SOLUTION BRIEF

SUPERMICRO AND KOVE EMPOWER LARGE IN-MEMORY APPLICATIONS RUNNING ON RED HAT OPENSHIFT

In collaboration with Kove:SDM ™ and Red Hat® OpenShift®

Supermicro Multi-Node Infrastructure Solutions

Executive Summary

TABLE OF CONTENTS

Executive Summary 1
Demand for Massive In-Memory Applications 2
Options for Effective Memory Use 2
What We Need is a Better Software-Defined Memory 2
Welcome to Kove:SDM 3
Why Supermicro 4
Why Kove 4
Why Red Hat OpenShift Container Platform 5
Massive Memory Technical Implementation 5
Proof of Concept Testing 6
Benchmark Results 7
Conclusion 7
References 8

Data is the new oil, and datasets are constantly expanding. AI/ML HPC, Database, and Virtualization applications can all benefit from in-memory processing, driving memory requirements beyond what a single server can provide. Further, the need to over-allocate memory for headroom in many applications leads to stranded and underutilized memory assets.

Supermicro, Kove, and Red Hat, have benchmarked a large in-memory application using Kove:SDM, spanning a multi-node cluster running Red Hat OpenShift on Supermicro BigTwin servers.

This turnkey solution enables a new class of applications and greatly improves productivity by negating the need to chunk datasets and associated programs to fit into constrained memory. Capex and Opex are further improved by making underutilized server memory in HPC data centers available to any application on demand.

The Increasing Demand for Massive In-Memory Applications

AI and Enterprise data sets are exploding in size. Processing that data is most efficient when all the data is fully resident in DRAM, avoiding the need to process the data in blocks or constantly page data to disk storage. However, while processor core counts are increasing, memory capacity and bandwidth are not scaling proportionally.

To enable the broadest set of virtualized applications, servers are being built with the maximum possible memory, limited only by the physical PCB real estate. However, there is a dichotomy: Increasing memory size addresses memory-intensive applications and increases costs and the probability of stranding memory. Meanwhile, to protect against possible server crashes, memory is often under-utilized, so you don't run out of it.

As a result, while many applications can benefit from more memory than can be provisioned on a single server node, many servers have underutilized memory, depending on workload.

Options for Effective Memory Use

Effective memory can be increased by accessing memory secondary memory on the server, often by paging to block storage. The problem: This results in a massive slowdown. It could be up to 125x slower, even using the fastest SSDs. The outcome: A 4-hour job running in SWAP could take more than 20 days to complete, rendering certain workloads impractical.

A common workaround is to rewrite the application to process the data in manageable chunks and then assemble the result. But this approach is fraught with numerous problems as it is error prone, creates developer costs, and may not solve the run time issues.

What is needed is a way to improve direct access to offboard memory by providing on-demand access to memory across servers. The industry has recognized this and has been working on a software-defined memory solution for many years in the form of CXL. However, CXL 3.0, which provides complete caching capability, is still several years away, will require new server architecture, and will only be available in forthcoming generations of hardware.

What We Need is a Better Software-Defined Memory

Workloads at the heart of everything from HPC to AI have significant memory requirements. But designers struggle to make use of the additional cores available in modern CPUs.

The leap forward in the number of CPU cores is mismatched with a lack of memory bandwidth. And it continues to worsen due to the limited physical space to incorporate more memory and the limited access to additional memory beyond the motherboard.

Concerns about latency compromises are surfacing, too. Even CXL 3.0 is still piggybacking on the PCI Express (PCIe) physical layer and relying on physical memory paired with PCIe, so one would ordinarily incur a penalty on a key critical metric—latency. Generally, the farther the memory is from the CPU, the higher the latency and the poorer the performance.

To embrace scaling, memory must be moved outside of the server. Yet, current options that include block storage and cloud services aren't viable solutions.

Meanwhile, software-defined memory helps ease pressure on DRAM while increasing computing efficiency and performance. This subset of software-defined technologies is unlocking a new age of disaggregated memory, mirroring the revolution that came with disaggregated storage.

Welcome to Kove:SDM – The Software-Defined Memory Solution

Your days of memory limitations are over. Kove:SDM is a breakthrough technology that gives enterprises the memory size and performance they need when and where they need it on any hardware. It empowers individual servers to draw from a shared memory pool, including amounts far larger than could be contained within any physical server, so each job receives exactly the memory it needs while reducing your power consumption.

With Kove:SDM, for instance, you can allocate 10x 64 GiB for ten containers on a compute node with only 64 GiB of memory, create containers with larger memory than the physical hypervisor, or burst to allocate 40 TiB to a single server for an hour. Control memory in real-time, on-demand, using easy-to-configure provisioning rules. Memory capacity scales up to CPU addressable limits beyond the limits of local physical DIMM slots.

Kove:SDM solves memory stranding by pooling memory into a global resource, shareable and reusable across the data center. In other words, it decouples or “disaggregates” memory from standard servers, aggregating memory into a shared memory pool resource. SDM policies then structure the access to the memory pool. After use, Kove:SDM securely zeros out and returns memory to the pool for reuse. As a result, Kove:SDM also provides strong security against attacks targeting memory penetration.

Kove:SDM Works with Any Server

Unlike CXL, no special chips or hardware are required to run Kove:SDM. Rather, it decouples memory from servers, pooling memory into an aggregate, provisionable, and distributable resource across the data center using unmodified Supermicro hardware.

Like a Storage Area Network (SAN) provisioning storage using policies, Kove:SDM delivers a RAM Area Network (RAN) that provisions memory using policies. Both provide a global, on-demand resource exactly where, when, and how it is needed. For example, an organization might:

- Use a Kove:SDM policy to allocate up to 2 TiB of need-based memory for any of 200 servers between 5 pm-8:30 pm;
- Develop a provisioning rule that would provision a virtual machine with larger memory than the physical hypervisor, such as a 64 GiB RAM physical server (hypervisor) hosting a 512 GiB RAM virtual machine; and
- Provision a 40 TiB server for a few hours or a 100 TiB RAM disk with RAID backing store for a temporary burst ingest every morning.

Kove:SDM uses three transparent software components: 1) Management Console (MC) that orchestrates memory pool usage; 2) Host Software that connects applications to a memory pool; and 3) XPD software that converts servers into memory targets to form a memory pool. Users and applications do not ever need to know that it is present.

As a result, Kove:SDM enables technology leaders and their enterprises to do things they could not have done before – including finally maximizing the performance of their infrastructure and people.

Why Supermicro

Supermicro, a Total IT Solution Provider for Cloud, AI/ML, Storage, and 5G/Edge, offers the industry's most extensive portfolio of workload-optimized servers to support the highest computing density across the broadest range of applications. Supermicro servers are used for demanding workloads across the entire IT industry, including 5G workloads, a range of AI, HPC, visualization, Big-Data Analysis, virtualization, and enterprise workloads.

Supermicro continues to innovate and design application-optimized solutions across the entire IT industry. By implementing Kove:SDM, Supermicro enables the class of very large dataset workloads to be efficiently and effectively processed in the shortest time possible. Even in some cases, enabling applications that would otherwise be impractical using standard methods.

The Supermicro Twin family of products has been designed for the most demanding applications while reducing OPEX through innovative design that reduces electricity usage and E-waste. Supermicro Twin Family Multi-node systems are designed so that all servers (nodes) are located within a single chassis with shared power supplies and fans. Supermicro has developed a product family that takes advantage of the Twin product family's latest computing, storage, and networking technologies. The Twin Family is ideally suited to workloads enabled by Kove:SDM as it provides a dense cluster of computing, memory, and storage while supporting a full range of networking. Kove:SDM enables the cluster to operate as a single, very large memory machine. The Supermicro BigTwin® infrastructure is perfect for this setup which can save 3-9x on infrastructure costs with OpenShift bare-metal. Additionally, each node can hold 6x NVMe to provide low-latency persistence storage for any workload.

Why Kove

After years of testing and validation, Kove premiered the world's first patented and mature software-defined memory solution (Kove:SDM). With Kove:SDM, your days of memory limitations are over. Users of Kove:SDM can now right-size memory to need, providing faster time to solution, better models, and the ability to do things they could not do before. Never again worry about fine-tuning datasets to fit hardware when you can dynamically size hardware to fit the datasets. With Kove:SDM, you can quickly scale up your memory for improved utilization, simplicity, efficiency, and performance and avoid the need to over-provision memory, risking underutilization and wasted resources. As a result, you can analyze any data size or computational need, no matter how large. With Kove:SDM, you can scale up your size and density on demand beyond server limits.

Kove:SDM on Supermicro Big Twin is applicable across a wide range of demanding computing, bringing the benefits of larger effective memory to:

- **AI/ML** - Provision resources to the model rather than forcing models to fit fixed resources. Enjoy deeper and faster lookups, analytics, and iterations. Improve your time on the solution and your return on your data scientists.
- **In-Memory Databases** - Analyse databases 100s of times larger than physical servers.
- **Containers** - Improve CPU utilization. Run more jobs in parallel on a single server, increasing workload capability by 20x. Gain the ability to run 7.5x more C3.ai containers.
- **High-Performance Computing** - Run big data analytics, genomics, and Monte Carlo computations entirely in memory. Build trading systems in Java with <11 µs risk exposure. Use standard servers to create any size memory server on demand (e.g., 32-256 TiB in a few seconds).
- **Enterprise, Cloud, and Edge** - Enables unlimited memory sizing. Any size computation can run entirely in memory. Achieve your green goals through 52% CO_2 reduction and 33% floor space reductions. Create a hybrid cloud to keep sensitive data on-premises without cost and scaling concerns. Reduce your power consumption needs by 50%. Greater utilization makes edge computing financially viable.

Why Red Hat OpenShift Container Platform

Red Hat OpenShift is a powerful container platform that offers several benefits when deployed along with Kove:SDM operator in bare metal environments such as the Supermicro systems. OpenShift's inherent scalability and flexibility for containerized applications make it ideal for managing applications in a software-defined memory environment. OpenShift's container orchestration capabilities enable seamless deployment and management of applications across the distributed memory infrastructure provided by Kove, ensuring optimal resource utilization, large amounts of memory available, and efficient workload distribution. This combination allows organizations to leverage the benefits of software-defined memory while maintaining the ease of managing containerized applications in a low-cost bare-metal environment.

The combination of Red Hat OpenShift with Supermicro BigTwin systems provides additional performance, reliability, and resource utilization advantages. The Supermicro BigTwin 2U 4-Node form factor offers the most reliable 3-node cluster with an extra node as a bastion node in one chassis. Supermicro's BigTwin system with high-density and high-storage options compliments OpenShift's capabilities by providing a robust and scalable foundation to run containerized workloads.

Massive Memory Technical Implementation

The Logical Diagram for our collaborative proof of concept is shown below. It shows two host computers running the applications and 3 Kove memory targets. However, Kove:SDM scales linearly and is limited only by the network interconnect infrastructure that customers provide. Kove:SDM software allocates memory on demand and as needed to the Application Hosts.

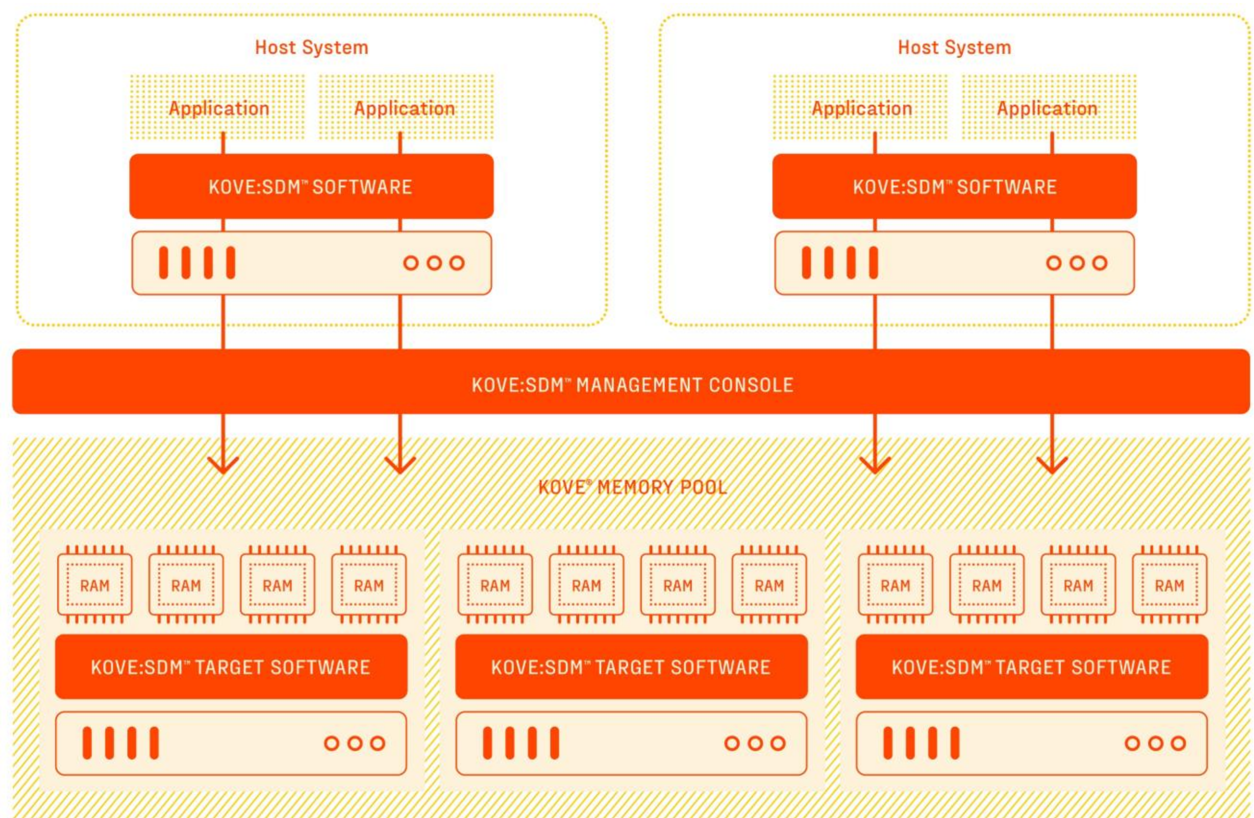

Figure 1 Kove:SDM Logical Diagram

The actual physical system for our collaborative proof of concept was implemented using two Supermicro 2U 4-Node BigTwin systems.

- The first Host was running Red Hat OpenShift bare-metal with three nodes as control plane/compute nodes for the application host and a single node as a bastion node to access the control plane/compute nodes and provision them with the DNS server for the rest of the nodes and run the Kove:SDM management Console.
- The second 2U 4-Node Supermicro BigTwin system ran the Kove® software. Each node was equipped with 1TB of memory per node, enabling Kove:SDM to leverage the combined memory pool across the four nodes. This system will provide additional memory resources to any control plane/compute nodes as needed.
- In the PoC, we used one 2U 4-Node Supermicro BigTwin for the targets. Additional targets can easily be added, as represented in the diagram.

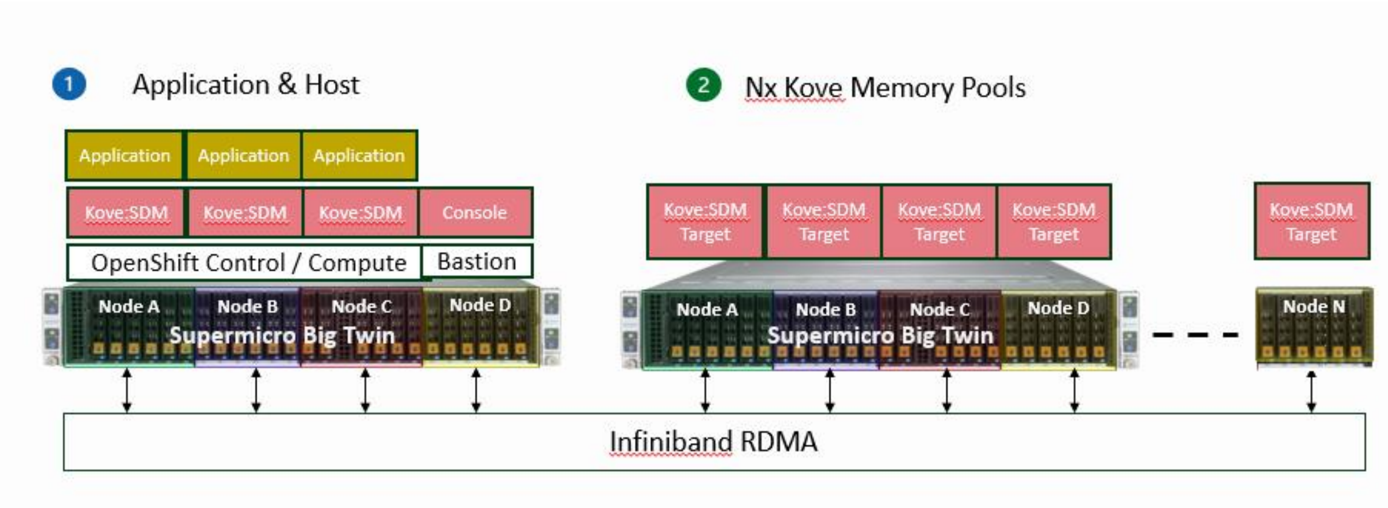

Figure 2 Proof of Concept Physical Diagram

1. Most reliable OpenShift cluster in a single chassis:

- The optimal number of nodes to create a small viable OpenShift cluster.
- Run once, run anywhere with Red Hat OpenShift.
- 1x bastion node to access the rest of the nodes and 3x control plane/compute nodes in a single chassis.
- Fully optimized containerized environment to build and deploy cloud-native applications.

2. High density for a larger pool of memory with the Kove® software:

- Four nodes are fully loaded with memory to provision more memory to client nodes.
- Kove® software running to monitor memory resources.
- Kove target memory will automatically provision memory to any workload if necessary.
- Workloads that need large amounts of memory will take full advantage of Kove's software-defined memory technology.

Proof of Concept Testing

During the testing phase, we subjected various scenarios to stress this environment using stress-ng. This tool facilitated targeted stress on the CPU and memory, enabling us to gather essential data for evaluating the compatibility and performance of Kove:SDM with Red Hat OpenShift.

Each scenario involved adjusting the CPU governor from performance to power save, with memory operated at different frequencies for each CPU governor setting.

The initial phase involved stressing the system using local memory, while the second phase involved reducing physical memory to allow dynamic memory allocation by Kove:SDM to the control plane/compute nodes requiring more memory than physically available. In total, the tests were executed with 16 scenarios, each repeated seven times to ensure accurate data collection across 2.165 quadrillion results averaged.

This proof of concept validated the seamless compatibility of Kove:SDM and Red Hat OpenShift in handling containerized workloads. Additionally, the test demonstrated the significant benefits offered by Kove:SDM in scenarios where workloads require memory capacities exceeding physical limitations while exhibiting minimal to no performance penalty.

Benchmark Results

Combined with Intel CPU governor settings, Kove:SDM provided 12 to 54% power savings, illustrated in Figure 3.

CPU Governor	Test	Memory MHz	Workload	Kove Benefit [Percent %]						
				Compute Node [64 GiB]				Kove	Totals	
				A	B	C	Total	1 TiB	Server + Target	256 → 64GiB
Performance → Powersave	1 → 2	3200	Idle	-34	-32	-32	-33	-26	-31	**-40%**
			Busy	-47	-46	-46	-46	-23	-42	**-50%**
Powersave	2 → 3	3200 → 2200	Idle	-5	-6	-7	-6	-6	-6	**-14%**
			Busy	-5	-6	-6	-6	-7	-6	**-14%**
Performance → Powersave	1 → 3	3200 → 2200	Idle	-38	-36	-37	-37	-31	-35	**-44%**
			Busy	-49	-49	-50	-49	-29	-45	**-54%**
Performance	1 → 4	3200 → 2200	Idle	-4	-7	-7	-6	-6	-6	**-14%**
			Busy	-3	-3	-2	-3	-6	-3	**-12%**

Figure 3 Power Savings with Kove:SDM

Conclusion

Kove:SDM is fully integrated with Red Hat OpenShift. The Supermicro BigTwin multi-node is the ideal and most reliable minimal Red Hat cluster. The Supermicro BigTwin system is very well adopted across industries, given its density and versatility, and is suited to broad applications. With the Supermicro BigTwin deployments, many nodes are typically concentrated in a single cloud instance. Many of these nodes will have spare available memory. Kove:SDM allows memory in nearby nodes to be dynamically allocated to a single Supermicro BigTwin node enabling it to perform very large in-memory workloads well over the memory capacity of a single node. The testing shows that memory can be dynamically and transparently assigned to a workload by Kove:SDM. The latency of memory across nodes is negligible compared to traditional memory management methods enabling applications that would otherwise be impractical given very long run times without Kove:SDM.

References

To learn more about Supermicro Server Infrastructure options, please visit:

- Supermicro Data Center Server, Blade, Data Storage, AI System
- Twin Servers: High-Density Multi-Node Server Solutions | Supermicro
- Networking Devices & Hardware Products For HPC | Supermicro
- Supermicro Solution for Red Hat® OpenStack | Supermicro
- AI Infrastructure Solutions | Supermicro
-

To learn more about Kove and Red Hat OpenShift, visit

- Enable Shared Memory Pools in Cloud, Edge, and On-Premise | Kove®
- Red Hat OpenShift enterprise Kubernetes container platform.

About Super Micro Computer, Inc.

Supermicro (NASDAQ: SMCI) is a global leader in Application-Optimized Total IT Solutions. Founded and operating in San Jose, California, Supermicro is committed to delivering first-to-market innovation for Enterprise, Cloud, AI, and 5G Telco/Edge IT Infrastructure. We are transforming into a Total IT Solutions provider with servers, AI, storage, IoT, switch systems, software, and services while delivering advanced high-volume motherboard, power, and chassis products. The products are designed and manufactured in-house (in the US, Taiwan, and the Netherlands), leveraging global operations for scale and efficiency and optimized to improve TCO and reduce environmental impact (Green Computing). The award-winning portfolio of Server Building Block Solutions® allows customers to optimize for their exact workload and application by selecting from a broad family of systems built from our flexible and reusable building blocks that support a comprehensive set of form factors, processors, memory, GPUs, storage, networking, power and cooling solutions (air-conditioned, free air cooling or liquid cooling).