Solution Blueprint

Cloud-Native Network Functions
5G Communications Service Providers

Transforming the Fixed-Access Network Using a Cloud-Native Approach

Red Hat and Intel have jointly developed a pre-validated solution blueprint for an optimized hardware and software stack supporting the evolution from infrastructure based on virtualized network functions (VNFs) to more agile cloud-native network functions (CNFs), specifically targeting broadband fixed-access transformation. This cloud-native architecture helps set the stage for high-performance, cost-effective telecommunication networks as the industry adopts fixed-mobile convergence for 5G and beyond.

Table of Contents

1 Networks in Transition

The migration to cloud-native infrastructure is one of the key digital transformative challenges facing the communications service provider (CoSP) community. The benefits are well understood and the transformation is well underway. 5G's massive increases in traffic rates and volume make dynamic network topologies and elastic capacity critical, while increased throughput and latency requirements are accelerating the transition toward deployment of additional new cloud services at the network edge.

Today's typical CoSP infrastructure runs a virtualized environment that may consist of both VNFs and CNFs, as in the Virtualized state depicted in Figure 1. Here, a hypervisor-based virtualization layer is positioned between the bare-metal or infrastructure-as-a-service substrate and the network functions. The VNFs run directly on virtual machines (VMs) to support decoupling the hardware from the software, enabling software-defined infrastructure and agile, on-demand deployment. CNFs can run on this infrastructure by means of a containers-as-a-service (CaaS) layer based on Kubernetes, for example, but they incur the overhead of the hypervisor and full guest OS in each VM. Deploying containers in VMs supports the CNFs developed by forward-looking CoSPs on existing infrastructure, alongside existing VNFs, but the operational complexity associated with stacking both platforms reduces the efficiency of CNF deployments and raises costs.

Particularly with 5G's massive increases in traffic and throughput requirements, the emerging competitive imperative is for CoSPs to transition toward fully cloud-native infrastructure, as shown in the Figure 1 Cloud-Native state. In this topology, the CaaS layer runs on bare metal, optimizing operations, performance, latency, and costs. The Hybrid state shown in the figure allows virtualized and cloud-native environments to operate side-by-side, modernizing CNF deployments while retaining functionality of VNFs for usages such as legacy services and a 4G core shared by both 4G and 5G.

Authors

Intel: Padraig Connolly, Andrew Duignan, Paul Mannion, Raghu Moorthy, Eoin Walsh

Red Hat: Franck Baudin, Hanen Garcia, Aaron Smith

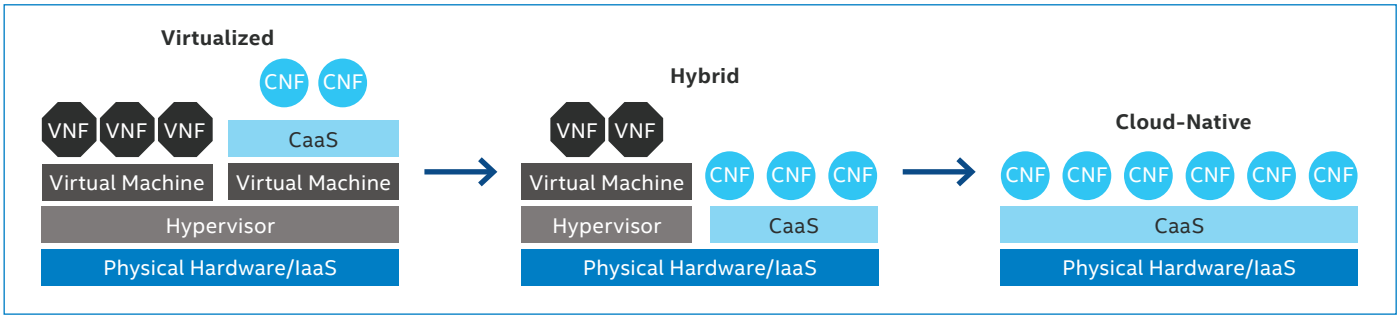

Figure 1. Deployment models for CNFs.

Aside from the advent of 5G, cloud-native infrastructures are also necessitated by fixed-mobile convergence, which is discussed in more detail in section 3.1, "Fixed-Mobile Convergence." For CoSPs delivering suites of services to every device, such as quad-play packages that combine mobile and wireline telephony, broadband, and pay television, operating separate infrastructures for each access network has become increasingly impractical. Just as eliminating the virtualization layer between CaaS and the bare metal is a medium-term efficiency mandate, so CoSPs are searching for ways to converge their access networks onto a common shared CaaS infrastructure.

Red Hat and Intel are enabling application vendors in the ecosystem-wide transition to CNFs with thought leadership and flexible, optimized platforms and tools. Full-stack involvement is helping increase developer efficiency, time to market, and cloud-native solution quality, as providers build out the software to enable fixed-mobile convergence and CoSP innovation more broadly.

This solution blueprint provides future-ready platform guidance for CoSPs as they develop cloud infrastructures robust and flexible enough to enable the transition to bare-metal deployment of CNFs. It identifies and illustrates the capabilities of a solution stack for cloud-native deployment of network functions based on specific Intel and Red Hat building blocks. It contributes to the development of open, standards-based architectures for a common approach to cloud-native support for multiple access technologies.

2 The Need to Retool for the Next Decade of Telecom

Even in the face of tremendous opportunity, the business environment in which CoSPs operate is challenging. Traffic throughput demands continue to increase—with further acceleration assured as 5G adoption gains momentum—while average revenue per user (ARPU) has plateaued or even begun to drop. Merger and acquisition activity is making the industry more competitive as carriers seek economies of scale, expanded scope of services, and existing customer bases.

The technology modernization requirements to handle the changing business environment are complex and multi-faceted, especially as provider networks expand in capacity and diversity of services. Most struggle to integrate large numbers of both proprietary and open-source software components into their networks. This is made more complex by their existing disjointed environments, where multiple access networks may each require separate integration and operations efforts. Moreover, they lack the ability to scale user-plane resources efficiently to support the full range of use cases, devices, and locations required, especially with the geographic distribution of services to the network edge.

As CoSPs build out their 5G networks, CNFs increase the already compelling benefits of deploying services to multi-access edge computing (MEC) infrastructures. Pushing distributed services and traffic processing to the edge can reduce backhaul to the network core, for lower bandwidth costs and transmission latency. Edge topologies therefore play a vital role in enabling the high throughput combined with low latency that are commonly required to realize the promise of 5G usage models. User-plane functions commonly being migrated to the edge in emerging 5G networks include the following:

- **Broadband Network Gateway (BNG)** for terminating fixed access subscribers
- **Access Gateway Function (AGF)** to adapt wireline access to the 5G core
- **Cable Modem Termination System (CMTS)** to enable cable modems to transmit packets over the Internet
- **User Plane Function (UPF)** for terminating 5G wireless subscribers

The inherent nature of legacy infrastructure adds cost inefficiency to the network. For example, to the extent that services remain tethered to specific equipment, upgrades and other scheduled maintenance to either software or hardware typically require resources to be taken offline. These sporadic planned outages can degrade or interrupt services unless adequate contingent resources are available. More broadly, providing failover redundancy requires duplicate systems to be maintained, on a one-to-one basis. Operationally, extended timelines are required for CoSPs to introduce new features and services because of the need to physically provision resources.

As these challenges develop and intensify, the industry is adopting new approaches to network design and operation that are designed to support emerging services and business requirements with greater agility, quality, and cost-efficiency.

3 Technology Trends Involved in Network Transformation

A range of transformative technology approaches are involved in the design and construction of future-focused CoSP networks. To some extent, these trends have emerged directly in response to changing requirements, while the trends themselves may also be the motivation for CoSPs to transform. In both cases, navigating the roles of these trends within future networks is a key strategic requirement.

3.1 Fixed-Mobile Convergence

Converging the separate fixed and mobile access networks that most CoSPs traditionally operate can dramatically simplify the network. Migrating both wireline and wireless access to the 5G converged core, as represented in Figure 2, allows a single control plane and a single user plane to span both fixed and mobile networks. In addition, it eliminates the need for duplication of services such as authentication, policy control, and lawful intercept. In that context, 5G is a unique business driver for CoSPs to adopt fixed-mobile convergence. Broadband Forum and 3GPP are defining mechanisms and patterns for CoSPs at various starting points to access the 5G converged core. For more details, see Broadband Forum publication M-464, "Migrating Fixed Access to 5G Core."

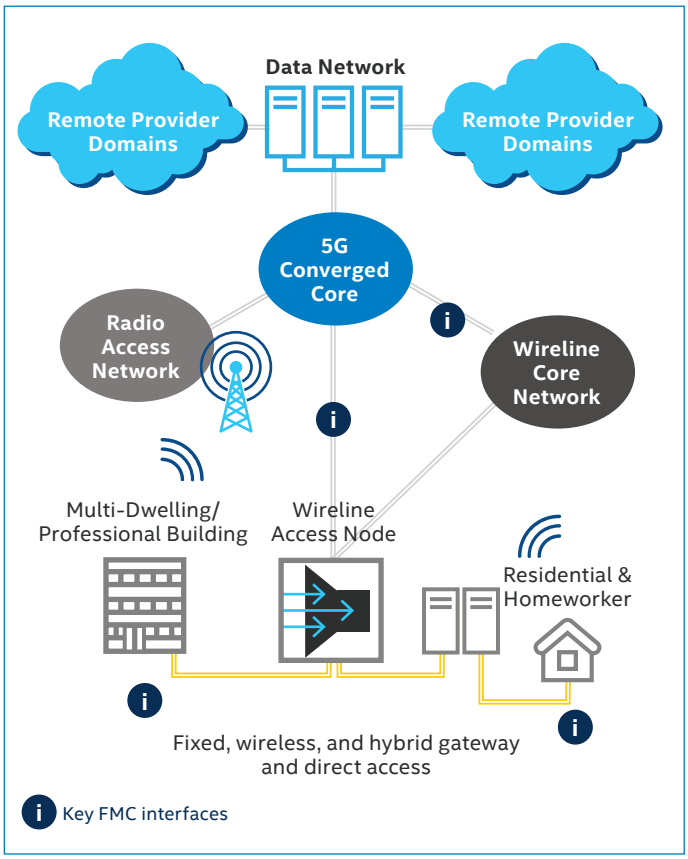

Figure 2. Fixed-mobile convergence.

Fixed-mobile convergence reduces infrastructure requirements and streamlines provisioning. As CoSPs move to take advantage of that potential cost savings, they particularly benefit from quad-play and other bundled service packages that draw together multiple channels to the subscriber, across devices as well as across connections such as fiber, cable, DSL, and 5G. The improved ability of 5G relative to 4G for high-bandwidth consumer applications such as media and gaming is also motivating fixed-line-focused providers such as cable companies to enter the mobile market.

3.2 Cloud-Native Architecture

Taking full advantage of cloud computing models requires that software is designed explicitly for the cloud, rather than simply being deployed there as an afterthought. Cloud-native architecture and deployment models are designed from the ground up to be lightweight, so applications can maximize agility, performance, and scalability to get the most value possible from infrastructure investments. At the same time, cloud-native principles can improve the ability of development organizations to serve the larger business, with higher solution quality and faster time to production for new products and services.

- **Microservices** construct applications as sets of small, autonomous, reusable modules on a shared fabric that communicate using standard protocols to instantiate composite functions
- **Containers** encapsulate microservices or other software entities along with all their dependencies in lightweight packages that can be deployed anywhere and seamlessly traverse multi-clouds with centralized orchestration
- **DevOps** is the modern standard for development teams to accelerate software development, testing, and release, with a high degree of automation to increase efficiency and deterministic repeatability
- **Continuous integration/continuous delivery (CI/CD)** is the concept of making many small, frequent releases rather than infrequent, monolithic ones, increasing the pace of change while reducing its incremental impact

3.3 Control- and User-Plane Separation (CUPS)

Separating control-plane functions from user-plane functions allows them to scale independently. For example, more user-plane resources can be deployed to respond to increased traffic without affecting the control plane. The CUPS approach allows for either distributed or centralized deployment. This blueprint calls for the control plane to be centralized, while the user plane is distributed to the edge, as shown in Figure 3. The network takes advantage of the distributed compute in the user plane to enhance network performance for latency-sensitive applications by intelligently choosing the best user-plane node—based on capability or proximity, for example. To that end, CUPS facilitates programmatic control of the user plane by software-defined networks to increase the efficiency of data delivery.

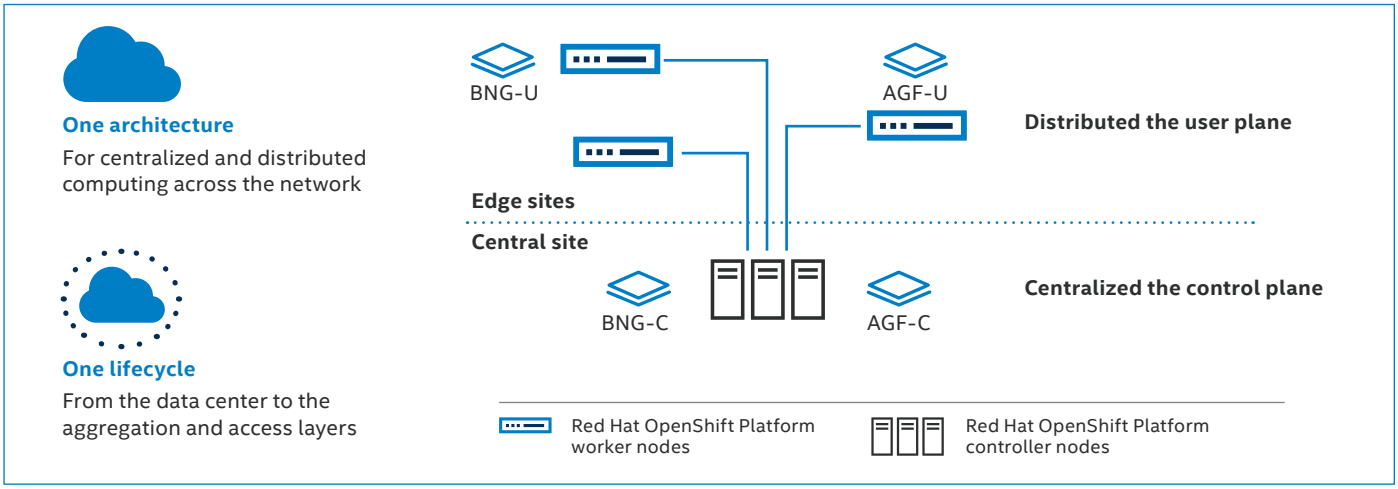

Figure 3. Distributed computing architecture for CUPS.

By abstracting network resources from the underlying hardware, CUPS is also instrumental to network slicing, which is the practice of creating multiple virtual segments of network bandwidth. Resources can be dedicated or prioritized for each network slice, providing differentiated quality of service among specific traffic types or customer service levels. That capability will become increasingly important as CoSPs address novel use cases in areas such as IoT, connected vehicles, and connected health. Red Hat OpenShift Platform provides a single container-based architecture across this entire topology, with a unified lifecycle that helps simplify the development and deployment of applications and services.

3.4 Network Disaggregation

The trend of CoSPs transitioning away from single-vendor solutions running on specialized, monolithic hardware continues, replaced by software-based network functions running on standards-based servers. In this model, bespoke physical appliances are being replaced by virtual ones. Communication between open components supplied by multiple hardware and software vendors allows the software-defined network to chain together CNFs (or VNFs) to dynamically create applications and services.

Network disaggregation supports an accelerated rate of change, because bringing a new product or service to market doesn't require deployment of new infrastructure. It also helps networks be future-ready, because updates and changes can be made using the software CI/CD paradigm, without taking any services offline. Particularly as the rate of change at the network edge continues to accelerate, disaggregation helps protect CoSPs from the cost and business impacts of shifting user-plane functionality requirements.

4 Platform Innovations that Enable CNFs

The past decade or so has seen a transition in the modes of delivering network functions from rigid, single-purpose architectures to flexible ones designed to integrate seamlessly with a dynamically changing environment, as illustrated in Figure 4. This journey began with network functions deployed as monolithic applications on custom, fixed-function hardware. The equipment was expensive and inflexible, and it tended to lock CoSPs into a single vendor for hardware and software, limiting options for new development.

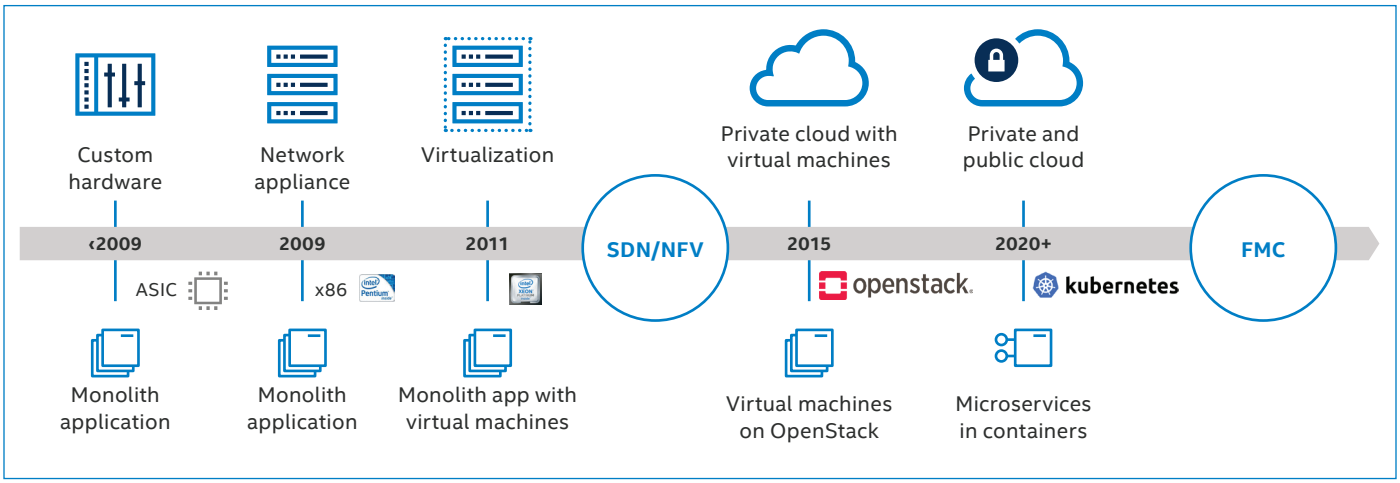

Figure 4. A decade of transition toward CNFs.

As commercial off-the-shelf computing platforms became more capable, the applications underlying network functions could be deployed on standards-based network appliances, and then to general-purpose servers using hypervisor-based virtual machines. The rise of software-defined networking and network functions virtualization set the stage for adopting private cloud topologies based on OpenStack. With each step in this transition, network functions became more agile and less tied to specific equipment or topologies.

This trend continues with cloud-native topologies that deploy CNFs as container-resident microservices, assembled dynamically as needed to provide services to the network. Red Hat OpenShift provides a consistent foundation for development and deployment of network functions across any combination of private and public cloud environments.

4.1 Cloud-Native Platform: Red Hat OpenShift Container Platform

Red Hat OpenShift Container Platform, illustrated in Figure 5, enhances Kubernetes for CoSPs by integrating components from Red Hat Enterprise Linux (RHEL) and other Red Hat technologies. The platform therefore benefits from broad hardening, testing, and certification initiatives by Red Hat. It is also developed and distributed using an open-source model to foster innovation. OpenShift is multicloud-ready, enabling clusters to be deployed within the data center and to a variety of public clouds.

Red Hat Enterprise Linux CoreOS is the operating system designed specifically for running containerized applications from OpenShift Container Platform. Deployment and maintenance of OpenShift Container Platform clusters are highly automated, enhancing the efficiency of IT operations.

Operators are software entities that represent fundamental sets of capabilities within the OpenShift Container Platform, as shown in Figure 6. They capture human operational knowledge in code, to encapsulate the processes for packaging, deploying, and managing Kubernetes applications. Operators foster repeatability in IT processes, perform ongoing health checks of system components, and handle over-the-air updates to OpenShift and third-party software. Some key Operators are described in more detail in the remainder of this section.

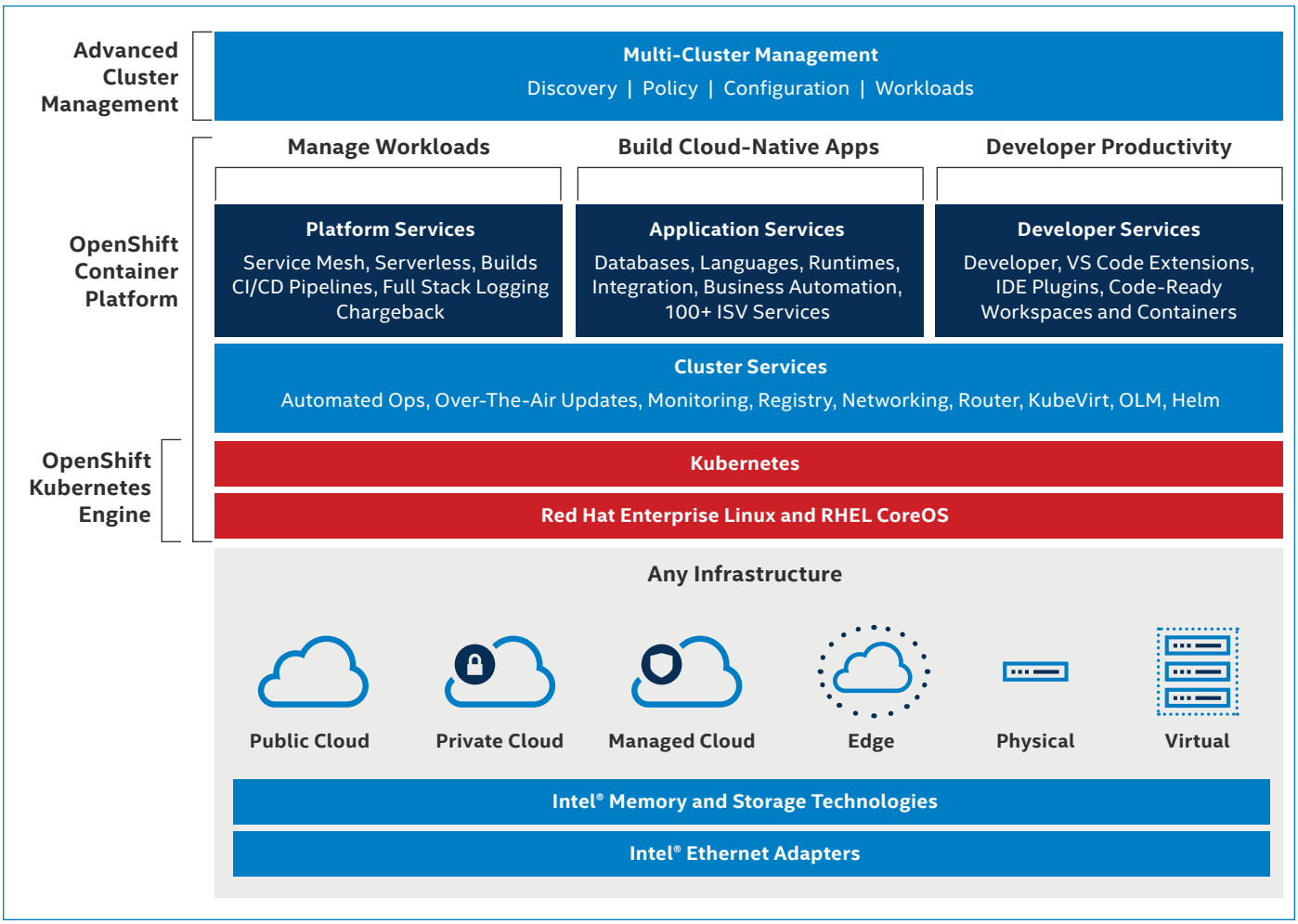

Figure 5. Red Hat OpenShift Container Platform—optimized for Intel technologies—helps customers develop, deploy, and manage innovative applications.

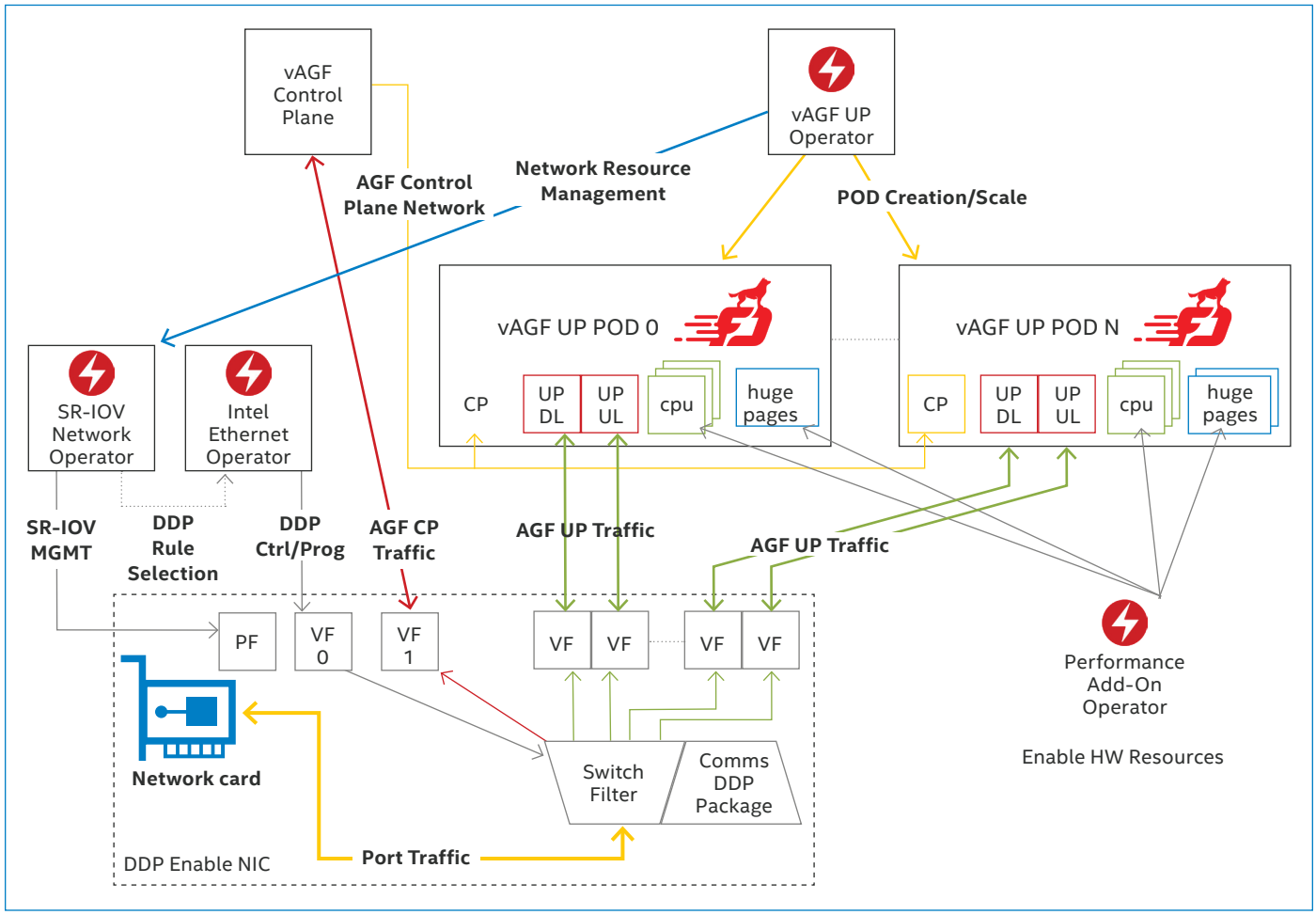

Figure 6. Cloud-native platform for fixed-mobile convergence.

4.1.1 Performance Add-On Operator

The Performance Add-on Operator (PAO) enables advanced node performance tunings on sets of nodes. The PAO controls the host allocation and configuration of performance-enhancing features such as CPU isolation/reservation, memory hugepages, and IRQ management. The PAO can also enable the use of a real-time kernel for certain low-latency environments. For fixed-mobile convergence, the PAO is used to enable CPU tuning and hugepage memory allocation.

4.1.2 SR-IOV Operator

The SR-IOV Network Operator creates and manages the components of the SR-IOV stack within the OpenShift Container Platform. It performs the following functions:

- Orchestrates the discovery and management of SR-IOV network devices
- Generates **NetworkAttachmentDefinition** custom resources for the SR-IOV Container Network Interface (CNI)
- Creates and updates the configuration of the SR-IOV network device plug-in
- Creates node-specific **SriovNetworkNodeState** custom resources

The SR-IOV Network Operator is used in fixed-mobile convergence usages to allocate SR-IOV VFs for use by BNG and AGF Pods. After creation of SR-IOV Networks, AGF Pods/instances only need annotations to allow connections to the converged networks.

4.1.3 Intel® Ethernet Operator

The Intel Ethernet Operator enables the deployment and management of Dynamic Device Personalization (DDP) functionality within the OpenShift environment (see discussion below in section 4.3, "Intelligent Networking: Intel® Ethernet 800 Series Network Adapters"). The Operator provides the following capabilities:

- Deployment of the appropriate DDP package on a per-node basis
- Management of firmware for DDP-capable network interface cards
- Reporting on the status of the DDP platform
- Automation based on the inclusion of DDP flow rules into the programmable packet pipeline for use by individual workloads; DDP flow rule and action pipelines become first-class citizens in the OpenShift platform

The Intel Ethernet Operator automates the use of hardware PPPoE packet processing to enhance session management and flow steering in the BNG and AGF user-plane application.

The SR-IOV Operator and Intel Ethernet Operator enable the inclusion of DDP flow rules during Pod deployment. CNF-specific flow definitions that have been created in the Intel Ethernet Operator can be referenced as annotations in a Pod specification in conjunction with a Custom Resource Definition. After an SR-IOV interface has been rendered by the SR-IOV Network Operator, a component of the Intel Ethernet Operator is called to enable the required flow definition. Figure 7 provides a high-level view of how the PPPoE and user-plane flows are programmed and how the packets are processed prior to entering the AGF Pipeline.

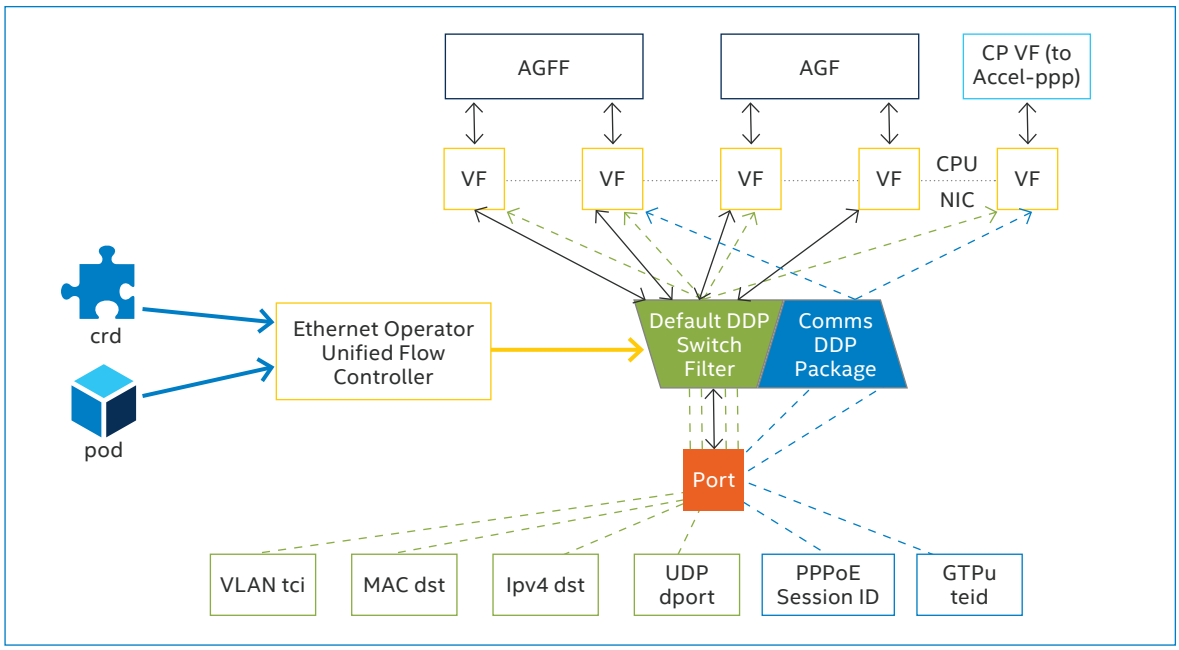

Figure 7. PPoE and user-plane flows.

4.2 Flexible Performance: 3rd Generation Intel® Xeon® Scalable Processors

The 3rd Generation Intel® Xeon® Scalable processor provides a balanced, scalable architecture that is built to be tailored to diverse implementations with a wide range of core counts, frequencies, and power levels. Architectural advances over its predecessor dramatically increase per-core performance, memory subsystem performance, and I/O bandwidth, to accelerate workloads and increase the concurrent workload capacity per server. Platform enhancements that accelerate diverse workloads from the edge to the data center include the following:

- **Increased core count and cache**, available in a flexible range of 8–40 powerful next-generation cores, L1 caches of 12–60 MB, and total design power of 140–270 watts
- **Expanded Intel® Advanced Vector Extensions 512 (Intel® AVX-512)**, accelerating bit processing by moving and reordering blocks of data faster within the signal-processing pipeline
- **Updated I/O subsystem**, including support for PCI Express Gen4, which provides double the bandwidth of PCI Express Gen3 for faster data movement
- **Enhanced memory subsystem**, with up to 1.60x higher memory bandwidth[1] and up to 2.66x higher memory capacity[2] compared to the prior-generation platform

Engineered for modern network forwarding plane workloads, the 3rd Generation Intel Xeon Scalable processor targets low-latency, high-throughput deterministic performance and high performance per watt. For today's software-defined environments, the platform is optimized for hybrid/multi-cloud deployments. It is built on open standards and APIs to enable infrastructures designed to accommodate current and future business needs.

Next-Generation Performance for Common CoSP Workloads

The architecture enhancements designed into the 3rd Generation Intel Xeon Scalable processor deliver results such as the following to CoSPs, compared to predecessor platforms:

- **21% boost in vBNG performance**, while enabling increased flexibility for fixed-mobile convergence, manageability, and scalability to expand emerging use cases[3]
- **72% better vCMTS platform performance**, as well as the potential for a further 10 percent improvement with additional Intel QAT offload[3]

The processor enhances cryptographic acceleration with new instructions that increase encryption throughput on the CPU without the need for dedicated hardware accelerators. As encryption requirements increase with the advent of 5G networking, this capability is becoming more critical to CoSPs. In addition, Intel® Software Guard Extensions (Intel® SGX) creates isolated memory enclaves that can help securely store encryption keys at the network edge.

Intel® Select Solutions for NFVI Forwarding Platform: Standard, Validated Forwarding-Plane Designs

The primary traffic load in 5G is associated with user-plane functions (the data, as opposed to the control or signaling associated with establishing traffic flow). Network implementations based on NFV infrastructure (NFVI) can accelerate those functions for greater speed and agility, improved time to market, and scale for future capacity and new services. CoSPs are expected to continue the transformation that began with purpose-built solutions and transitioned to virtualized appliances, to a model of fully virtualized and cloud-native based networks. To accomplish this transformation, the journey to a more fully disaggregated deployment model must be realized.

The Intel Select Solutions for NFVI Forwarding Platform are predefined, workload-optimized designs that help reduce the challenges of infrastructure evaluation and deployment for user plane workloads such as BNG, UPF, AGF, and CMTS. Solutions are validated by OEMs/ODMs, certified by ISVs, and verified by Intel. Intel develops these solutions in extensive collaboration with hardware, software, and operating system vendor partners such as Red Hat and with the world's leading data center and service providers. All Intel Select Solutions are tailored combinations of Intel data center compute, memory, storage, and network technologies that deliver predictable, trusted, and compelling performance.

Currently in version 2, the Intel Select Solutions for NFVI Forwarding Platform incorporates the 3rd Generation Intel Xeon Scalable processor and Intel Ethernet 800 Series Network Adapters.

Ingredient	Plus Configuration	Base Configuration	Controller Node Configuration
Processors	2x Intel® Xeon® Gold 6338N processor	2x Intel Xeon Gold 5318N processor	2x Intel Xeon Gold 5318N processor
Memory	512 GB DDR4	256 GB DDR4	
Intel® Optane™ Persistent Memory	Recommended		
Discrete Network Adapters	4x Intel® Ethernet Network Adapter E810 2CQDA2	2x Intel Ethernet Network Adapter E810 2CQDA2 or 4x Intel® Ethernet Network Adapter E810 CQDA2	2x Intel Ethernet Network Adapter E810 CQDA2
Local Storage	2x Intel SSD D3-S4510 Series or higher @ 480 GB or larger		
LAN on Motherboard	10Gbps or 25 Gbps port for Pre-boot Execution Enviroment (PXE) and Operation, Administration and Management (OAM)		
	1/10 Gbps port for management		

Configurations within the Intel Select Solutions for NFVI Forwarding Platform v2.

4.3 Intelligent Networking: Intel® Ethernet 800 Series Network Adapters

The Intel Ethernet 800 Series Network Adapters provide network I/O that complements the 3rd Generation Intel Xeon Scalable processor's compute and memory advances. The adapters use PCI Express Gen4 for improved bandwidth to the system board with network throughput up to 100 Gbps per adapter port. It delivers standards-based networking performance across NFV and CFV workloads through a combination of sophisticated packet processing, intelligent offloads and accelerators, and high-quality open-source drivers for data-plane processing. In addition to optimizing throughput, the adapters are designed to enable broad interoperability and agility.

Dynamic Device Personalization (DDP) allows multiple personalization profiles to specify optimizations and packet-handling parameters for individual traffic types, increasing throughput and enabling sophisticated traffic prioritization. The DDP programmable packet-processing pipeline provided by the Intel Ethernet 800 Series Network Adapters supports on-demand reconfiguration of network controllers at runtime, enabling workload-specific optimizations. DDP is enhanced in the Intel Ethernet 800 Series Network Adapters with greater programmability than its predecessor, as well as workload-specific protocols for added flexibility.

The enhanced Data Plane Development Kit (DPDK) is an open-source set of libraries and drivers supported by the Intel Ethernet 800 Series Network Adapters that accelerates packet processing in the data path. It also facilitates building packet forwarders designed to operate on general-purpose, standards-based servers. DPDK technology is incorporated as a bundled feature with RHEL, enabling Intel Ethernet 800 Series Network Adapters to be controlled entirely in user space. That approach accelerates operations by allowing network packets to bypass the kernel network stack entirely. With 3rd Generation Intel Xeon Scalable processors, CoSPs can enhance DPDK L3 Forwarding performance by up to 88 percent compared to the prior generation.[3]

5 Conclusion

The application of NFV and cloud-native principles that swept through the 5G network core and RAN is now taking hold in fixed-access networks. The application of cloud-native principles to cable and broadband access aggregation sites will allow the CoSP community to secure needed technical flexibility and afford them a more efficient pay as you grow, software like approach to network provisioning. Combining modern software approaches with open source operators that expose Intel hardware assists using a standard approach enables the VNF/CNF provider ecosystem to build and pre-validate performance-optimized solutions. These innovations are vital for CoSPs working to replace and improve on their current fixed-access TCO models.

This forward-looking approach combines open-source cloud expertise from Red Hat and Intel with their leadership on high-speed enterprise systems. It lays a future proof path to cloud-native fixed-access networks and 5G fixed-mobile convergence. The emerging transition will make current services more efficient and enable CoSPs to bring new services to market faster and more efficiently on the same fixed-access cloud platform.

For a deeper look, see

Red Hat's Cloud-Native Container Platform for Hybrid-Multicloud NFV

Intel Select Solutions for NFVI Forwarding Platform

[1] 3rd Gen Intel Xeon Platinum 8380 CPU: 8 channels, 3200 MT/s (2 DPC) vs. 2nd Gen Intel Xeon Platinum 8280 CPU: 6 channels, 2666 MT/S (2 DPC).

[2] 3rd Gen Intel Xeon Platinum 8380 CPU: 8 channels, 2 DPC (256 GB DDR4) vs. 2nd Gen Intel Xeon Platinum 8280 CPU: 6 channels, 2 DPC (128 GB DDR4).

[3] Performance varies by use, configuration, and other factors. See [91,92] at https://www.intel.com/3gen-xeon-config.

0621/RKM/MESH/341974-001US