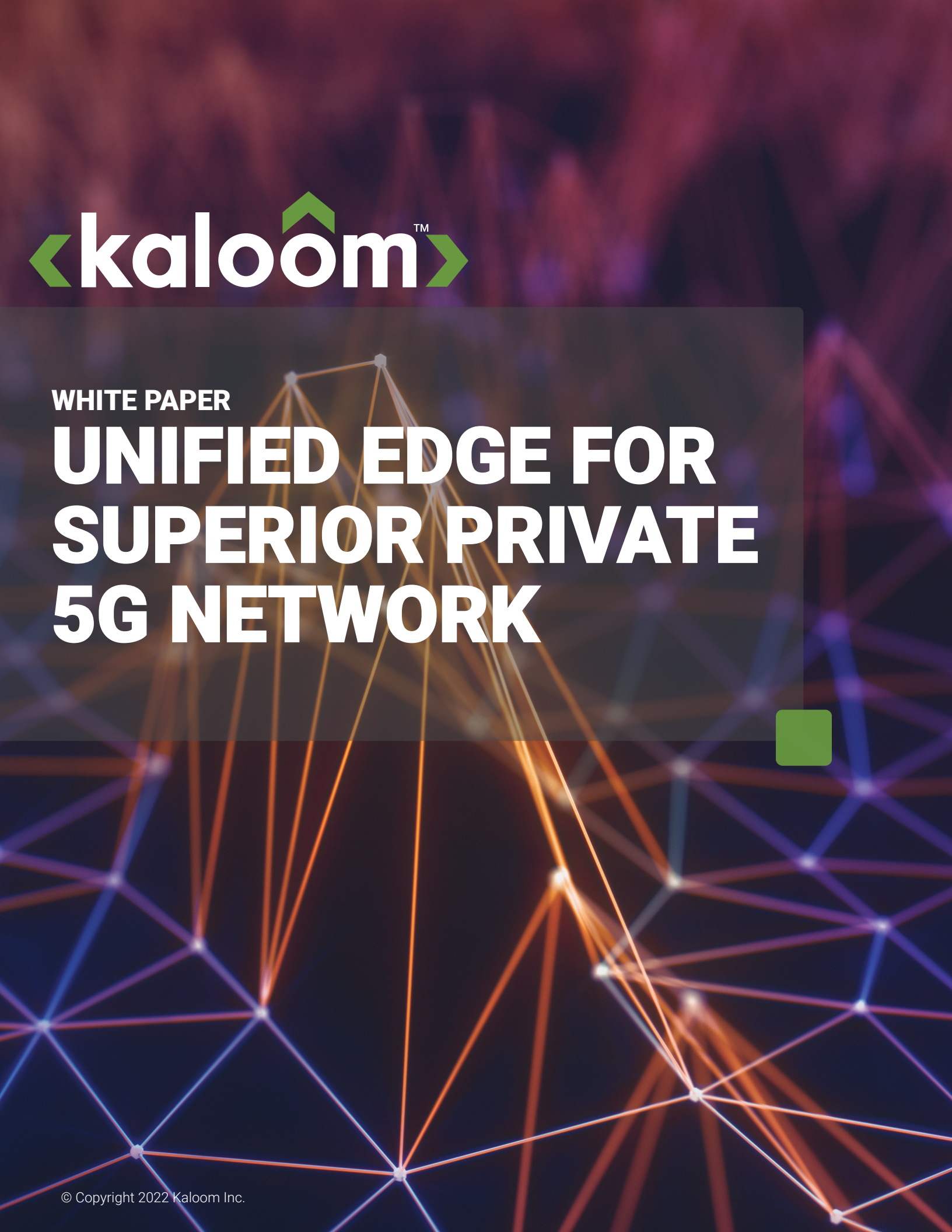


WHITE PAPER

UNIFIED EDGE FOR SUPERIOR PRIVATE 5G NETWORK

Introduction

5G creates a conducive environment for enterprise transformation and digital innovation for bolstering data throughput, significantly reducing latency and jitter for applications such as industrial 4.0. In a private 5G network environment, enterprises can exercise precise control on picking and choosing network and IT services at the edge for customizing network design to meet unique business needs.

Enterprises have a pressing need for impromptu marshaling of network resources and application services to respond to transient demands for business-critical applications. They need programmable network fabrics that create an open rubric to consolidate, share and reuse network resources and intelligence for multiple uses. By contrast, current networks are severely impeded by hardware, proprietary architectures, and fixed protocols preventing autonomous responses, based on real-time intelligence, for flexible responses to highly dynamic business needs.

5G Private cellular networks and programmable network fabrics have a symbiotic relationship—together they create consolidated and shareable networks able to serve multiple customized application needs.

Unified Edge for coordinated network response

Kaloom's Unified Edge™ solution accelerates network responses with the synchronization of network, compute and storage functions. Kaloom and Red Hat have collaboratively envisioned and developed Unified Edge solution that meets the unique needs of emerging 5G edge data centers. Kaloom Unified Edge is a highly automated and virtualized network fabric with integrated switching, routing, and 5G data processing. The solution is based on open networking principles and is built with Red Hat® OpenShift® Container Platform as the foundation of a unified edge solution for network, compute, and storage. Together, they create a programmable infrastructure that meets automation and digital innovation needs.

The Unified Edge enables a seamless flow of unimpeded data traffic and unencumbered network operations governed by a common containerized virtual infrastructure for all edge network functions and applications. Its native multi-tenant operations provide services in isolated network slices and calibrate networks for scarce resource use at the edge whether it's a customer premise or collocated at DC service provider where space, power, and hardware capacity are in short supply.

Digital Applications and Private 5G Enables Bespoke Edge Network

Today, most business applications are running in the cloud, which is not always scalable and sustainable as performance suffers with expanding number of applications. Enterprises need control over their network to cope with the digital transformation and continuous innovation markedly accelerated post-Covid. They want to move business-critical applications to the distributed edge infrastructure where latencies are much lower and give them better control to ensure quality and data security.

In addition, new industrial automation applications are data-intensive and demand extremely reliable wireless networks with sub-4 milli-second latencies, which neither the current Wi-Fi nor public 5G networks can support. Private 5G networks, at the enterprise edge, provide a viable environment for enterprises to innovate IT/OT applications.

In the private 5G edge infrastructure, enterprises can exercise control to pick and choose network and application elements for customizing and executing their evolving solutions. It opens the way for bespoke enterprise network and IT infrastructure that they need in a multi-use, multi-location, multi-vendor, and multi-technology environment to meet moving targets for application and service performance.

Bespoke Private 5G Edge Needs Elastic Network Infrastructure

A prerequisite for adopting 5G networks and edge technologies in the enterprise is that its execution should not disproportionately increase risk. In the past, turnkey network projects were the preferred choice for the enterprise. They relied on experienced incumbents to take complete responsibility for the design and execution of projects while managed services providers were responsible for operating them. Unfortunately, such an approach leaves very little room for the enterprise to achieve flexible network solutions with the favorable attributes of 5G.

Enterprise CIOs must leverage network solutions that dynamically adapt to application needs and partner with credible integrators to ensure the success of the 5G-enabled edge applications. To support the evolving needs of business-critical applications, major network design changes are needed to streamline the complexities at the distributed edge.

Ensuring resource efficiency at the edge

At the edge, space is limited, and resources cannot be scaled at the demand rate. Currently, demand growth leads to a disproportionate increase in resources due to silos created by proprietary equipment.

Resource utilization efficiency is much more daunting in the 5G world with diverse performance requirements. Mission-critical applications such as robotics, telesurgery, and X-Reality have expanded and made exacting demands on throughput and latency. By contrast, the resource requirements of non-mission-critical applications are relatively modest.

As the diversity of applications grows, it becomes increasingly costly to deploy dedicated network resources. Multi-use programmable network fabric with isolated virtual network slices is needed for full capacity utilization of assets and to lower their per-unit cost. However, the resources need to be in a shared pool and have a standard operating platform for their redeployment for multiple uses.

Adapting to dynamic application demands

Enterprises IT are coping with a diversity of needs springing from hybrid workforces, digital customer engagement, and operational improvement initiatives from multiple departments of global companies. Unfortunately, current enterprise network infrastructure is not supple enough for rapid reconfiguration to meet the needs of a raft of video and X-Reality applications, flexible automation, reams of IoT data for automated and intelligent management of work processes, and pervasive cybersecurity threats.

Traditional network switches operate with embedded software with little scope for programming to increase their data processing capacity or adapt to the changing traffic mix. Current Software-Defined Network (SDN) methods, with tools like OpenFlow, are limited in adjusting forwarding rules for directing data traffic to conform to desired outcomes.

The forwarding of data across the diversity of environments in fluid traffic conditions with fixed protocols is fraught with the risk of misalignments in traffic flow. New methods for higher-level data plane programming, such as P4[1] , are needed as the heterogeneity of network applications expands.

Handling data growth cost-effectively at the edge

The edge is potentially a choke point as data volumes grow and the networks serving them. User Plane Function (UPF), the workhorse of the 5G network, ensures that packets get to the right destinations, with the right priority, after being processed or modified as necessary. For applications requiring ultra-low latencies that 5G promises, 5G UPFs are deployed at the edge.

The performance expectations for the user plane have risen exponentially to hundreds of Gb/s per second throughput, latencies lower than 4 milliseconds while supporting a large number of connected devices. However, the UPFs implemented in proprietary hardware are inflexible. The ones implemented entirely as software on general-purpose compute servers cannot scale efficiently, putting new digital services' quality and overall growth at risk. Linear scaling of servers to support data growth is not always an option at the edge.

A recalibration of data processing network functions is warranted to get higher throughput with programmable hardware acceleration and modular virtualized software for utilizing resources proportionate to the needs.

Optimizing data flow at the edge

At the edge data centers, it is essential to remove the redundant movement of traffic to avoid congestion. Traditional switches do not have in-built computing capability and send data to servers elsewhere to perform compute-intensive network services such as load balancing, denial-of-service attack detection, and 5G UPF.

While the additional servers add to the capital cost, operating costs and network latency increase with the back-and-forth movement of data between switches, routers, and other network functions. Likewise, networks isolated from computing and storage increase the complexity of data flows between them and increase operating costs for their interoperation. By contrast, enterprises can consolidate data processing for multiple network functions on programmable white box switches, thus significantly improving network latency and costs.

Ensuring rapid response with real-time intelligence

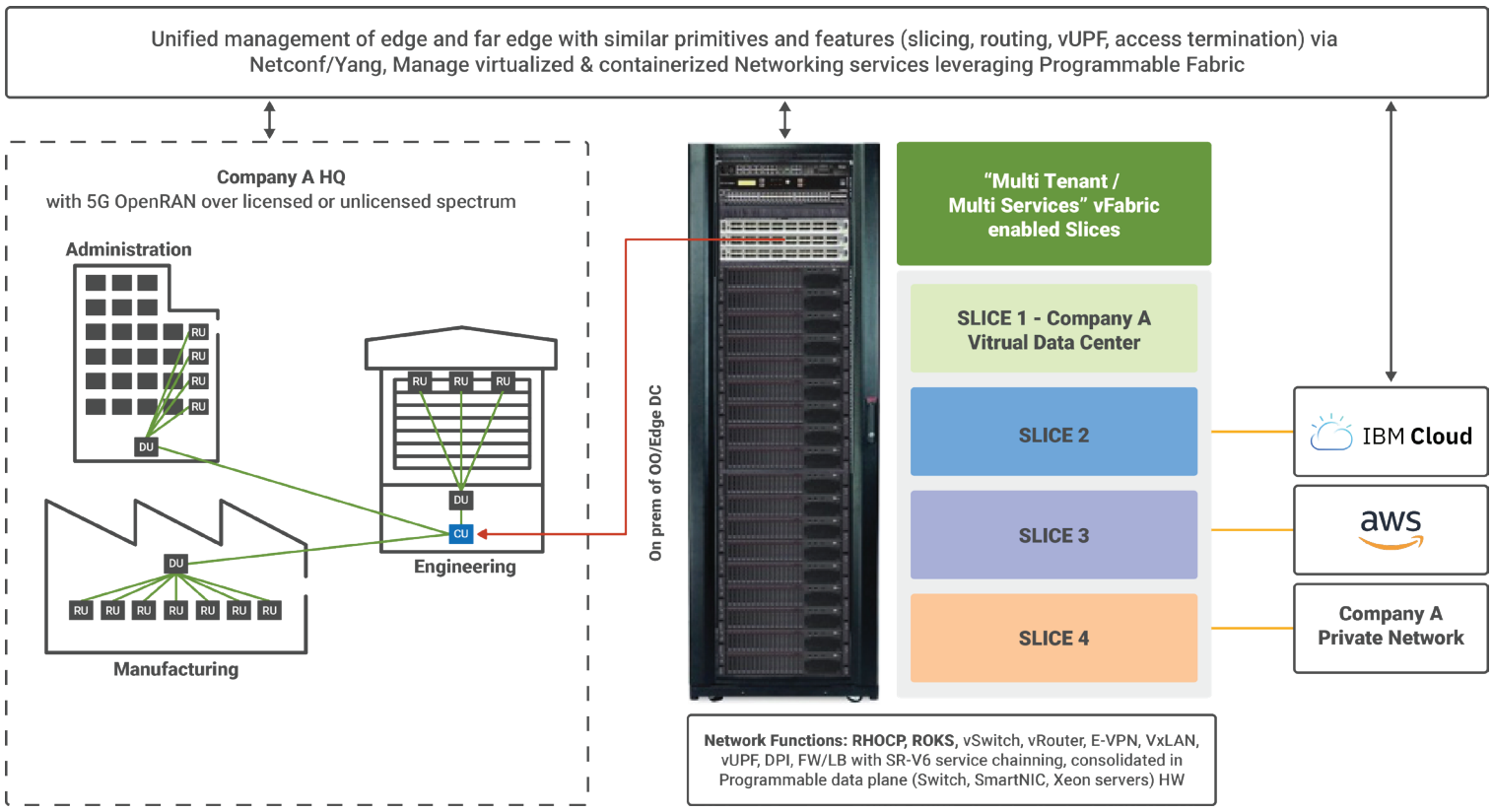
Industry 4.0 applications evolve faster as they learn continuously from data analytics to meet business needs better and improve their performance. Periodic intelligence gathering in current networks is delay prone with intermittent data gathering by polls. Consequently, the responses for corrective action are slow. Data is not gathered at a granular level closer to data flows and variances detected in real-time.

Additionally, open interconnected network mesh is susceptible to Distributed Denial of Service (DDoS) cybersecurity attacks that snowball into their catastrophic collapses. In the absence of granular intelligence, detecting the early signs of a mounting attack is tricky.

Networks should be able to respond to in-band telemetry in real-time and make changes in the flow of data to close the gap in latency metrics. Elasticity is needed to make course corrections to achieve agreed service quality metrics, especially for mission-critical applications in the 5G environment.

Unifying network fabric for the distributed cloud

The edge nodes in the enterprise network could be on-premises at hosted edge cloud, enterprise data center, individual branches, or even on moving objects such as ambulances and harvesters. Traffic flows not only to and from the cloud and the edge but also between edge nodes. Furthermore, traffic is distributed across multiple and hybrid clouds. It is possible to lose sight of data flows when they span heterogeneous networks and IT architecture.



The enterprises need unified software-based fabric allowing for workloads to transition and communicate across a single fabric from the edge to the data center, facilitating an end-to-end design that can traverse different transports. It must additionally support programmability to effortlessly route data using real-time intelligence to avoid congestion and a deterioration in the quality of service.

Kaloom's Unified Edge Provides Programmable Networking for the Private 5G Edge

Kaloom believes that data processing in the 5G world calls for an entirely new network architecture that disaggregates software from the hardware of telecom equipment, is cloud-native, fully programmable, self-forming, and automated.

Kaloom's Unified Edge is a highly automated and virtualized network fabric with integrated switching, routing, 5G UPF, and Red Hat OpenShift container platform for the accelerated application deployment at the private 5G edge. It uses only three OCP-compliant white box switches equipped with Intel XEON processors (used for control and management plane functions) and fully programmable high-performance packet processors (Intel Tofino/Tofino2) for P4 data plane applications thereby significantly reducing capital and operating costs for edge deployments. For session scalability, the Unified Edge fabric includes as well Intel Stratix10 FPGAs, bringing the number of sessions up to 2M per FPGA, which can be combined with the Tofino (or later with the Tofino2) in a smaller form factor (server switch).

Key Benefits of Kaloom Unified Edge

> 10x Cost Savings

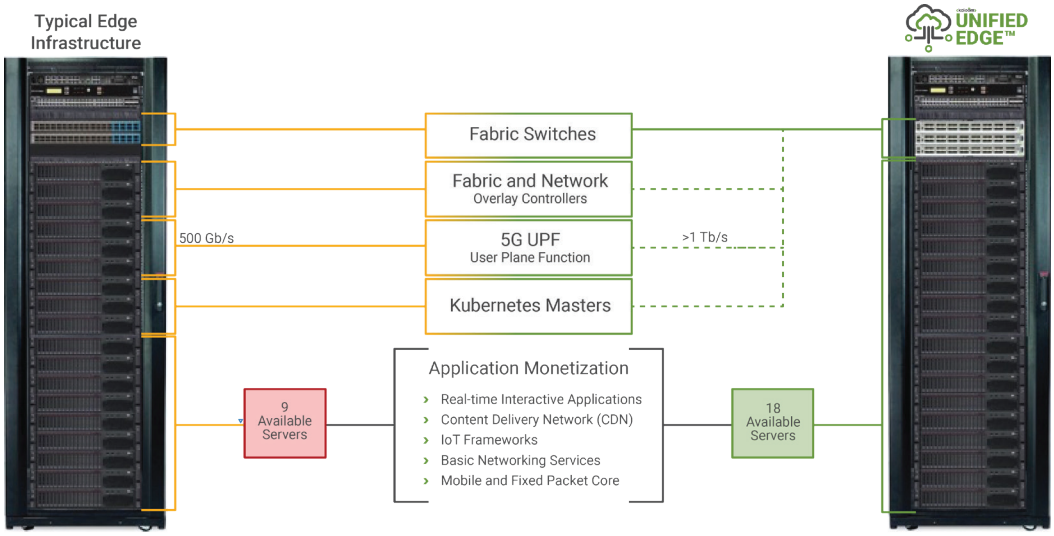
The consolidation of multiple control and management functions onto the three white-box switches and leveraging capabilities of the programmable ASIC for data processing, the solution reduces the space used by up to 8 RUs, enabling over 10x cost savings.

2x Throughput and 7x Reduced Latency

By offloading virtual switch functions from the server NICs onto the switching ASIC, Kaloom can achieve up to twice the throughput and reduce latency by up to seven times, increasing the overall system's performance while consuming fewer resources. Reduced latency will also help meet the needs of private 5G mobile networks that are deployed for latency-sensitive industrial workloads that require more certainty and performance in network connectivity

Up to 2.3 Tbps Throughput

The integrated solution supports the elastic scaling up of workload nodes and maintains the flexibility to add more hardware assistance as needed to handle workloads with increasing complexity.



Lowest cost of ownership. Accelerated ROI with most resources available for monetization.

Programmability adds flexibility, agility, and scalability

By taking a software-centric and programmable approach that leverages different hardware options (P4 programmable ASIC, FPGA, Smart NICs, DPUs, CPUs), Kaloom's Unified Edge easily adapts to dynamic application needs, including new requirements of 5G (and evolving wireline) networks. Its integrated virtual router and UPF can sustain over 1.3 terabytes of data and sub-4 milli-second latency that CPU-powered computing devices cannot achieve.

Cloud-Native fabric unifies compute and network

Kaloom's cloud-native network fabric software increases the efficiency of space utilization at the edge by using containers to replace virtual machines. As Kubernetes has become the standard for the orchestration of containers on computing, Kaloom has applied the same technology to the network component by integrating Red Hat OpenShift into its Unified Edge solution. This approach allows enterprises to use the same Kubernetes platform and the underlying Linux distribution over servers and switches.

Using the same orchestration layer for these components makes resources at the distributed edge and hybrid cloud better streamlined and optimized, significantly lowering the friction for rapid deployment of network resources for edge applications.

Virtual Fabrics Enable Secure 5G Network Slicing

The Kaloom network fabric is the first product with full support for network slicing. The Kaloom Unified Edge is designed for multi-tenant operations, providing isolated network slices for a 5G network. Once instantiated, the Unified Edge creates a self-forming virtualized fabric (vFabric), ensuring the complete separation of the services down to the hardware level for better security and quality of experience fitting the diverse application needs of the enterprises, yielding additional cost savings. Within each slice, it is possible to deploy several instances of virtual network functions such as virtual routers, virtual firewalls, or virtual 5G UPFs. Further, these slices can be created dynamically via automation and orchestration through API calls, supporting rapid and flexible 5G slice deployment

Conclusion

At the edge, the enterprise needs higher levels of reliability, security, and performance with less space, assets, and resources. Kaloom meets the seemingly conflicting demands by making greater use of software-defined networks, intelligence, and programmability to raise performance while reducing costs.

Kaloom's programmable and unified edge contributes to the elasticity of networks by dynamically configuring its components to serve the evolving demand of applications. The accent is on consolidation and unification of networks to remove the friction and redundancies that lead to sub-optimal utilization of networks. It's programmable 5G UPF optimizes data processing on heterogeneous hardware assets for the best performance and lowest costs.

For more information please visit: www.kaloom.com or contact our sales representatives at sales@kaloom.com