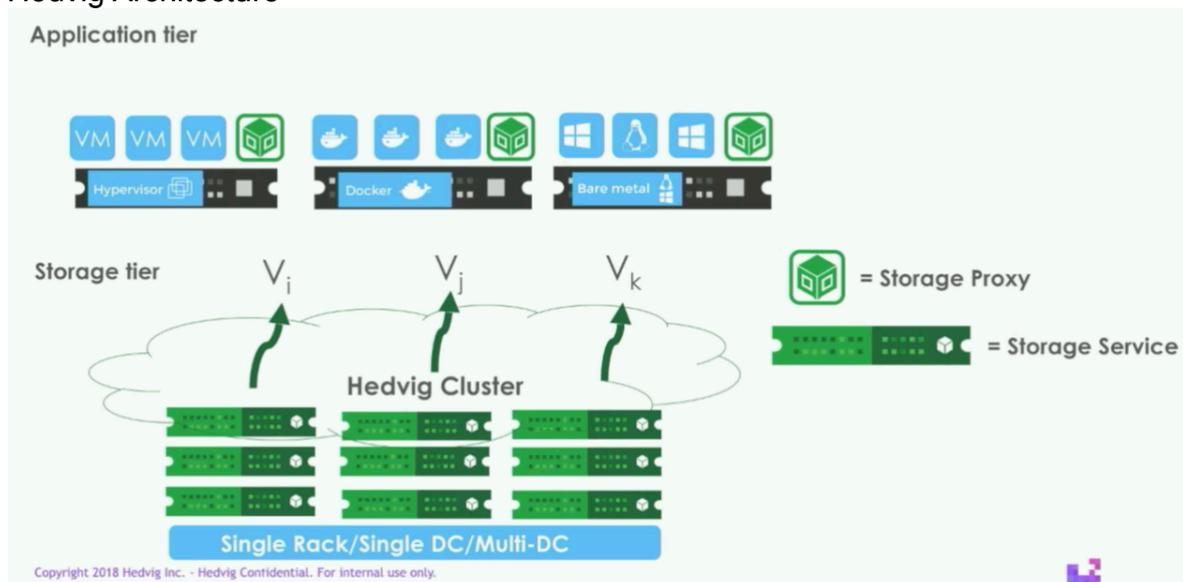Declaratively state that the volume shouldn't cross boundaries

OpEx
Want to take on 75 - 80% of primary and secondary data in the DC
"Most comprehensive data management solution the world has ever seen"

*Bharat Naik
Distributed Storage Platform

Hedvig Architecture

Hypervisor agnostic
Proxy on bare metal deployed as KVM instance
Each host requires a proxy

2 proxies per host (active / passive) for HA

Protocol consolidation on a single platform

Deduplication, Compression and encryption at a virtual disk level

Workload maps to a virtual disk

Every node has a proxy using a local address to expose iSCSI or NFS

Global deduplication (toggled on / off at a virtual disk level)

Default replication policy is "Agnostic" (let the system decide where to put the data)
Rack Aware is an option
DC Aware is also an option

Declarative data sovereignty

The same policies apply whatever protocol you want to use

Containers:
10-node cluster
vDisk: 1TB, 3, Agnostic
Ctr1 - Ctr64 - virtual disk spread all over the cluster
In this example, each container sits on 3 different nodes

Storage Pools:
Containers get assigned to storage pools
Containers get striped across 3 disks

Demarcation between metadata and data

Data Process:
1. Local data persistence
2. Replication

Metadata Process:
Global knowledge of everything happening in the cluster

Block-level deduplication, 4K granularity

Can integrate with external KMS infrastructure

Correctness is paramount in this system

Abhijith Shenoy - hybrid cloud demo

1. Tiering Layer
2. Stretched cluster

Suhani Gupta - demo of Kubernetes integration

Srividhya - converged backups demo
Data convergence - pri and sec data in a single distributed platform
Achieved through scheduled snapshots
Take snapshots in pri and go back to any snapshot in pri
What are VMware snapshots? Point in time copies of a VM's disk files
Primarily used for restoring a VM

Create VMware snapshot
Hedvig snapshot -> Each VM and its disks are treated as virtual disks at Hedvig
Consolidated VMware snapshot
Hedvig snapshot is a metadata only operation

"GitHub for Storage"
How are Hedvig snapshots implemented?
Every virtual disk has a notion of version
VirtualMachine <==> Code Repository
Snapshot <==> Branch
Hedvig Version Tree <==> Git Version Tree
Deletion of a Snapshot version <==> Merge of a branch
Clone from a snapshot version <==> Clone repository

Gaurav Yadav - Hedvig as a backup target
Backup via NFS

Shufan Ge (https://www.linkedin.com/in/shufan-ge-b8717975/) Object Storage
Each object is a virtual disk - deduplication and encryption can be performed at an object level?