Solution Brief

Communications Service Providers
Service Assurance

Intracom Telecom Machine Learning Boosts Energy Efficiency of Red Hat OpenShift NFV Workloads

The Intracom Telecom NFV Resource Intelligence™ (NFV-RI™) and Intel® Xeon® Scalable processors with Enhanced Intel SpeedStep® Technology predict virtual network function traffic loads to optimize power savings. Intracom Telecom has certified NFV-RI for Red Hat® OpenShift® to provide a streamlined path to deployment—including for 5G networks—offered through the Red Hat Ecosystem Catalog.

Table of Contents

The benefits of network functions virtualization (NFV) are well understood by communications service providers (CoSPs) and are a key enabler as mobile networks expand from network core to edge. Many CoSPs have reported significant total cost of ownership (TCO) reductions.

But at the same time, there are important challenges. In contrast to legacy non-virtualized IT workloads, cloud-native network functions (CNFs) based on the Data Plane Development Kit (DPDK) or Vector Packet Processing (VPP) stack have very stringent key performance indicators (KPIs) that govern deterministic performance, low latency, and zero packet drops. One effect of this stringency is that the CPU in a server will always operate at the highest frequency to provide the requisite performance for these NFV workloads.

Specifically, a property of CNFs based on DPDK or VPP is that they keep the servers constantly running at a high power-up state, as if the servers are always operating at peak demand. The reason is that DPDK and VPP rely on polling mechanisms to achieve zero drops, low latency, and high throughput in packet processing. This polling keeps CPUs at the highest utilization at all times, even during idle periods, thus consuming the maximum possible power and elevating operational expense (OpEx).

DPDK and VPP enable fast networking functions and accelerated packet processing. Existing system software such as the Linux operating system feature power-management policies to dynamically manage power for CPUs when they are not busy, but because DPDK and VPP CNFs always appear to be at 100 percent CPU utilization, the OS is unaware of their actual load at any given time. Finding ways to leverage knowledge of the actual CPU utilization allows the implementation of mechanisms to adapt CPU frequencies according to the CNF's load.

Intel® Network Builders ecosystem partner Intracom Telecom has added a frequency feedback loop (FFL) workflow to its Network Functions Virtualization Resource Intelligence™ (NFV-RI™) solution to predict CNF traffic levels and dynamically adjust the frequencies of each CPU core used by DPDK and VPP CNFs according to their incoming load, while promoting zero packet drops.

Intracom Telecom NFV-RI Overview

NFV-RI is a resource intelligence platform that runs on Intel architecture-based servers and delivers enhanced and deterministic performance for CNFs, increased server density, and optimal utilization of NFV infrastructure resources. To this end, it leverages key technologies such as Intel Resource Director Technology (Intel RDT) and Enhanced Intel SpeedStep Technology to control performance-critical resources including last-level cache (LLC), DRAM bandwidth, and CPU frequency for each of the co-located workloads. Resource decisions are driven by a variety of AI-powered workflows, ranging from simpler grid-search explorations to more advanced, dynamic, and continuous optimization processes based on self-learning cognitive agents.

NFV-RI Frequency Feedback Loop (FFL)

To overcome the power waste associated with DPDK and VPP CNFs constantly running at a high-power state, even when not necessary, Intracom Telecom has added the frequency feedback loop (FFL) workflow to its NFV-RI. The FFL uses machine learning to predict CNF traffic levels to enable reduced power usage during light or off-peak traffic periods, without compromising performance. The solution can dynamically adjust the frequencies of CPUs processing DPDK and VPP CNFs according to their incoming load, while helping meet the goal of zero packet drops, in a fully automated way.

FFL has a real-time, closed-loop mechanism that adapts the frequency of the cores the CNF is running on to match their actual traffic load, while helping ensure that frequency is high enough to support zero packet drops. This means that CNFs are operating at maximum CPU frequencies during peak hours, and moderate or minimum frequencies during off-peak or light traffic hours.

FFL Working with Enhanced Intel SpeedStep Technology

FFL makes use of Enhanced Intel SpeedStep® Technology in Intel processors to dynamically alter the operating frequency. This feature allows the system to dynamically adjust processor voltage and core frequency, decreasing average power consumption and heat production. Combined with existing power-saving features, Enhanced Intel SpeedStep Technology can help operators achieve a sophisticated balance between performance requirements and power consumption. Enhanced Intel SpeedStep Technology uses design strategies that include separation between voltage and frequency changes and clock partitioning and recovery.

FFL Machine Learning

At the heart of the FFL is a machine learning module that predicts how busy CNFs will become in the next short-term period. To achieve this, it uses network interface card (NIC) statistics, CPU event counters, and metrics from the CNFs themselves, from current and previous time windows. Therefore, depending on the current power level and the predicted busyness level, the mechanism proactively scales a CNF's frequencies up to prevent packet drops or down to save power. FFL handles two types of traffic variations using machine learning techniques:

- **Predict imminent overload situations** (CNF capacity saturation while operating at a certain CPU frequency) and scale up frequencies well in advance, at a suitable level.
- **Detect underload situations** (operating with more CPU frequency than is sufficient to sustain a certain traffic level) and scale down frequencies at a suitable level, but in a more gradual manner.

In case of abrupt, sudden traffic bursts that are impossible to predict, FFL reacts fast at sub-second, user configurable intervals, setting the CPU frequencies to their maximum level directly.

FFL is able to handle open CNFs, those that expose their busyness in terms of the load on receive (RX) queues via the DPDK Telemetry API, or through custom RESTful element managers. FFL can also work with closed CNFs that do not expose anything. In the latter case, FFL employs machine learning techniques to infer the busyness level of a CNF indirectly, leveraging platform metrics. Because this approach lacks the observability of the open CNFs, in most cases the FFL is forced to operate more conservatively (with higher frequencies), accepting power-optimization losses to avoid packet loss.

Three Phases of FFL

Figure 1 shows the three phases of FFL operation—configuration, training, and closed-loop operation. In the configuration phase, the user provides all the parameters needed for configuring FFL in its subsequent phases, such as the suitable data sources and the mapping of CNFs to FFL instances. Note that with FFL, the user can simultaneously run multiple independent instances of the feedback loop on the same platform, with each instance controlling a group of one or more CNFs. The CNFs within a group are typically tightly coupled, such as a chain of packet processing functions that accept their own traffic. As such, the decisions for scaling CPU frequencies are made independently for them, according to their current traffic load.

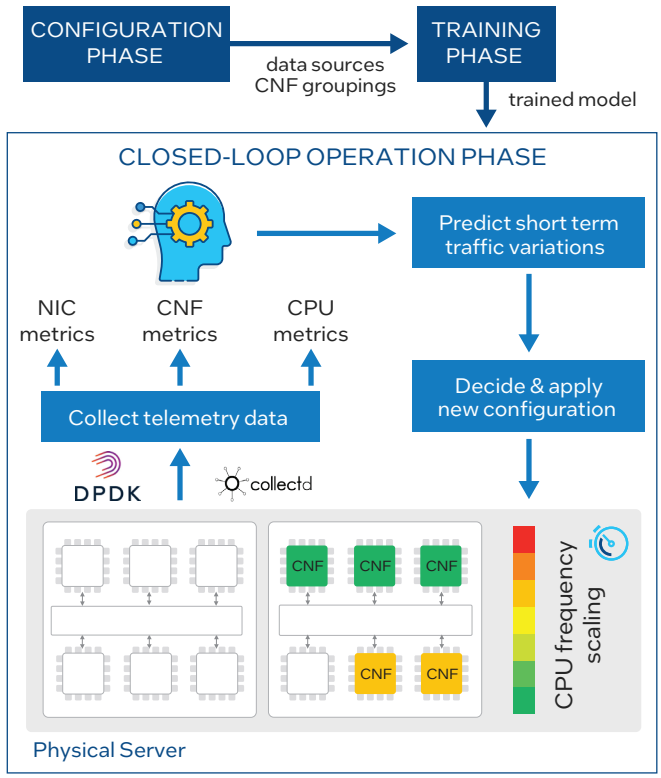

Figure 1. FFL operation at a glance.[1]

In the training phase, FFL is trained to find ideal frequencies for different traffic levels. The ideal CPU frequency for a certain traffic level is the minimum CPU frequency that yields zero packet drops. In general, the higher the traffic levels, the more CPU frequency is required to keep packet drops at zero. Note that different CNFs may have different ideal frequencies, even for the same traffic and on the same platform. In this phase, the user is initially asked to specify the maximum known capacity of the CNF, which corresponds to the maximum amount of traffic that the CNF is able to sustain in production without errors.

Subsequently, FFL will attempt to identify the ideal frequencies for a certain number (N) of traffic steps, uniformly distributed in the range between 0 and Cmax, asking the user to feed CNFs with the corresponding traffic rate in each step. The outcome of this phase is a trained model able to predict how traffic will vary in the subsequent short-term period and decide the best CPU frequency for the predicted level.

In the closed-loop operation phase, the user launches one or more FFL instances on the machine, specifying which trained model should be used for each. This phase may run indefinitely in production. It dynamically detects changes in traffic load and decides automatically which CPU frequency should be applied to the CNFs.

Burst Tolerance

Intracom Telecom has evaluated[2] how effective FFL is in handling sudden traffic bursts in an effort to understand the main factors that affect it. Intracom Telecom has used the sample DPDK L2 forwarding application[3] extended with busyness metrics such as those of vEPC (e.g., full polls and empty polls). The packet size is 64 bytes. The packet rate ranges between about 3.9 Mpps (2 Gbps) and about 31.2 Mpps (16 Gbps).

The key conclusion is that the CNF's tolerance to packet drops on sudden bursts depends not only on the burst size, but also on the starting and ending traffic levels. Per Table 1, the tests show that significant bursts can be processed between 2 to 4 and 2 to 10 Gbps, and this is an acceptably large burst. In scenarios when bursts are even larger than that, for example a burst size of 10 Gbps that yields about 60 drops when starting from 2 Gbps, but yields about 20,400 drops when starting from 6 Gbps. The larger the ending traffic of the burst is, and the smaller the current CPU frequencies are, the faster the RX queues, including network interface card queues, fill up, resulting in packet drops.

Table 1. Number of packet drops during sudden bursts. Left column shows bursts in terms of starting data flows and ending data flows in Gbps. Right column shows number of packets dropped during the burst.

BURST (GBPS)	AVG DROPS
2 -> 4	0
2 -> 6	0
2 -> 8	0
2 -> 10	0
2 -> 12	60
2 -> 14	6,300
2 -> 16	22,000
6 -> 8	-
6 -> 10	-
6 -> 12	-
6 -> 14	3,200
6 -> 16	20,400
10 -> 12	-
10 -> 14	-
10 -> 16	291
14 -> 16	-

FFL Improves Energy Efficiency at Tier-1 CoSP in Greece

Intracom Telecom collaborated with a Tier-1 CoSP in Greece in order to optimize the energy consumption in user plane components of the CoSP's mobile network.

The top portion of Figure 2 shows a representative downlink traffic pattern and the server power consumption for a virtualized EPC prototype Intracom Telecom evaluated, reproducing real traffic patterns measured in the Greek Tier-1 CoSP's mobile network during an average 24-hour period. According to the top right plot, by default, the power consumption on the server running the CNFs is constantly high, even though traffic varies during the day.

For light, off-peak traffic periods (for example, between 22:00 and 08:00), this suggests overprovisioning, because CNFs could be operated at lower frequencies, consuming less power, without experiencing packet drops or increased latencies.

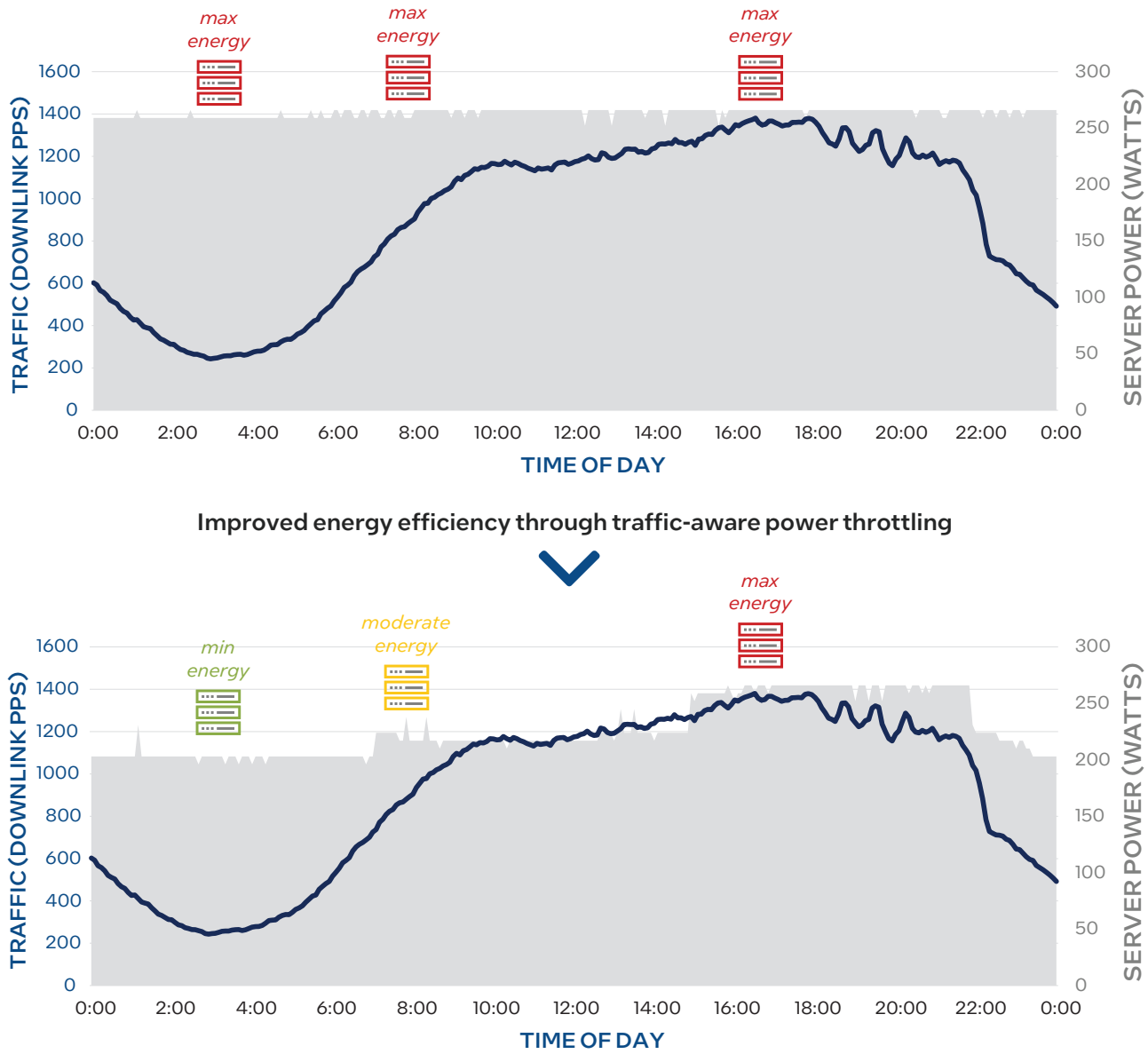

Figure 2. Reduced power consumption of the server using FFL.

As illustrated in the bottom portion of Figure 2, Intracom Telecom's FFL was able to adjust the power consumption of the vEPC server in a way that consistently followed the traffic pattern.

Certified for Cloud-Native Infrastructure on Red Hat OpenShift

Intracom Telecom is expanding its outreach to CoSPs for cloud-native infrastructure by validating NFV-RI on Red Hat OpenShift, which is illustrated in Figure 3. OpenShift enhances Kubernetes with continuous integration/continuous delivery (CI/CD) pipeline tools, automation, and security capabilities. CoSPs benefit from testing, hardening, and integration with other Red Hat open source offerings. The platform is built for multicloud environments, and enables clusters to be deployed on demand in the data center, at the edge, or on public cloud infrastructure.

OpenShift is designed to provide a consistent operating environment from on-premises to the cloud and edge, for running and managing workloads across hybrid and multi-clouds. The platform can be deployed using any combination of cloud, bare-metal, and virtualized infrastructure. It can also be consumed as a managed service from the major public cloud service providers.

To help boost operational efficiency, OpenShift provides automation across the stack, with platform services to manage workloads, application services for building cloud-native applications, and developer services to enhance developer productivity. Self-service provisioning for developers streamlines the software lifecycle, from development to production. Advanced orchestration and workload management respond effectively and quickly to fluctuating system utilization, in tandem with power-management functions enhanced by Intracom Telecom NFV-RI.

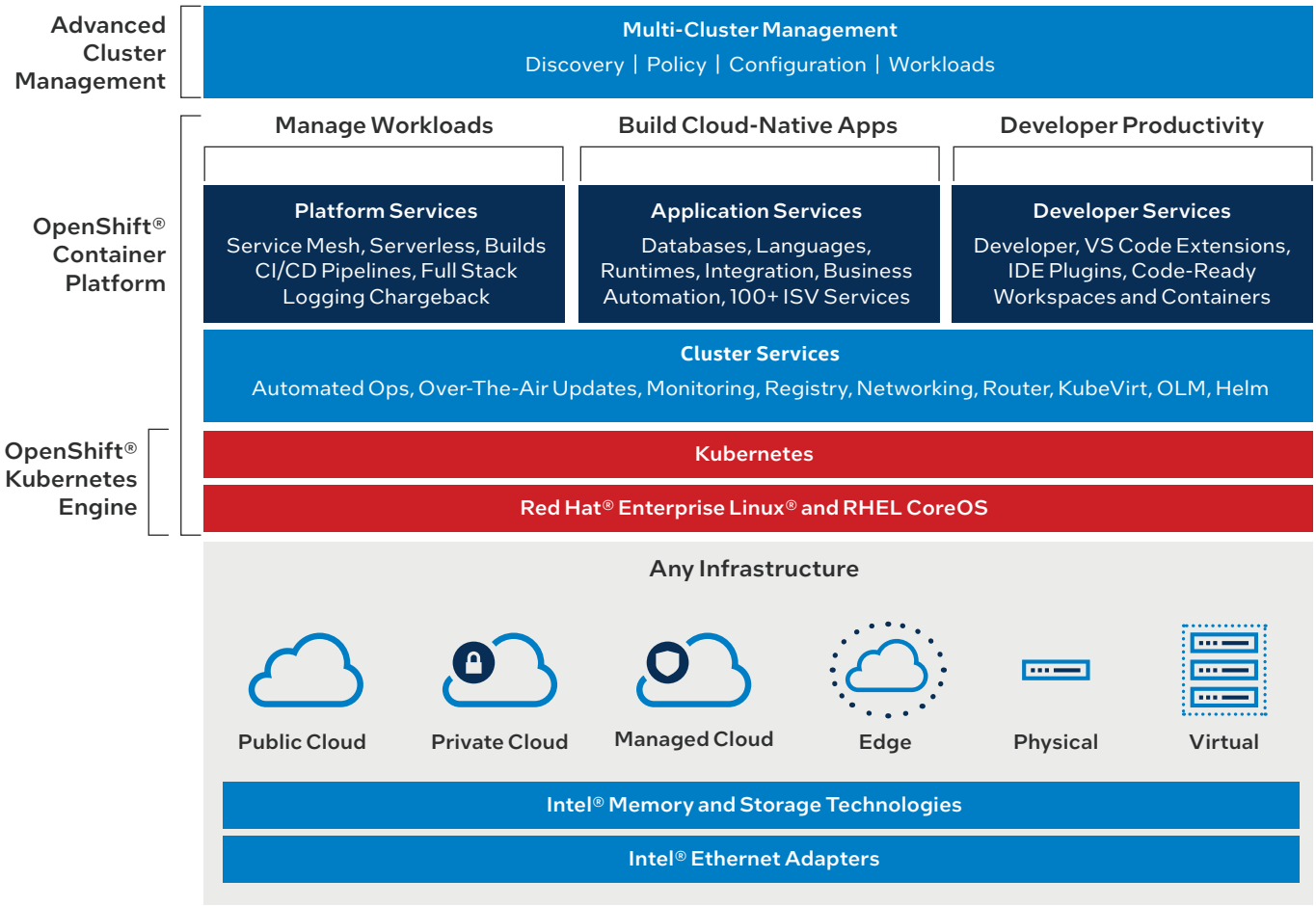

Figure 3. Red Hat OpenShift builds services and capabilities on top of Kubernetes.

Red Hat leadership for the cloud-native ecosystem includes enablement through open source initiatives and with commercial software vendors. This work helps facilitate and validate the stable interoperation of various software components with the OpenShift platform to reduce risk and simplify implementation by end customers. To help streamline deployment of NFV-RI on Red Hat OpenShift, Intracom Telecom offers container images and a Helm chart that have been certified by Red Hat through the Red Hat Ecosystem Catalog.

The certification provides assurances to customers that the implementation is stable and enterprise-ready. Helm is the package manager for Kubernetes, and Helm Charts are bundles of Kubernetes manifests used to define, install, and upgrade applications. They provide pre-validated single-command installation of NFV-RI for OpenShift.

Managing Server Power on 5G User Plane Workloads

OpenShift clusters provide highly automated operation with 5G CNFs, and NFV-RI enables robust power management for polling workloads associated with those NFVs. For example, hardware resources associated with the 5G user plane function (UPF) can be considered in three general categories: idle resources, resources processing the control plane, and resources processing the user plane.

Power governors and other management mechanisms for the first two types are well established and supported by commercial solutions. Telemetry can be used to identify idle infrastructure and low-utilization control-plane resources, and the server hardware can automatically be tuned for power consumption using Enhanced Intel SpeedStep Technology. Operating at the lowest possible power state to meet service requirements enhances energy efficiency, which reduces costs and carbon footprint.

Using NFV-RI on Red Hat OpenShift enables UPF workloads to take full advantage of the power management capabilities provided by the Intel infrastructure, helping optimize overall energy consumption. The Intracom Telecom solution leverages telemetry to dynamically configure the hardware power state to match the needs of fluctuating workloads for each 5G UPF pod instance on a server node. NFV-RI integrates with the workload to fine-tune the associated FFL machine learning algorithms, improving both OpEx and climate sustainability while protecting service level agreements (SLAs). The overall architecture and solution components are depicted in Figure 4.

Conclusion

NFV provides CoSPs great benefits, but requires stringent KPIs. This results in CNFs keeping servers at a high-power state constantly, as if they were always operating at peak demand. This leads to CPU frequency overprovisioning, which results in unnecessary energy use.

Intracom Telecom's NFV-RI FFL workflow leverages machine learning to predict how busy CNFs will become, with a Helm chart certified by Red Hat for use with OpenShift. Using Enhanced Intel SpeedStep Technology for Intel processors, the FFL workflow can dynamically adjust the CPU frequencies of DPDK-based or VPP-based CNFs according to their incoming load, while maintaining zero packet drops. With the NFV-RI FFL workflow, CoSP infrastructure can operate at low power during off-peak or light-traffic periods and CNF service level objectives (SLOs) can remain assured and unaffected with the potential to reduce power consumed by the system.

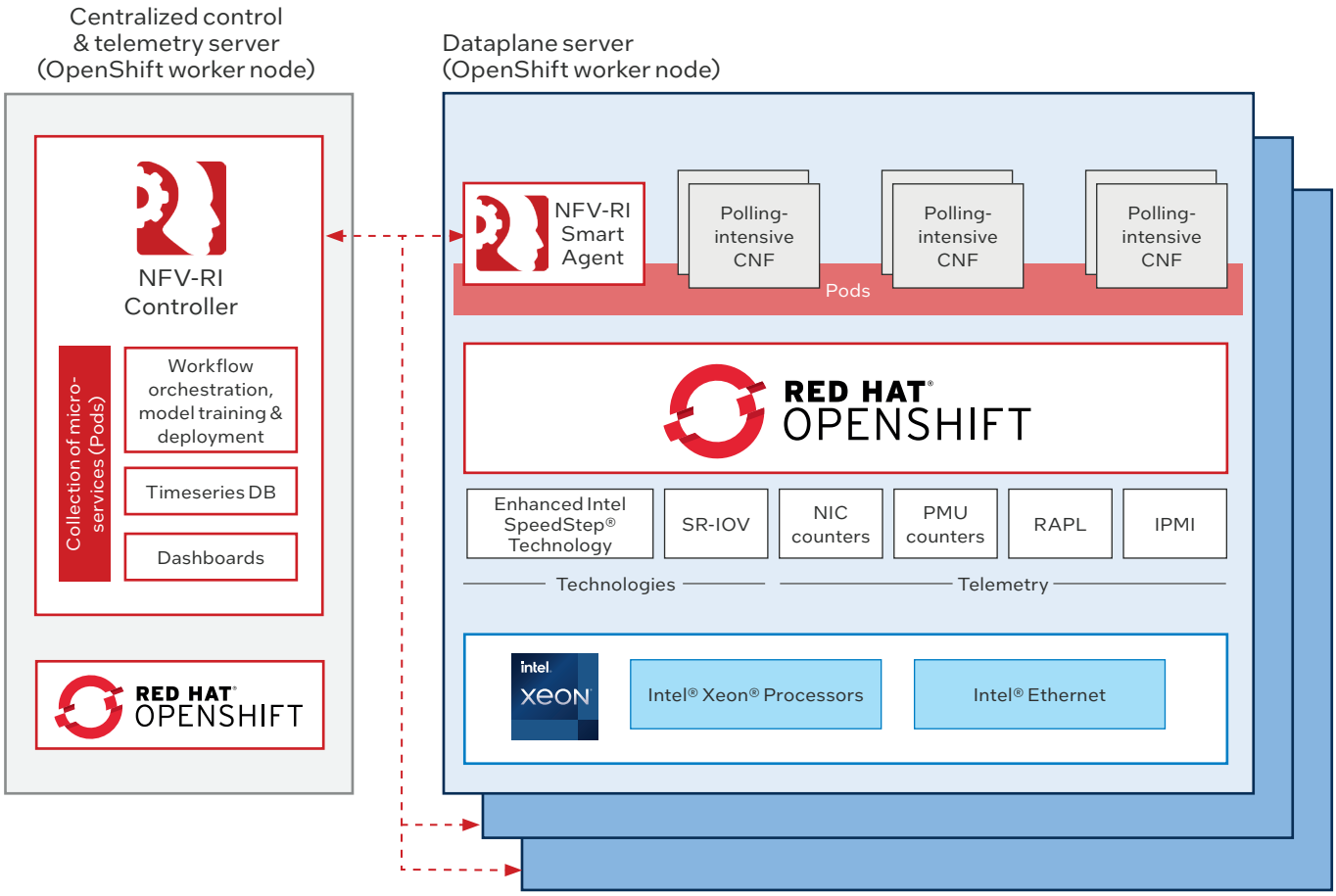

Figure 4. NFV-RI's integration in Red Hat OpenShift.

Learn More

Intracom Telecom NFV-RI:
https://www.intracom-telecom.com/en/products/telco_software/sdn_nfv/nfvri.htm

Intracom Telecom NFV-RI Certified OpenShift Helm Chart (Red Hat Ecosystem Catalog):
https://catalog.redhat.com/software/helm/detail/625fc71c6251947aa2a44701

Intel® Xeon® Scalable processors:
https://www.intel.com/content/www/us/en/products/processors/xeon/scalable.html

Enhanced Intel SpeedStep Technology Chapter 14:
https://software.intel.com/content/www/us/en/develop/download/intel-64-and-ia-32-architectures-sdm-combined-volumes-3a-3b-3c-and-3d-system-programming-guide.html

Solution provided by:

[1] Figures provided courtesy of Intracom Telecom.

[2] Testing conducted by Intracom Telecom in July 2020: SUT server utilized dual 24-core Intel® Xeon® Platinum 8168 processors (microcode: 0x2006906) supporting Non-Uniform Memory Access (NUMA). Intel® Hyper-Threading Technology was turned on, and Intel® Turbo Boost Technology was turned off. BIOS version was SE5C620.86B.0X.01.0076.101320171718. Total RAM for node 0 was 196 GB and for node 1 was 198 GB. System storage totaled 960 GB and was provided by an Intel® SSD SC2BB960G7. 25 GbE network connectivity was provided by an Intel® Ethernet Controller XXV710. The system also featured a Platform Controller Hub integrated 10 Gigabit Ethernet Controller with integrated SATA controller. Operating system was Ubuntu 18.04.2 LTS with kernel version 5.4.0-45-generic. Workload 1 was l2fwd traffic forwarding app built into DPDK v19.11. Workload 2 was OMEC's Next Generation Infrastructure Core RTC evolved packet core. Compiler was gcc version 7.5.0 and libraries included collectd v5.9.2. Client server utilized dual Intel Xeon Platinum 8168 processors (microcode: 0x200005e) supporting Non-Uniform Memory Access (NUMA). Intel Hyper-Threading Technology was turned on, and Intel Turbo Boost Technology was turned off. Total system RAM was 384 GB. System storage totaled 960 GB and was provided by an Intel® SSD Data Center 3520 SC2BB960G7. 25 GbE network connectivity was provided by an Intel Ethernet Controller XXV710.

[3] L2 Forwarding Sample Application (in Real and Virtualized Environments): https://doc.dpdk.org/guides/sample_app_ug/l2_forward_real_virtual.html

Performance varies by use, configuration, and other factors. Learn more at www.Intel.com/PerformanceIndex.

Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See backup for configuration details. No product or component can be absolutely secure.

Your costs and results may vary.

Intel technologies may require enabled hardware, software, or service activation.

Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy.

0722/RKM/MESH/PDF 350488-001US