White Paper

Network architects, Next-Generation Network Decision Makers
Mobile Network Operators

Intelligent Fabric and xPUs Accelerate Network Services End-to-End

Kaloom's Unified Edge solution powered by Intel® programmable switches, IPUs and CPUs, and Red Hat-OCP provide a single execution platform for Kaloom* and 3rd Party containerized networking applications for Edge to Multi-Cloud networks

The availability of virtualized telecom solutions that can support multiple services from a single server is helping wired and wireless telecom providers to move from siloed networks that each need their own infrastructure which can lead to high costs, operational complexity, and low service provisioning agility.

Additionally, with the advent of software-defined networking, 5G and edge computing, cloud native container network functions (CNFs) can be distributed across the cloud and the edge bringing compute and networking resources closer to users reducing transport latency to meet stringent performance expectations.

These two networking trends are enabling the network infrastructure to scale to economically serve the needs of new digital applications and devices, in addition to reducing the cost per connected device and per gigabit of IP traffic to profitably support emerging applications.

Cloud and edge data centers must evolve to provide the cloud-native programmable infrastructure that:

- Elastically scales end-to-end network applications and lowers the overhead costs without losing performance and security.
- Pares the number of networks dedicated to applications and data processing that are spread across multiple network functions.
- Consolidates the network sprawl by sharing infrastructure across multiple applications from the core to the edge, and
- Provides consistent, automated operation across different types of data centers – far edge, edge, and cloud.

This paper aims to explain how the use of programmable hardware available from Intel combined with Kaloom's cloud-native Unified Edge solution can help meet the demanding needs of business-critical applications for 5G. By creating a common execution environment, the Kaloom's Unified Edge solution simplifies complex distributed edge-to-cloud networking.

Table of Contents

The Challenges

5G makes seemingly contradictory demands: a leap in network agility and performance, and at the same time, costs that must decline sharply to meet slowing or decreasing average revenues per user/device. Both are achievable with a far more productive programmable network that eliminates redundancies and synchronizes the use of diverse resources.

Unified Edge solution simplifies complex distributed edge-to-cloud networking, accelerates time to market of new services and reduces the total cost of ownership of edge infrastructure.

The challenges of current networks are:

- **Balancing performance and costs:** In a highly virtualized environment, nearly 30% of CPU cycles are consumed to support networking functions[i]. Network operators facing cost pressures must evolve from a server optimized infrastructure to a function-based infrastructure that leverages hardware acceleration devices best suited for individual workloads such as applications, networking, and infrastructure. However, the increasing variety of devices creates a challenge to intelligently decide the best use of servers, switches, and new hardware acceleration solutions to cost-efficiently scale networking and data processing.
- **Switching evolution:** Fixed Ethernet switching is the current market leading solution. Existing data centers rely on non-programmable switch fabrics using fixed-function Ethernet ASICs. Consequently, the latencies are high, and these switches have low energy efficiency resulting in poor cost/performance per gigabit of IP traffic.
- **Need more robust networking functions in Kubernetes:** 5G and associated emerging applications will be deployed at the edge as cloud native applications based on Kubernetes. Currently, Kubernetes platforms encompass compute and storage functions but lack an adequate networking solution to support high-availability container networking applications.
- **IPv4 falls short for end-to-end traffic engineering:** Existing IPv4-based transport networks struggle to provide automated service function chaining for end-to-end traffic engineering. They use linear methods for moving packets from one point to another. As the number of applications expands over distributed networks, the number of network nodes increases. It complicates traffic engineering aggravated by the need to maintain a per-application, per-flow routing state for each network slice.
- **Service quality is not deterministic:** Current edge and cloud data center networks are not equipped to ensure service-quality agreements for 5G applications. They need end-to-end network slicing with fine-grain visibility to find root causes of performance lapses and automated means for troubleshooting. In dense shared networks with multiple service and technology providers, it is essential to pinpoint the exact source of the problem.

How Kaloom's Unified Edge solution is addressing challenges

Kaloom delivers a fully programmable and automated cloud-native edge networking software solution disrupting how distributed cloud and edge data center networks are built, managed, and operated by fixed and mobile telecom operators, data center, and cloud service providers.

Kaloom's Unified Edge solution unifies data processing of all its containerized network functions – switches, routers, gateways, and 5G UPF – and even third-party network functions. It creates a common containerized operating environment for the edge IT infrastructure leveraging Intel hardware, including computing, storage, and networking.

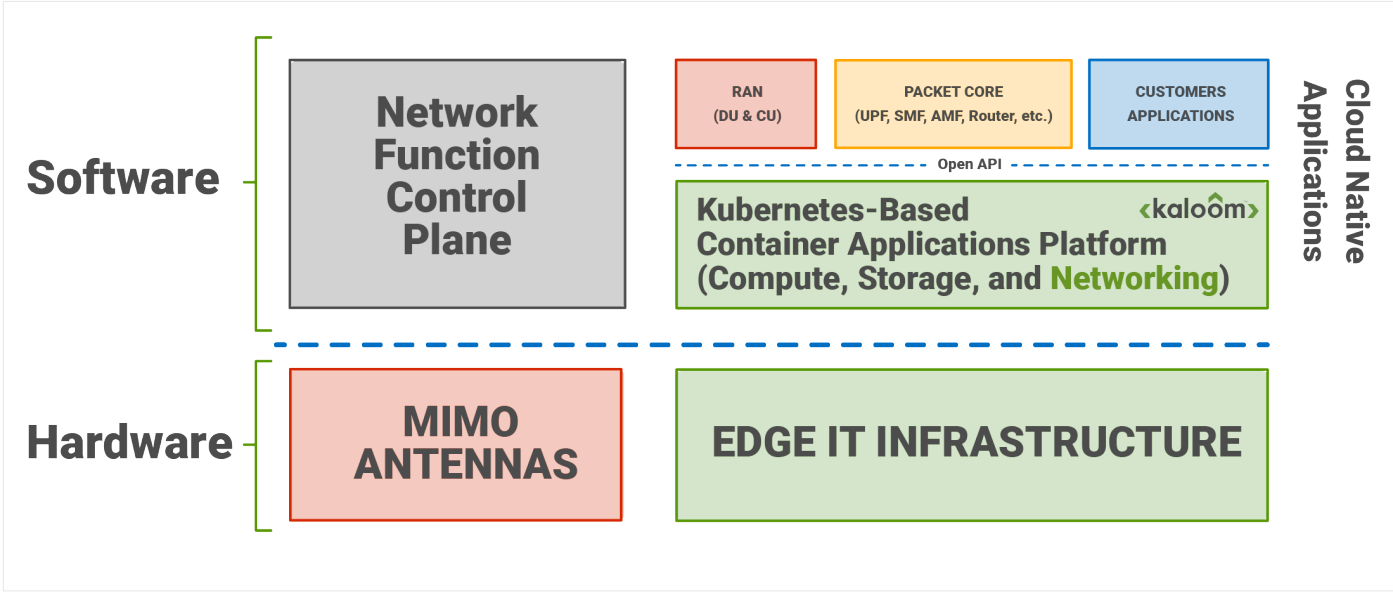

Figure 1. Common containerized operating environment for the edge IT infrastructure

Furthermore, Kaloom's Unified edge solution consolidates multiple networks with fabric virtualization to create fully isolated network slices over a shared infrastructure – each supporting the unique need of 5G applications.

Intelligent Hardware Acceleration Scales Data Processing and Computing at the Edge

The stringent demands of data-intensive applications have led to the introduction of a new class of devices such as the P4 programmable Intel® Tofino™ Intelligent Fabric Processor, Intel® Stratix® 10 FPGA, and Intel® Infrastructure Processing Unit (Intel® IPU). Each of these specializes in an aspect of data traffic processing combined with programmability and acceleration. However, the diversity of devices creates a challenge to intelligently leveraging the best of servers, switches, and new hardware acceleration solution capabilities and cost-efficiently scale compute and data processing.

The bedrock of the solution to close the gaps is a common fabric created by the convergence of an intelligent fabric and heterogeneous hardware acceleration. The Kaloom solution eliminates network redundancies and dynamically selects the appropriate hardware to adapt to application performance needs. By leveraging programmable switching fabric, networks are consolidated with programmable switches and hardware acceleration devices such as Intel Stratix 10 FPGA and IPUs for data processing, while CPUs concentrate on computing for control plane and customer applications.

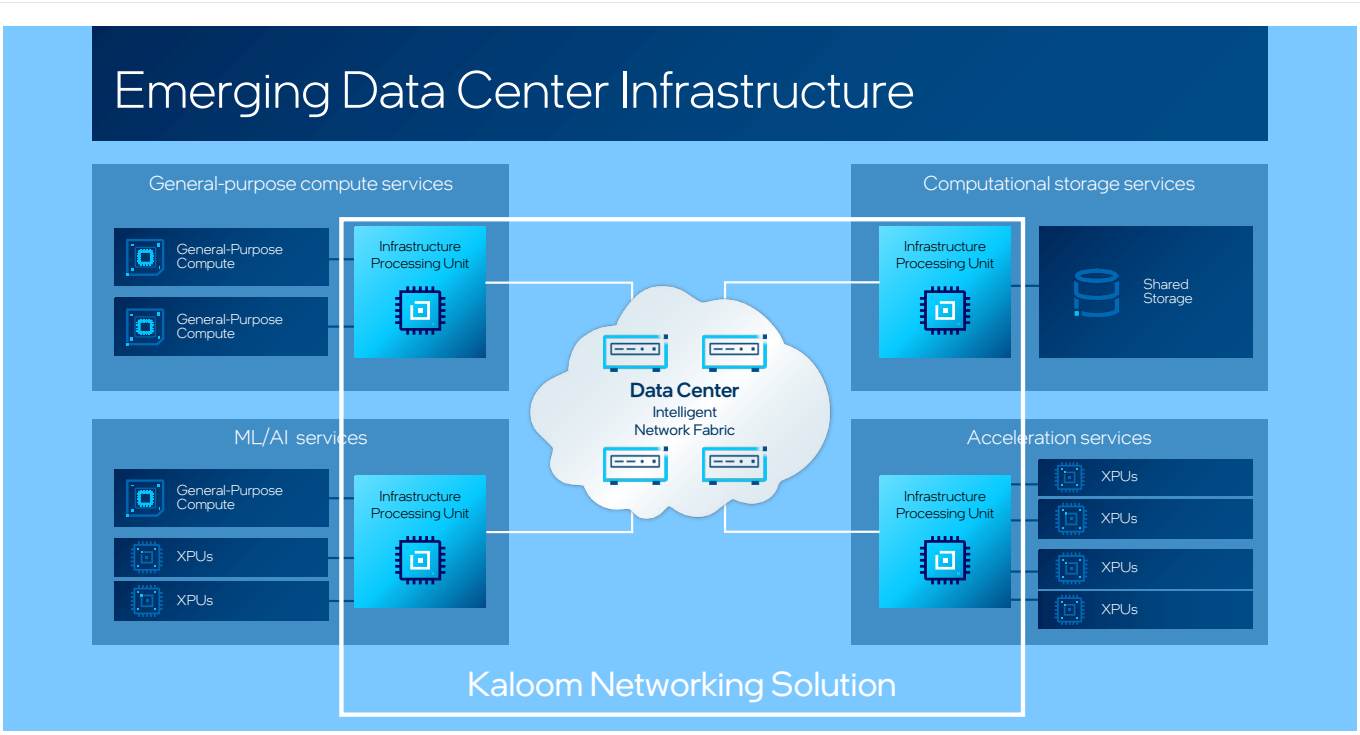

Figure 2. Kaloom programmable fabric with interconnected IPUs

Intel® Infrastructure Processing Units (IPUs) are an advanced networking device with hardened accelerators and Ethernet connectivity that accelerates and manages infrastructure functions using tightly coupled, dedicated, programmable cores. An IPU offers full infrastructure offload and provides an extra layer of security by serving as a control point of the host for running infrastructure applications.

Unified Edge Solution Consolidates Networking Functions for a Leaner Edge

Edge computing has become a new revenue battleground to monetize 5G applications. However, edge data centers are constrained by rack space, cooling, and power. They need more real-estate to accommodate increasing numbers of application servers for generating revenues while processing more data with less network infrastructure.

Kaloom's Unified Edge solution extracts the most value from limited infrastructure assets that can be accommodated at the edge. It embeds the Red Hat OpenShift Container Platform (RH-OCP), including the Linux OS and Kubernetes. It significantly raises rack space utilization at the edge using RH-OCP supervisors and worker nodes for network operations and customer applications. The same Supervisors running on Intel® Xeon® Scalable processor-based bare-metal servers can be used to orchestrate the deployment of Kaloom, 3rd Party applications and customer cloud-native applications.

Additionally, by running the network fabric controller, the 5G UPF control and data plane components, and RH-OCP Supervisors on Intel Tofino IFP-powered programmable switches, the Unified Edge solution frees up processors on application servers for customer revenue-generating applications.

Containerized Networking Aligns IT and the Network World for Lockstep Operations and Scalability

The use of OpenShift, including Red Hat Enterprise Linux CoreOS, provides a single OS and Kubernetes environment across servers, programmable switches, and IPUs. Furthermore, Kaloom augments RedHat OpenShift creating a complete network-optimized application container platform, including computing, storage, and networking. Kaloom provides an advanced Kubernetes container networking interface (CNI) with support for multiple interfaces and dynamic network-change discovery. These mechanisms are a must for carrier-grade and high-availability applications.

The network-optimized application container platform enables the management of switches at scale, precisely like mainstream application servers.

Programmable White-Box Switches Increase Agility and Enhance the Cost-Performance Ratio

Kaloom's Unified Edge solution is a highly automated and virtualized network fabric (see Figure 2). Kaloom CNFs are defined with a clear separation between the control plane and the data plane. The control plane is several interworking cloud-native applications consisting of software components written in Go, deployed as containers over a Kubernetes framework provided by RedHat OpenShift. Examples of CNFs include Kaloom Virtual Router (KVR), Kaloom VXLAN Gateway (KVG), and Kaloom 5G UPF control plane. Kaloom has embraced P4 as a data plane programming language. One P4 application typically comprises several P4 components running over various chipsets in a white box server. A particular P4 component could run on an Intel Tofino IFP, Intel Stratix 10 FPGA, Intel IPU, or Intel Xeon Scalable processor (e.g., P4 to DPDK).

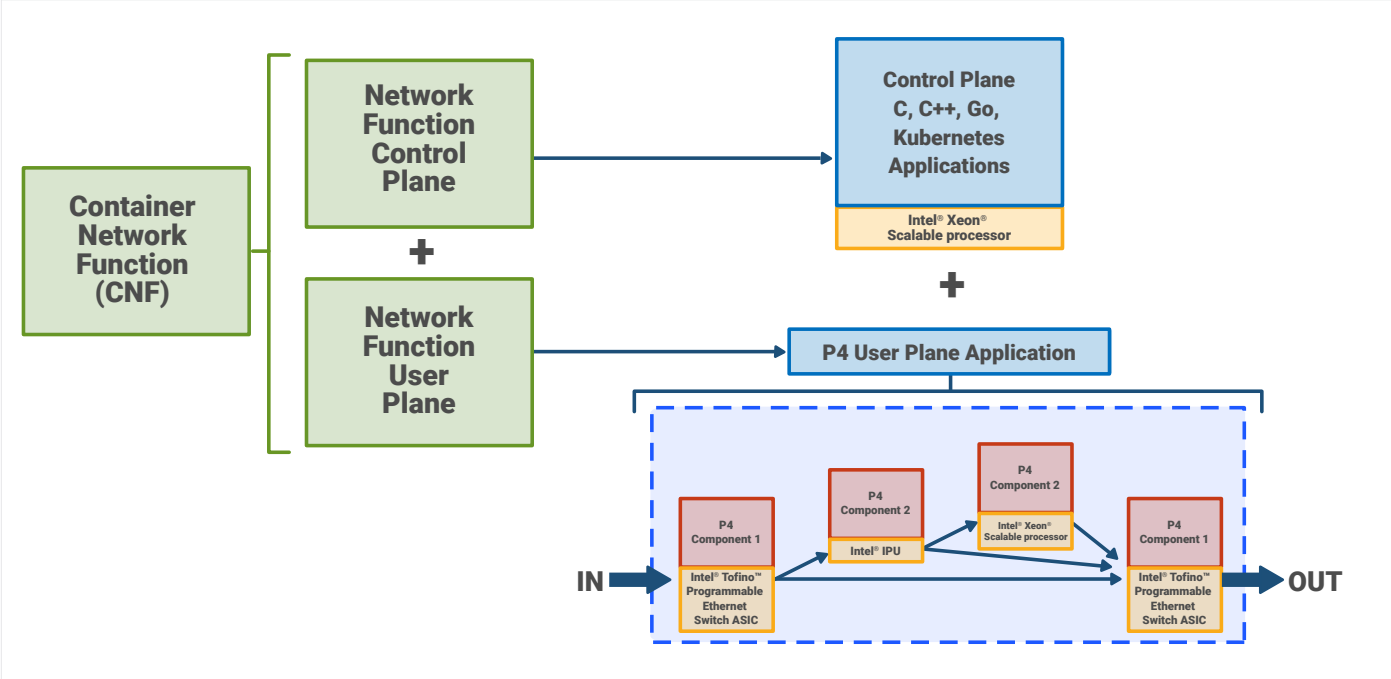

Figure 3. Heterogenous hardware acceleration

The physical infrastructure is programmable to keep pace with the scalability and quality of service requirements of diverse applications. Additionally, it dynamically responds to changes in protocols and new standards. Using P4 programmable white-box switches lowers the fixed costs of hardware and increases the overall performance, thus requiring less hardware and power. All Kaloom virtual network functions can sustain Terabits/second throughput at or near sub-microsecond latency.

Unique Fabric Virtualization Creates Fully Isolated Network Slices

Kaloom's Unified Edge solution natively supports the concept of "slicing." It enables partitioning a physical data center into multiple independent and fully isolated virtual data centers, each with its virtual fabric (vFabric). It allows elastic allocation of resources to a virtual data center, creates fully isolated broadcast domains, and permits the deployment of multiple private clouds. Furthermore, each domain can be assigned different attributes for varying use cases or to meet unique SLAs. For example, they can be part of a virtual data center (vDC) for operators to offer various cloud services, and each of its vFabrics can simultaneously host millions of cloud service users (e.g., tenants).

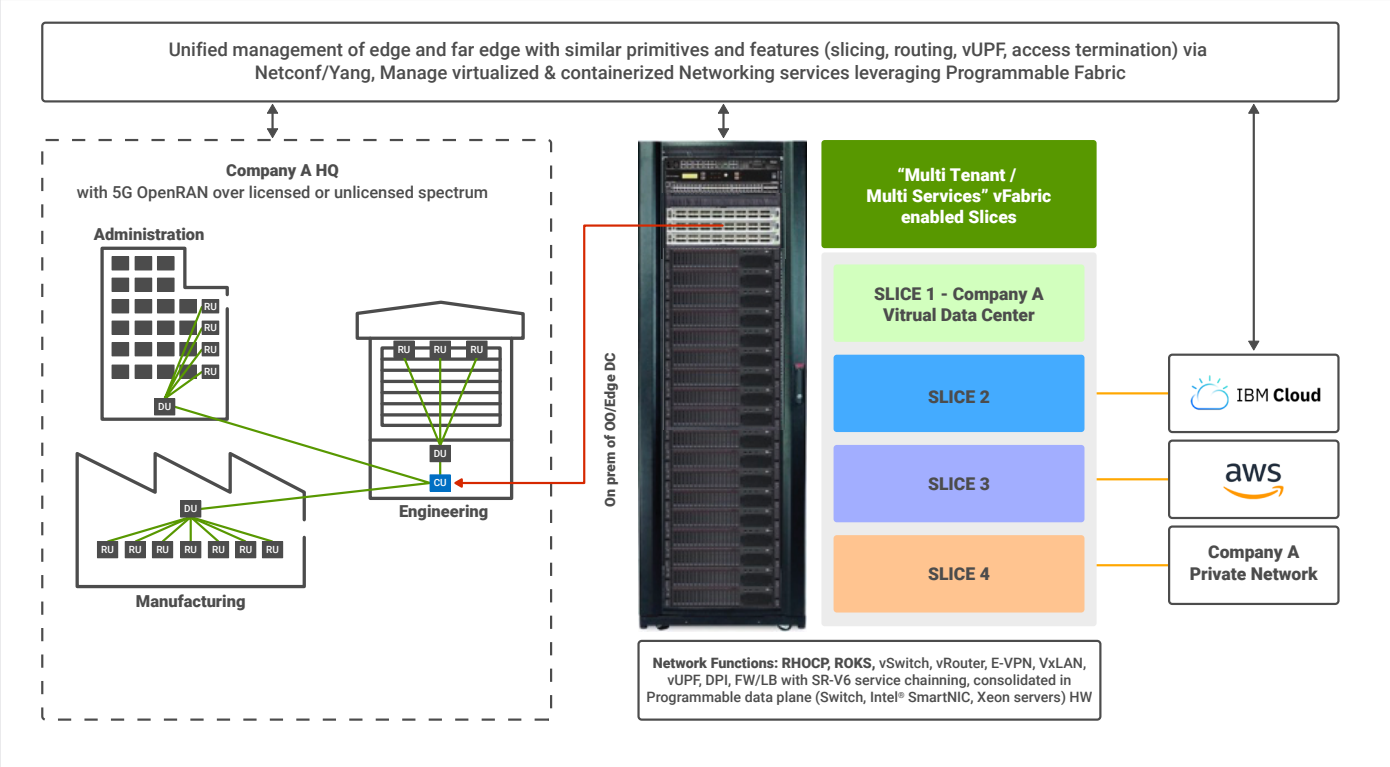

Figure 4. Fully isolated slicing

The Kaloom Unified Edge solution is designed for multi-tenant operations, providing services in isolated network slices (see Figure 4). P4 programmability at the data plane layer enables the data flows to smoothly meet each slice's performance and quality standards. The formation of a network fabric is fully automated with integrated self-discovery and self-forming mechanisms. vFabrics are scalable because they don't rely on VxLAN identifiers with an inherently limited capacity to expand the number of nodes in a network.

Srv6-Based Service Chaining Marshals End-To-End Networking and Quality of Service

To host 5G applications, data centers need the simplicity and scalability offered by the next-generation segment routing to interconnect distributed transport network domains seamlessly. Kaloom uses Segment Routing over IPv6 (SRv6)-based intelligent traffic engineering for automated end-to-end service chaining. Kaloom's fabric service chains all its CNFs – switches, routers, gateways, and 5G UPF – and even third-party network functions. It uses a single intelligent source and relieves the remaining intermediate nodes from the task of calculating the required path through the network.

Furthermore, networks need to exercise strict control over the quality of service for mission-critical applications. Operators require fine-grain visibility to identify the root cause of every single problem. Some transient problems have proven even more complex and elusive to debug, making legacy solutions like ping, traceroute, sFlow, mirroring, or even packet captures incapable of identifying the root cause. In-band Network Telemetry (INT) changes the game by providing visibility and answers to the following questions: 1) what is the path used by this packet? 2) what was the state experienced by this packet at each hop? 3) what other flows contributed to this specific state?

SRv6 with hop-to-hop telemetry for in-band operations, administration, and maintenance (IOAM) lays the basis for network slices to achieve the service quality goals of 5G communications. In the event of failure of some network functions, real-time in-band network telemetry data helps to make automated corrections in service chaining and keep packets flowing using the best possible path. Network operators gain visibility into the state of health of network assets and can use AI to spot early signs of failure and promptly make corrections.

Kaloom offers in-band telemetry with the concept of "postcards," enabled by the support of the P4 programmable pipeline, allowing real-time fine-grained visibility for simplified troubleshooting and reduced OPEX in the context of not only debugging and day-to-day management but also capacity planning.

Conclusion

The overall efficiency of 5G networks requires the optimal functioning of its hardware and software resources. Kaloom believes in a heterogeneous hardware acceleration architecture delivering end-to-end network applications leveraging Intel Xeon Scalable processors, P4-programmable Intel Tofino IFPs, Intel Stratix 10 FPGAs, and Intel IPUs. Each can be used optimally to provide differentiated service attributes in network slices.

Intelligent fabric virtualization is the bedrock to improve efficiencies by consolidating network islands and data processing. It reserves a higher proportion of servers for driving applications, segments the network to increase its utilization, and sparsely utilizes scarce resources at the edge. P4 programmability provides the means to consolidate and configure network resources for the chosen objective.

The Unified Edge solution has a single execution environment for rapidly assembling computing, storage, and networking resources. The containerized network fabric software, along with Intel programmable hardware, lends flexibility to the transport network for dynamically adapting to the changing application needs of evolving cloud and edge data centers.

Learn More

Kaloom Home Page

Kaloom Unified Edge Fabric

Intel® Xeon® Scalable processors

Intel® Tofino™ IFP

Intel® Stratix® 10 FPGAs

Intel® Infrastructure Processing Unit

Intel® Network Builders

Notices & Disclaimers

[i] IPU Based Cloud Infrastructure White Paper
https://www.intel.com/content/www/us/en/products/docs/programmable/ipu-based-cloud-infrastructure-white-paper.html

Intel technologies may require enabled hardware, software or service activation.
No product or component can be absolutely secure.
Your costs and results may vary.

0223/LV/H09/PDF ♻Please Recycle 354331-001US