2,800 VMs in 5 servers: Hyperconverged Cloud with CloudStack and StorPool

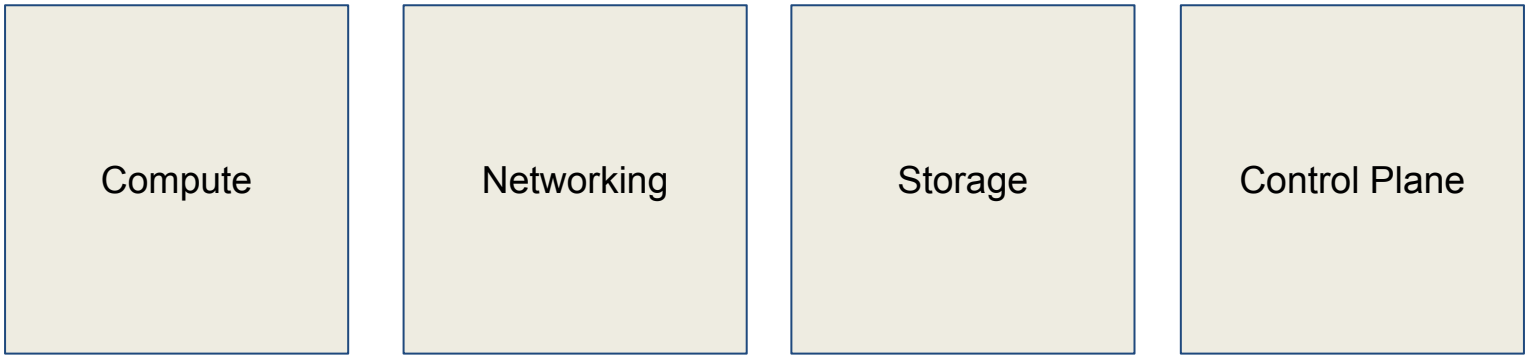
Simple, compact HCI infrastructure

Agenda

- The "standard" way to build a cloud
- Hyperconverged Infrastructure
- Beyond HCI
- How control plane, secondary storage and backups work in the StorPool HCI solution
- Some numbers

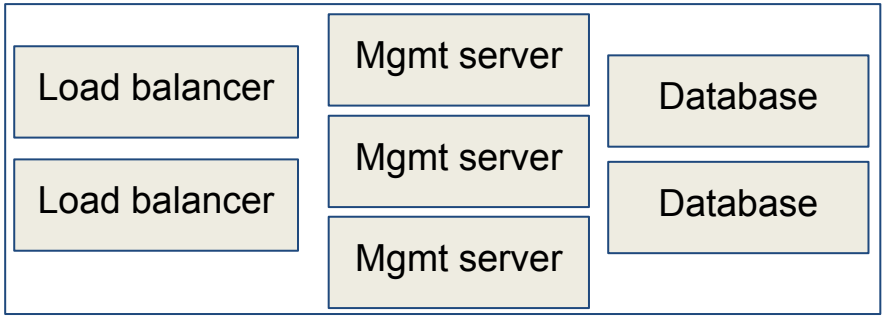
StorPool
DISTRIBUTED STORAGE

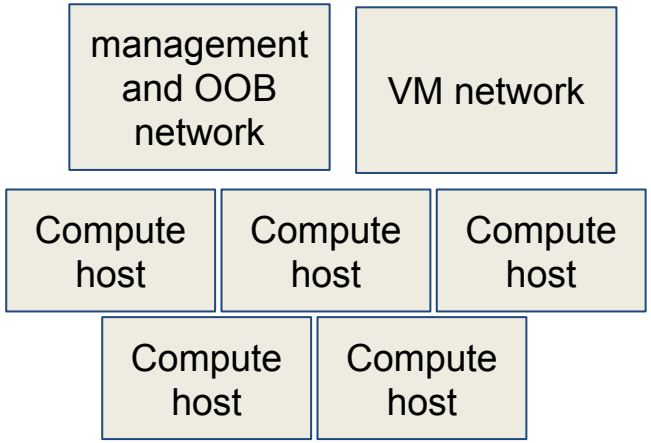
Cloud = 4 Core Components

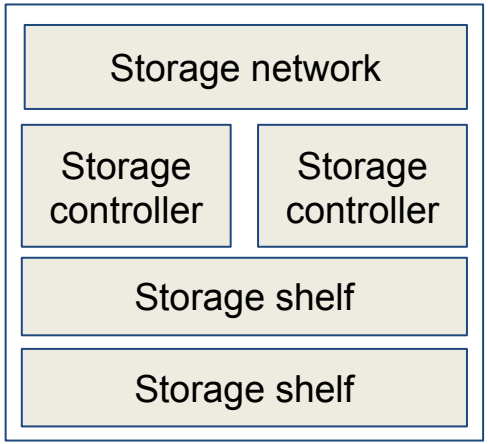


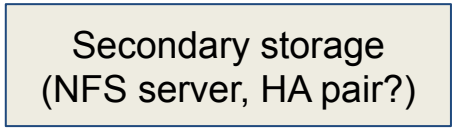
StorPool
DISTRIBUTED STORAGE

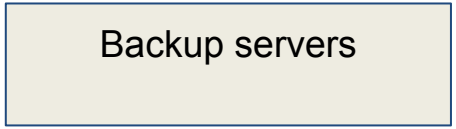
The Problem: The Heavy Cloud





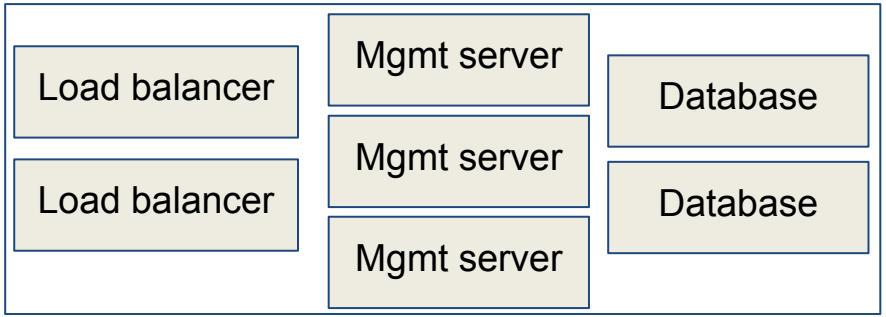


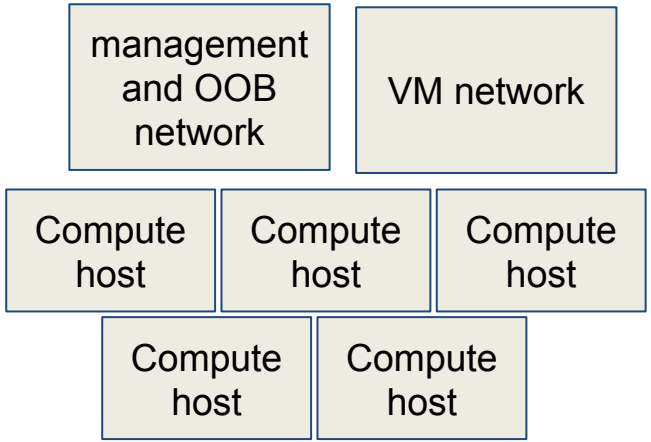


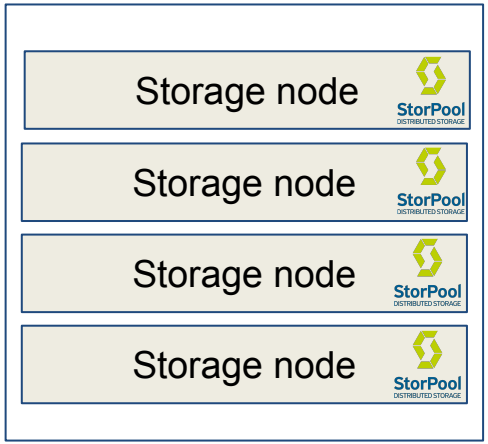


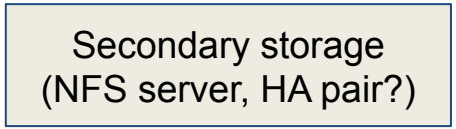
StorPool
DISTRIBUTED STORAGE

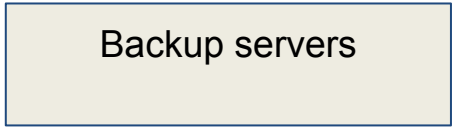
The Problem: The Heavy Cloud





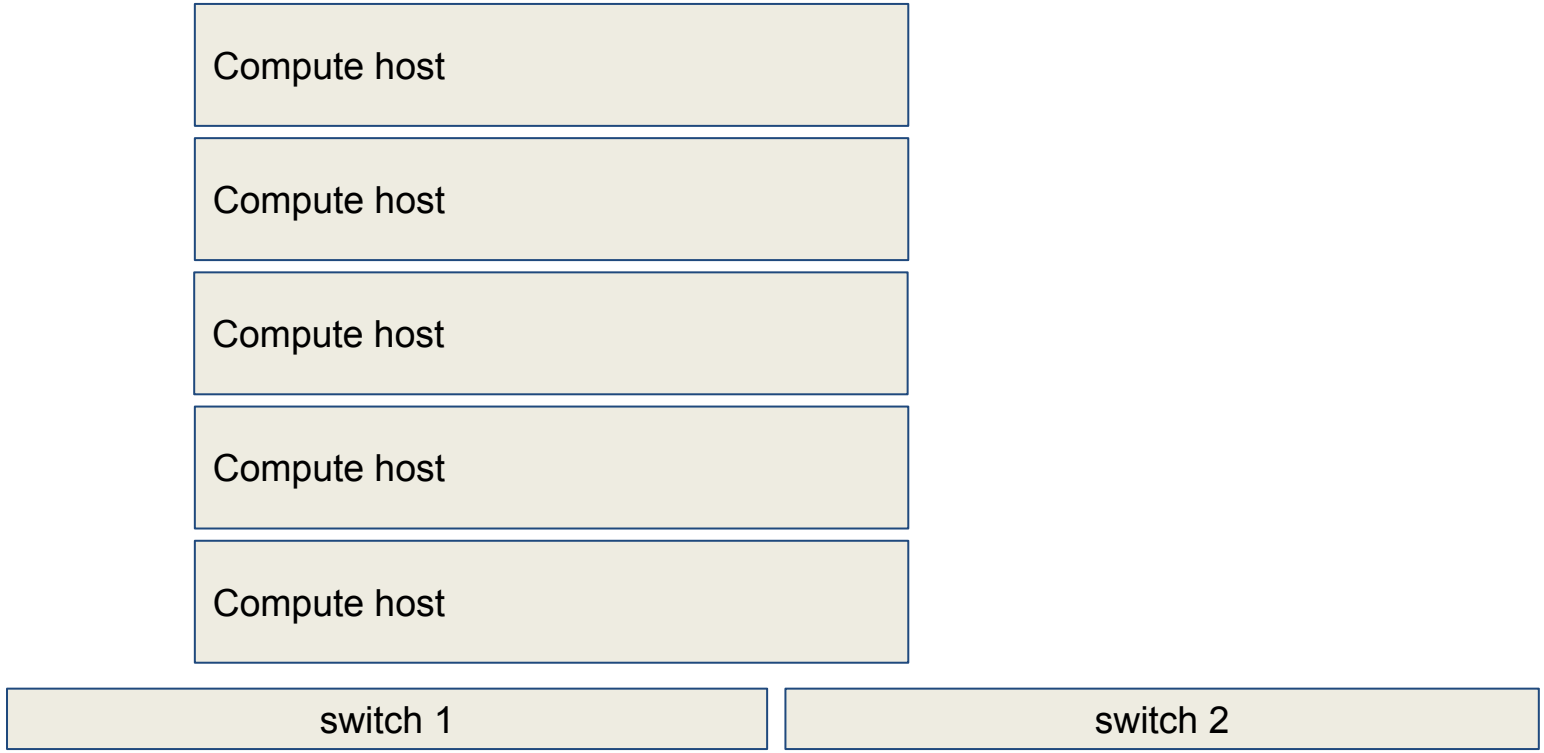






StorPool
DISTRIBUTED STORAGE

From 20+ components to 7 components

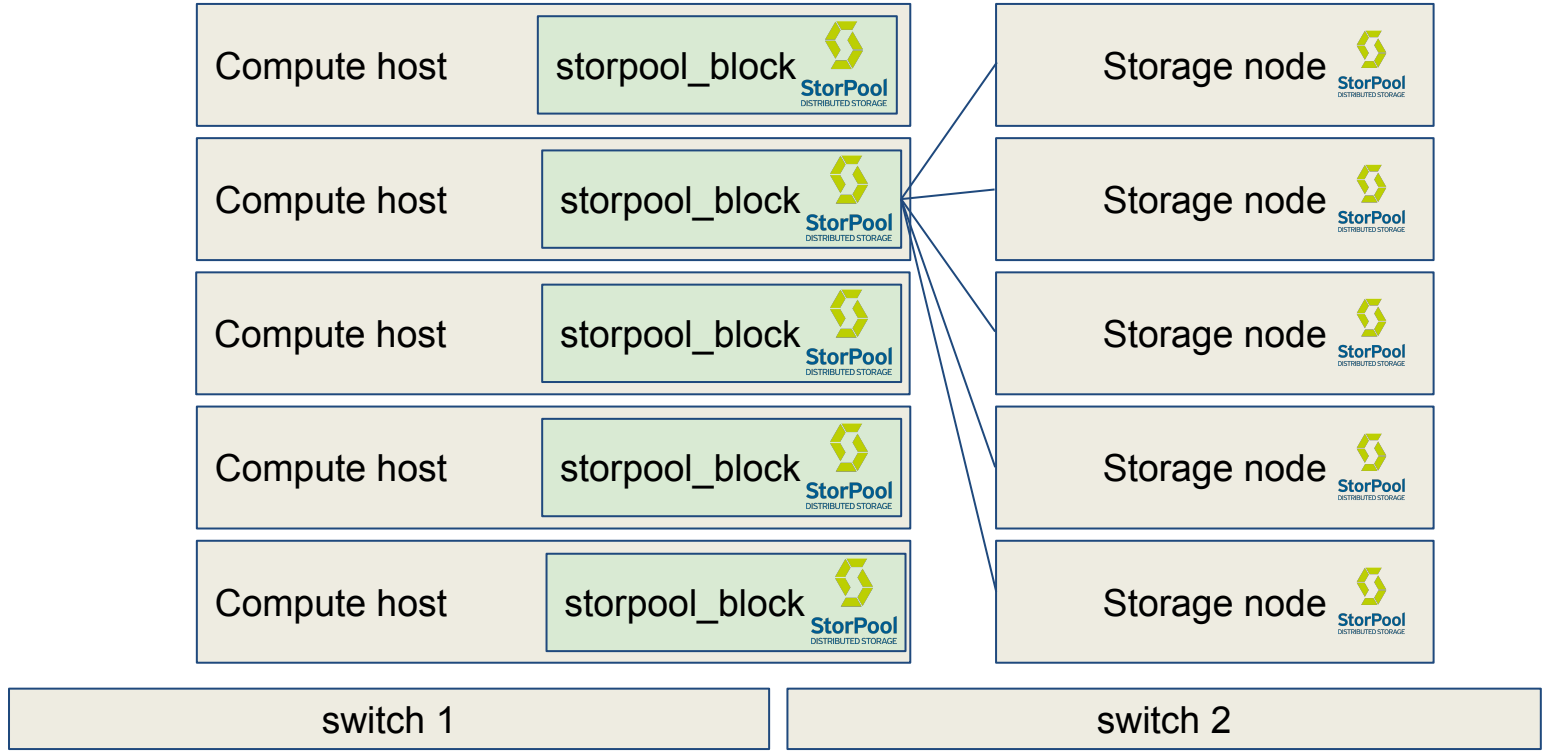


StorPool
DISTRIBUTED STORAGE

From 20+ components to 7 components

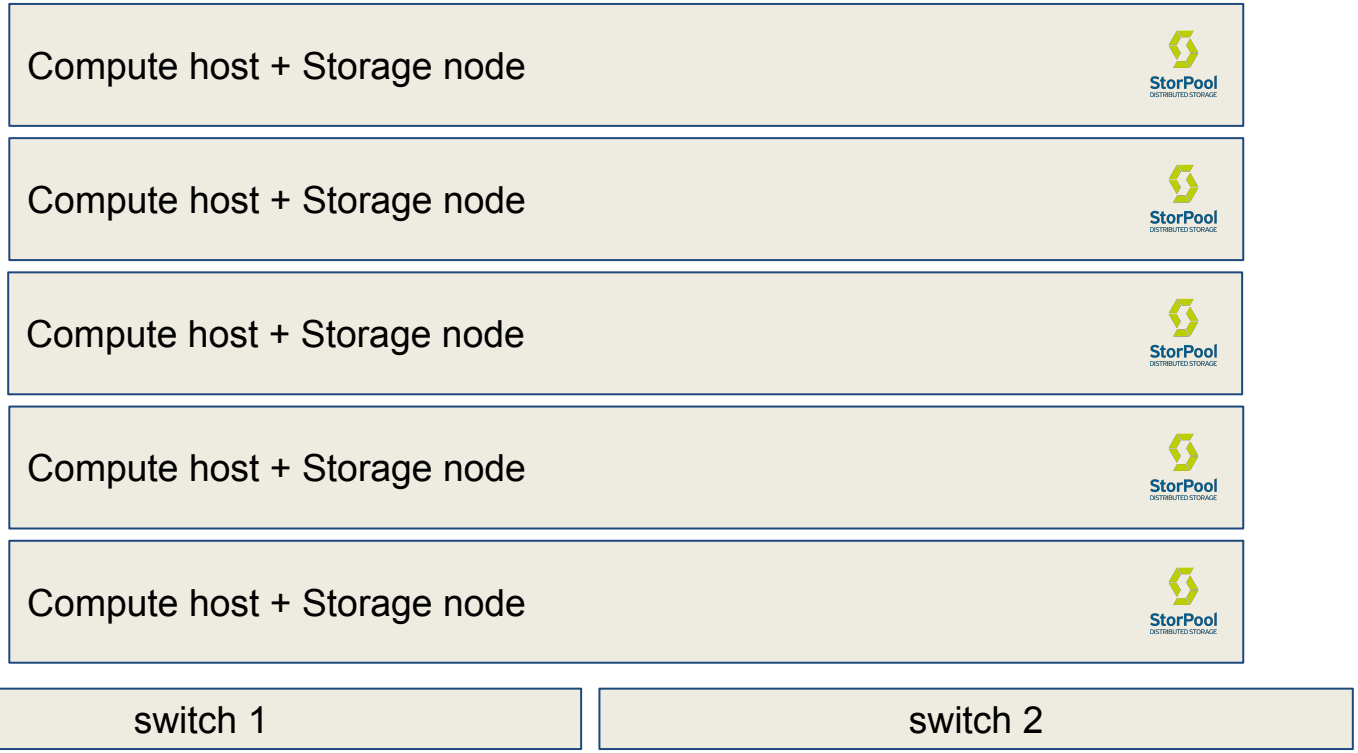
StorPool
DISTRIBUTED STORAGE

From 20+ components to 7 components



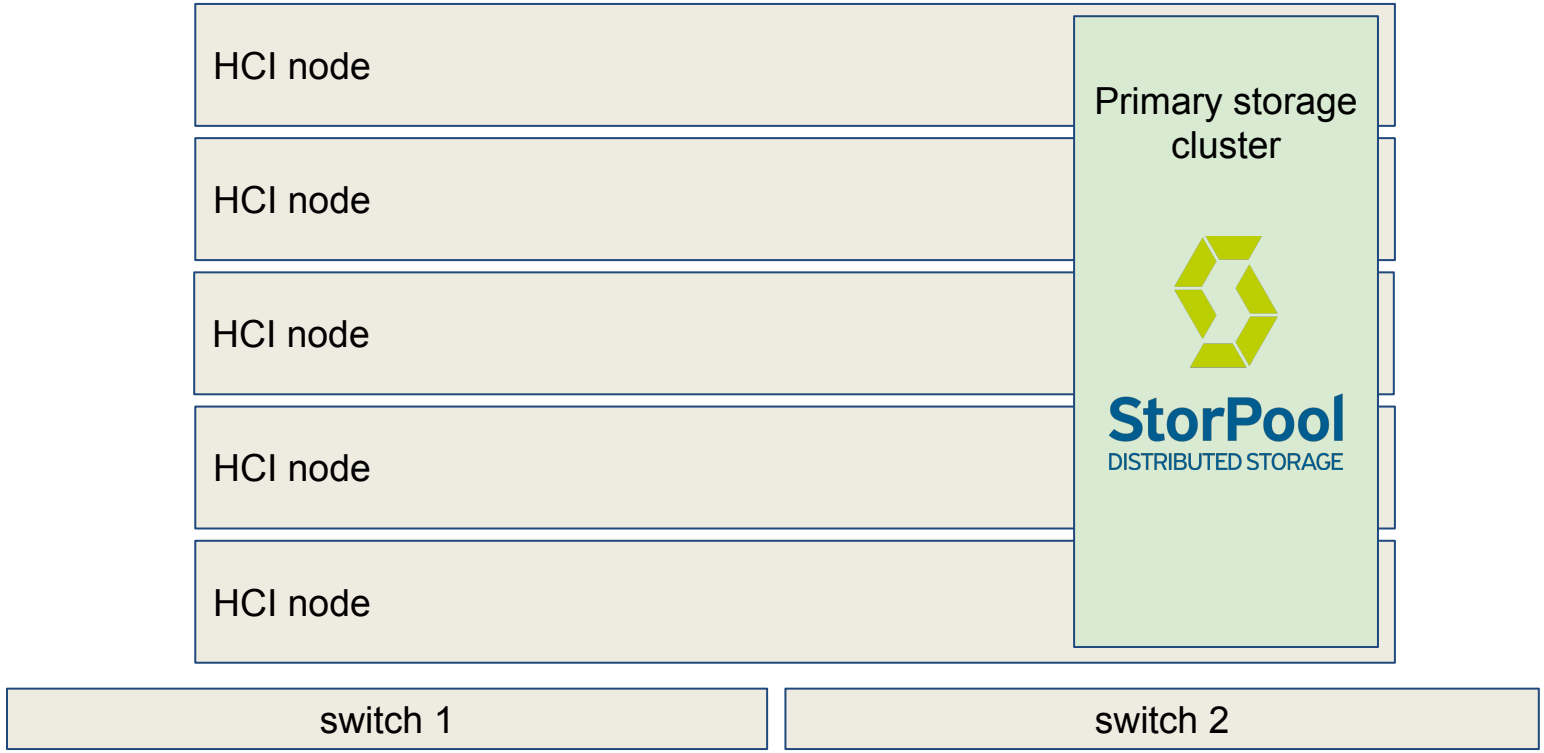
StorPool
DISTRIBUTED STORAGE

From 20+ components to 7 components



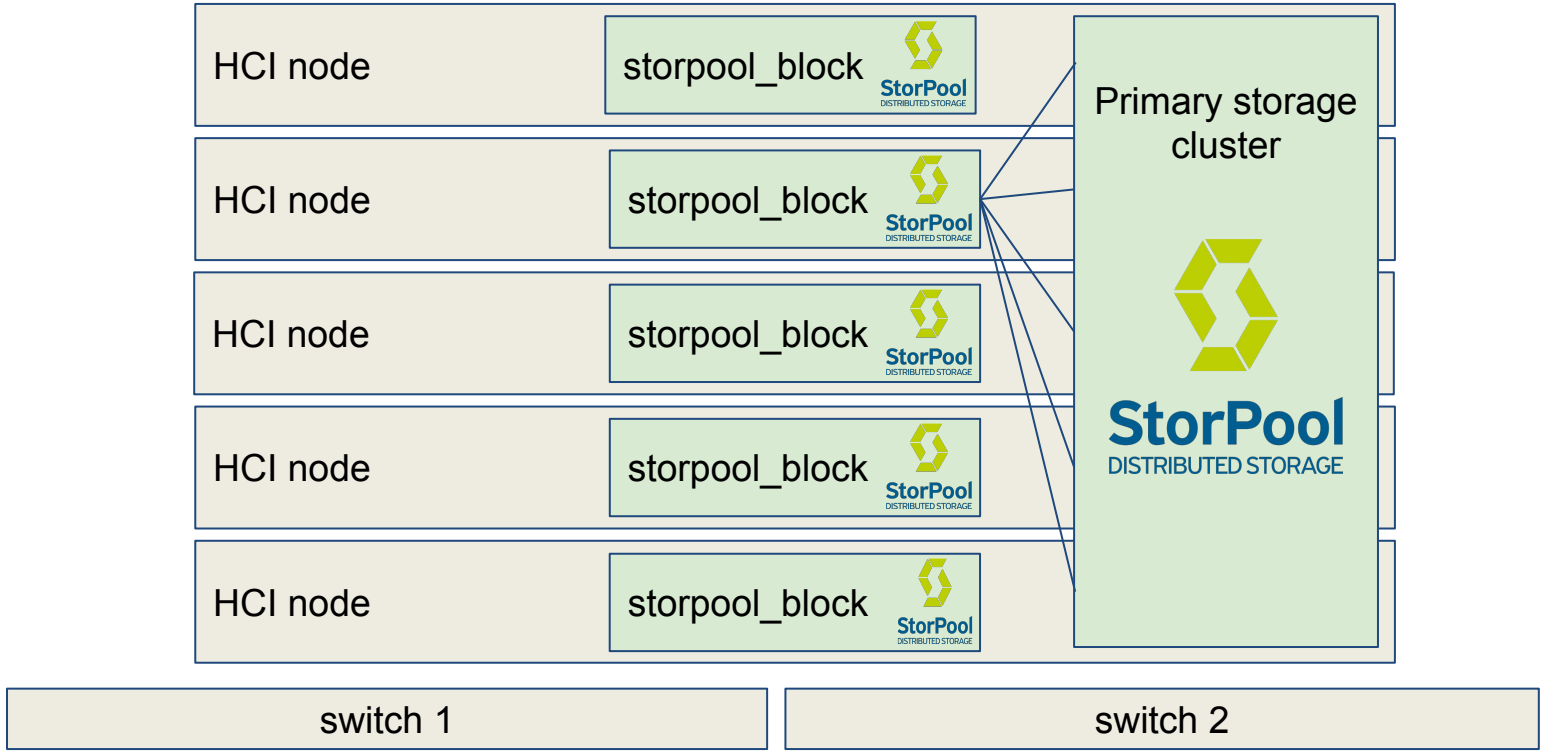
StorPool
DISTRIBUTED STORAGE

From 20+ components to 7 components



StorPool
DISTRIBUTED STORAGE

From 20+ components to 7 components



StorPool
DISTRIBUTED STORAGE

Hyperconverged Infrastructure Done Right

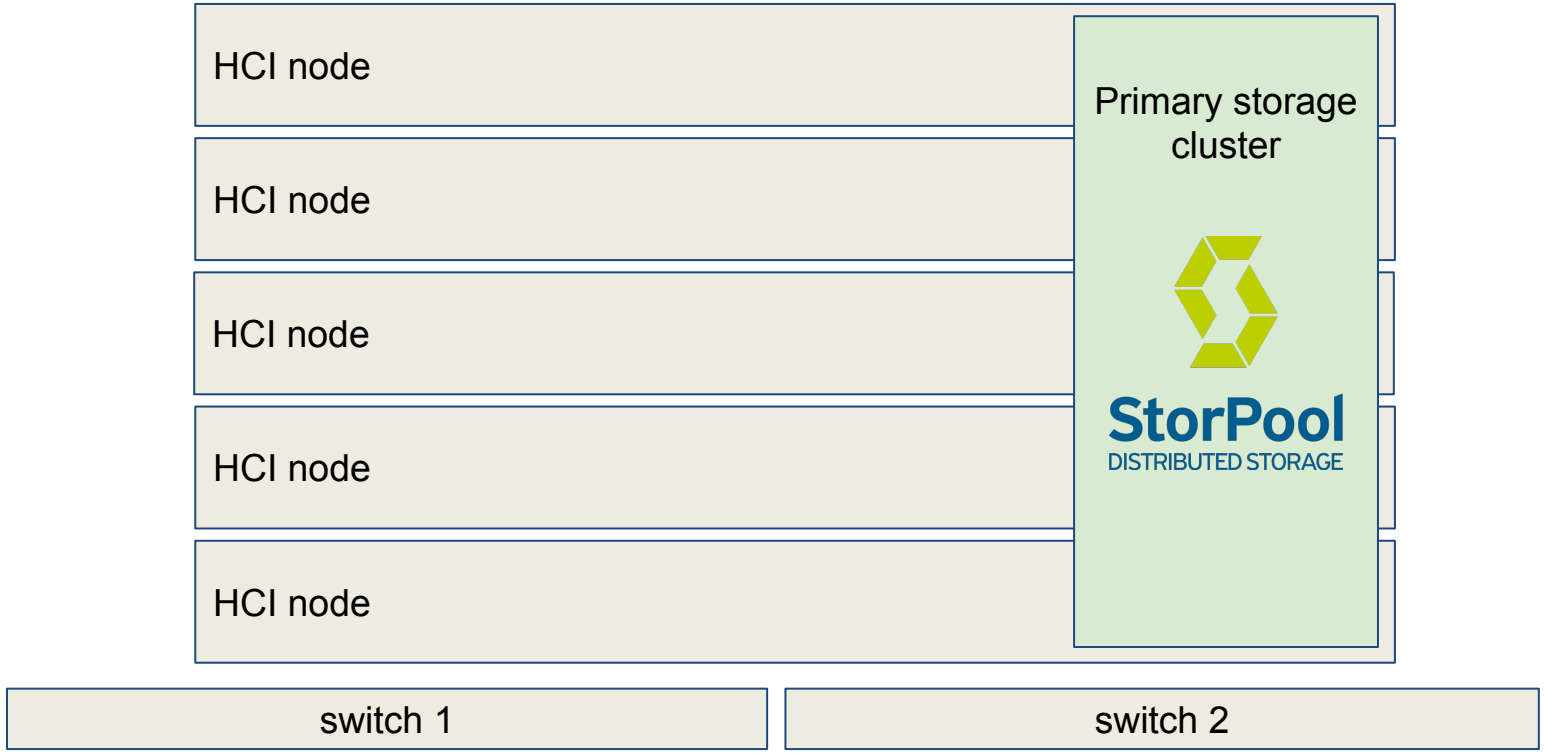
1. Primary storage runs on the compute hosts
    - no special storage network
    - no dedicated storage boxes
    - just a few NVMe drives, some CPU, some RAM
    - extremely efficient - a lot of performance from a little hardware

"Beyond HCI"

2. HCI StorPool systems have all the features and capabilities of stand-alone StorPool storage systems
3. Control Plane runs on the compute hosts. Simple HA VM.
4. Secondary storage (NFS) runs on the compute hosts. Simple HA VM.
5. On-site Backup services integrated in the primary storage

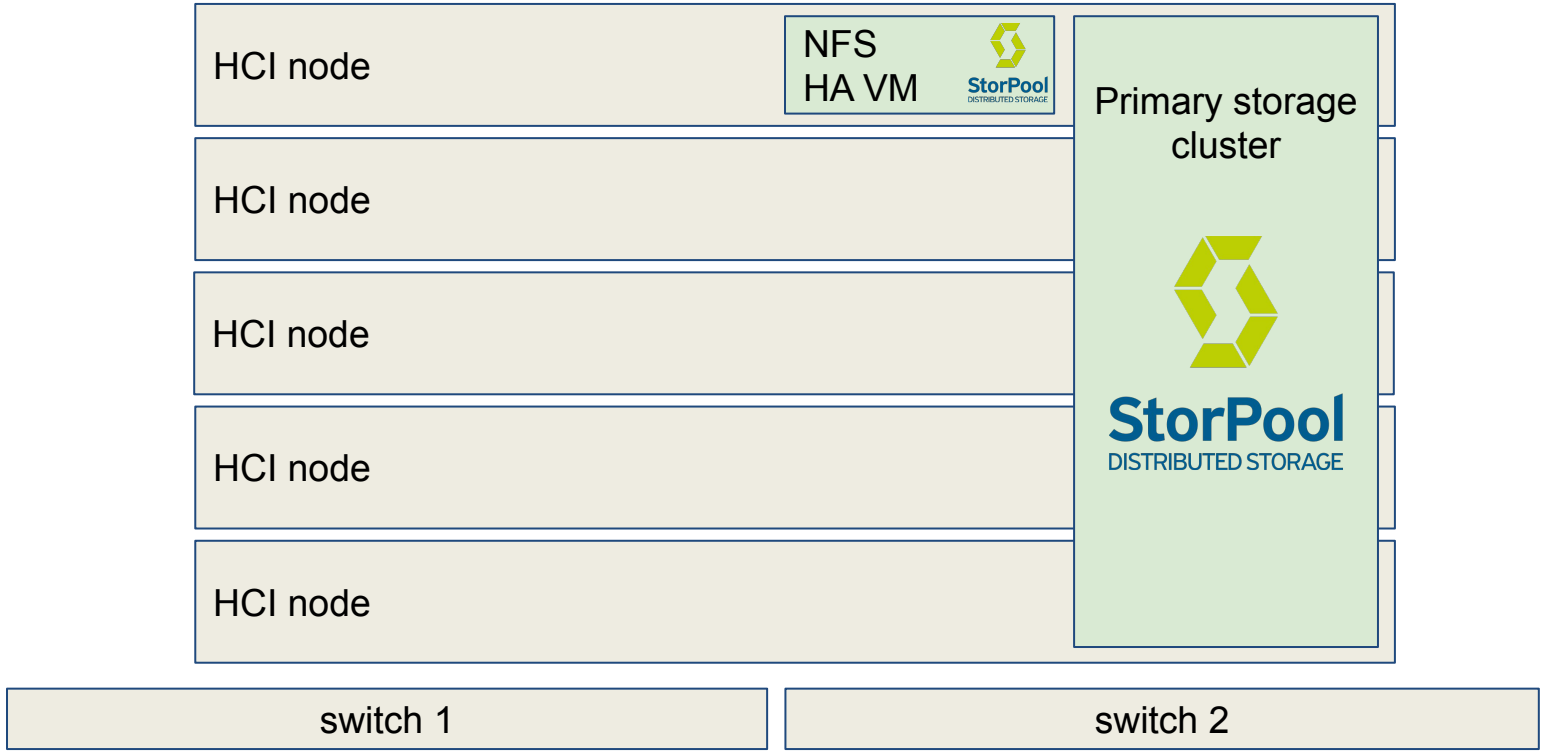
StorPool
DISTRIBUTED STORAGE

From 20+ components to 7 components



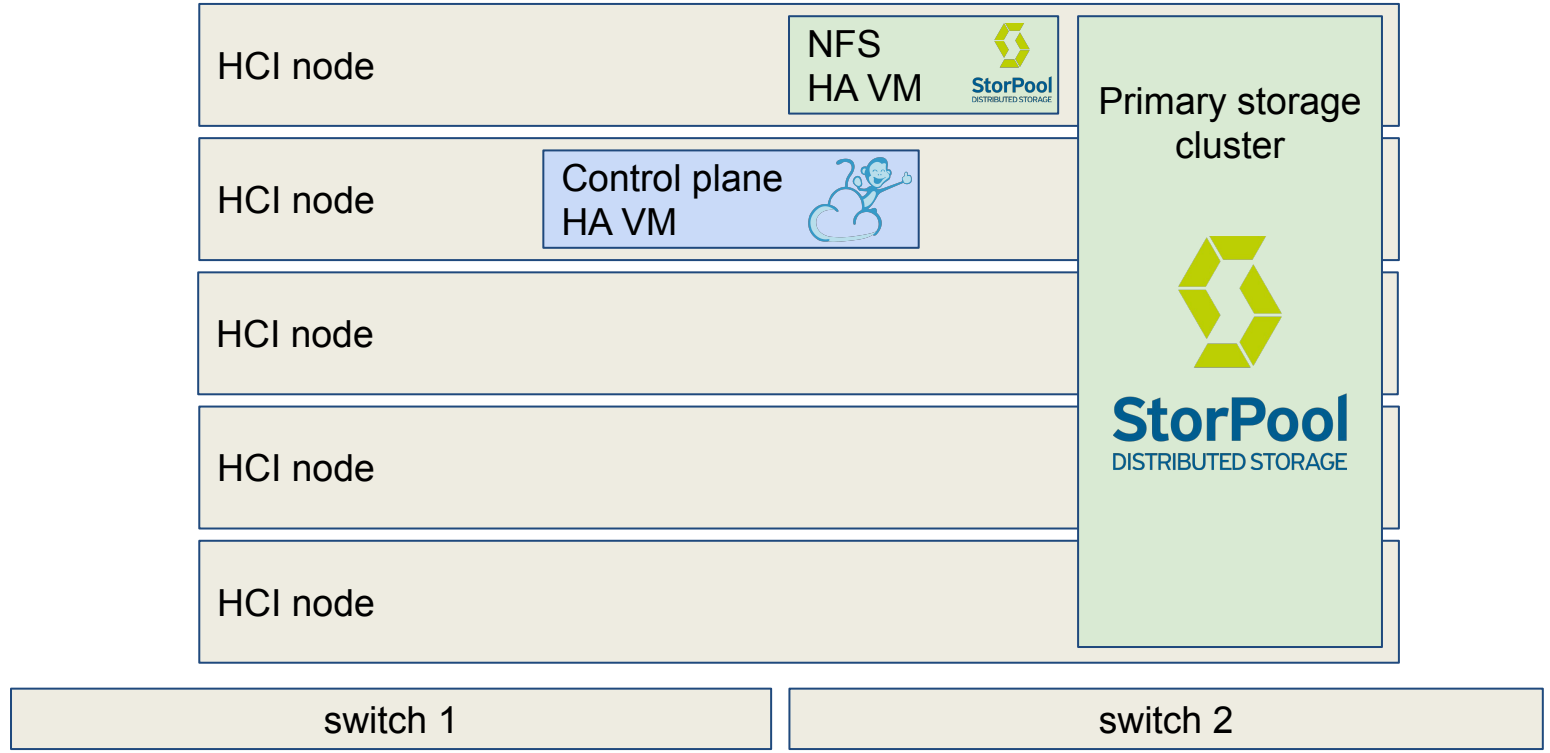
StorPool
DISTRIBUTED STORAGE

From 20+ components to 7 components



StorPool
DISTRIBUTED STORAGE

From 20+ components to 7 components



StorPool
DISTRIBUTED STORAGE

Hyperconverged Infrastructure Done Right

5. Resource allocation and isolation
6. Redundancy and HA schemes
7. Large environments are built with multiple reasonably-sized clusters
    - simplicity and maximum availability within each cluster
    - minimize dependencies between clusters
    - still permit workload migration and balancing

StorPool
DISTRIBUTED STORAGE

Control plane

Instead of 3 tiers for a redundant control plane, implement a simple "HA VM" redundancy scheme.

The StorPool cluster keeps the HA VM up.
No extra cluster software/configuration needed.

In a way, like popular undercloud/overcloud architectures of OpenStack, but much much simpler.

May have similar or better availability than the standard control plane redundancy scheme.

StorPool
DISTRIBUTED STORAGE

Great Secondary Storage solution

Most operations are instant

- create snapshot
- create volume from snapshot
- create volume from template (if template is cached)
- etc.

Small NFS footprint

Eliminate the NFS box(es)

StorPool
DISTRIBUTED STORAGE

On-site Backups

Instead of storing on-site backups on a dedicated backup server…

Use snapshots in the primary storage system with proven extremely high data retention rates.
Many snapshots per VM. No overhead for creating, deleting or keeping snapshots.

Users love instantaneous restores.

Separate from off-site backups and Disaster Recovery (!)

StorPool
DISTRIBUTED STORAGE

The numbers:

Different optimization goals - max density, **lowest TCO per delivered resource**, etc.

5x 2 RU servers each with

- 1.5 TB RAM - 6.0 TB total with N+1 redundancy
- 96 CPU cores - 960 hardware threads
- 4x 7.68 TB NVMe drives -- 70 TB usable in 5 nodes with StorPool (2+2 erasure coding)
- power density <300W per rack unit - easy to cool (cheap and efficient)

1450x VMs, each with

- 4 GB RAM
- 2 vCPUs
- 100 GB SSD-based virtual disk

StorPool
DISTRIBUTED STORAGE

The numbers:

1450x VMs, each with

- 4 GB RAM
- 2 vCPUs
- 100 GB SSD-based virtual disk

N+1 redundancy for RAM capacity. No oversubscription
3:1 CPU oversubscription
2:1 storage oversubscription (ratio of provisioned size to stored data)

Very high hardware utilization -- StorPool, Control Plane, Secondary Storage are all together just a few percent of the resources.

StorPool
DISTRIBUTED STORAGE

The numbers:

Possible paths to even better efficiency:

- Clusters with 11+ nodes - storage redundancy schemes, less overhead for redundancy of RAM capacity
- higher CPU oversubscription (depends on use-case)
- higher storage oversubscription with cloned provisioning (depends on use-case)
- RAM oversubscription with KSM, ballooning, zswap(?), oomd(?) - needs research and testing

Sufficiently advanced technology! The efficiencies stack up.
We do this with StorPool and KVM.

StorPool
DISTRIBUTED STORAGE

Thank you!

StorPool Storage

The best storage solution when building a Leading Cloud

www.storpool.com | info@storpool.com

cloud STACK COLLABORATION CONFERENCE

Fill in the CloudStack User Survey

Help us understand the CloudStack Ecosystem