www.pipelinepub.com
Volume 18, Issue 10

Transforming Enterprise IT with On-Demand Connectivity

By: Ihab Mahna, David Pollington

Agility in the provision of cloud resources like AWS, GCP, and Azure and application development on mobile devices has created a multibillion-dollar consumer app business.

Modern enterprise apps targeting new revenue opportunities, cost efficiency, automation and resiliency now require similar agility, and not only from the cloud but also from connectivity.

In particular, the trend toward automation in Industry 4.0 sectors such as smart manufacturing, warehousing, mining, and ports is driving increased demand for connectivity that is high performance but also more configurable to support rapid application development and evolution.

Seeking the best of both worlds

Agility is required not only in defining and configuring connectivity to match a particular task, but also to be able to dynamically update the intent of connectivity to meet the transient app needs.

For example, in a manufacturing or warehouse environment, an enterprise might need a mixture of connectivity tailored to different groups of devices: ultra-reliable high bandwidth connectivity for controlling machinery on a production line, and high-density interconnectivity for a multitude of service robots, cobots, and autonomously operated forklift trucks.

Essentially, enterprises are seeking the best of both worlds—the functionality, performance, and reliability of 5G twinned with the flexibility, control, and ease of use that they've grown

accustomed to with their existing Wi-Fi deployments and cloud services. One-size-fits-all over subscribed connectivity is no longer fit for this new type of enterprise application.

5G slicing and control

While a public 5G network may meet some of these needs with respect to high throughput and low latency, it's essentially best effort with no quality of service (QoS) guarantees and no ability for the enterprise to stipulate and control the connectivity provided.

To address this shortfall, 5G introduces the notion of slices that enable the public 5G network resources to be segmented by the mobile network operator (MNO) to provide what is essentially a virtual network to the enterprise customer. These slices, however, are predefined and provide little flexibility for customization, and hence may be too constraining for some enterprises. The slices are also statically defined, whereas what many enterprises really want is an ability to control and manage their connectivity and QoS requirements on a more dynamic basis to map resources to an application or device as circumstances change (as in adaptive slicing). There is no way to adapt these slices without giving the application using it the control and automation tools to drive it.

Advancing agility through 5G MPN

For 5G to be relevant and useful to enterprises, and especially for supporting a wide range of Industry 4.0 applications, it will be necessary to provide fine-grained control over slice configuration and management to ensure predictable performance and reliability matched to device and application needs, and an ability to make changes on the fly. In addition, the concept of slicing needs to expand end-to-end to include edge service, IoT services, Wi-Fi, and private LAN. Some of these domains are administratively outside the scope of MNOs' operations.

This can be achieved through the deployment of localized 5G infrastructure, dedicated to the enterprise application DevOps teams—in essence, a 5G mobile private network (MPN) on-premises that is owned and operated by the enterprise.

Not only does a 5G MPN give an enterprise more control and flexibility, but it also affords them higher levels of privacy and security by allowing them to operate their own infrastructure rather than sharing infrastructure within a public network. For some mission-critical applications, this is especially paramount, but for manufacturers in general, high security is a key driver for choosing an MPN over utilizing a public 5G network slice. Deploying private 5G is now the top priority of IT decision-makers in medium-sized enterprises according to a recent analysis by Technalysis Research. Additionally, 76 percent of manufacturers plan to deploy 5G MPNs by 2024, according to Accedian.

Enterprises have a range of options open to them and the flexibility to decide whether to go with an MNO or managed services provider (MSP) that can fully design, deploy and configure the MPN for them (for instance, for school campuses), or a partner that simply builds and deploys the MPN infrastructure dimensioned to the enterprise's requirements and leaves configuration to the enterprise to manage along with day-to-day operations (for example, in smart manufacturing, ports, mining, and so on).

On-demand connectivity intents

Many if not all of these enterprise application development groups are already accustomed to using the cloud (public, private or hybrid) for their computing needs. CTOs and CIOs are inspired to modernize these applications by moving to on-cloud and on-demand connectivity, and they will expect a similar orchestration capability for managing connectivity, one that extends across all their domains and assets used in application and service delivery. Very few of these enterprises, however, will have teams with telco expertise and they are looking to reduce on-prem networking infrastructure. Therefore, a more abstracted approach needs to be provided; this is achieved through the notion of on-demand connectivity intents.

Put simply, intents form the basis of a zero-touch system in which the 5G MPN and associated network assets and managed services are configured autonomously based on application behavior to meet the enterprise's requirements without the enterprise needing to manually configure the network and services itself.

This is like OTT video platforms performing adaptive video and caching automation at the edge. Now this concept needs to extend to all modern enterprise applications. The automation must happen at the edge and with awareness of application.

Intents is a declarative representation of connectivity requirements for an application endpoint to be connected via network slice with expected QoS. This can also take into account any enterprise preferences and policies around power efficiency and cost, while handling the complexity caused by continuous changes to the system to optimally allocate network resources. In short, an intent defines the expectations on service delivery and network operation through the expression of a set of goals, functional requirements, and constraints, thereby giving enterprises wide-ranging agility in how they specify and manage their connectivity—and by extension, their private 5G network and edge services.

Managing intents

At a practical level, intents can be managed in a number of ways depending on the skillsets of the enterprise. For those enterprises with limited resources, a set of easy-to-use self-management tools (low- or no-code) can be provided for defining intents, app and device group administration, and monitoring network and application performance as well as end-to-end security within the MPN. These tools can also extend to some basic automatic detection and isolation of faults.

Conversely, for those wanting more fine-grained control, orchestration could be provided to network administrators and DevOps teams programmatically via a set of RESTful APIs with dynamic control over throughput, latency, packet error rate metrics, network segments, IP domains and so on, and optionally bootstrapped via a number of Infrastructure as Code (IaC) templates (like boilerplate builds, for example 5G slice types such as ultra-reliable low-latency communications (URLLC) and massive machine-type communications (mMTC)).

In short, enterprises should be able to configure and manage their 5G MPNs and scale them up and down using DevOps-friendly interfaces as easily as they can using Kubernetes with cloud resources for their application and service deployments.

Moreover, this agility should ideally extend to having one DevOps CI/CD process (pipeline) for both cloud and connectivity with the ability to manage resources across all domains within service delivery (5G, Wi-Fi, cloud, mobile edge computing (MEC), software-defined wide area network (SD-WAN), and more).

Building an agility ecosystem

In the case of the cloud, an ecosystem has grown up around it supplying tools and supplementary services to make the consumption of these resources easy, cost-efficient, automated, and assured. Some were provided by the cloud providers, some were invented outside and acquired by cloud providers, and some grew independently to become unicorns.

To enable enterprise 5G MPNs to be built on-demand and engineered for performance and security as simply as is now enjoyed with cloud resources will require a similar ecosystem of tools and services to emerge.

Achieving this will open up 5G MPNs to enterprises of all sizes—in essence, similar to the democratization of telco APIs brought about by the introduction of developer-friendly platforms (and RESTful APIs) from the likes of Twilio a decade or so ago.

This approach was trialed and honed within 5G-ENCODE, a £9 million collaborative project aiming to develop clear business cases and value propositions for 5G applications in the manufacturing industry. The tool essentially acts as a ‘Kubernetes for MPNs,’ but also extends into other technology domains (4G, 5G, Wi-Fi, SD-WAN, MEC, public 5G slices and more) to provide comprehensive management end-to-end and all exposed via an intuitive 'single pane of glass' and set of DevOps-friendly interfaces.