

Independent market research and competitive analysis of next-generation business and technology solutions for service providers and vendors

**HEAVY
READING**
**WHITE
PAPER**

Virtualization: A Critical Capability for Service Provider Success in IoT, 5G & Beyond

A Heavy Reading white paper produced for Wind River



AN INTEL COMPANY

AUTHOR: STEVE BELL, SENIOR ANALYST, HEAVY READING

INTRODUCTION

The Internet of Things (IoT) is a huge opportunity for communications service providers (CSPs), because of the sheer scale of investment it will attract and the business impact it will have globally. This white paper looks at developments in the IoT landscape and the critical factors and obstacles that CSPs need to be aware of as they consider the development of their network architecture to address the market and meet customer expectations.

THE IOT LANDSCAPE

Over the last three years, the IoT has become one of the hottest areas of interest for consumers, corporations and governments. As technology and software companies have introduced smart devices for home automation and security, enhanced peoples' health regimes with fitness trackers and smartwatches, and collaborated with automotive manufacturers to make cars more connected and intelligent, consumers have gradually become aware of the benefits of an instrumented and data-driven world.

Figure 1: Companies With M2M Solutions in Place, 2016

Industry	Percentage
Retail	17%
Transport & logistics	17%
Health care/life sciences	19%
Manufacturing	20%
Automotive	28%
Energy & utilities	28%
Consumer electronics	29%

Source: Vodafone M2M Barometer Report

At the same time, corporations and governments have recognized the transformational potential of these technologies for competitive capabilities, business models and entire economic value chains. Increasingly, boardrooms and government chambers are focusing attention on the way in which the IoT is shaping the global landscape. Companies are just beginning to recognize how the IoT and the data it generates can reduce cost, improve supply-chain efficiencies, enhance product design and allow predictive maintenance. Governments are looking at how industrial policy needs to be reshaped to ensure that national industries remain competitive, and the degree to which regulations need to be adjusted or created to address such fundamentals as individual privacy, data security and a world that could soon be full of autonomous vehicles.

M2M to IoT & Industrial IoT

In reality, IoT is not a new market or technology. Machine-to-machine (M2M) and telemetry technologies have been in existence for well over 20 years. Many of the M2M communications and operations protocols are proprietary and unique to specific industries. Telemetry systems used in fleet management, asset tracking and infrastructure monitoring have used cellular

GPRS and CDMA technology because of the low cost of communications modules, and the almost ubiquitous coverage.

The IoT has emerged from a combination of various technologies that have been continuously developing over the last decade. The ongoing simultaneous reduction in cost and increase in semiconductor compute capability has resulted in very small-footprint modules and systems on chips. This allows devices such as actuators, controllers, wearables and gateways to be much more intelligent, particularly when the microprocessors link to a huge variety of sensors that enable the device to determine its state and context within its operating environment.

This same increase in semiconductor capability has fostered improved performance in wireless technologies. Short-range and personal-area connectivity has been enhanced with continuous developments in Bluetooth and WiFi. Additionally, multiple wireless technologies have been built upon 802.15.4, including ZigBee, Z-Wave and Thread. In the wide-area wireless arena, cellular has continued to develop as a mobile broadband solution that delivers large amounts of data at very high speed.

Over the last 18 months, the wireless industry has been frantically developing low-power wide-area cellular solutions (LTE NB-IoT) to compete with alternative offerings from companies such as SigFox, LoRa Alliance, Silver Spring Networks and Ingenu. All of these competing alternative wireless systems use a variety of proprietary and standards-based technologies and operate in unlicensed spectrum. They have been developed within the last seven years, and recognized the fundamental need of industrial customers that cellular operators were ignoring. This need was for simple low-cost networks connecting to devices that are low power, operate on batteries that can last for 10 years or more, and transmit very small amounts of data infrequently.

The development of all of the wireless technologies for both short and long-range communications has paralleled the increasing adoption of the Internet protocol within the industrial and enterprise workspace. Increasingly, production processes have been linked with capillary networks to gateways that can connect into internal IT systems. At the same time, buildings became more automated, with heating, ventilation and air-conditioning systems connected to building management and security systems. However, most of these connected systems were data silos, where data was flowing north-south in a proprietary or industry standardized format.

Increasingly, there was interest in linking and consolidating data from discrete silos into actionable information for building efficiency, and optimizing production and supply chains. This is where three other technologies actually enabled the development of the industrial IoT. The rapid adoption of cloud technologies, the use of big data analytics and the application of machine learning have facilitated the aggregation of data to deliver fundamental insights into virtually every aspect of business operations.

The combined effect of all these technologies is fueling the growth of the IoT domain. The use of the word "domain" is intended to symbolize the coming together of technologies and markets, which combined create exponential outcomes that some refer to as the "Fourth Industrial Revolution." No matter where you look, there are forecasts of between 20 and 50 billion connected devices by the beginning of the next decade, with resultant growth in business and GDP in the trillions of dollars. These billions of devices will form capillary networks that have different communications formats, operating protocols and addressing mechanisms that drive a tide of data in varying formats. Although cellular and LPWA devices may be in

the range of 2 to 6 billion connections initially, understanding these capillary networks is vital to know how to design the network to handle the diversity of data streams.

The impact of the IoT will be felt across every aspect of life as we currently experience it. Clearly, the continued development of smartphones, tablets and wearable technology will allow these devices to be the remote controls that interface to the smart systems that will be found in the workplace, home and our smart cars. However, these are just the direct points of contact; increasingly, there will be a fabric of smart systems that envelop everyday life, from smart public transportation systems and smart utilities to smart health systems, all of which will operate within smart cities and communities.

Critical Applications

The significance of this tapestry of interconnected smart systems cannot be underestimated, and individuals and societies will become increasingly dependent on the interlocking solutions and services they provide. Underpinning these systems and services has to be a high-reliability backbone of communications networks, comprising both fixed and wireless technologies. In many cases, this requires that the critical applications not only have a primary connection, but also redundant and failover options in the event of connectivity loss.

Not all applications will require this level of communications infrastructure integrity, but in cases where it is vital that systems do not fail, it requires a different type of systems thinking approach. There has to be recognition of the end-to-end nature of the design, identifying the interlocking dependent systems and specifying which are nice-to-have applications within the design, which require low latency, and which require resilient guaranteed control feedback. As an example, in a building or facility maintenance scenario, a loss of power will result in building automation and management systems triggering backup generators to support designated high-priority systems.

However, within this scenario, there are subsystems with critical applications that absolutely cannot fail, such as elevator power-pack controls and brakes. Essentially, anywhere there is a potential risk of loss of life, due to a breakdown in a control feedback loop, there has to be redundant and backup solutions with millisecond or sub-millisecond kick-in. Other examples of such situations are mining operations that deploy large autonomous trucks, connected hospitals where systems in intensive care are vital for life support, and the emerging autonomous car scenario, where vehicles will have to react to careless pedestrians and cyclists.

No discussion of critical infrastructure would be complete without mentioning security, which has taken on new significance with the use of massive numbers of unsecured IoT devices as components in enormous, weaponized distributed denial of service (DDoS) attacks. Increasingly the requirement is for security in depth, starting with trusted and authenticated devices, through secure communications to encrypted data. The leading service providers are using machine learning and other artificial intelligence techniques to monitor IoT device signatures and develop threat intelligence and automated response capabilities.

Edge & Fog

In the enterprise and industrial space, there is increasing focus on the need for rapid localized decision-making for devices that rely on constant information feedback. As M2M silos became connected, the north-south flow of data to centralized cloud repositories yielded significant insight but did not address the requirement for high-speed decision control loops. There is a

requirement for intelligence to be close to these time-critical decision loops in order to reduce latency. Most of these decisions take place at the edge of the enterprise, which is also where business interacts with the edge of the wide-area network.

There are several challenges associated with this intersecting edge. In order to facilitate rapid decision-making, the enterprise has to be able to identify, address, authenticate, manage and secure a broad range of devices, many of which have differing address mechanisms as well as communications and operating protocols. A growing trend is the use of smart gateways that can connect these differing capillary networks, and provide translation capability. This can result in multidirectional information flows across systems that enhance both the quality and speed of decision-making.

These smart gateways coexist in the world of network premises equipment that has become increasingly complex as the number of connectivity options, services and applications have increased. This has resulted in a plethora of dedicated boxes from multiple suppliers, covering wireline control, routers, VPN access, Internet access, firewalls and LAN accelerators. Trying to interwork these smart gateways into these systems can prove very difficult, since the boxes tend to have vendor-specific proprietary management approaches that can be difficult to navigate in order to create an integrated IoT solution.

In 2009, the concept of cloudlets, or fog, resulted from the work of a group of researchers from institutions such as Carnegie Mellon University, Lancaster University, Microsoft, Intel and AT&T. The concept utilizes mobile edge computing (MEC) to distribute cloud compute and storage capability at the network edge, most often at the customer premises or in CSPs' points of presence. Since then, this fog concept has evolved to the extent that network edge servers can be either standalone gateways or blades in servers. Fog computing takes advantage of enhanced device compute capability, as well as developments in network functions virtualization (NFV). This allows software versions of routers, firewalls and security protocols to be loaded in servers at the network edge, and makes it much easier for smart gateways to interconnect. The servers also allow for easy provisioning, and provide the opportunity for real-time analytics of data in motion that enhances the ability to provide multidirectional time-critical decision loop feedback.

CSPs & IOT

The requirement for a resilient and highly reliable communications network, with redundancy and failover from wireline to wireless, means that fixed and wireless CSPs are integral to the development of IoT. The question is: Will they be solely focused on providing the connectivity, or will they be providing incremental value-add and managed services to enterprises and consumers? Based on research conducted in late 2015 and market observation throughout 2016, it is fair to say that most CSPs are at an early stage of business development in the IoT space, and that the business is currently minimal for most providers, with only a small number of leading players claiming significant revenue.

There is a strong indication that, by 2018, IoT will become mainstream for most CSPs. This would coincide with the timeframe when fully commercial market introductions of 3GPP standardized low-power wireless technologies, including NB-IoT, LTE Cat M1 and EC-GSM-IoT, are likely to occur. These technologies cover a range of different use cases and can be deployed in several different ways, either standalone or in-band using existing LTE resource blocks that provide mobile operators with significant versatility to address emerging IoT opportunities.

Figure 2: Weighted Ranking – Tier 1 CSPs by Use Case

Use Cases	Weighted Ranking
Smart cities	2.56
Health	2.52
Consumer	2.4
Industrial	2.39
Smart home	2.39
Smart meters (all utilities)	2.27
Connected car	2.24
Retail/beacons	2.2
Energy (production, distribution)	2.1

Opportunities

The cellular operators with carrier-grade reliability can ensure mobility and availability almost ubiquitously. Coupled with the benefits of an established mobile ecosystem that can drive volume and economies of scale, they are a natural partner for emerging IoT markets. The opportunity is for CSPs to deliver more than just connectivity. One of the challenges already identified is how enterprises manage millions of devices, and the requirement to authenticate, secure and maintain them over a lifetime (often 10 to 15 years) is a significant issue. Mobile operators bring years of experience of doing this for phones and smartphones, and being able to apply this capability as a globally managed service could add significant value to companies. The ability to monitor devices and upgrade software over the air means that security patches and enhanced capabilities can be installed to extend the life of the device.

The collection and storage of data, combined with the application of analytics, is another value-add service that CSPs can provide. This becomes particularly attractive where anonymized customer data is combined with network data to provide contextual and proximity-related information. This unique combination of information, coupled with the capability to do sophisticated analytics, or to apply machine-learning mechanisms at scale and deliver this to customers, is something that could provide competitive advantage to enterprise customers.

The ability to deliver these managed services requires a change in mindset by most CSPs. They need to focus on identifying and figuring out how to deliver the guaranteed outcome that customers want, rather than just delivering services and solutions. As an example, the customer may desire enhanced customer service, predictive maintenance for field installations and reduced costs in operations. In designing a solution to meet this set of requirements, the CSP would need to understand the customer's overall system, including any capillary networks supplying data. Based on this understanding, a determination would need to be made about which critical decisions must be made where, how much data is being generated and how to tier it. This will define where it needs to be stored, and what needs to be transmitted across the network, taking into account latency requirements for decision control loops.

The operator would then have to design a network solution that ensures that it provides end-to-end security, doesn't overly burden the network and can be priced competitively for customers with the flexibility to deliver the outcomes required under changing circumstances.

Operators need increased flexibility and agility in their network to be able to effectively design these solutions. For example, AT&T, Telefónica and BT are all partnering with various cities to define and design optimized solutions and business models to deliver smart cities.

Challenges of IoT

As great as the opportunity is, there are significant challenges for CSPs to serve this market. First, it is not a single market nor is it a static one. From smart homes to smart cars and smart health, there is a clear alignment with existing experiences in serving smartphone subscribers, providing combined cellular and WiFi services. However, in all three use cases there are distinct and separate ecosystems of devices, applications, systems integrators and third-party partners that have to be accommodated. This means that revenues generated have to be shared across an extended value chain, making margins extremely slim. This puts pressure on the CSP's network to be as cost-effective and efficient as possible.

In the case of industrial and enterprise IoT, together with smart cities, the diversity of segments and applications means that there are very few "one size fits all" solutions that can be developed and repeatedly deployed. Compounding this, multinational companies want their specific solution deployed across all of their global operations, requiring the solution to be nuanced to meet local requirements, and for CSPs to handle all of the connectivity and service interoperability challenges.

Traffic Patterns

Since IoT is in its early stages of development, most CSPs don't yet know how they will have to optimize their networks to deal with different traffic patterns. They are likely to be dealing with everything from trickle-type data from NB-IoT devices to highly critical, high-bandwidth, low-latency traffic that may be directly from connected autonomous cars, or multiple streams of data and signaling from vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) scenarios. When used for IoT applications, the challenge is that LTE has not been particularly efficient, resulting in high signaling traffic, even though data rates are relatively low. The introduction of three new IoT-related technologies alleviates some of this signaling storm but requires increased flexibility in the core network to handle different access technologies.

SDN & Virtualization

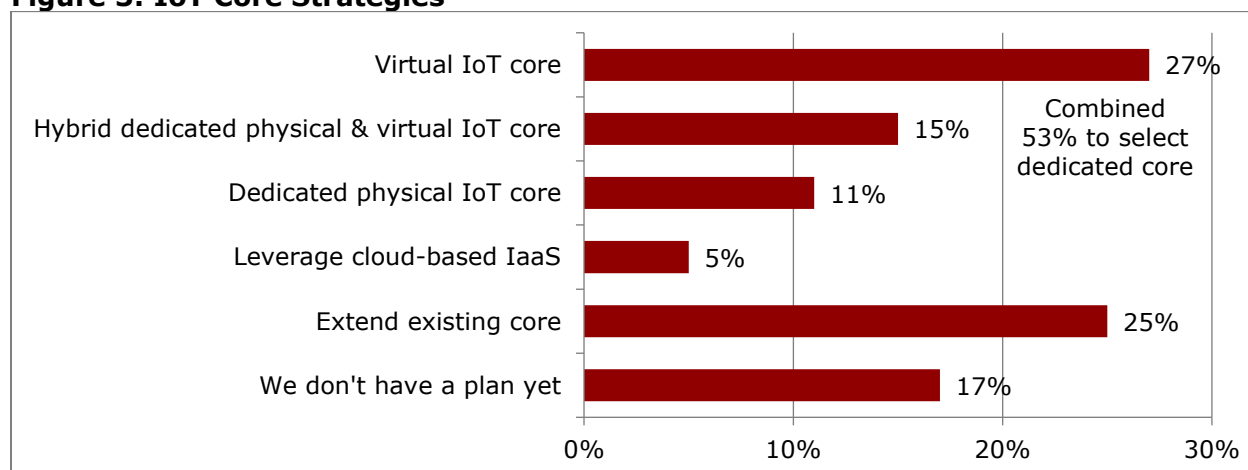
The challenges of flexibility, agility and cost control have encouraged mobile operators to embrace software-centric cloud-hosted core networks. The transition from hardware-based mobile cores is at an early stage, but the industry has made good progress around virtual evolved packet core (vEPC). The transition process is all about learning, and one of the lessons to date is that virtualization of the existing hardware product portfolio is only part of the solution. A more significant step over the long term will be toward a software-based core in which lean virtualized network functions (VNFs) designed to run on virtual machines are deconstructed into smaller components and operators can take advantage of extreme automation processes, designed to run in the cloud, sometimes referred to as "cloud native."

The diverse IoT market segments that require many types of service, each with their own specific performance and economic requirements, are causing operators to consider a dedicated core network for IoT. The progression toward virtualization makes it operationally easier to dedicate core networks for multiple services or users than when the core was comprised of physical hardware. The idea is to create virtual cores for different types of services, to

optimize this core according to the traffic profile, and to address specific requirements such as quality of service for mission-critical applications. These dedicated core networks are sometimes referred to as network slices, and they give the operator the opportunity to dedicate a network to a specific type of service, such as meter reading, or to provide virtual private networks (VPNs) for large utility companies. In the case of a virtual private network, where there are specific network requirements or the data being transmitted is sensitive, it is possible to logically separate service with a private virtual packet gateway that allows the customer to manage their own users and policies.

In a survey conducted by Heavy Reading at the end of 2015, a broad cross-section of global operators shared their perspectives on their plans for deploying IoT/M2M cores, as shown in **Figure 3**. A combined 53 percent of respondents identified that the plan was to deploy a dedicated core of some form. This includes 27 percent of respondents that are looking at a virtual IoT core, whereas 11 percent are looking at a physical IoT core, which could be a "virtual core in a box"-type solution that may be suitable for dealing with dedicated industrial customers, such as utilities or critical infrastructure providers.

Figure 3: IoT Core Strategies



Source: Heavy Reading's Operator IoT Survey, January 2016 (n=131)

The responses also capture the early-stage thinking of many operators. This is evident in the diversity of options being considered, including 15 percent looking at a hybrid combination of physical and virtual that may enable a quick start but provide flexibility and opportunity to innovate. A quarter are looking at extending their current existing core, which could reflect that they have an established M2M business, and it just makes sense to continue that model. The 5 percent of operators that are considering using cloud-based infrastructure as a service could be attempting to address both cost and flexibility requirements, without having to acquire new skills and talent to deploy a virtual core. The 17 percent that do not yet have a plan is consistent and indicative of the "wait and see" strategy that many operators have adopted. Their challenge will be to move very rapidly if this market starts to take off, and it is probable that cloud-based calls could be the least resistant route to market.

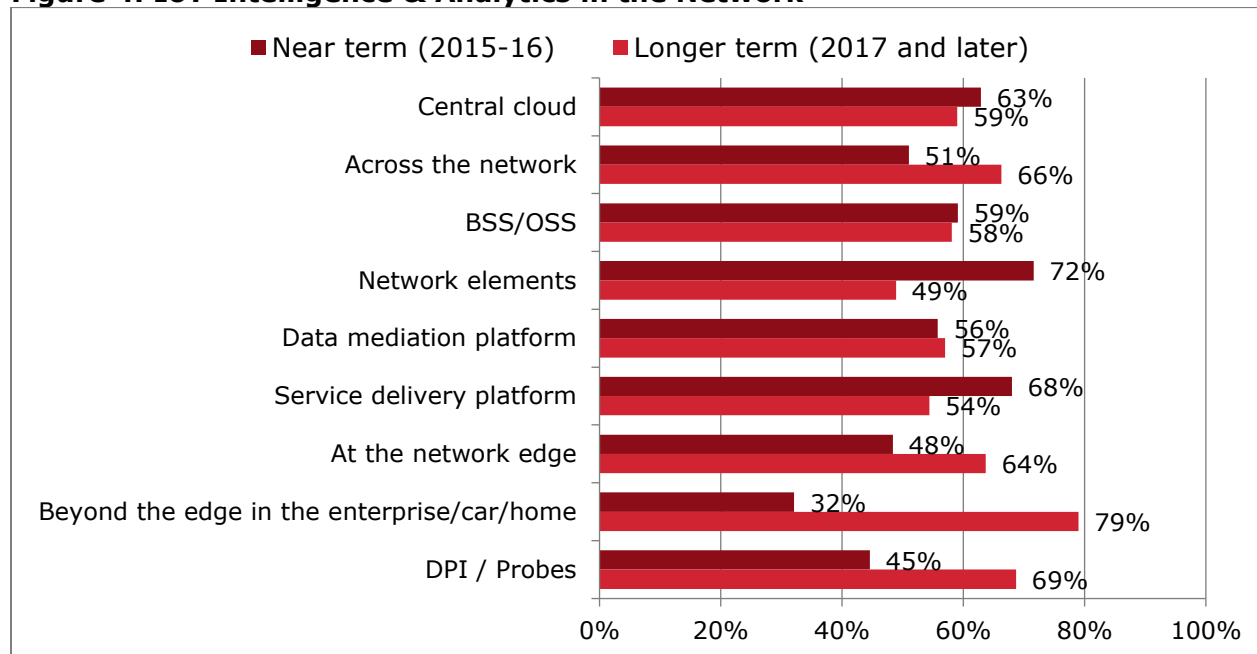
Fog & Cloud

The virtual edge is a major focus for enterprise customers seeking to understand the implications of fog computing on their operational performance, network options and hardware requirements. Depending on the network architecture chosen, it can provide a little more

tolerance to latency if the design is optimized for edge analytics. Many times, this decision hinges on determining what data is necessary for low-latency control loops locally, and what needs to be transmitted to the cloud. This decision will invariably be use case-specific, trying to seek a balance between the efficiencies of IoT and overall network costs.

In the same 2015 operator survey, we asked them their opinion on where IoT intelligence and analytics will reside within the network. As shown in **Figure 4**, intelligence in the network definitely seems to be shifting away from the center. In the near term, network intelligence will reside in network elements, according to 72 percent of respondents, closely followed by service delivery platforms (68 percent), central cloud (63 percent), BSS/OSS (59 percent) and data mediation platform (56 percent).

Figure 4: IoT Intelligence & Analytics in the Network



In 2015, 48 percent of operators thought intelligence would be at the network edge; by 2017, this has risen to 64 percent. Additionally, only 32 percent of operators believed that intelligence and analytics would reside beyond the edge of the enterprise/car/home. This area shows the biggest shift, with 79 percent believing that intelligence and analytics will reside beyond the edge of the network in the enterprise by 2017.

Strategies

Overall, there is a lot of optimism about the potential of IoT, but few CSPs have formed concrete plans, and still fewer have implemented them. Clearly, Tier 1 operators such as AT&T, Verizon, Orange Vodafone and Telefónica are pressing ahead, but there is divergent thinking on how business models and networks will evolve to meet the requirements of IoT. Equally, the market is still at an early stage of development with a great deal of fragmentation, and this is likely to continue. This will increase pressure on CSPs to focus their organizations and networks, in order to scale and provide the agility that enterprise customers, in particular, will be looking for. Security is undoubtedly the pivotal capability that will establish and sustain the CSPs as the primary provider of communications, and potentially value-added services.

Analytics and intelligence will play a prime role going forward, both from a network operations perspective and as a service to customers. However, the way in which this is organized, serviced and delivered is still very fluid. Those mobile operators actively embracing software-centric and cloud-hosted core networks will be better positioned to match infrastructure to services, as well as pursue market opportunities at relatively low incremental cost.

Positioning for 5G

5G is very important from a service provider perspective, as it is about seamless management of heterogeneous networks across multiple frequencies, and obviously the efficient handling of huge amounts of data traffic and storage. The most progressive operators are exploring cloudification of VNFs designed to run in the cloud, and be deployable in a horizontal shared infrastructure that will allow functions to be installed, without the need of manual involvement in the onboarding process. This is pushing the boundaries of cloud application design, but the technologies and architectures being considered are well aligned with the 5G roadmap and the deployment of next-generation core by 2020.

Network slicing, which will be a key enabler of IoT, is in the early stages of implementation in LTE, and is a core component of the 5G vision, since it enables "the many services on one network" concept. The proposed next-generation core architecture anticipates that slicing will extend across the radio network as well as the core, and one of the standards being developed is the network slice selection function that will determine which core instance a device will connect to and how.

Despite operator enthusiasm for 5G, at this stage most industrial clients are either ambivalent or negatively disposed toward it. This follows on from the mismatch of expectations and requirements, in terms of timeframe and performance, between current cellular network capabilities and industrial needs. Industrial customers are coming to grips with the imminent closure of GPRS and CDMA networks, due to the fact that operators are deploying more efficient LTE networks to replace them. Consequently, it is incumbent on the operators to start convincing customers of the value and flexibility that virtualized software-defined networks can deliver, in both 4G and 5G.

WIND RIVER & IOT

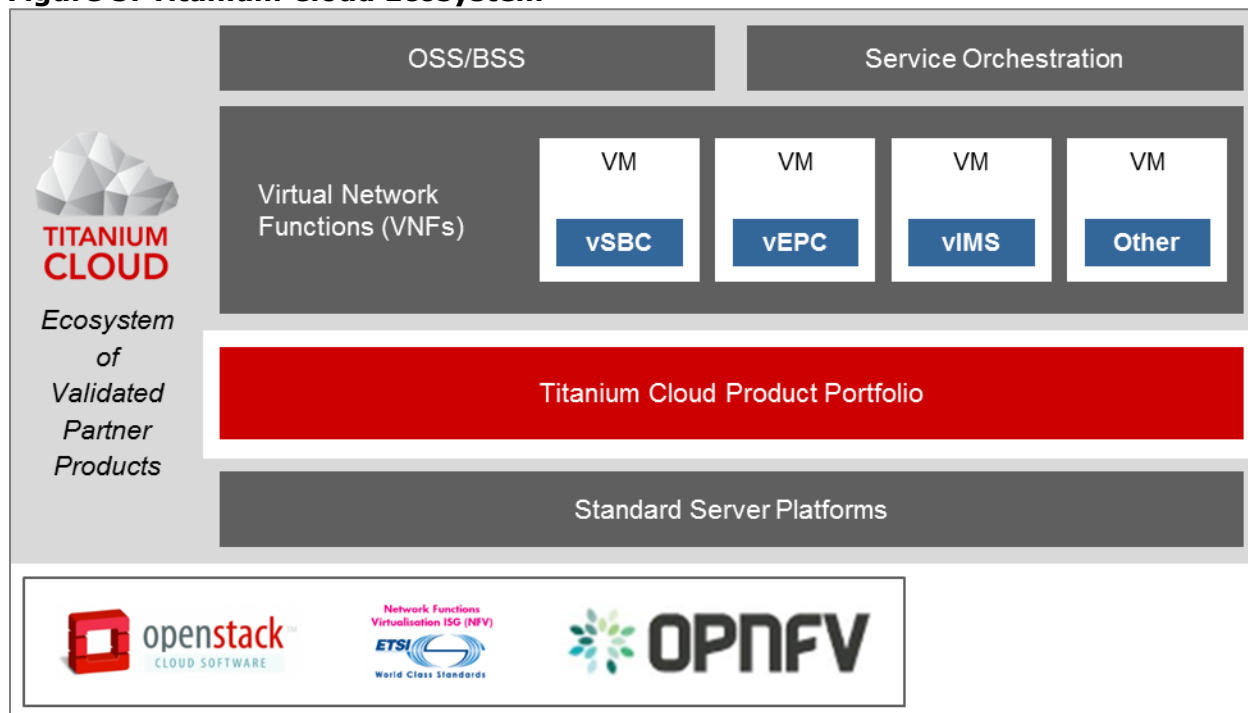
Wind River recognizes that, in order for service providers to maximize ROI on IoT, they will have to introduce new services and new customers quickly to increase top-line revenue growth. At the same time, it is essential to have service agility to manage these services and devices as efficiently as possible in order to optimize operational costs. Supporting this imperative, Wind River has introduced two complementary solutions: Helix Device Cloud and the Titanium Cloud product portfolio. The Titanium Cloud family of carrier-grade virtualization solutions provides a common platform that delivers IoT services and VNFs. Helix Device Cloud is the ready-built platform that enables operators to safely and securely monitor, manage, service and update devices in the field.

Titanium Cloud Product Portfolio

Wind River has a strong history of delivering platform software and tools for deployment in applications where failure is unacceptable for business or safety reasons. The products within

the Titanium Cloud portfolio are the industry's only open, commercial NFV software platforms that deliver the carrier-grade reliability required for telecom networks. The software-based fully integrated, ultra-reliable and deployment-ready virtualization platform enables service providers to deploy virtualized services faster, at lower cost and with guaranteed uptime.

Figure 5: Titanium Cloud Ecosystem



Source: Wind River

Titanium Cloud products run virtual functions with carrier-grade reliability, and accelerates the performance of the VNFs to maximize opex savings and to provide:

- The flexibility to scale services rapidly and efficiently
- The ability to deploy new services dynamically
- Performance capability to maximize the number of subscribers supported on each server with minimized operating costs

The Titanium Cloud portfolio of solutions includes: Wind River Titanium Core for CSP data centers, central offices, and PoPs; Titanium Edge for small-footprint telco edge applications; and Titanium Control for highly reliable industrial command and control applications. The portfolio is supported by the Titanium Cloud ecosystem of validated hardware and software products from industry partners, and is complemented by professional services expertise to accelerate deployment.

Helix Device Cloud

For most enterprises, the compelling case for the IoT is the ability to access the valuable data being generated by the multitude of field devices. That can happen only if the devices delivering that data, and the gateways that direct data to enterprise systems, are continually

performing as expected. The overall performance of a system often hinges on the health of these field devices. If a device, sensor, embedded agent, or gateway begins to falter, the consequences can be dire.

The challenge of maintaining devices may sound basic compared with aggregating and analyzing data, but it is fundamental to a successful IoT strategy. Helix Device Cloud automatically collects and integrates data from disparate devices, machines and systems, enabling operators to track device status, share data and proactively determine when updates are needed. Using an embedded software agent, device properties and operating data can be transmitted securely to the cloud. Operators can easily view device information through a Web-based management console, perform diagnostics and take prompt corrective action.

The cloud-based platform is also designed to integrate with enterprise systems that utilize or analyze data from IoT networks. Device Cloud data and event forwarding ensures that device health issues will signal other systems of potential problems, enabling them to respond accordingly and potentially prevent ingestion of bad data.

The combination of Wind River's cloud services and platforms provides a holistic support system for service providers as well as enterprise customers to effectively and reliably manage the continuum between devices and the cloud across diverse networks.

CONCLUSION

There is a degree of wishful thinking going on in the mobile industry, with laggard operators believing that they can go into IoT catchup mode at a later date, and still win. If they do this, they will almost definitely miss the IoT window of opportunity.

One of the most significant requirements is for operators to have a virtual core as a mechanism to accelerate IoT services. Too many of them are still apprehensive about embracing this technology for IoT; consequently, this will prove to be a barrier to their success. Many are trying different approaches, including hybrid setup, and it appears now as if the two IoT industry leaders – Vodafone and AT&T – are the only ones with a clear line of sight.

If service providers embrace virtualization – including the virtual core and virtual edge – they will be able to move beyond connectivity and provide other services. What are the services that they can sell and monetize? One is the versatility and adaptability of service creation to meet the changing business models of customers, which, in turn, will create new opportunities that haven't previously been considered. These opportunities are going to be diverse and unlikely to be homogenous, across either one specific industry or a group of industries.

Other services could include managing the portfolios of devices, applications and services going forward, or providing endpoint to cloud security, which most enterprises don't have the wherewithal to do. The provision of holistic security continuity is becoming increasingly difficult with the introduction of fog. Operators could also deliver integrated analytics capability, and data management: where it is stored, how it is tiered and what portions are transmitted. This would include dataflow analysis and insight analysis as managed services. Additionally, if they have an identical high-reliability virtual server operating on the industrial side of the edge, there is the possibility for end-to-end carrier-grade enhanced applications and tailored capability.