Solution Overview

Highlights

- Distributed, controllerless SDN fabric easily spanning geographically dispersed centralized and edge data centers so they appear as one logical entity
- Comprehensive network slicing with full isolation across management, control and data planes
- Rich per-tenant services including Layer 2/3 VPNs, VXLAN-based Layer 2 extension and Layer 1 VirtualWire™
- Automated, fabric-wide provisioning for configurations, services and policies with a single command via RESTful APIs, ensuring consistency and reducing errors and provisioning complexity
- Pervasive fabric-wide visibility of attached devices and traffic flows for real-time troubleshooting and proactive network planning
- White box economics reduce CapEx by 30-60%

Pluribus NETWORKS

The Network Fabric for Distributed Cloud and 5G

Simplifies the networking of multiple edge compute locations through an open networking fabric based on distributed SDN intelligence and VXLAN virtualization, implemented on white box switches

Historically, most data center, private cloud and public cloud implementations have been built around a centralized architecture where storage, compute and networking resources reside in one or two locations, or at most a handful. With the advent of new technologies such as 5G, artificial intelligence (AI), machine learning (ML) and the Internet of Things (IoT), there is an emerging class of applications that has a new set of requirements that cannot be met by this centralized cloud architecture. This new class of applications demands that compute resources be deployed at the network edge, closer to users and things, so they can deliver on these new requirements.

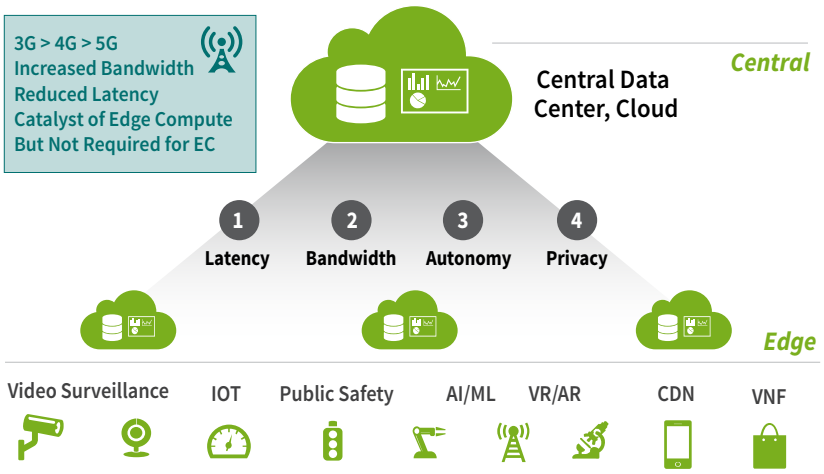

Figure 1: There is a set of emerging applications with requirements that cannot be met with centralized data centers and cloud architectures, requiring that compute and storage be deployed at the edge, closer to users and things.

There are four fundamental drivers of edge compute:

1. **Reduced Latency.** Many emerging applications demand low round-trip latency. AR and VR, for example, target round-trip latency of less than 20 milliseconds (ms).
2. **Bandwidth Cost.** The cost of bandwidth drives a need to process data at the edge, such as filtering video surveillance footage from hundreds or thousands of high-definition cameras.

3. **Full Autonomy.** When sensors and actuators work together in an IoT traffic safety application, those interactions and analytics must continue even if connectivity to a central cloud is lost.
4. **Data privacy.** Data sovereignty and the European Union's General Data Protection Regulation (GDPR) require that organizations that store user data be able to contain that data in a known location and be ready to show it immediately upon user request.

Distributed Cloud

Distributed cloud is a fusion of edge compute with the cloud consumption model. Edge computing deploys resources closer to users and things into new locations outside the traditional data center – central offices, modular data center containers, remote environments and more. The cloud consumption model delivers the benefits of spinning up and using infrastructure and services as needed, scaling up and down to meet variable demand and only paying for what you use. 5G is a catalyst that brings ubiquitous bandwidth, lower latency, network slicing and architectural enhancements that will help accelerate distributed cloud. However, 5G is not required for distributed cloud; these two technologies will progress in parallel and accelerate as they come together.

Distributed cloud will have many edges, and the placement of workloads will be determined by the location where the application's requirements can be met at the lowest cost. There will be an increase in the number of mini and micro data center locations, possibly by a factor of ten or more. With this comes increased complexity and a need for a counter balance – a highly automated network fabric that can make multiple edge locations appear as one logical unit in order to simplify the management of multiple remote sites.

Next-generation Software-defined Networking Fabric for Distributed Cloud and 5G Edge

Pluribus Networks' Netvisor® ONE network operating system (OS) and the Adaptive Cloud Fabric™ (ACF) deliver a controllerless software-defined networking (SDN) fabric that provides a VXLAN overlay and is ideal for simplifying the networking required for distributed cloud. The Linux-based Netvisor ONE OS runs on open networking hardware that provides cost-effective white box economics to mitigate the growing expense of deploying infrastructure at multiple edge locations.

Layered on top of Netvisor ONE is the Adaptive Cloud Fabric, which uses distributed intelligence to create a controllerless SDN fabric that easily spans multiple edge sites and makes multiple switching nodes appear as one logical entity, which can dramatically reduce operational complexity and associated costs. The fabric is architected with no single point of failure and delivers extremely low latency to support distributed cloud services regardless of geographical location. The fabric is also deeply sliceable across the management, control and data planes, as mandated by the 3GPP standards body for 5G, providing multi-tenant scale with rich services and granular telemetry per slice.

Open Networking with White Box Economics

Netvisor ONE is a Linux operating system based on the open source FRRouting routing project and is instantiated in one or more lightweight containers on bare metal leaf and spine switches. The switches are Open Compute Project (OCP) and Open Network Install Environment (ONIE) hardware-compliant, providing an open, secure and programmable next-generation network OS. They include hardware from well-known vendors and suppliers such as Dell EMC, D-Link and Edgecore. These open, disaggregated hardware solutions avoid vendor lock-in and can achieve savings of up to 60% over branded, vertically integrated, proprietary solutions. The capital deployment cost is particularly important in multi-site distributed cloud deployments where costs can quickly grow.

Adaptive Cloud Fabric, a Controllerless Software-defined Networking Fabric

First-generation SDN is based on a centralized controller that holds the state of the network and uses the out-of-band (OOB) management channel to program switch nodes in the fabric. However, with multi-site data center and distributed cloud deployments there are a number of drawbacks to this approach, including the cost of the controller licenses and supporting servers, latency to communicate back to the controller impacting new flows and reconvergence, as well as multiple single points of failure associated with OOB communications. The Adaptive Cloud Fabric is a next-generation SDN implementation. It boasts the benefits of a centralized view of the network, but the network state and the associated SDN intelligence are distributed to all switch nodes in the fabric – there is no central controller. This results in the ability to span multiple sites across campus, across town or even across the world with no latency penalties and a highly resilient fabric.

Distributed Cloud with Adaptive Cloud Fabric

Seamless scalability across multiple locations and providers

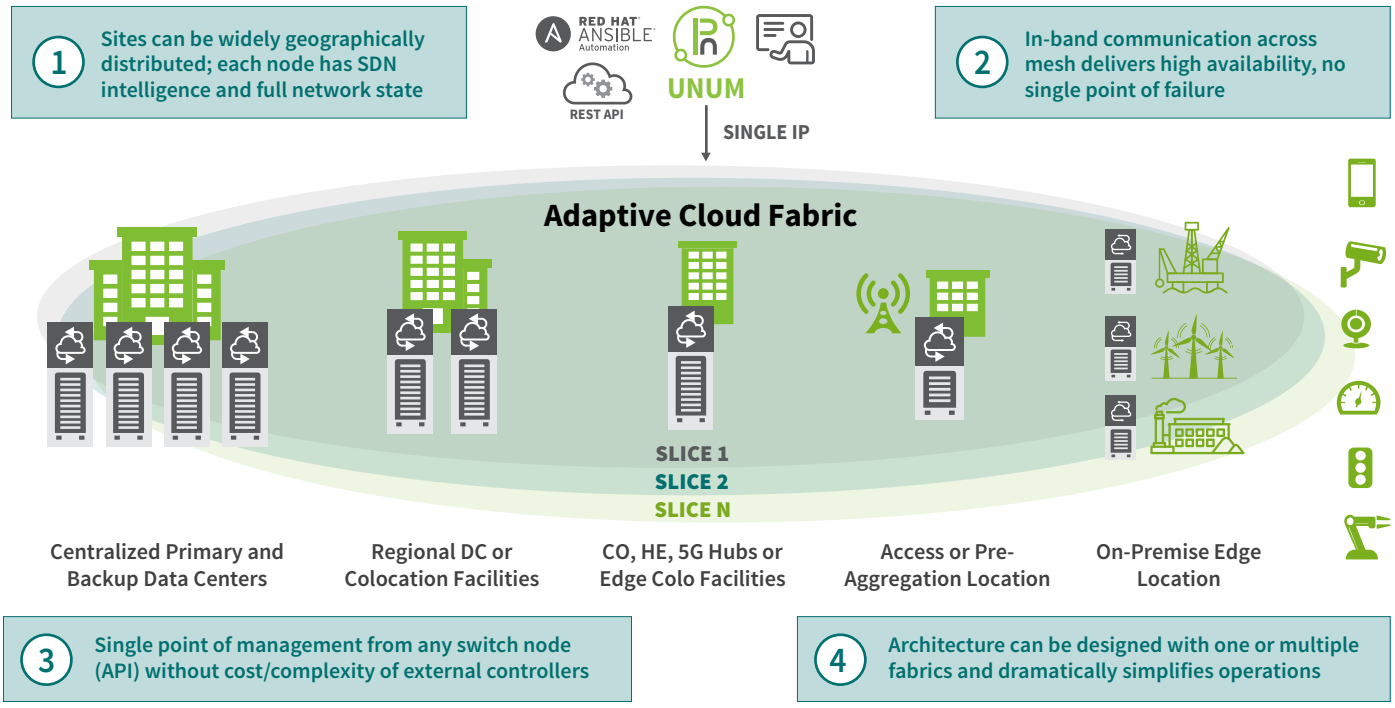

Figure 2: The Adaptive Cloud Fabric stretches across all edges, enables dramatically simplified operations and provides rich services, programmability and telemetry per slice.

The deployment of the fabric is simple and highly automated – all that is needed is simple Layer 3 connectivity. First, all of the switches federate in a SDN management plane fabric and share full state across every switch node. Second, an underlay topology is built that can work with open networking spine switches running Netvisor ONE, or third-party spine switches in a brownfield deployment scenario, including legacy spine switches from vertically integrated vendors. The brownfield deployment is a challenge with controller-based solutions but can be achieved with the controllerless SDN design because standard Layer 2 and Layer 3 protocols are run in the underlay to connect with the existing infrastructure. The third and final step is the automated provisioning of a VXLAN overlay in the desired topology that virtualizes the network and distributes network service capabilities to all nodes in the fabric.

Simplifying Operations

The Adaptive Cloud Fabric can simplify a multi-site deployment because it can be used to create a single logical fabric (or multiple fabrics if desired).

Once the fabric is deployed, every switch in the fabric is aware of the full state of the fabric and communicates with every other switch regardless of geographical location, and any switch in the fabric can control all switches in the fabric. That means the network operator can make a configuration change via a CLI command or a REST API call to a single switch in the fabric and the distributed fabric intelligence in turn will populate that change to all switch nodes. If for some reason a node cannot accept a change, no node in the fabric will accept the change until the blocking problem is resolved, ensuring consistent configuration and reducing human error across the fabric.

Both NetOps and DevOps automation tools, such as Ansible or the Pluribus UNUM™ management platform, are available to provision either individual nodes or the fabric. In addition, traditional NetOps interfaces and a wide array of Linux tools are supported for scripting and automation. As a result, workflow automation reduces configuration time by up to 90% over traditional box-by-box management, lowers the risk of configuration errors and dramatically improves service velocity and operational agility.

Agnostic to Compute Layer Virtualization

Many network virtualization approaches are executed in the compute layer, which creates a lock-in between the compute virtualization layer and the network virtualization layer. The Adaptive Cloud Fabric terminates the VXLAN tunnels in the networking switch using hardware-based virtual tunnel end-points (VTEPs). This allows the compute layer to be implemented with any compute virtualization technology desired and use any orchestration solution – VMware with vSphere, KVM with OpenStack, containers with Kubernetes, bare metal implementations and so on. Offloading the network processing of the VXLAN tunnels to the switch also frees up compute cycles on the server to better scale workloads, allowing network traffic be processed at wire speed by the switches.

Industry-leading Network Slicing

Network slicing is mandated by 3GPP for 5G. Also known as segmentation, network slicing has been designed into Netvisor ONE and the Adaptive Cloud Fabric from the ground up. The network can be comprehensively sliced across not only the data plane like many other implementations, but also across the control and management planes. This provides multi-tenancy in which each tenant is completely isolated and can manage their slice independently using whatever tools they desire, with full isolation from all other tenants. Slicing can be used to separate lines of business, applications with different performance requirements or untrusted traffic that may expose a large attack surface.

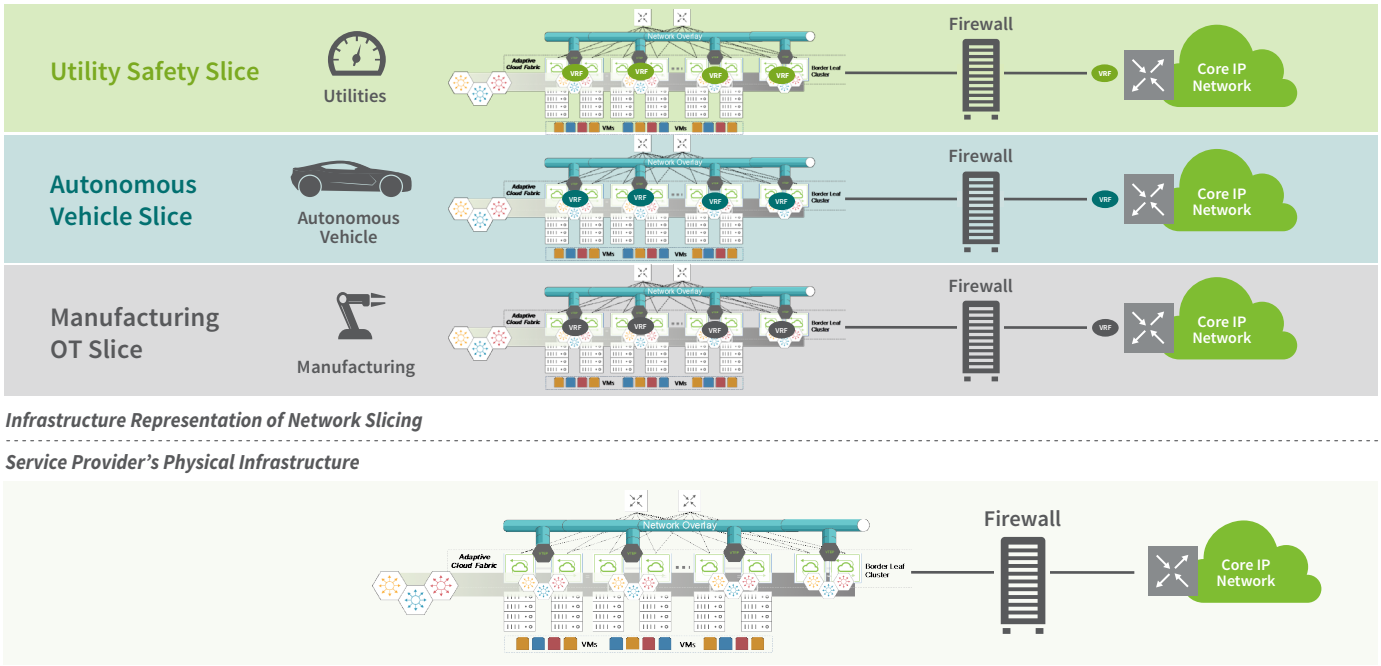

Figure 3: Network slicing has been designed into Netvisor ONE from the start, providing industry-leading slicing/segmentation across the data, control and management planes.

Pervasive Security

The fabric is inherently secure. Slicing can be used to completely isolate untrusted traffic. In addition, the fabric is designed to accept centralized security configurations with fully consistent, distributed policy enforcement. Dynamic service insertion enables the efficient sharing of physical or virtual services and resources such as firewalls, intrusion protection, security services, application delivery controllers (ADC) and load balancers across the interconnected fabric. This allows the centralization of policy and enables the dynamic assignment of services across the fabric regardless of end-point location or traffic origin.

Centralization simplifies appliance integration and administration, as all traffic across the fabric dynamically shares access to pooled resources, eliminating appliance sprawl and reducing the number of appliances needed to support operations. In addition, availability and resiliency for appliances is improved as multiple appliances can be shared to meet service demands across the fabric. Service insertion enables new appliances and services to be quickly added and provisioned into existing networks, reducing deployment time from days to minutes.

Latency-optimized, Distributed Services

ACF offers a robust architecture whereby per-slice IPv4/IPv6 anycast gateways with up to 2,000 VRFs per slice are distributed to every node in the fabric to provide very efficient east-west routing. Containerized vRouters sit at the border leaves to support north-south routing, and there can be multiple vRouters per slice. Rich Layer 2/Layer 3 VPN services can be offered per tenant and are distributed to each node in the fabric for an efficient, low-latency architecture. Layer 1 VirtualWire is also offered, providing a transparent Layer 1-like tunnel service (similar to pseudowires) that connects any two ports through the fabric and ensures all packets pass, regardless of geographical location, as if they were directly connected by a wire. Many other standard network services are available on a per-tenant basis.

Rich Telemetry and Integrated Visibility

The Adaptive Cloud Fabric embeds monitoring telemetry on every port to monitor ports and TCP flows as well as all devices connected to the network. The embedded telemetry exposes important service behavior characteristics such as application type, connection state and end-to-end connection latency. The embedded telemetry monitors all application flows, including traffic within VXLAN tunnels. Pluribus offers a database that can hold up to 2.5 billion flows, and performance metrics can be viewed through standard RESTful APIs, IPFIX or command line queries, or can be visualized by the Pluribus Insight Analytics™ platform. The Insight Analytics platform provides comprehensive application-aware network performance management (NPM) and operational intelligence to improve real-time and historical visibility to assure service availability, performance and quality. Pluribus vProbe technology extends visibility into VMware servers to expose the performance-related characteristics of application traffic traversing the hypervisor.

A New Way of Networking for Distributed Cloud and 5G

The Pluribus Networks approach to next-generation data center architectures delivers an open, virtualized and programmable network fabric that ensures optimal performance and availability across centralized and regional data centers as well as distributed mini and micro data centers, with simplified management and powerful performance analytics. Enabling freedom from legacy network constraints, the Pluribus Adaptive Cloud Fabric is powered by a wide range of open networking switches, including devices from Dell EMC, D-Link Systems and Edgecore, as well as the Pluribus Freedom™ Series network switches. These next-generation data center switches are purpose-built for software-defined and virtualized data centers of all sizes and deliver a cost-effective, high-performance and highly scalable network foundation for demanding data center deployments and virtualized workloads.

The combination of open networking hardware and the Pluribus Adaptive Cloud Fabric delivers a capability set that is designed to empower any size organization to do more with their next-generation data center architectures and enable them to move to a distributed cloud edge compute architecture while eliminating complexities, reducing risk and speeding the time to value for their edge investments.

Use Cases:

- **Next-generation networking fabric for the service provider distributed cloud** – As service providers turn central offices and other strategic real estate into mini data center locations, the Adaptive Cloud Fabric can unify these distributed sites to simplify operations and increase agility. The fabric can be sliced for multi-tenant operations. VLANs and IP addresses can be reused across slices and open APIs can be used to integrate the fabric with orchestration systems and management tools.

- **Interconnect centralized enterprise data centers to infrastructure at colocation facilities, including far edge colos** – In addition to modernizing the centralized data center with open networking and an SDN-controlled fabric, enterprises or managed service providers can now expand their geographical footprint by interconnecting centralized data center locations to colocation sites much more easily and holistically. The Adaptive Cloud Fabric can unify distributed data centers, regardless of location, to simplify operations and bring agility to distributed cloud deployments.

- **Data center interconnect fabric for colocation facility operators** – Design high-scale and easy-to-deploy SDN-controlled interconnect fabric to provide automated cross-connect capabilities. The fabric can support Layer 3, 2 or 1 (transparent) connections through the fabric connecting enterprises to SaaS, public cloud services or other partners, shrinking interconnection lead times from months to minutes. The fabric can be sliced, with services delivered and managed per tenant. Colocation providers can build a portal that controls the SDN infrastructure through APIs with a few mouse clicks and has links up and running in no time.

- **Segmentation for untrusted IoT data** – IoT can present a large attack surface. The Adaptive Cloud Fabric can be used to slice the network to segment untrusted traffic from business and production traffic and send IoT traffic to a security operations center. With the Adaptive Cloud Fabric, enterprise customers can have highly scalable network segmentation done by the VRFs as fabric objects, whether for a single data center or multiple geographically dispersed data centers, including at edge locations.

Pluribus Networks, Inc.
6001 America Center Dr.
Suite 450
San Jose, CA 95002

+1-650-289-4717
+1-855-438-8638
www.pluribusnetworks.com