SUSE

White Paper

Future-Proof Telecom Modernization with an Open and Flexible Infrastructure Platform

SUSE has built an adaptable edge computing infrastructure platform to support the telecommunication applications of today and tomorrow.

SUSE Adaptive Telco Infrastructure Platform (ATIP) is a massively scalable management solution for telco-grade cloud native bare metal and cloud edge infrastructure. It allows zero-touch secure device onboarding as well as automated bootstrap and lifecycle management of operating systems, Kubernetes clusters and applications. SUSE ATIP is based on SUSE's profound experience with edge customers across various industries and adds telco-grade features and performance in the relevant elements of the system architecture.

Edge is predominantly a management at scale challenge

While in the typical large data center the number of Kubernetes clusters under management is typically in the range of several tens or at most hundreds, the number easily goes into the thousands, tens of thousands or even millions at the edge – and this is a general trend across all industries, including the Telco industry. This requires a shift away from "manual" cluster management with a graphical user interface, towards an API-driven programmatic approach.

With Rancher Prime SUSE provides a market leading Kubernetes management solution that is famous for its simplicity, robustness and outstanding user experience.

To address edge's management at scale challenge, we are making two key additions to Rancher.

1) **Vendor-neutral programmatic edge cluster lifecycle management.** We integrate CNCF's Cluster API to offer a vendor-neutral integration point for programmatic management of the entire edge stack, including bare-metal and private cloud infrastructure, operating system, Kubernetes clusters and applications. Extensions of Rancher's graphical user interface make handling of edge infrastructure easy and provide intuitive visualizations to understand operational state, resource utilization and inventory. Under the hood, the entire edge infrastructure is managed in the same declarative way that made Kubernetes so successful in the first place. Every management operation is done through creation and manipulation of Kubernetes Custom Resources on a central management cluster.
2) **Gitops based infrastructure management.** The scale of edge requires a more efficient and rigorous approach to managing desired infrastructure state. GitOps is emerging as the goto solution to this problem. Rancher's GitOps component Fleet, in combina-

tion with the abovementioned adoption of Cluster API, enables a GitOps based workflow for managing edge infrastructure. Desired state - such as, which OS and Kubernetes version are used, how many control plane and worker an edge cluster should contain, which ingress controller and other additional Kubernetes components should be deployed - is defined and stored as YAML files or Helm charts in a Git repository. SUSE ATIP can of course also be combined with 3rd party GitOps tools such as Flux or ArgoCD.

Using the well-established workflows of version management through Git provides a very robust approach for managing state definitions.

- Changes can be peer-reviewed before they get merged,
- Automated linting and compliance checks can be implemented through well-understood CI/CD principles and toolchains,
- Complete history of changes is available by default, and
- Rollbacks are easy and 100% predictable.

The GitOps engine then takes care of synchronizing the state definitions from a Git repository to the management cluster, which turns them into effective declarations of wanted state. The SUSE Edge management stack then reconciles actual state of edge infrastructure with the declared wanted state, for example by rolling out operating system and Kubernetes clusters to the bare metal hosts in a newly onboarded edge site, changing the configuration of an edge cluster or installing an application update on a group of edge clusters simultaneously.

Built for Edge from the ground-up, so you can get optimal performance and integrated security at minimum cost

Container optimized OS: The role of the Linux operating system in a Kubernetes

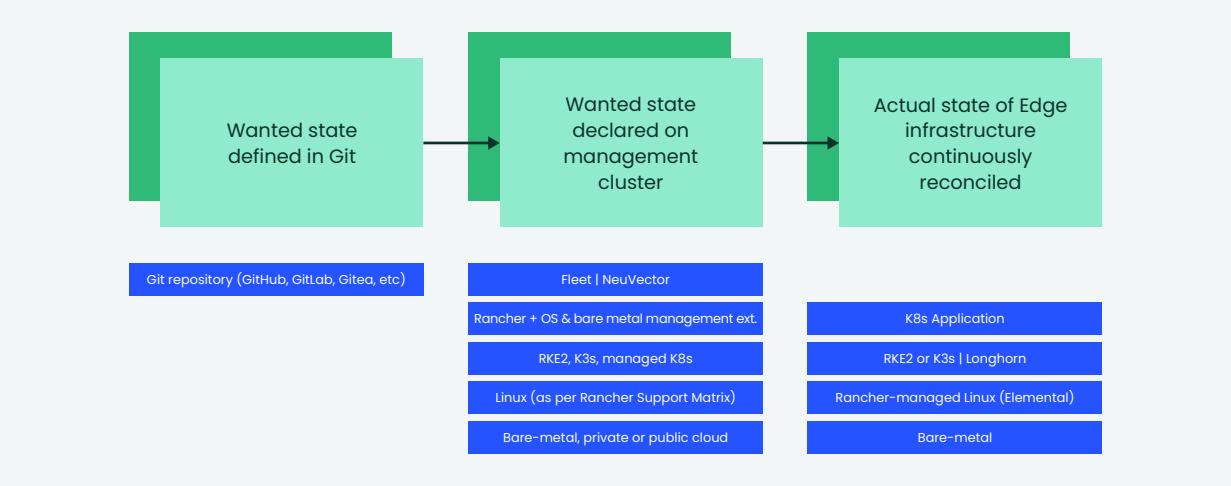

Figure 1: GitOps based workflow for managing Edge infrastructure and how it maps to the SUSE Edge system architecture

stack requires only a small subset of the functionality typically provided by enterprise Linux distributions. SUSE Linux Enterprise (SLE) Micro has been designed specifically to serve as a host OS within a Kubernetes stack.

Secure. A container-optimized OS is also more secure, as it has a much smaller attack surface. In addition, SLE Micro has an immutable root file system, making it virtually impossible for attackers to modify system settings, add or replace binaries in standard system paths.

Reliable. SLE Micro's transactional update mechanism makes updates atomic, i.e., they either succeed or fail, in which case the system automatically gets rolled back to its previous state. Situations where the system is in a corrupt state after a failed update do no longer occur.

Zero-trust cloud native security with NeuVector. Our market leading cloud native zero-trust security solution NeuVector does not only provide traditional vulnerability scanning and security policy compliance tools, but its unique capabilities to observe and police the run time behavior of cloud native applications (which is also known as zero-trust security) provide an unprecedented level of security against a much larger set of attack vectors that would remain undetected by traditional security tools.

Telco-grade OS and Kubernetes At The Edge Are Made Easy and Reliable

Cloud Native Network Functions (CNFs) have special needs in terms of high packet throughput and reliable access to system resources such as CPU and memory. Sometimes applications need to be scheduled for specific hardware nodes within a cluster since they require access to GPUs, accelerator boards or similar devices.

Addressing these needs goes beyond the typical capabilities and configurations of off-the-shelf enterprise Kubernetes and Linux. **SUSE ATIP is designed to include a set of dedicated technologies to facilitate the typical use cases of the telco industry.**

Node Feature Discovery

Node Feature Discovery is a Kubernetes add-on developed by the Cloud Native Computing Foundation (CNCF). It detects features of the underlying hardware and labels Kubernetes nodes accordingly, so that the Kubernetes scheduler can consider these features when scheduling workloads. Node Feature Discovery is available in the Rancher Manager Cluster Tools and can easily be installed on each downstream cluster, either manually through the GUI or via the cluster YAML definition.

Multiple IP addresses per Kubernetes pod with Multus Container Network Interface (CNI)

Vanilla Kubernetes supports exactly one IP address per Kubernetes pod. The Multus CNI allows assigning multiple IP addresses to Kubernetes pods. Multus only works in combination with an additional CNI that serves normal intra-

cluster networking such as connectivity between control plane and worker nodes. Whereabouts is used to ensure overlap free IPAM for Kubernetes pods when using Multus. Multus is a supported CNI in Rancher Kubernetes Engine (RKE2) and can automatically be installed at cluster creation. Whereabouts can be installed at the same time via a feature flag in the cluster definition.

Single Root I/O Virtualization (SR-IOV)

SR-IOV is a technology that allows a single physical PCIe device to appear as multiple virtual PCIe devices. It is typically used to allow direct access to network interfaces from Kubernetes Pods without the overhead of Kernel interrupt handling, but can support other types of PCIe devices too (e.g., accelerator boards, FPGA boards, etc.). The SUSE ATIP solution uses the upstream Kubernetes SR-IOV operator, usually in combination with the Multus CNI, to create and expose SR-IOV VFs to Kubernetes pods. Rancher can deploy the SR-IOV operator to downstream clusters through the cluster tools marketplace and/or the Helm chart controller at install time. The operator supports creating SR-IOV Virtual Functions (VF) through Kubernetes Custom Resources.

Telco-grade Performance Is Made Available

CPU Pinning and IsolCPUs

Telco workloads need to process high volumes of user traffic and carry out computationally heavy operations with deterministic performance. A normal userland process in Linux does not receive fully deterministic performance at all times, since its execution speed can be impacted by the process being moved to different CPU cores by the Linux kernel scheduler, and may have to share access to physical system resources with other processes or the kernel itself. The most important feature in this context is CPU Pinning, which ensures that a process is always scheduled to a particular CPU core, avoiding cache misses and other performance impacting effects. A related feature is CPU Isolation (also known as IsolCPUs), which prevents other processes (including the kernel itself) from running on that same CPU core. The Linux kernel used in all SUSE Linux variants supports both features - CPU Pinning and IsolCPUs.

Huge Pages

Telco CNFs require low-latency access to large amounts of memory when reading from and writing to SR-IOV Virtual Functions for packet I/O. Huge Pages significantly reduce memory page lookup times and Transaction Lookaside Buffer (TLB) misses. The Linux kernel used in all SUSE Linux variants has support for huge pages built in. Kubernetes automatically discovers huge pages and provides them as schedulable resources the same way as standard memory and CPU requirements.

Topology Aware Scheduling in Kubernetes

Sometimes it is important to make sure a Kubernetes workload has the right set of resources allocated across the vari-

ous resource types (CPU cores, Memory, PCIe bus). For this purpose, Kubernetes provides the concept of QoS classes and resource requests.

Due to the Non-Uniform Memory Access (NUMA) architecture of modern multi-processor systems, certain resources can be accessed with higher bandwidth from certain CPU sockets and cores. In simplified terms, this can be imagined as certain resources being "closer" to each other than others. In other words, the resources in a system have a *topology*.

Topology aware scheduling, such as putting NICs exposed as SR-IOV virtual function into the same NUMA zone a given Kubernetes pod runs in and making sure to allocate huge page memory from that zone, is an important element in guaranteeing that performance-sensitive applications can run successfully on Kubernetes.

As of Kubernetes version 1.22, the Kubernetes Memory Manager and the CPU Manager support generating NUMA-aware affinity hints for the Kubernetes Topology Manager, which then allocates pods to resources within a Kubernetes worker node based on those hints and its configured allocation policies. So, setting up NUMA awareness correctly is merely a matter of configuring the various layers of the system correctly – including the application manifests of course. The behavior of these managers can be configured through kubelet flags, which can be freely configured through Rancher when creating new clusters.

Real-Time Kernel

Some telco applications require guarantees in terms of how long it takes to get scheduled after a triggering event has taken place. In a normal Linux system, the delay between a trigger event (e.g., a software interrupt) depends on everything else running on the system and is therefore largely non-deterministic. Such systems may provide sufficient performance, but the performance is not guaranteed. The PREEMPT_RT patch for the Linux kernel provides the ability to schedule tasks in a way that reliable upper bounds for getting access to a CPU core can be calculated. A Linux kernel including the PREEMPT-RT patch is typically referred to as a realtime kernel. SLE Micro provides the option of using the realtime flavor of the SUSE Linux Enterprise kernel.

Telco specific protocols and precision timing are supported

Precision Time Protocol and Synchronous Ethernet Support

When multiple local systems, such as the bare metal hosts in an edge Kubernetes cluster, need to be time synchronized in the sub-microsecond range, typical timing protocols like Network Time Protocol (NTP) are not sufficient. For this purpose, the IEEE standard IEEE1588, also known as Precision Time Protocol (PTP), has been created. SUSE Linux Enterprise provides packages for all relevant PTP implementations available in Linux.

Synchronous Ethernet is an ITU-T standard (defined in ITU-T Recommendations G.8261, G.8262 and G.8264) for transmitting clock signals over the Ethernet physical layer. It is used to provide a time synchronization source to all nodes in an L2 broadcast domain. Synchronous Ethernet (SyncE) support in the Linux kernel is currently underway, with an expected availability in the 6.2 release, which is expected in the first half of 2023. SUSE will then backport the code into the SUSE Linux Enterprise kernel.

Stream Control Transmission Protocol (SCTP) Support

SCTP support is available as a stable feature in Kubernetes as of 1.20. This requires Kubernetes to be deployed with a CNI that supports SCTP. Within the range of supported CNIs in RKE2, CNIs supporting SCTP are Calico and Cilium.

Support for Data Plane Development Kit (DPDK) Based Container Workloads

Currently SLE Server 15 and SLE Micro 5.3 support DPDK 1.19. Starting with SLE 15 Service Pack 5, which is targeted to be released in Q2/2023, SUSE will support DPDK 1.22.x. Other DPDK versions are available in the SUSE Package Hub and in openSUSE.

Summary

Our vision for the next generation of telecom infrastructure is that 5G and Edge computing will enable a new class of novel applications and use cases. SUSE Adaptive Telco Infrastructure Platform (ATIP) is purpose built to enable telecom operators in transforming their networks by delivering flexible, adaptable infrastructure to support this new class of applications. SUSE helps solve the challenges of managing large numbers of diverse and varied clusters scattered across bare-metal infrastructure, public and private clouds. We do that while delivering the feature set and security posture you would expect from a vendor who has been delivering business-critical Linux solutions for more than 30 years.

To learn more, visit Telco Industry Solutions, Edge Solutions

Or

Contact us at https://www.suse.com/contact/

SUSE

2023

SUSE Software Solutions
Germany GmbH

Frankenstraße 146
90461 Nürnberg
Germany

www.suse.com

For more information, contact SUSE at:

+1 800 796 3700 (U.S./Canada)

+49 (0)911-740 53-0 (Worldwide)

Thank You