SOLUTION BRIEF

THE WORLD'S FIRST TPCx-HCI BENCHMARK RESULTS USING VMWARE® HCI AND SUPERMICRO A+ SERVERS POWERED BY AMD EPYC™ 7003 SERIES PROCESSORS

Supermicro AS -1114S-WN10RT

TABLE OF CONTENTS

SUPERMICRO

Supermicro is a global leader in high performance, high efficiency server technology and innovation that develops and provides end-to-end green computing solutions to the datacenter, cloud computing, enterprise IT, big data, HPC, and embedded markets. Our Building Block Solutions® approach allows us to build and provide a broad range of SKUs that are optimized to individual customer needs and workloads.

Executive Summary

Optimum performance requires optimized hardware and software. TPCx-HCI is a powerful tool for measuring and optimizing HCI performance. This rigorous HCI benchmark exercises both hardware and software when measuring an HCI cluster's scheduling, load balancing, and overall performance. It also tests failover performance when the cluster suddenly loses a node. Supermicro is partnering with AMD and VMware to publish the world's first TPCx-HCI benchmark result using the VMware Hyper-Converged infrastructure (HCI) solution consisting of VMware vSphere® 7.0 Update 2 virtualization, VMware vCenter® management, and VMware vSAN™ 7 storage.

This test used a cluster of four single-socket Supermicro WIO A+ (model AS 1114S-WN10RT) servers powered by 64-core AMD EPYC™ 7713 processors and 1TB of main memory. Enabling

Simultaneous Multithreading (SMT) yielded 512 threads on the 256 processor cores. The VMs ran RHEL 7.7 and the PostgreSQL 10.6 Database Management System (DBMS). This solution achieved an industry-leading score of 4,790.18 tpsHCI @ $49.60 USD / tpsHCI. Supermicro WIO A+ servers are currently available for sale.

About the TPC Express Benchmark HCI (TPCx-HCI)

The Transaction Processing Performance Council (TPC) developed the TPC Express Benchmark HCI to measure HCI performance under a demanding database workload that stresses the virtualized hardware and software compute, storage, and networking resources.

TPCx-HCI has two unique characteristics:

- It has an elastic workload that varies the load delivered to each VM by as much as 16x while maintaining a constant load at the cluster level. Sustaining optimal throughput for this elastic workload on a multi-node HCI cluster typically benefits from frequent VM migrations to maintain load balancing across all nodes. TPCx-HCI measures VM migration efficiency and uniformity of data access from all nodes.
- The Data Accessibility test abruptly powers down a node while continuing the benchmark on the other nodes. The test sponsor (Supermicro) must include a throughput graph for this test that demonstrates both the performance impact and recovery time required to regain resilience.

Achieving good TPCx-HCI performance requires many optimizations, such as:

- Software-defined storage (e.g., vSAN)
- Live migration (e.g., VMware vSphere® vMotion®)
- Load balancing (e.g., VMware vSphere® Distributed Resource Scheduler™ (DRS))
- Hypervisor scheduler
- Compute performance
- Storage performance
- Networking performance

Testbed Configuration

Each tested Supermicro AS -114S-WN10RT server included two disk groups. The first disk group used one mixed-use group of Kioxia CM6 3.84TB NVMe devices for the cache layer. The second group used read-intensive Kioxia CD6 3.84TB NVMe devices for the capacity layer.

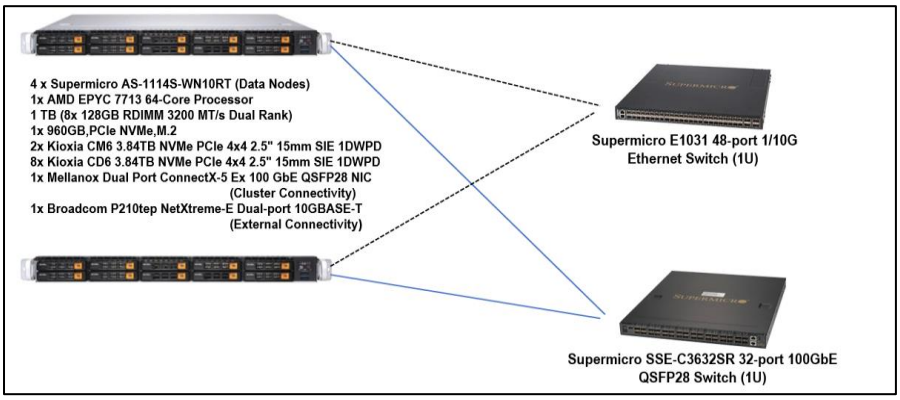

Figure 1 - Supermicro Cluster Configuration

The test cluster included a dual-port Broadcom BCM57416 NetXtreme-E 10GBASE-T RDMA Ethernet Controller with one port configured for external access. The vCenter Server Appliance accessed the nodes via the management network on the other port for the management network. Additionally, each server included a dual-port Mellanox ConnectX-5 EN 100GbE card with both ports connected to the 100GbE Switch, where:

- One port carried the vSAN traffic and the transactions coming from the driver.
- The other port carried the vMotion traffic.

vSphere 7.0 includes many performance optimizations relevant to this workload. The following two subsections describe how this result used some of these optimizations.

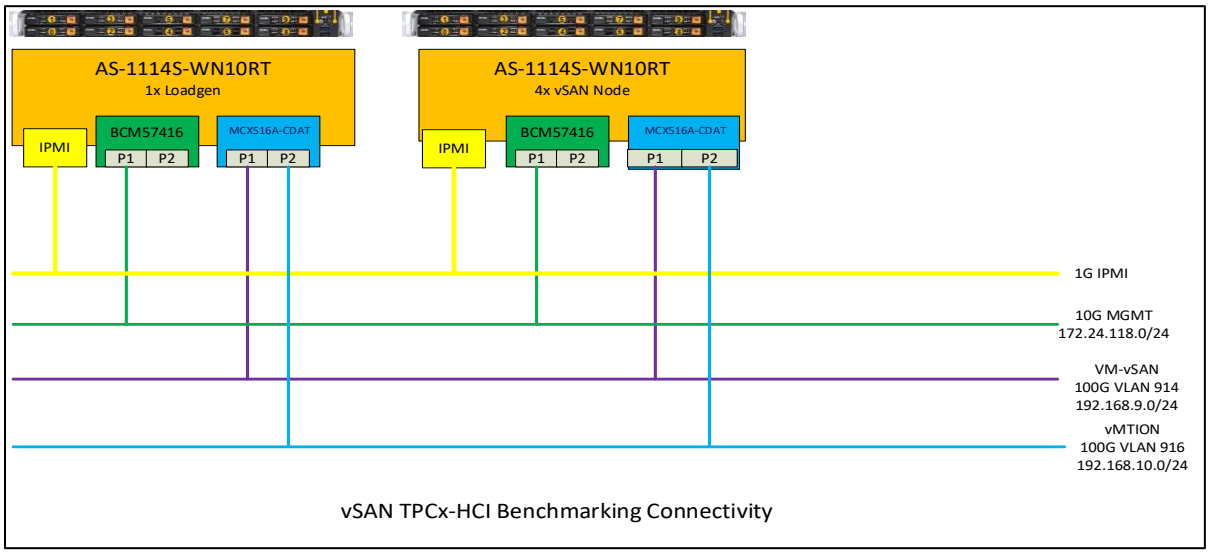

Figure 2 – Testbed network connectivity

Inter- and intra-node Scheduling and Load Balancing

TPCx-HCI workload tests data access uniformly and live VM migration speed by evaluating inter-node load balancing and VM migration. The TPCx-HCI Specification neither requires nor choreographs how, when, or where load balancing should occur. Instead, it fosters load balancing and live migration technologies by allowing a sponsor to benefit from an efficient load balancer. These tests have the following prerequisites:

- A four-node implementation must use five tiles (see Benchmark Implementation on page 6). The benchmark run must start with all 60 VMs in the five tiles present on three of the four nodes, with the fourth node having no VMs. Midway through the warm-up period, the test sponsor may enable load balancing on the cluster to allow VMs to float to the idle node.
- The benchmark elasticity feature changes the proportion of the overall load sent to each VM every 12 minutes while maintaining a constant overall load, as shown in Figure 3[1].

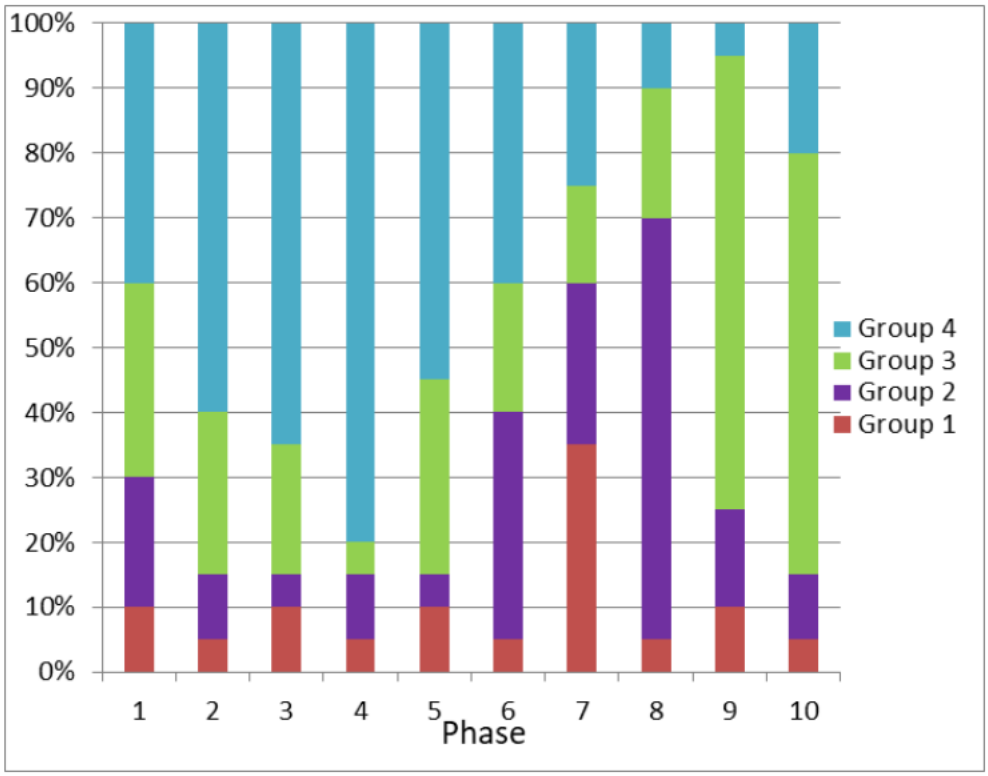

Figure 3 – Each group's load varies from one phase to the next while the overall load remains at 100%.

Imagine a use case where a private cloud serves multiple tenants whose workloads peak at different times of the year. The overall cluster load remains relatively constant, with VM resource allocation varying depending on which tenant workload is peaking. The hypervisor on each node must therefore choose the best resource allocation scheme to maximize the performance of all VMs in the node. However, having five tiles on four nodes causes per-node load fluctuations and requires inter-node load balancing to maintain good performance. Confining all five tiles to three nodes at the start of the test severely impacts performance without efficient load balancing.

Figure 4 graphs the Supermicro FDR performance[2]. There are 60 VMs with load levels that vary by as much as 16x every 12 minutes. These VMs resided on three nodes when the test began. VMware DRS was enabled after 12 minutes, and some VMs migrated to the fourth node. DRS continued to migrate VMs throughout the two-hour run to maintain load balancing. Supermicro's testing shows that:

- The 12-minute moving throughput average deviated from the reported throughput by as much as 7.5%.
- The maximum deviation of the 1-hour moving average was only 1.2%.

These results demonstrate the efficiency of the VMware ESXi™ scheduler and VMware DRS.

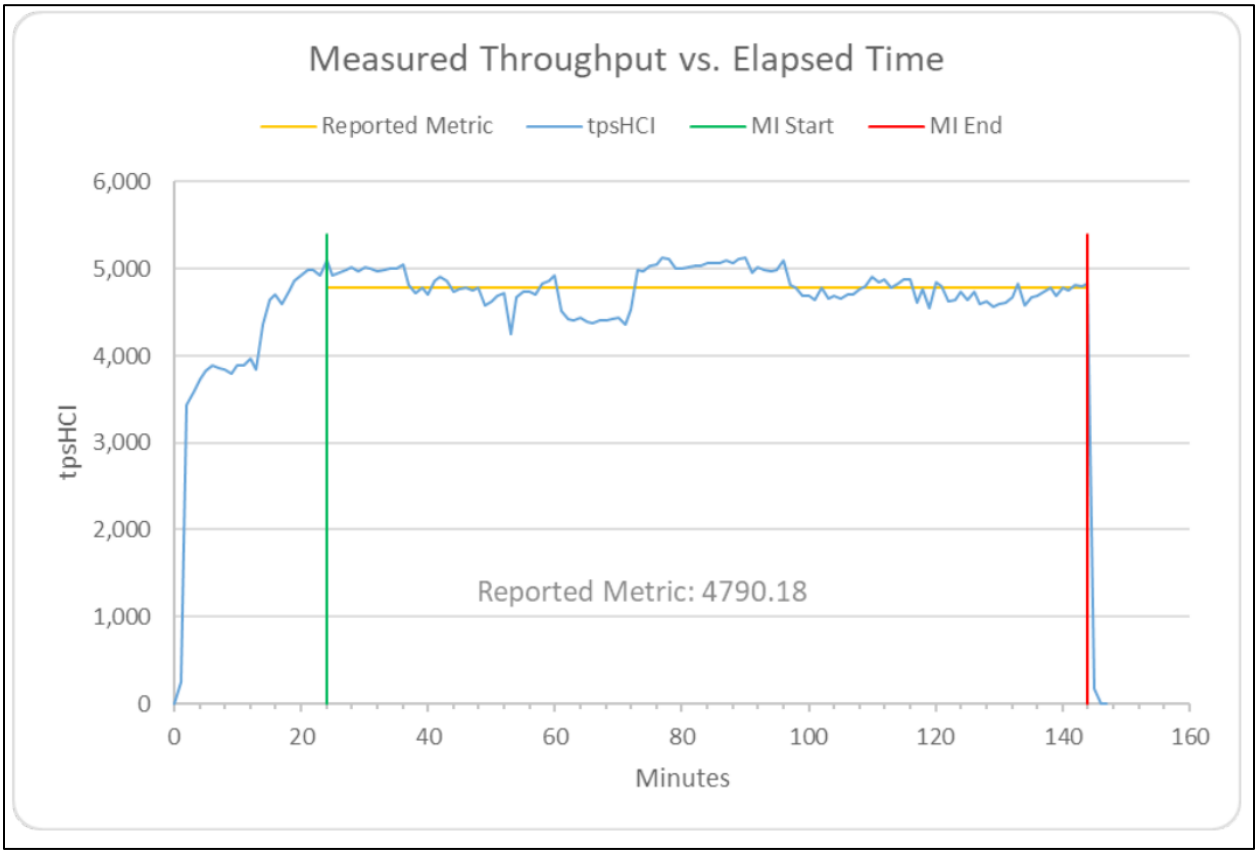

Figure 4 - tpsHCI throughput during the ramp-up period and two-hour measurement interval

Recovery From a Lost Node

The TPCx-HCI Data Accessibility test requires the cluster to continue processing after a node suffers an immediate, ungraceful shutdown. Figure 5 (reproduced from the Supermicro FDR) shows that processing continued with negligible performance after node loss. However, the cluster completed recovering to the pre-loss state after 2:05:38, easily surpassing all TPCx-HCI Data Accessibility requirements. VMware vSAN 7 Standard Edition was key to this rapid recovery[3].

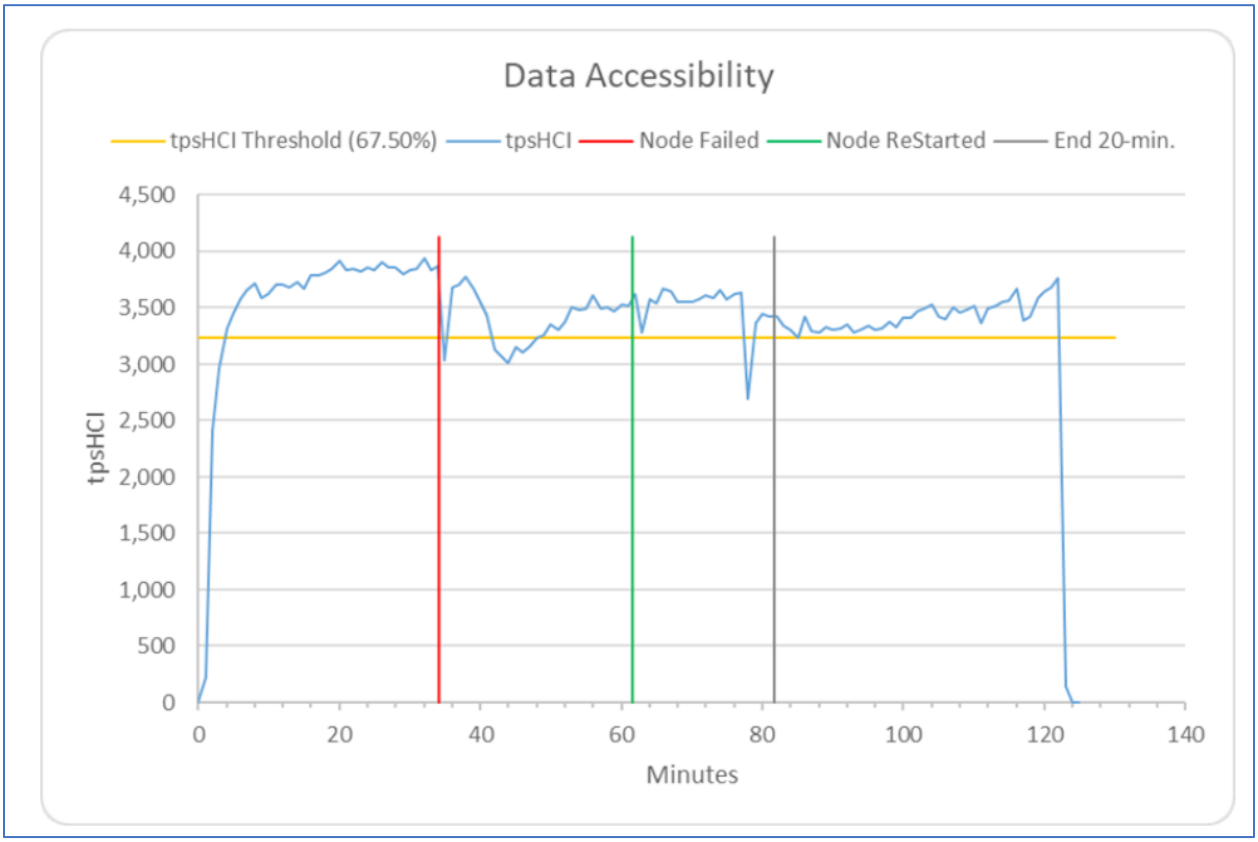

Figure 5 – Transactions continue with slight performance loss following a node failure

Benchmark Implementation

TPCx-HCI is similar to other virtualization benchmarks in that it uses a tile-based architecture, where each tile is a unit of configuration and load distribution replication. A TPCx-HCI tile has 12 VMs with different characteristics. A tile consists of four groups, where each has one Tier A VM and two transaction-specific Tier B VMs, as shown in Figure 6[4].

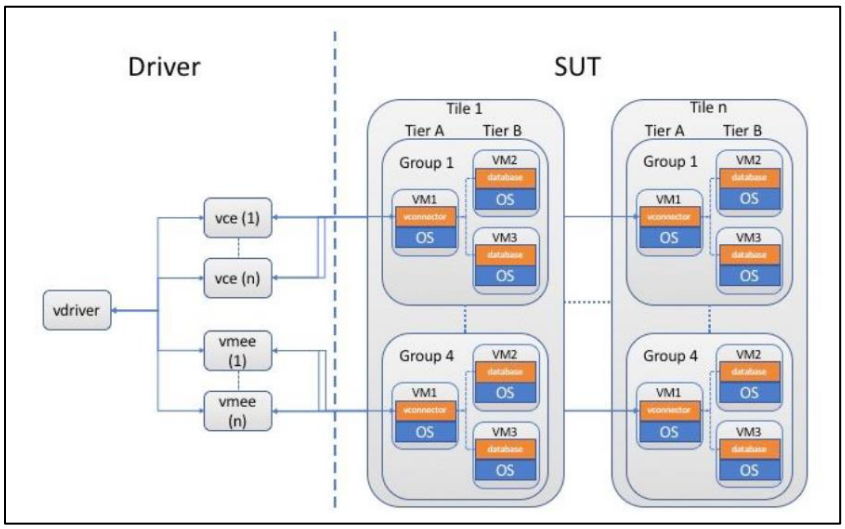

Figure 6 - Architectural layout of the driver and SUT

In the published report[5], the driver consists of 650 customer emulator threads (VCEs) and 200 market exchange emulators threads (VMEEs). The SUT consists of nine tiles. VM1 of each group contains that group's Tier A, which runs the business logic application (database front end). VM2 is the Tier B VM that holds the DSS database and accepts the two storage-intensive DSS transactions. VM3 is the Tier B VM that holds the OLTP database and accepts the nine CPU-intensive OLTP transactions.

The four groups have the same architecture but different load levels. Groups 1, 2, 3, and 4 contribute an average of 10%, 20%, 30%, and 40% of the tile's total throughput over the two-hour measurement interval. Each group's load varies over the measurement interval, as shown in a run time of the benchmark, as shown in Figure 3 above.

Supermicro H12 Generation A+ Servers:

Supermicro H12 Generation A+ servers support 3rd Gen AMD EPYC™ Series Processors up to 8TB of DDR4-3200MHz memory in 1U or 2U form factors that deliver high performance, flexibility, scalability, and serviceability that power mission-critical enterprise workloads in demanding IT environments. Users can select a wide array of storage and networking options to optimize performance and efficiency for their specific datacenter virtualization needs.

Footnotes

1. Supermicro WIO cluster ran on four Supermicro AS 1114S-WN10RT servers, each with one AMD EPYC™ 7713 processor and 1TB of main memory. Overall, in the cluster, the processors had a total of 256 cores, with hyper-threading enabled for a total of 512 threads. Benchmark result using the VMware HCI solution vSphere 7 U2 virtualization, vCenter management, and vSAN storage. The VMs ran the Red Hat Enterprise Linux 7.7 operating system and PostgreSQL 10.6 database management system (DBMS). Achieving the score of 4,790.18 tpsHCI @ $49.60 USD per tpsHCI, is available immediately. World's first TPCx-HCI benchmark result in price and price-performance http://tpc.org/5801; see also: https://www.amd.com/en/processors/epyc-world-records and https://blogs.vmware.com/performance/2021/12/tpcx-hci-benchmark-with-vmware-hci.html
2. TPCx-HCI User's Guide
3. TPC Download Current Specs/Source
4. TPCx-HCI Result Highlights
5. Supermicro TPCX-HCI Executive Summary

Related Information

- Supermicro TPCX-HCI Full Disclosure Report (includes test data)
- Supermicro TPCX-HCI Supporting Files
- Supermicro A+ Products