SOLUTION BRIEF

DELIVERING THE NEXT GENERATION SERVICE DELIVERY PLATFORM FOR MISSION CRITICAL WORKLOADS AT THE EDGE

Complete Edge Cloud Platform Provides A Zero-Touch Deployable, Autonomously Managed, and Hardware Agnostic Solution for Running Workloads at the Distributed Edge

TABLE OF CONTENTS

Executive Summary 1

What types of workloads are deployed at the Edge? 1

Complete Edge Cloud Platform Provides a Zero-Touch Deployable, Autonomously Managed, and Hardware Agnostic Solution 3

Major Cruise Line Deploys Supermicro servers and NodeWeaver software 4

Benefits of the Supermicro/NodeWeaver Solution: 6

Summary 6

For More Information 6

Executive Summary

NodeWeaver is a software-defined “nano-cloud” platform that simplifies infrastructure and applications' deployment, management, and orchestration at the distributed edge. It installs on the bare metal of any x86_64 hardware. It enables the deployment of resilient and scalable compute clusters capable of running multiple applications, whether they run as VMs, containers, K8s, and/or real-time. Servers running NodeWeaver automatically combine at each edge location to form highly available compute clusters, delivering a cloud-native experience with reliable and scalable compute and storage for applications.

NodeWeaver executes workloads on as few as one or two or scales up to N number of servers, providing reliability and high availability without reengineering the application. Thanks to the integrated remote management and monitoring service, the customer can update and control all the deployments, even with ephemeral internet connectivity. The autonomous management system eliminates the cost of downtime related to software or hardware failures. NodeWeaver’s real-time mode allows even latency-critical workloads to run alongside traditional applications, removing the need for specialized hardware. The solution is optimized for virtualization, cloud, hosting, content delivery, hyper-scale/hyper-converged, and other general-purpose computing workloads.

What types of workloads are deployed at the Edge?

With advancements in artificial intelligence, deep learning, and machine learning technologies, companies are unlocking the value of data at the edge to improve insights and enable real-time responses and decision making. This leads to an explosion of

the number of devices at the edge, which generate massive amounts of data that are too large to transfer to the cloud. Sending data to the cloud for processing can mean increased latency, processing and storage costs, and security risk in transferring sensitive information through the entire network. Additionally, local compute resources are essential in situations requiring near real-time data processing and response.

Most solutions that include inferencing need to be updated frequently (or at least more frequently than traditional applications) because of changing environments, advances in state-of-the-art, and models that no longer perform the way they did when first deployed (model drift). This requirement drives the need to be able to remotely deploy, manage and monitor edge solutions that include an inferencing component.

Another unique aspect of these inferencing workloads is that they often run on heterogeneous computing platforms (GPUs, NPUs, FPGAs, and ASICs), and the infrastructure needs to be able to orchestrate, monitor and monitor these different computing devices.

In addition, a significant use case for Edge Computing is called Secure Access Service Edge (SASE). SASE is a architectural model that combines network and security-as-a-service functions together and delivers them as a single service. Conceptually, SASE extends networking and security capabilities beyond where they're typically available. This allows remote workers (or guests on ships at sea, for example) to take advantage of firewall as a service (FWaaS), secure web gateway (SWG), zero-trust network access (ZTNA), and a variety of threat detection functions.

SASE solves several problems that traditional network security solutions face, such as:

- High latency and poor performance due to backhauling traffic through centralized data centers or VPNs
- Increased complexity and cost of managing multiple point solutions for different security functions
- Lack of visibility and control over traffic and devices
- Inability to scale and adapt to changing business needs and threats

Additional SASE benefits include:

- Improving user experience and productivity by reducing latency and enhancing performance
- Simplifying network and security management by consolidating multiple functions into a single platform
- Enhancing security posture by applying consistent policies and protection across all users, devices, applications, and locations
- Reducing operational costs by eliminating the need for hardware appliances, VPNs, and MPLS circuits
- Increasing agility and scalability by enabling rapid deployment and adaptation to changing business requirements and threats

SASE deployment architecture consists of a global network of SASE points of presence (PoPs). The SASE PoPs provide network optimization, routing, encryption, and security services, such as firewall, secure web gateway, zero trust network access, identity and access management, and security broker.

Traditionally, these network and security functions were deployed as separate, dedicated hardware devices for each service, however today, these network services are run as Virtualized Network Functions (VNFs)

A typical enterprise might have dozens, hundreds or thousands of geographically distributed locations running SASE services – all critical to operations. Businesses must maintain maximum efficiency and resiliency for these services to optimize operations,

reduce downtime, and drive greater profit margins. Running as single, standalone appliances, these "traditional" platforms lack redundancy or the ability to scale as requirements grow. Consequently, when there is a failure, disruptions occur, resulting in downtime and disruptions to processes, slowing down business and increasing the cost of ownership. Additionally, the inflexibility of an appliance-based approach presents numerous operational challenges related to upgrading, scaling, and future proofing systems, so having an autonomous, easy-to-deploy-and-manage platform which enables the delivery of these applications in an efficient, scalable, and highly available manner is crucial to success.

Complete Edge Cloud Platform Provides a Zero-Touch Deployable, Autonomously Managed, and Hardware Agnostic Solution

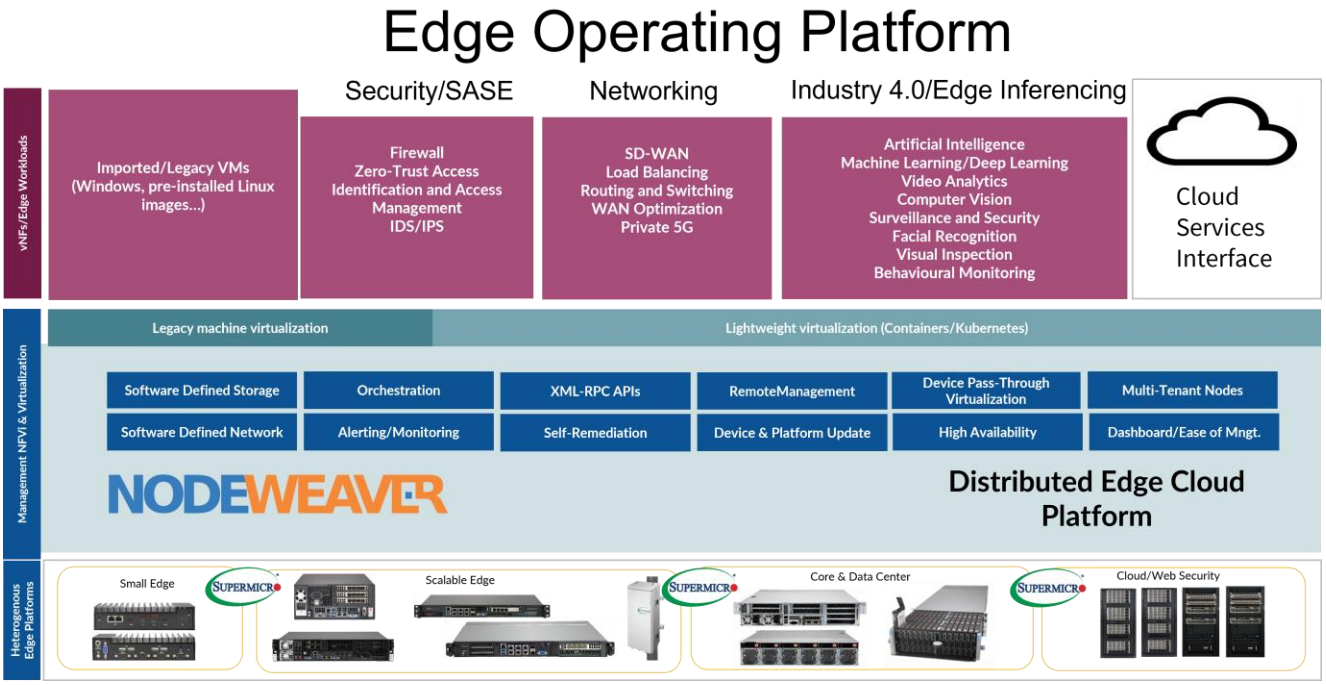


Figure1:Edge Operating Platform

NodeWeaver is an edge operating platform optimized for running any type of application (VMs, containers, K8s, realtime) on the distributed edge and manageable at mass scale. It is a fully integrated system that includes multiple hypervisors, software-defined storage, software-defined networking, orchestration, provisioning, monitoring, remote access, high availability, and remediation, all in a small footprint.

In addition to delivering cloud-native functionality at the edge, NodeWeaver addresses the challenges of deploying, operating, and managing infrastructure and applications at the distributed edge.

- Zero Touch Deployment: NodeWeaver provides an integrated, secure, and patent-pending zero-touch configuration system called DNSOps, which combines the power and ubiquity of DNS with a USB key pre-downloaded installer that brings any hardware from zero to a fully operating private edge cloud in minutes without requiring customization or user intervention. It enables the deployment of autonomously run compute clusters capable of running multiple virtual

machines and container-based workloads reliably and cost-effectively. One or more servers running NodeWeaver automatically combine at each edge location, delivering a cloud-native experience with reliable and scalable compute and storage for applications.

- Autonomous Operation: NodeWeaver's Autonomic Engine empowers edge nodes to handle issues and task execution without user intervention, even when connectivity is limited or non-existent. All the features of a cloud are available out of the box: high availability, load balancing, storage replication, software-defined networks, scripting, and API access. Businesses can significantly lower costs by reducing the need for IT expertise or physical intervention.

- Simplified Scalability: NodeWeaver adapts itself to the available hardware, so it is possible to create clusters with heterogeneous hardware without needing to adapt or configure anything. Live swapping means you can replace a faulty system with any available system: no need to be restricted to identical hardware forever. Users can mix hardware within a cluster of nodes or even replace a specific node with new hardware, and the system automatically recognizes the processor, memory, and storage and optimizes/rebalances the cluster to meet the needs of the applications. The extremely lightweight footprint of the NodeWeaver platform – requiring only one physical core and 1 GB of memory - means that you can select the optimal hardware for your use case without wasting resources.

Major Cruise Line Deploys Supermicro servers and NodeWeaver software

One of the world's largest cruise lines uses Supermicro servers and NodeWeaver as the foundational platform for a global distributed edge infrastructure solution across all ship and shore-based operations to support a SASE (Secure Access Service Edge) rollout across their enterprise. The customer needed to have virtualized network functionality globally across all of its ships and in land based locations.

Challenge

The customer used the legacy approach of deploying standalone physical appliances to deliver network functionality such as firewalls, load balancers, access/management, and WAN optimization. These physical appliances are:

- Expensive – the proprietary nature of purpose-built hardware adds significant cost.
- Don't scale – additional capacity or compute power requires replacing with a new appliance.
- Consume rackspace and power, and require cooling
- Inflexible – proprietary appliances only do one thing.
- Represent a single point of failure – if there is a hardware failure, services go down until another hardware appliance can be shipped. Or, the customer has to purchase 2x of each appliance for redundancy.

Solution

Nearly all modern network functions can be run as virtual instances (VMs and/or containers). By virtualizing multiple network functions and consolidating them onto a single platform, the customer was able to achieve several significant benefits:

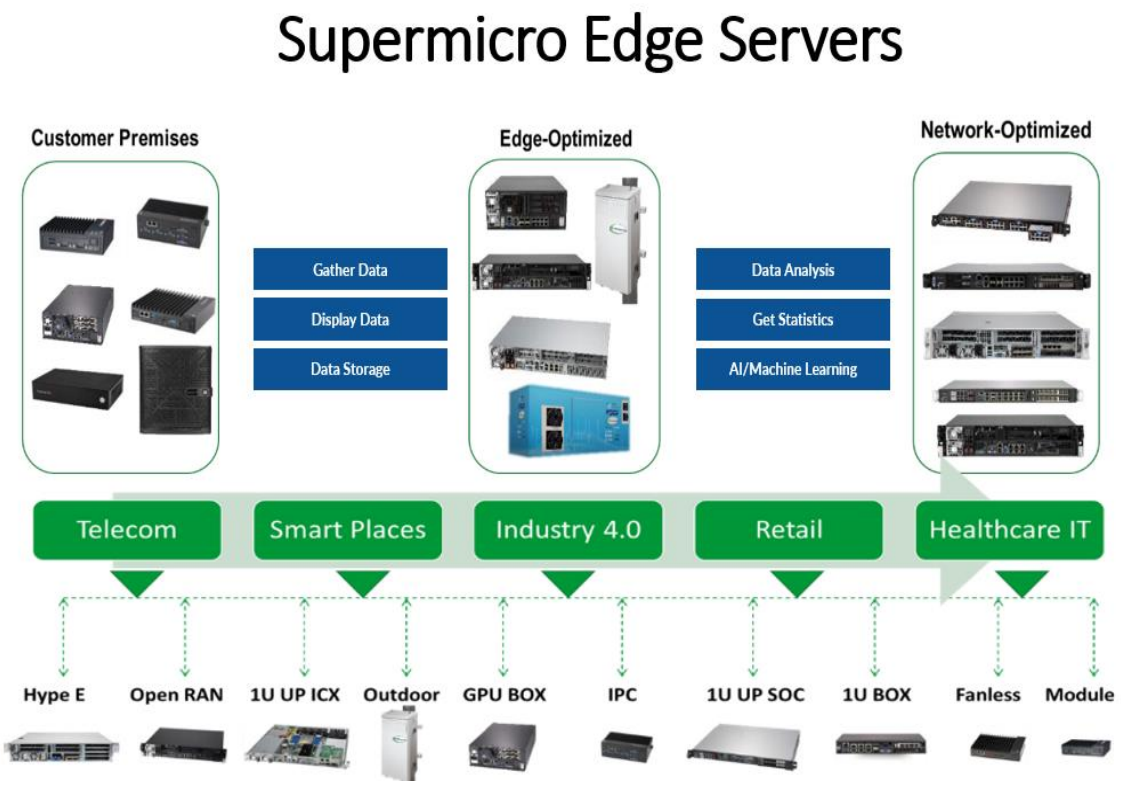
- Efficiency – no longer must run multiple standalone appliances, each occupying space, consuming power, and generating heat.
- Resiliency – by running as a highly available, two node cluster, the VNFs will automatically restart on the surviving node in the event of a hardware failure. Given that these ships are at sea for days or weeks at a time, resiliency is critical as replacement parts aren't readily available.
- Scalability – additional nodes can be added, one at a time, as they grow and add additional services
- Flexibility – clusters can be comprised of heterogeneous hardware, so they don't have to worry about maintaining a spares pool of different servers over time.

- Management – Deploying, managing, and operating many geographically distributed locations at scale requires capabilities such as zero-touch deployment, autonomous operation, self-healing, self-optimization, remote management, and monitoring.

While the initial use is focused on VNFs, this is a universal service delivery platform that acts as the foundation for delivering many other high-value services in the future across their entire global footprint. A vital example of this is Artificial Intelligence (AI). AI is being used today by companies of all sizes across all industries, and these applications are based on the collection and analysis of vast amounts of data from various IoT devices, such as video cameras and sensors. To take advantage of the actionable insights these AI applications can deliver, appropriate compute infrastructure must be in place to store and analyze this data. With Supermicro and NodeWeaver, the customer has a future-proof platform to run any application and incorporate any type of specialized processor, such as GPUs. The critical criteria for this project were the ability to create flexible, scalable, and highly available nano-clouds at the edge, which autonomically addresses failures, and the ability to easily manage hundreds of distributed locations worldwide with minimal human intervention. After evaluating other solutions in the market - Supermicro server PN: SYS-110P-FRN2T, which integrates compute, storage, and networking on a commodity platform, was selected. Supermicro and NodeWeaver were selected due to their simplicity, flexibility, reliability, and time to value.

An Example of Edge Products: Supermicro SYS-111E-FDWTR

Supermicro has the most extensive edge portfolio in the industry. An example of a core edge platform is the SYS-111E-FDWTR, one of the most powerful systems in its form factor. This compact system supports the latest 4th Gen Intel® Xeon® Scalable processors (up to 32 cores - 205Wor 350W with Fan Upgrade) and can support up to 2TB in DDR5 memory. This 1U short depth system measures in at 16.9" (D) x 17.2" (W) x 1.7" (H), weighing in roughly over 15 lbs. This IoT system has front I/O network options for you to use the onboard 2 10GbE LAN ports with options of adding three additional front I/O add-on cards supporting 2 X PCIe 5.0 full height full length and 1X PCIe 5.0 HHHL. Key applications for this server include multi-access edge computing, flex-RAN, Open-RAN vBBU, artificial intelligence (AI) on edge, and machine learning.



Benefits of the Supermicro/NodeWeaver Solution:

- Zero touch deployment capability enables customers to deploy the full environment with all applications and network settings without requiring skilled personnel. This is highly important given that these are being deployed on ships and other land-based locations worldwide.
- Utilizes the latest technology with 4th Gen Intel® Xeon® Scalable processors (up to 32 cores - 185W) and can support up to 2TB in DDR5 memory with onboard 2 X 10GbE BaseT ports for network connectivity.
- Network services are resilient to failure since they are virtualized on this platform.
- Highly resilient 2-node clusters that run multiple VMs can replace multiple dedicated appliances at customer sites and provides a scalable platform on which other high-value services can be delivered in the future.
- Front I/O makes it easy to deploy.
- Redundant 800W AC Power Supplies with Power-on mode for AC power recovery help to keep the platform running without disruption.
- The server supports the creation of nano-clouds at the edge, providing redundancy and continued service without disruption.

Summary

In today's world, processing large amounts of data at the edge is critical in supporting application demands at the edge. Processing data closer to where it is being generated decreases latency, increases processing speed, and improves results for real-time applications. Supermicro's building block solutions enable us to provide you with a wide selection of products to satisfy your requirement, delivering the best solution fully optimized with low latency and high availability.

Simplicity, flexibility, reliability, and time to value from Supermicro and NodeWeaver provide a great Edge Cloud Platform to meet today's challenges and requirements at the Edge.

For More Information

Supermicro : https://www.supermicro.com/en/products/system/iot/1u/sys-110p-frn2t

Nodeweaver: https://nodeweaver.eu

SUPERMICRO

As a global leader in high performance, high efficiency server technology and innovation, we develop and provide end-to-end green computing solutions to the data center, cloud computing, enterprise IT, big data, HPC, and embedded markets. Our Building Block Solutions® approach allows us to provide a broad range of SKUs, and enables us to build and deliver application-optimized solutions based upon your requirements.

Visit www.supermicro.com