

# Accelerating Apache Cassandra™

## Solution Overview

### Key Benefits

- Disaggregate physical resources to independently scale compute, storage and network resources
- Improve Performance by avoiding node rebuild operations
- Simplify operational challenges
- Deliver better application performance
- Instantly provision new nodes with storage resources
- Improve Cassandra performance

### Pavilion Benefits

- Latency and Performance of direct-attached SSDs, but with the operational benefits of shared storage
- Data resiliency & high availability
- Up to 20 active-active storage controllers
- Thin Provisioning
- Snapshots & Clones
- Independently scale for performance or capacity in the same 4U footprint
- Deploy less flash resources
- Up to 920 TB of NVMe flash storage in a single 4U appliance

## Deploy Low-Latency Rack-Scale shared storage in large-scale Cassandra deployments

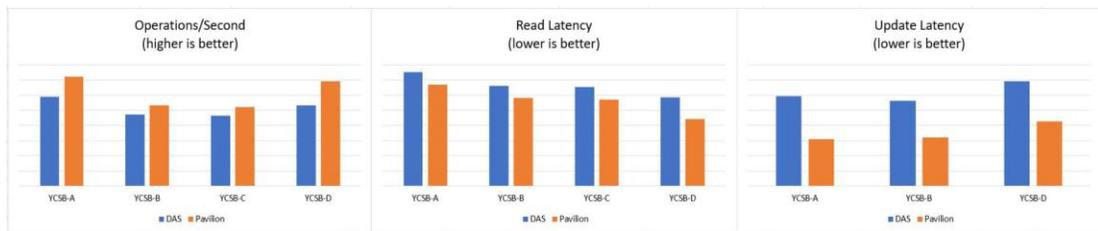
Cloud, big-data analytics, mobile, and cloud-delivered applications using Apache Cassandra are driving a new paradigm in IT infrastructure design. Resources need to be flexibly deployed so that ever-changing requirements can be satisfied on a day-by-day or even minute-by-minute basis. Compute, network, and storage resources all need to be able to scale independently to meet an ever-increasing and diverse set of application requirements.

This shift is creating a demand for disaggregated storage resources that can scale to new levels, where performance tuning can be eliminated so any workload can be satisfied at any time. In addition, rapid response to infrastructure changes or failures is also required to improve operational agility.

In order to successfully gain the benefits of Rack-Scale Design, several requirements need to be satisfied.

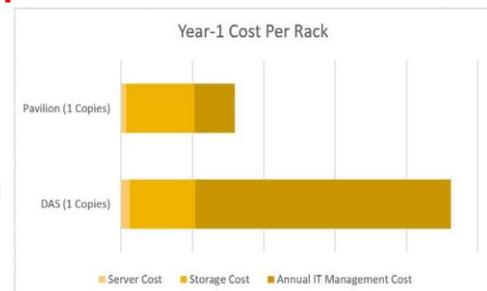
### Improve Cassandra Performance

The Pavilion array can offer 100s of Terabytes of low-latency logical flash storage from a disaggregated 4U storage appliance. Racks of Cassandra nodes can be supplied low-latency storage capacity from a central storage appliance that can deliver up to 120 GB/second bandwidth and 20 million 4K read IOPS. In addition, the Pavilion array offers important data management features that lower the cost of deploying large clustered applications significantly. As a result, it is now possible to get the same performance advantages of direct-attached SSDs with flexible shared storage. Below is a performance comparison of direct-attached SSDs and shared Pavilion storage using the Yahoo! Cloud Services Benchmark (YCSB). The results show that Pavilion will improve performance and provide lower latency in Cassandra deployments. This DAS-based configuration consisted of 16 2U servers, each with 12.8 TB of NVMe SSDs.



### Reduce Costs and Infrastructure Sprawl

The performance test above demonstrated that Pavilion enables reduction in the number of servers based upon improved performance per node. Eliminating DAS SSDs also allows 1U 'disk-less' servers to be deployed, which saves on rack space, power and cooling, which delivers capital and operational costs savings, as shown in this graph to the right.



## Reduce Raw Flash Capacity Requirements

With direct-attached storage (DAS), flash resources are under-utilized and stranded in individual servers across a cluster, resulting in up to 80% wasted flash capacity.

Pavilion allows administrators to decide how much storage to provision to each Cassandra node at deployment time as opposed to when the servers are ordered. With thin provisioning, you only use what the database actually requires at that time, and avoid over-provisioning of resources.

## Improve Resiliency and Reduce Overhead

By leveraging Pavilion Rack-Scale shared storage centralized data management features, it is now possible to dramatically reduce the response time required to rebuild Cassandra nodes or recover from failures.

Pavilion leverages RAID6 internally to eliminate the impact of individual SSD failures, by allowing applications to continue without interruption in the event of SSD failures.

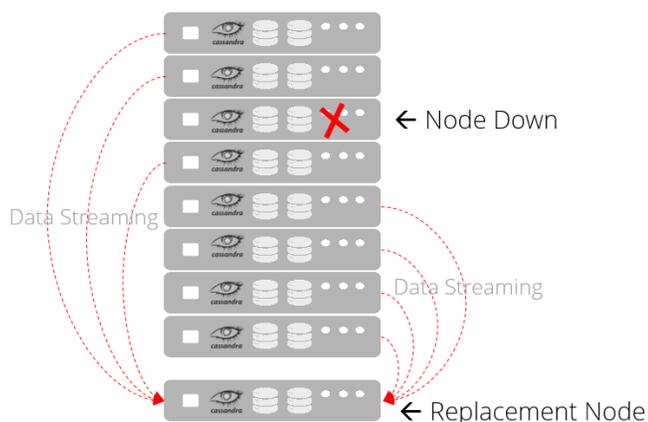
In addition, by leveraging Pavilion instant Clones, new nodes can be brought online without excessive data copying over the network, which is typically required to initialize a replacement node in the event of a failure. Using DAS, a failure of a node requires that the new node be brought online and initialized by copying data from other nodes over the network. This results in excessive network bandwidth usage that impacts response times, as well as a lengthy amount of time spent prior to the cluster becoming fully-operational and performant.

By leveraging Pavilion instant Clones, you can bring up a replacement node without copying all of the data over the network. The Pavilion system can present the node with a complete copy of the data instantly, which allows the cluster to become fully-operational in a minimal amount of time.

The diagrams below outline the differences related to reacting to building a new Cassandra node and bringing it online when using DAS vs. shared rack-scale flash storage from Pavilion:

### Hyper-Converged Infrastructure using DAS

- ✗ Node Loss Impacts Compute and Storage resources
- ✗ Replacement node requires a full data sync
- ✗ Impacts performance of other nodes
- ✗ Re-Sync consumes network bandwidth



### Pavilion Rack-Scale Storage Solution

- ✓ Use 'disk-less' nodes with no storage installed
- ✓ Minimize down time in the event of node failure
- ✓ Instant data availability on the new node after re-mount
- ✓ Avoid copying data over the network from other nodes
- ✓ Run repair to do small delta-resync instead of full copy

