SOLUTION BRIEF

DOUBLE THE CACHE CLUSTER BANDWIDTH WITH INTEL® OPTANE™ PMEM 200 SERIES

Deploy a high-availability cache cluster for real-time applications with Red Hat® OpenShift & Redis® Enterprise Operator on a Supermicro X12 BigTwin® System

TABLE OF CONTENTS

Executive Summary 1
Solution Overview. 1
Value Proposition 2
Sub-millisecond Response 3
Easy Button for Hybrid Cloud 4
Baseline Configurations. 6
Performance Metrics 8
Performance Monitoring 9
Summary of Redis Performance. 10
Conclusion 11

Executive Summary

As the market demands more bandwidth and real-time online services, the challenge to support new technology such as 5G, IoT, and real-time applications has dramatically changed data center infrastructure requirements. Cloud IaaS (Infrastructure-as-a-Service) providers constantly optimize caching services to reduce the “Last Mile” latency at strategically located Edge Data Centers.

Real-time applications are prevalent in various markets, especially in advertising, media & entertainment, online gaming, e-commerce, mobile apps, healthcare, higher education, and government. Supermicro has partnered with Intel, Red Hat, and Redis to create a high-availability cache cluster in a compact 2U 4-Node system for easy deployment and scalability with hybrid cloud capabilities. With its award-winning multi-node design, in-memory performance, low-latency NVMe storage, and redundant Titanium level power supplies, one Supermicro X12 BigTwin system can reliably deploy and support real-time applications with the power of Red Hat® OpenShift and Redis Enterprise.

The solution was inspired by the growing demand for cloud-native applications that stream static and dynamic content, including rich media. These applications must be optimized for viewing on mobile devices, AR/VR glasses, televisions, and web browsers on tablets and laptops of all shapes & sizes.

SUPERMICRO

Supermicro (Nasdaq: SMCI), the leading innovator in high-performance, high-efficiency server and storage technology is a premier provider of advanced server Building Block Solutions® for Enterprise Data Center, Cloud Computing, Artificial Intelligence, and Edge Computing Systems worldwide. Supermicro is committed to protecting the environment through its "We Keep IT Green®" initiative and provides customers with the most energy-efficient, environmentally-friendly solutions available on the market.

Solution Overview

Red Hat® OpenShift is an Enterprise-ready, hybrid-cloud Kubernetes platform built to run and scale container-based applications to provide a consistent platform for managing hybrid cloud, multi-cloud, and edge deployments.

Redis is an in-memory data structure store used as a database cache and message broker. It is an ideal database for highly interactive, scalable, low-latency geo-distributed apps.

To reduce CAPEX & OPEX, as well as accelerate time to value for Enterprise DevOps teams, Supermicro has introduced the REC-Optane Series, based on popular X12 SuperServers, powered by 3rd Gen Intel® Xeon® Scalable Processors & Intel Optane™ PMem 200 Series, as an integrated solution with Red Hat OpenShift & Redis Enterprise.

Value Proposition

With Supermicro's high-density BigTwin system, Redis was benchmarked across a 3-node cluster to highlight the power of the REC-Optane product series.

The top 3 design decisions to deploy a Supermicro X12 BigTwin as a Redis cluster are as follows:

1) Easy to reliably deploy hybrid cloud infrastructure on bare-metal servers
    - Red Hat® OpenShift makes it easy to manage hybrid cloud infrastructure with an Enterprise-Ready Kubernetes platform
    - Red Hat® OpenShift Data Foundation provides persistent storage for hybrid cloud and multi-cloud container deployments
    - The Supermicro X12 BigTwin is a 2U 4-Node system, allowing Cloud IaaS Providers to simplify the logistics of hardware deployments at strategic Edge Data Centers
    - The HW RAID 1 NVMe Switch on each node allows administrators to automate the configuration of redundant boot drives for Red Hat® OpenShift with Redfish APIs, enabling zero-touch provisioning

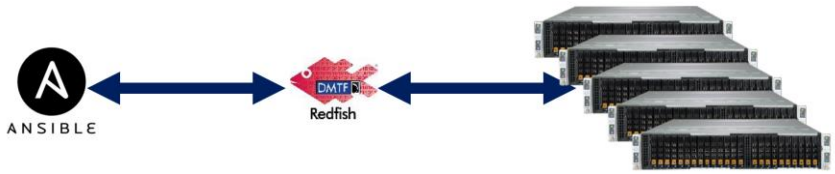

Figure 1 – Zero-Touch Provisioning

2) Performance per Dollar
    - Optimal value of K8s on bare-metal (using a max core count of OpenShift subscription of 64C per node)
    - Double the memory capacity with Intel® Optane™ PMem 200 Series compared to DRAM-only configurations
    - Power-efficiency of the Supermicro X12 BigTwin (Optimized OPEX with multi-node design, shared power & cooling)

3) Seamless support for real-time applications with Redis Enterprise Operator
   - Easily configure a high-availability cache cluster with Redis Enterprise Operator to support real-time applications
   - Stay up to date with the latest Redis features through the Red Hat® Operator Framework
   - Run once, run anywhere with Kubernetes-native applications

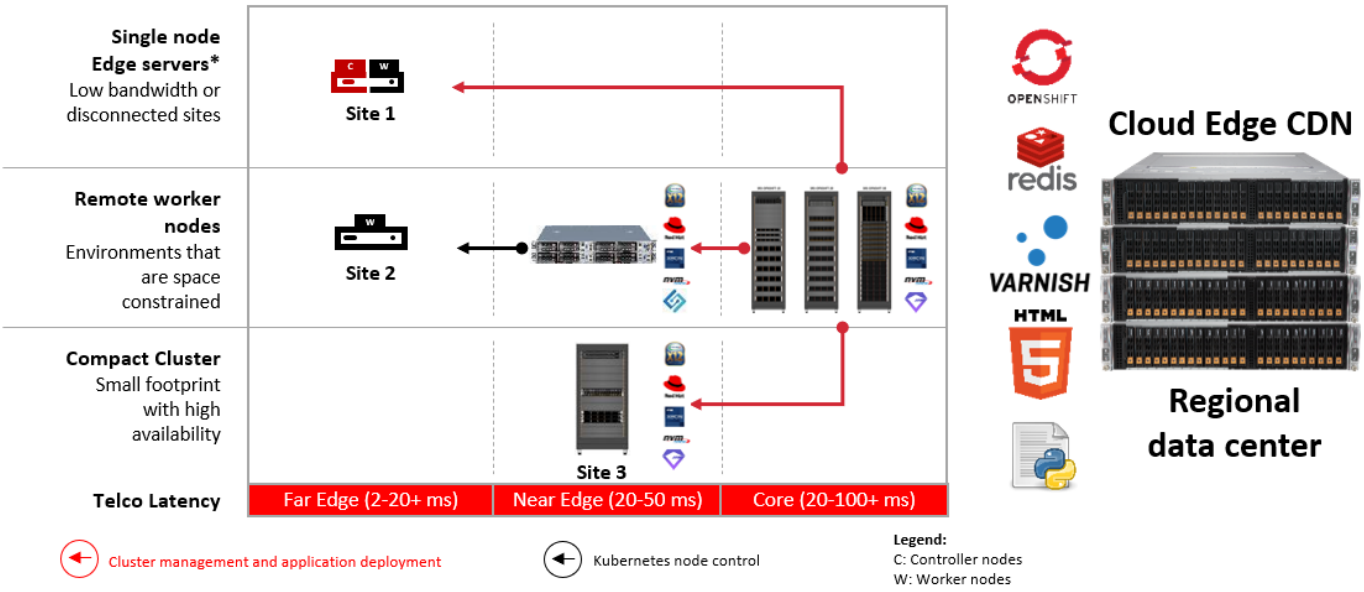

Figure 2 – OpenShift Container Platform Support: Run Once, Run Anywhere

Sub-millisecond Response with Reduced Cost from Edge to Cloud

Supermicro's patented Twin Architecture is the foundation for the most energy-efficient and advanced server platforms in HPC, Data Center, Cloud Computing, and Enterprise IT applications. This high-performance, high-density system features optimum airflow for energy-efficient cooling and easy maintenance with hot-swappable nodes and redundant PSUs.

The Red Hat OpenShift cluster is deployed on the Supermicro BigTwin® 2U 4-Node chassis containing 3rd Gen Intel Xeon Scalable Processors. The Supermicro X12 BigTwin®, SYS-220BT-HNTR supports up to 4TB of memory per node with Intel® Optane Memory. Based on Supermicro customer surveys, 512GB of DRAM per node is a popular memory capacity to support static and dynamic web content. Still, more memory is needed to support rich media. 1TB per node would be ideal, but it is cost-prohibitive due to the high costs of 128GB RDIMM-3200 modules. For the same price of 512GB of DRAM, system memory may be doubled with Intel® Optane PMem on the Supermicro X12 BigTwin, as shown in Table 1 below.

By using Intel® Optane™ PMem 200 Series, the Supermicro X12 BigTwin can potentially double the memory bandwidth without significant degradation in latency compared to a DRAM-only setup. This solution brief will examine the relative P99 latency of Intel® Optane PMem vs. DRAM and ensure a typical SLA of 1ms can be met.

Based on the benchmarks results, Cloud IaaS Providers may consider using Intel® Optane PMem to deploy enlarged high-availability cache clusters with fewer servers. With the Supermicro X12 BigTwin, lowering the cost of larger memory

capacities becomes much more practical, especially with its shared power, cooling, and storage backplane to save snapshots periodically of the Redis database. In addition, this allows for hardware consolidation of in-memory caching and All-Flash NVMe storage platforms – resulting in lower server count and product mix to simplify deployments and lower TCO.

SYS-220BT-HNTR	Description	Qty	Function
DevOps-in-a-Box	Based on X12 BigTwin® 2U 4-Node	1	Redis Enterprise Cluster: **Optane-REC-BT**
Intel Processor	ICX 8352V 36C @ 2.1G, 195W	8	2 CPUs per Controller + Worker Nodes
Memory	32GB DDR4-2933 RDIMM	32	256GB per Node (64 RDIMMs for Baseline)
Persistent Memory	128GB Intel® Optane™ PMem 200 Series	32	1028GB in Memory Mode per Node
Bare-Metal Network	100GbE 2-port QSFP28 PCI-E 4.0 LP Card	4	Bare-Metal Network Interfaces
Provisioning Network	10GbE 2-port RJ-45 PCI-E 3.0 AIOM	4	Isolated Provisioning Network Interfaces
Optional Boot Controller	M.2 NVMe HW RAID Controller	4	HW RAID Protection for OpenShift
Boot Devices	1TB M.2 NVMe PCI-E 3.0 Devices	8	Redundant Boot Devices
Storage	3.84TB U.2 PCI-E 4.0 Drives	8	Local Storage
Kubernetes SW	Red Hat® OpenShift Container Platform	3	Hybrid Cloud, Platform-as-a-Service
Persistent Storage SW	Red Hat® OpenShift Data Foundation	3	Software-Defined Storage for Containers
Optional Bastion SW	Red Hat® Enterprise Linux	1	Bastion Host: Deployment & Management

Table 1 – Red Hat OpenShift Compact Cluster Specifications

Figure 3 – Supermicro OpenShift Compact Cluster

Hit the Easy Button for Hybrid Cloud Infrastructure Deployment & DevOps Management

Key benefits in deploying and managing these clusters:

1) Streamline the deployment of a hybrid cloud infrastructure with Red Hat OpenShift for enhancing DevOps

   One Bastion node within the enclosure may be used as a host for deploying and managing the Red Hat® OpenShift cluster, allowing for zero-touch provisioning without needing an external system for deploying the 3-node compacter cluster. The Bastion node is optional but offers powerful orchestration support with Ansible playbooks.

2) OpenShift has built-in CI/CD support for cloud-native applications. Many mobile applications require high-availability caching services. The cache cluster size and location of a CDN PoP (Points-of-Presence) or Edge Data

Center can significantly impact a mobile application's response. Therefore, it is critical for app developers and cloud architects to monitor and manage the performance of these applications early in the development cycle through UAT (User Acceptance Testing), as well as production environments (core, regional, and/or edge data centers) with different SLAs (service-level agreements). Popular CDN edge servers typically respond to queries in less than 30ms.

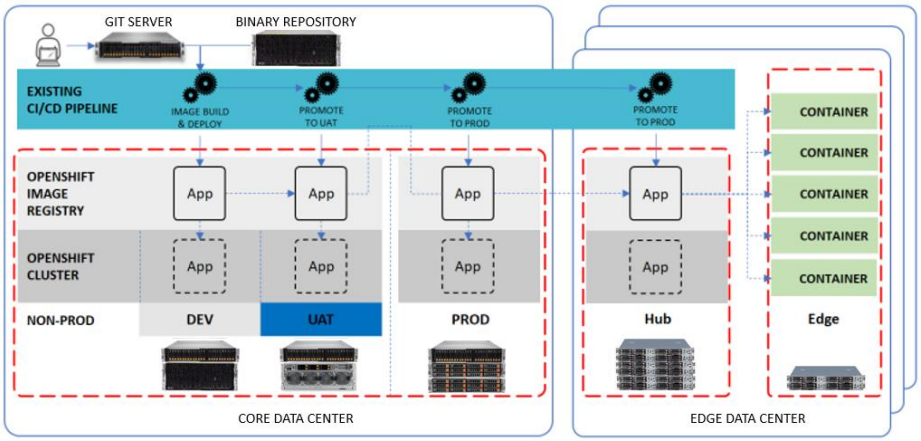

Figure 4 – Continuous Integration & Deployment with Supermicro Compact Cluster

3) The Redis Enterprise Operator was deployed across three nodes to support a high-availability cache cluster.

Supermicro has run several performance tests with the following setup. Figure 5 below shows the Kubernetes Abstractions and how a Job with Pod and ConfigMap can be used to run the memtier_benchmark[1] across the Redis Enterprise Cluster deployed by the Redis Enterprise Operator.

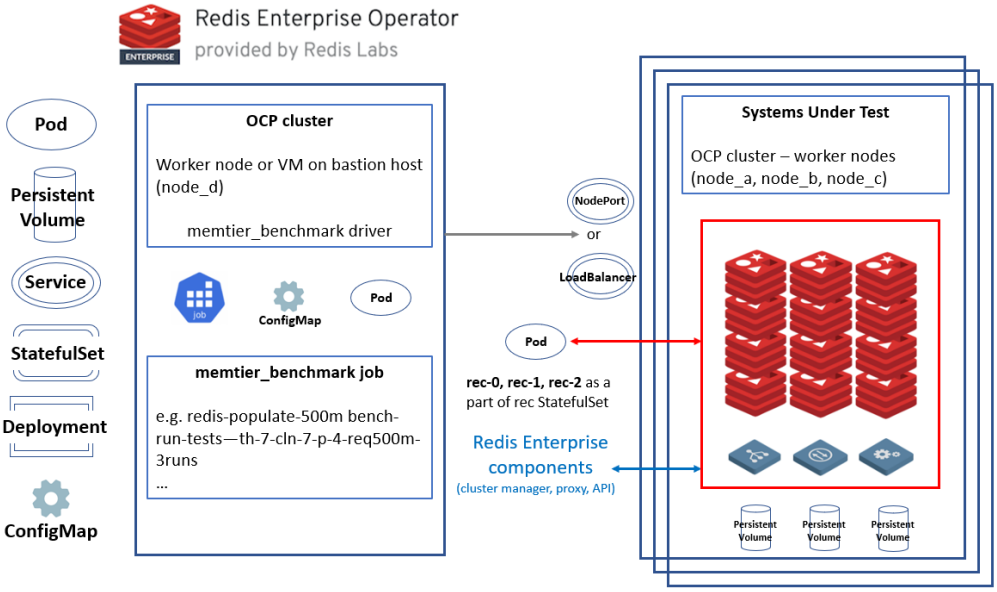

Figure 5 – Kubernetes Abstractions

Baseline and PMem Use Cases Configurations

The tables below provide details for the hardware and software configurations used to benchmark. The objective was to show that the DRAM performance would be similar to the Intel® Optane™ PMem performance but double the cluster's memory capacity. As a result, the parameters can significantly impact throughput and latency. Those details are included below for repeatability and transparency.

	Configuration 1 – Baseline (DRAM-only) 1TB Dataset	Configuration 2 – Memory Mode (DRAM + Intel® Optane™ PMem) Same Dataset Size used in Baseline	Configuration 3 – Memory Mode (DRAM + Intel® Optane™ PMem) 2x Dataset Size
Total Memory (per node)	16 x 32GB DRAM = 512GB [1.5TB across 3 nodes]	8 x 32GB DRAM + 8 x 128GB PMem = 1 TB in Memory Mode [3 TB across 3 nodes]	8 x 32 GB DRAM + 8 x 128 GB PMem = 1 TB in Memory Mode [3 TB across 3 nodes]
Dataset size	1TB (500GB master shards + 500GB replica shards) with replication across 3 nodes	1TB (500GB master shards + 500GB replica shards) with replication across 3 nodes	2.1TB (1.05TB master shards + 1.05TB replica shards) with replication across 3 nodes

Table 2 – Baseline (DRAM) and Optimal (Optane) Memory Configuration

Software Component	Version
OpenShift Container Platform	4.8.27
Redis Enterprise Operator	6.2.10-4

Table 3 – Red Hat OpenShift Software Versions

Before starting the tests, Supermicro worked closely with Intel to define the parameters used in the synthetic benchmark to simulate containerized application workloads.

Component	Parameter	Value
Redis Enterprise Operator	Version	6.2.10-4
	Deployment	3 Nodes
Redis Enterprise Cluster	Memory & CPU Resource Allocation	redisEnterpriseNodeResources: limits: cpu: 64 memory: 450Gi shards_placement: sparse redis_version: 6.2 proxy: mode: dynamic scale_threshold: 50 threads: 3 max_threads: 36
Redis Enterprise Database	Proxy Policy & Shards Count	proxy_policy: all-master-shards oss_cluster: true memorySize: 1050GB replication: true shardCount: 56 → 112 with replication
memtier_benchmark	Version	1.3.0 (edge)
	Threads	6
	Clients	5
	Pipelines	4
	Ratio	SET:GET 1:4
	Test Time	3 x 600sec = 15min
	Requests	3 000 000 000
	Object Size	100b

Table 4 – Redis Enterprise and Benchmark Tool Configuration

Performance Metrics - Results with DRAM and Intel Optane Persistent Memory

The following sections provide the results of synthetic benchmark performance for the 3-node Redis cluster. Several tests were conducted with different configurations and parameters (see Figure 6 below for details). Three configurations were tested, one test with DRAM-only as a baseline, and two tests were run with Intel® Optane™ PMem to measure the impact in latency with different datasets. Memory Mode is the default configuration for the Intel® Optane™ PMem modules. Synthetic workloads were generated using the memtier_benchmark.

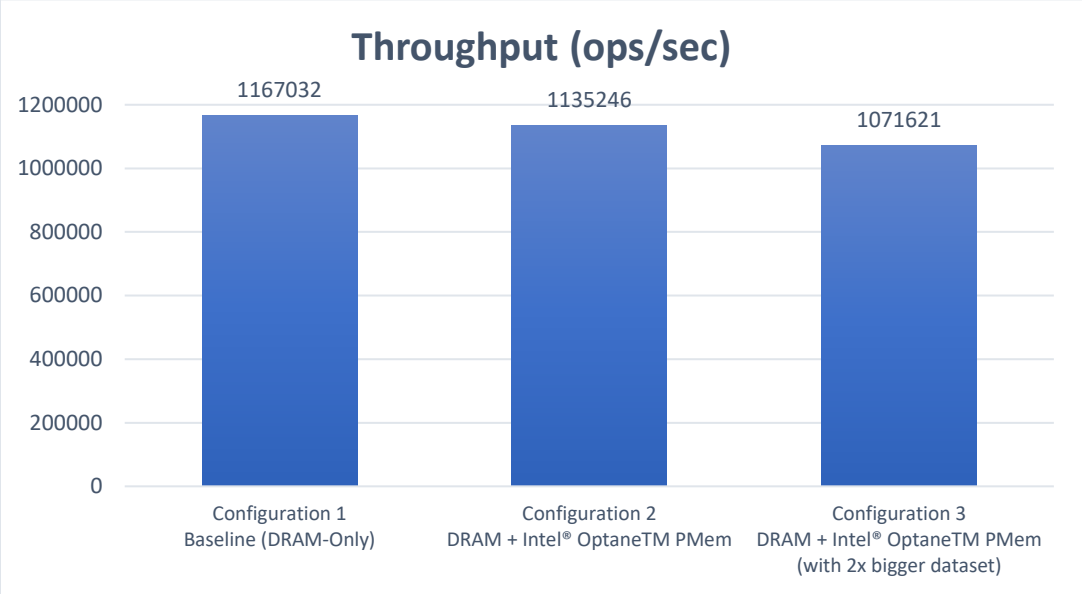

Figure 6 – Performance of DRAM vs. PMem

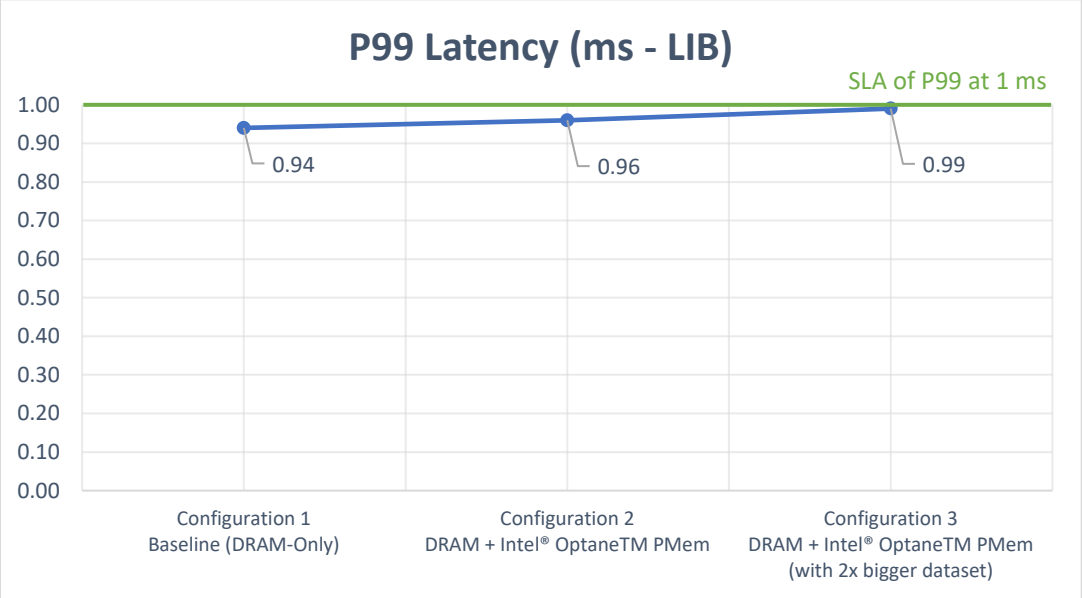

Figure 7 – P99 latency of DRAM vs. PMem

Performance Monitoring

DevOps teams may want to monitor infrastructure and/or application performance to ensure applications are meeting SLAs. Grafana is a very popular open-source monitoring tool used to build interactive dashboards configurable to suit a DevOps team's needs. Grafana may be customized further to monitor application performance to gain a deep understanding of user patterns and integrated microservices. These dashboards are critical for CSPs with CDN PoPs to help make more intelligent decisions to manage resources, configure load balancers, and even identify DDoS threats.

Figure 8 shows a snapshot of the Grafana dashboard, customized to monitor the performance of the 3-node Redis cluster simply, in response to various synthetic benchmarks outlined above in Table 2. This dashboard was focused on monitoring throughput (ops/sec), latency (ns), total keys, and memory usage. This custom Grafana dashboard is modeled from the default view of the Redis telemetry service, shown in Figure 9.

Figure 8 – Redis Telemetry on Grafana dashboard

Figure 9 – Redis Telemetry (from REC UI – 5min scope)

Summary of Redis Performance

Below is a Redis performance summary for the three different tested configurations. The total throughput and latency of the Intel® Optane™ PMem configurations were close to the performance of the DRAM-only configuration. This result is remarkable as the configuration has twice the amount of memory across the 3-node cluster with PMem, operating in Memory Mode. In addition, configurations 1 and 2 used the same dataset size of 1TB, which had a delta in P99 latency of 2.1%, with a 2x bigger dataset of 2.1TB, the P99 latency delta of 5.3% against the baseline configuration.

Traffic generation and benchmarking tool: memtier_benchmark	Core Metrics		
	Total throughput (ops/sec)	P99 latency (ms)	Average latency (ms)
Configuration 1 – Baseline (DRAM only)	**1,167,032**	**0.94**	**0.30**
Configuration 2 – DRAM + Intel® Optane™ PMem	**1,135,246**	**0.96**	**0.31**
Configuration 3 – DRAM + Intel® Optane™ PMem (2x bigger dataset)	**1,071,621**	**0.99**	**0.33**
Redis with PMem in MM vs. Baseline	97.3%	102.1%	103.3%
Redis with PMem in MM (2x bigger dataset) vs. Baseline	91.8%	105.3%	110.0%

Table 5 – Redis Performance Summary: DRAM baseline and PMem Test Results

Conclusion

In partnership with Intel, Red Hat, and Redis, Supermicro offers a performance-optimized hybrid cloud building block to scale-out Redis Enterprise with Red Hat® OpenShift, running on a Supermicro 2U 4-Node BigTwin® (SYS-220BT-HNTR). Building scalable caching infrastructure has never been easier, thanks to the Operator Framework, which Supermicro has proven with the Redis Enterprise Operator, certified by Red Hat, as a Kubernetes-native application. With 3rd Gen Intel® Xeon Scalable Processors, 8352V CPUs, DDR4-2933, plus Intel® Optane™ PMem 200 Series in Memory Mode across 3 nodes, the solution delivers game-changing performance with double the memory capacity for the same cost of a typical DRAM-only configuration. The DRAM-only and PMem configurations achieved over 1 million ops/sec with P99 latency below 1ms. Additionally, the PMem configuration processed twice the amount of data than the DRAM-only configuration while staying within a typical SLA of 1ms P99 latency.

With a balanced memory topology across two CPUs and support for 64+ cores per node, the Supermicro BigTwin offers tremendous performance per dollar. The solution maximizes the value of OpenShift, as each bare-metal subscription supports up to 64 cores or 2 sockets per node. In addition, the Supermicro BigTwin supports up to 6 NVMe PCI-E 4.0 drives per node, providing low-latency persistent storage for the Redis database. With Red Hat® OpenShift Data Foundation, data can be replicated or striped across the local 3-node cluster or with external S3 object storage services via the Multi-Cloud Object Storage Gateway.

Scaling resources and predicting application performance at the edge are challenging, requiring a sophisticated cache-tiering solution to optimize OPEX. Now, these challenges can be addressed with this solution. With Red Hat OpenShift®, doubling the memory capacity per node becomes exceptionally valuable in the context of a microservices architecture. The Kubernetes Engine simplifies resource management by elastically allocating resources to support a wide variety of applications leveraging Redis. This unlocks the true potential of PMem in a 2U 4-Node system, supporting twice the amount of memory capacity of a typical 1U server.

References

1. https://github.com/RedisLabs/memtier_benchmark

Additional Resources

For more information, please visit the following links:

- Solution Landing Page: Redis Enterprise Ready Solutions – https://www.supermicro.com/en/solutions/redis
- Thought Leadership Blog: Validating Redis Enterprise Operator on Bare-Metal – https://community.intel.com/t5/Blogs/Thought-Leadership/Big-Ideas/Validating-Redis-Enterprise-Operator-on-Bare-Metal-with-Red-Hat/post/1386198
- More Details about Intel PMem – https://www.intel.com/content/www/us/en/products/docs/memory-storage/optane-persistent-memory/optane-persistent-memory-200-series-brief.html
- Contact: _cloudnativesys@supermicro.com