DELL EMC AND 5G

Analysis and strategy to capture the 5G mobile opportunity

Dell EMC Service Provider Strategy and Solutions
January 2018

Revisions

Date	Description
September 2017	Initial internal release Authors - David Lake, Tord Nilsson, Kevin Shatzkamer

Table of contents

Executive summary

Imagine the future – a scalable, composable and automated wireless network infrastructure that meets the high-performance needs of tomorrow's demanding consumer, Internet of Things (IoT), and mission-critical services and makes it easy to build, deploy, manage, operate, and assure new end-to-end applications. The new era of 5G networks will see technology innovations and operational innovations, but also business model innovations that result in intelligent devices and applications consuming and generating data like never before. These intelligent applications will introduce a new set of requirements – latency, bandwidth, capacity, coverage – that require transformation of the entire end-to-end architecture, from radio access network (RAN) to Operational Support Systems (OSS). The entirety of this software-defined infrastructure – from Cloud RAN (C-RAN) to virtualized network functions to software-programmable switches and routers, is built on common building blocks of compute, storage, and networking. What used to be possible only in science fiction movies – flying drones, driverless cars and planes, machine-to-machine interactions, seamless communication around the globe – is fast becoming a reality.

The wireless industry has always experienced accelerating demand and innovation, from the early days of cellular voice mobility of 1G to circuit-switched data of 2G to high-speed ubiquitous data access of 3G and 4G. Every decade, the mobile industry goes through a major upgrade cycle of their network architecture – from the Radio Access Network (RAN) to the Packet Core – to deliver technology innovations that meet the insatiable demand of mobile consumers and increasing proliferation and capabilities of the smart mobile devices and the new generation of applications and services.

The impending 5G transition, with significant advances in bandwidth and improved latency and quality of service (QoS), will enable a new wave of services including enhanced mobile broadband, connected cars, drones, smart retail, industrial robots, and much more.

In the last several years, Service Provider networks have begun a journey towards software-defined infrastructure, leveraging the capabilities of compute, network and storage virtualization to drive new capital and operational models, deliver new services, and improve overall service delivery economics. The new operational imperatives for Communications Service Providers (CoSPs), can largely be captured in three core technology shifts:

- Leveraging increasing disaggregation of hardware and software stacks to shift workloads towards general purpose compute, such as x86.
- Decoupling core infrastructure and networking services from applications and protocols, and exposing those services as a platform to applications.
- Developing a set of information models, data models and APIs to transform operations from bespoke processes and associated infrastructure scripts, to more unified automation frameworks that allow service providers to develop DevOps-style operational processes.

5G networks will be the first true end-to-end network built around these paradigms, extending virtualization into the radio access network and network edge, virtualizing the network core, and extending end-to-end network overlays for network and service slicing.

Defining the Dell EMC 5G opportunity

These three core technology shifts provide clear opportunity for Dell EMC to participate in a meaningful way in the evolution of the cellular network. In short, the principles upon which Dell EMC is executing on a 5G strategy can be summarized into the following ten points:

1. 5G is a foundational new architecture, part of the continuous once-in-a-decade re-architecture of cellular networks. The changes are broader

than a new access technology. **5G represents an important architectural pivot which Dell EMC, and the broader Dell Technologies, has the opportunity to not only capture, but lead in defining the underlying infrastructure technology.**

2. **5G will be the first end-to-end architecture which is fully software-defined**, from the radio through the core. Software-defined leaves the data center.
3. Even though 5G is likely 2-3 years away, **investments in prototype, research, and trials for 5G by vendors, operators, industry fora, and governments are happening now**. Dell EMC is involved with a number of partners and CoSP in 5G, including AT&T, Verizon, Telefonica, Intel, Ericsson, and Nokia.
4. **The new services and applications will drive new architectural, technical, and operational models**, forcing things to be done differently from before – from design to implementation to operations.
5. **The new operational model will leverage APIs and software-programmability to a level not yet seen in networking**, beyond that defined in SDN today, with the separation of control and user planes happening not just at the macro level (network), but at the micro-level (VNFs).
6. **This new paradigm shift will result in an increased need for de-centralization of the infrastructure, and causes the lines between a "network switch" and a "server" to disappear**. What we currently think of as the data plane of a virtual network function running on compute will move to the network switches, and what we currently think of as the network data plane will extend (virtually) into the servers. This will yield network switches which need even more "open-ness" and programmability than they have today, and network-integrated servers to incorporate increasingly network-orientated accelerators (FPGAs, switch fabrics, etc.).
7. **The level of programmability in network switches will continue to increase beyond the current SDN-defined switch abstractions, down to the forwarding plane itself.** The network data plane is likely to become a commodity, and to leverage open source for well-defined functions. Much like the internet has its own domain-specific languages (WikiML, HTML, etc.), the network will also develop its own domain-specific language, with common higher-level scripting languages (Python, Go, Java) across the compute virtualization and network domain.
8. **Operations will be driven by data, and the need to capture, process and react to network data in real-time will give rise to machine learning**. Innovation in machine learning for network data is still in its infancy. Anomaly detection – establishing a baseline of the network performance, traffic flows, and user mobility; and reacting to both gradual changes to the baseline and to anomalies – will enable predictive understanding not just of the network itself, but of macro events occurring in and around the network.
9. **The "open-ness" required to make this feasible extends across all of Dell Technologies**, from the infrastructure (compute, storage, networking), to the compute and network virtualization layers (VMWare), the intersection with public cloud (Virtustream), and the service and application platform (Pivotal). Bringing together the assets across all of Dell Technologies for 5G creates a clear "Better Together" opportunity.
10. **With all the foundational changes to the network and operations, it is still not expected that the buying paradigm of CoSP changes drastically**, especially in initial deployments. The dependency on the Network Equipment Providers has persisted

over a century, and the barriers to entry in the cellular industry have not changed drastically. There has been limited success amongst startups at penetrating either the cellular radio or core. Partnership with the NEPs remains an important part of an overall Dell Technologies strategy, and the threat from non-partner NEPS (i.e., Huawei) to capture the infrastructure is significant.

1. Introduction

While much of 5G is yet to be defined or designed, the rapid roll-out of the internet has shown that, in order to be successful, infrastructure needs to be flexible to cater with future use-cases, many of which cannot even be guessed at today.

In parallel, the shift from fixed to wireless access with the growth of the smart-phone has taken place, fueled by the movement from mobile networks being predominantly voice to exclusively data. At times, the pace of change demanded by the Internet generation has been difficult to reconcile against the need to roll-out new infrastructure at national-scale. The old, telco-focused models of design and deployment are no longer able to deliver the flexibility required at the speed of change demanded by new, Internet applications.

This paper seeks to give some historical background to the development of mobile services explaining the move from 1G to 5G, discuss the known future demands for 5G, the changes happening in the telecom market, the changes in the landscape of equipment suppliers, the development of new technologies relevant for 5G and concludes with a discussion on how the open, flexible approach from Dell can deliver the 5G demands.

2. Evolution of Mobile Networks

Since the invention of the first mobile phone by John F. Mitchell and Martin Cooper of Motorola in 1973, the rise of mobile devices has seen a rapid rise with over 8.0 billion connections by 2016 (Cisco, 2017) and predictions of over 11.6 billion devices by 2020 (Mobile Future, 2016).

The way in which mobile devices are used has changed as well with initial usage being on mobile voice connections to predominantly fixed lines on the Public Switched Telephone Network.

While mobile data has now overtaken desktop or fixed-line data as the predominant method of Internet access (Titcomb, 2017) it wasn't until some 20 years after the first mobile devices were used that any data service was available – this in the form of very slow-speed GPRS connections, slow even by comparison with contemporary fixed-line dial-up services.

It took the roll-out of 3G in 2001 (and, more recently, 4G) before data access came to be the norm for mobile devices; in fact, the traditional voice and related service are now more likely to be delivered by an over-the-top solution using that data connection than by the services of a telecoms company.

As new demands are made of the network, 5G gives the industry a chance to reflect on what has gone well for what is the most ubiquitous form of connectivity on the planet and consider how the technology needs to adapt to meet future needs, some of which we can't even begin to consider.

2.1 Journey from 1G to 4G

Mobile telephony dates back to the late 1940s when the UK Post Office (Storno, 2017) introduced the radiotelephone service in the UK and AT&T a similar system in St. Louis, MI. Both were limited to coverage within single metropolitan area, initially using a single channel but expanded to allow up to 9 simultaneous calls within a large geographic area.

It took the invention of the re-use of frequencies in a regular, cellular pattern by both AT&T and NTT, plus the inclusion of intelligence in the handset to select and move channels as needed, for truly country-wide mobility to become possible. Of these systems, introduced in Japan in 1979, the Nordic countries in 1981 and the US in 1983, AMPS and TACS became the default standards. These became the

1G standards - both using analog radio modulation systems and neither considering any form of data transmission.

Voice quality was poor and the handsets were very power-hungry, leading to the development of 2G systems in the 1990s. Two competing systems emerged: TDMA and CDMA, the latter predominantly in North America and the former in most of the rest of the world. Eventually, disparate TDMA standards converged and became known as the Global System for Mobile (GSM).

A by-product of the digital modulation systems was the ability to carry some data services. Initially, this was limited to the 160-character Short Message Service (SMS), but was soon expanded to a Packet Radio system carried adjunct to the original 2G data. While the technical possibility existed to send data at up to 117kbit/s using later systems, GPRS over 2G (often referred to as 2.5G) was inefficient, slow and prone to latency issues in the data stream.

However, during the mid-1990s, popularity for use of the Internet at home grew and people started demanding the same connectivity while mobile. The technology was unable to provide this as it was based on circuit-switched voice carrying packet-based data.

The ITU developed a framework for 3rd Generation mobile connectivity in the IMT-2000 (ETSI, 2017) specification, creating the 3rd Generation Partnership Project, 3GPP, the standards body now responsible for all radio and core network standardization in mobile systems.

Two main air-interfaces were designed in the 3G UMTS specifications; W-CDMA used in most of the world except China (which used TD-SCDMA) and EVDO (part of CDMA 2000) developed in the US by a related organization, the 3GPP2. Each of these system offered an evolutionary approach to improving data rates, up to a theoretical maximum of 14.7Mbit/s in the case of UMTS.

However, the circuit-switched element of voice was retained in UMTS as voice services was initially seen as the main use-case for mobile devices. IMT-2000 also defined a "3G" capable service as one which could deliver 200kbit/s of data; even by the time of the first live networks in the early 2000's, this data-rate seemed pedestrian by comparison with the developments in the fixed networks.

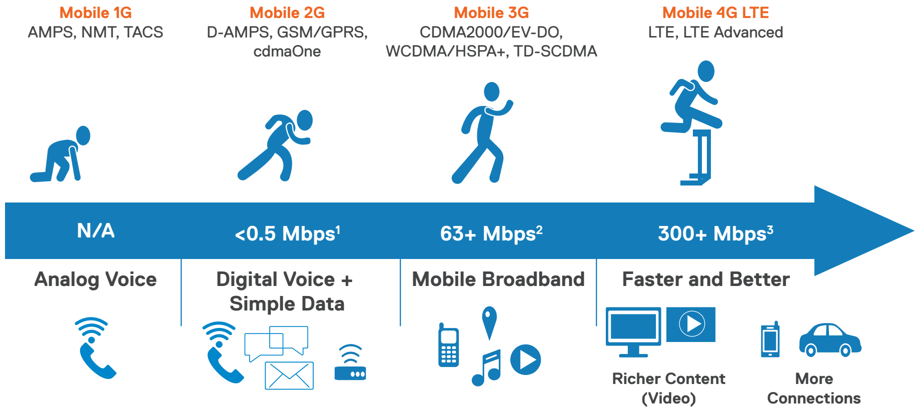

Figure 1: From "1G" to LTE-Advanced (LTE-A) Mobile Technology Evolution

[1] Peak data for GSM/GPRS, latest Evolved EDGE has peak DL data rates capable of up to 12 Mbps;
[2] Peak data rate for HSPA + DL 3-carrier CA; HSPA + specification includes additional potential CA + use of multiple antennas, but no announcements to data:
[3] Peak data rate for LTE advanced Cat 6 with 20 + 20MHz DL CA; LTE specification includes additional potential CA + additional use of multiple antennas, but no announcements to data

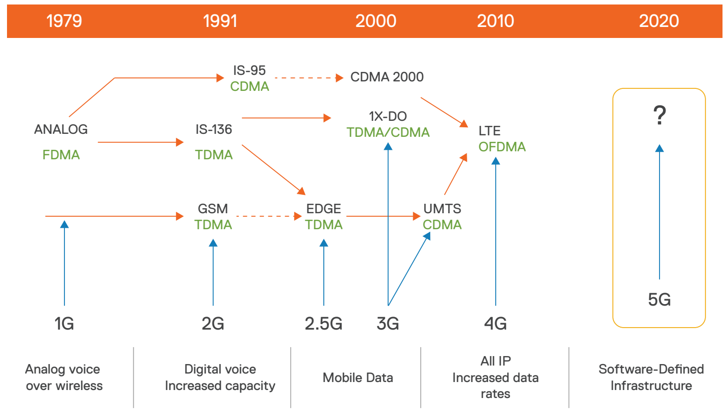

Figure 2: Evolution of the Mobile Access Technologies

Other problems with 3G systems included the incompatibility between the air-interfaces and corresponding complexity required in handsets. The investment issues following the dot.com crash of 2001 also meant that many 3G deployments were left uncompleted, leaving users on very slow 2.5G networks.

In 2009, the 3GPP began to consider moving to an all-IP mobile network with new air-interfaces more appropriate for data. In the meantime, work had been progressing on two rival standards, WiMax with a heritage in the IEEE 802 committees and LTE from the cellular vendors.

The ITU issued a study on future mobility solutions, IMT-Advanced (ITU, 2017b), which defined 4G connectivity as 100Mbit/s downlink while mobile and 1Gbit/s fixed.

A battle between LTE and WiMax took place over the next 4 years; several operators had tried to gain an early market advantage by deploying WiMax networks while others had opted for LTE. Although LTE became the de-facto standard in 2013, it wasn't until late 2015 that the last public mobile WiMax network was switched off and LTE became the only operational 4G standard.

Of particular note in 4G is that the entire network is focused on data-access; "legacy" voice is delivered over IP-based solutions such as VoLTE. There are no traditional voice-switching components.

2.2 IMT-2020 moving towards 5G

Initially published in 2012 and subsequently revised many times, the ITU IMT-2020 "IMT for 2020 and beyond" (ITU, 2017a) discussion paper caused 3GPP to start considering the next generation of mobile devices and associated infrastructure, later to be dubbed 5G.

This ongoing work led to publication in early 2017 (and expected ratification in November 2017) of a number of key requirements for a network to deliver in order to be 5G-compliant (TelecomTV, 2017):

- Minimum requirements for peak data rate (Downlink:i20Gbit/s, Uplink: 10Gbit/s)
- Target requirements for "user experienced data rate" (Downlink: 100Mbit/s, Uplink: 50Mbit/s)
- Targeted peak spectral efficiency (Downlink: 30bit/s/Hz, Uplink:15bit/s/Hz)
- Minimum requirement for user plane latency for enhanced Mobile Broadband (eMBB): 4ms
- Minimum requirement for user plane latency for URLLC: 1ms
- Minimum requirement for control plane latency is 20ms (10ms encouraged)
- Minimum requirement for connection density is 1,000,000 devices per km^2

- Requirement for bandwidth is at least 100MHz below 6GHz, 1GHz above 6GHz
- Four classes of mobility:
  - Stationary: 0km/h
  - Pedestrian: 0km/h to 10km/h
  - Vehicular: 10km/h to 120km/h
  - High speed vehicular: 120km/h to 500km/h

These requirements represent significant capability improvements over 4G LTE networks.

3. 5G new demands

That the ITU requirements have taken nearly 4 years to reach the point where they can be published as base-5G needs testifies to the level of discussion there has been in the industry trying to qualify the need for a new network and quantify the magnitude of the problems which 5G seeks to address. As such, the ITU specification is really only useful in a type-specification for vendors and to understand why these metrics have been arrived at, it is necessary to examine the use-case behind them.

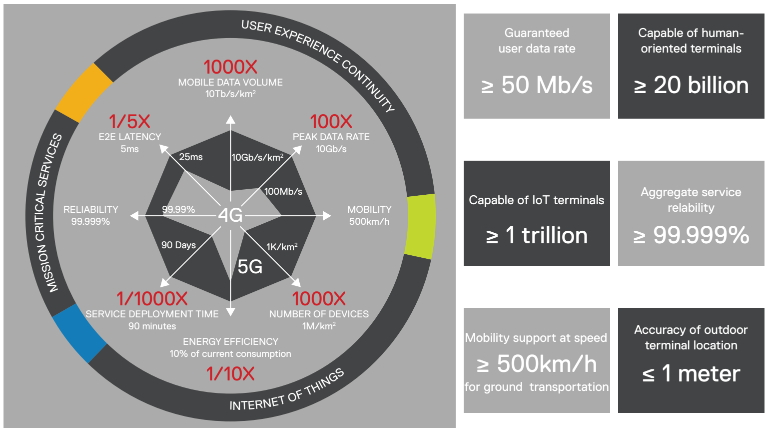

Figure 3: New Service Characteristics & Capabilities Enabled By 5G

Delivering these requires changes to the entire network – both from an architecture and a technology perspective. As such, 5G standards reflect changes in air-interface, the manner by which the RF connection is deployed, to the current packet-core architecture, and to the surrounding operational support systems.

What is obvious from the work of the 3GPP on 5G is that the one-size-fits all approach to both air-interface and transport core network which has been a consistent feature of all past mobility solutions is no longer viable; the disparate demands for the future 5G network and the goal to provide device-specific, customer-specific, and application-specific experiences cannot be delivered without impacting the cost and complexity of operating in the status-quo.

Standardization work towards 5G radio aspects is due for completion in Rel 15 which will be “frozen” in September 2018.

Important decisions towards this goal were made in March 2017. 3GPP released discussion papers on a new 5G Core in January 2017 and moved to a formal System Architecture in June 2017 with the latest versions of 23.501 and 23.502. Work is expected to take at least two years to complete the 5G Core standards.

3.1 New traffic types

Mobile networks were initially built with one use-case in mind: voice. In fact, the architecture, operation and optimization of mobile networks from first-generation analog to 3G mirrored that of the development of the PSTN-based networks.

DELL EMC

The introduction of data services in 2.5G and later 3G challenged the manner by which packet-based data is carried over circuit-switched networks; RF links suffer from fading, multi-path effects and retransmissions which in-turn cause the predominant Internet protocol, TCP/IP, to behave poorly.

4G, as the first all-IP network driven by data services and optimized for long-lived high-bandwidth streaming video, has provided some relief, but there remains a disconnect between the operation of the air-interface and packet throughput performance. Consequently, even though “headline” data rates on LTE appear to be close to wired connections, actual throughput (goodput) can seem to be significantly lower, accompanied by high-levels of latency.

In LTE, voice-traffic is simply another data service, significant complexity had to be introduced to ensure that voice-quality can be maintained under all conditions, including segregating data-traffic associated with the voice service on specific air-interface modulation types, partitioning the network traffic into low-latency forwarding schemes, enacting methods for monitoring radio performance and taking evasive action should RF quality drop below given thresholds, and also deploying dedicated core networks that were optimized for voice traffic models.

It is for these reasons, along with the initial lack of LTE ubiquity, that voice over LTE (VoLTE) did not appear immediately, with operators instead opting to deploy 4G networks for data and continue to leverage existing 2G/3G networks for the delivery of voice service. This has prolonged the life of both 2G and 3G networks, and led to the early appearance of IoT services on these increasingly under-utilized networks

3.2 IoT

Perhaps the widest set of uses cases and the most challenging for any network infrastructure to carry efficiently, IoT presents challenges in both the amount of data, the immediacy of the data and the sheer number of devices, with estimates ranging from 6 to 50 billion devices in the next 3 years (IEEE Spectrum, 2017).

Due to the range of potential applications and services, it is not possible to arrive at a single definition of IoT. Instead, it is imperative to look at some diverse IoT use-cases and their impact on the network to have an idea of just how disparate the requirements might be:

Factory Automation: Manufacturing plants are getting smarter. Machines will host hundreds of sensors and actuators giving visibility and control not only to local staff but to remote operators. This is often the case in large manufacturing plants where the machines will be owned/operated by the OEM but located in the customer premise. The issue here becomes both the sheer number of devices and the control-loop latency; the presence of the OEM in the control-loop in current mobile architecture would add many milliseconds of delay to the machines’ control functions. 5G networks, in this instance, are expected to have the ability to connect, collect and process device functionality close to the edge, but under control of a distant service owner, the OEM. The key to success here is including the OEM in a manner by which the local functionality is unaware of the extra connections, yielding the creation of what is known as the IoT Gateway.

Key technologies in 5G include:

The ability to connect to devices over non-standard RF interfaces and the ability to position software/application functions close to the edge using local compute functions, such as those specified in ETSI Multi-Access Edge Computing (MEC).

Networks will also need to be sliced (logically separated) to ensure sufficient isolation and security between public and private application instances. For the operator, the key is to be able to deploy functionality locally to the end devices from a single management instance. Network slicing is inherently built on the principles of SDN.

Sensor Networking: Often seen as the bedrock of IoT development, sensors are being embedded in everything and "cloud-based" solutions being offered to consume and present this data. Mostly, the data produced by these sensors is small, not overly time or delay sensitive and sent at regular times but with a long period. This presents a number of issues in today's networks which will be required to be addressed in 5G:

Sheer number. RF capacity is a finite resource and techniques to re-use and share spectrum plus identify new areas of spectrum is required.

Geographical spread. Sensors networks can be anywhere but the main available spectrum is in the short-range >6GHz bands. Deploying new base-stations is expensive and increasingly difficult to do; 5G imagines an "ultra-densification" of RF connectivity which may prove to be impossible to achieve given market economics and lack of available sites. Therefore, flexible solutions where-by local-area coverage networks can be built to serve a particular area or group of sensors will need to be developed. Flexibility is the key here – being able to use existing or common compute assets to provide RF connectivity; transparent management operation for the SP, etc.

Traffic Profile. Simple sensors will produce minimal levels of data – a temperature reading may only be a few Bytes of data. In today's networks, transporting even a few Bytes requires many more of overhead and this adds to the overall cost of transport. In an Internet predominantly concerned with carrying high-value, large amounts of data, the overheads can be justified; this is no longer the case with IoT.

Device Power. Many IoT sensors are expected to be small, long-lived (20+ years) battery-powered devices located in remote, inaccessible areas or deployed once. Current mobile technology is built predominantly around the smart-device which measures its battery life in hours. While some of this is due to the intensive processing on the device itself, a large proportion is due to the nature of serving the RF interface. Technologies such as LoRA and services such as SIGFOX have concentrated on optimizing battery life through the use of for-purpose RF techniques which do not integrate to current mobility systems. 5G will seek to provide connectivity to suit this market.

Identity, Security and Management. Aligned to the above issues, one-time deploy devices where power needs to be managed are designed in such a way that their identity is expressed in a simplified manner compared to the protocol-intensive methods used in today's mobile systems. The primary method of providing identity and security in cellular systems is via the SIM card; these will often cost more than the entire IoT sensor. Likewise, software patches can be applied to a smart-device with sufficient power and compute functionality over-the-air; this is unlikely to be the case with Iot sensors. This impacts both the sensor and the network supporting it – software on the device at build-time should really be seen as existing for the life of the device and therefore for many years. The network will need to be aware of these complexities as well; changes which would potentially disconnect thousands of IoT sensors (for example due to a protocol change) are going to prove troublesome to manage.

3.3 AR/VR

VR/AR has the potential to swamp the current networks with its demand both for raw bandwidth and very stringent latency requirements (Qualcomm, 2017).

Ultra-high quality immersive video already requires bandwidth in the order of several hundred Mbit/s although it can be quite delay-tolerant (depending on the application).

DELL EMC

However, video when applied to remote control and tactile IoT applications such as remote diagnosis, tele-medicine or hazardous environment operations makes demands of both bandwidth (in the order of 1 Gbit/s per video stream) and latency. Kings College London's Tactile Lab paper (Aijaz et al. 2017) shows the effect of adding latency to the control loop of a haptic actuator and associated robotic arm demonstrating that the entire control loop, both in terms of RF delay and application processing overhead, needs to be less than 10ms and closer to 1ms in such applications. This is beyond the 50+ms minimum latencies seen on today's mobile networks.

More problematic for the headset/immersive VR applications is the effect on the wearer of poor-quality video coupled with excessive latency (over about 10ms); Virtual Reality Sickness.

Given the physical limitations of the air-interface and spectrum availability, it is obvious that some of these issues can only be solved by positioning as-much of the content and control close to the user.

Key technologies which will enable VR/AR in 5G:

Local placement of services via the use of pre-positioned content to reduce latency. MEC is an ideal candidate solution.

Orchestration of services across different access media. The ability to manage the provision of software from 3rd parties (e.g. gaming software, remote diagnosis, etc) into the edge-compute system.

High-bandwidth services. VR/AR solutions call for high definition and ultra-high definition video. Densification of the RF layer will be required. Technologies such s C-RAN whereby area-wide baseband processing will be necessary to meet this need.

3.4 Mission-critical

As mobile coverage becomes ubiquitous and the "connection method of choice," services currently delivered on for-purpose networks which have been engineered to provide some form of tolerance to failure and/o availability will migrate to the mobile network. While some of these services will simply have service levels in excess of a consumer-grade service, others will be engaged carrying safety-critical traffic, for example vehicle-to-vehicle, rail transport, fire/flood alerting systems, etc.

This requires that the network has the ability to deliver to against defined service levels; redundant systems may also be required on the network with data shared/stored in manners that it retains availability without compromising security. There is no means today to declare that a particular data service is critical and requires specific handling through the infrastructure; 5G will provide methods by which data context can be determined. Additionally, in times of emergency, networks may become partitioned; 5G will have the ability through local compute functions to place elements close to the affected area ensuring that critical processing continues.

Key technologies within 5G which enable the handling of mission critical communications are:

Control/User Plane Separation; the ability to steer traffic which carries criticality into a network slice which has been engineered to provide a different level of availability and/or redundancy than other parts will enable SLAs to be defined

C-RAN; densification of the access network in areas where mission critical data is connected requires rapid fan-out of RF connectivity.

Edge Services; in order to provide redundancy in the event of network loss, some mission critical applications will need to be relocatable to the edge of the network; for example, entire local cellular service should be able to survive in the event of a failure to a central core network in the case of a weather-related event. MEC will enable such functionality.

3.5 Enhanced mobile broadband

The current data services available on LTE networks is a "one-size-fits-all" – it is not uncommon for the goodput to vary across a large range, especially while moving at speed. Additionally, all data travels over the single packet core and discriminating traffic into discreet user-groups is complex. While local break-out services such as LIPA and SIPTO are available, these are rarely used as defining which connection a particular session should use is difficult.

Economically, it is difficult for operators to provide comprehensive data services in low-usage and unusual locations; it is not uncommon for urban areas to have a choice of 4 or more operators with good data rates but rural areas to have one or even no available operator. It is thus difficult for the consumer to choose the optimal operator given such variability in coverage. Other environments such as planes of trains are notoriously difficult to provide coverage to in a multi-operator environment.

The goal of 5G is to provide ubiquitous connectivity at high data rates in all locations, whether moving or at rest. 5G will also allow user-group access for specific data – for example, the enterprise email service on a user's handset will always be routed over a specific connection for that traffic and not over a default "Internet" connection. This means that specific traffic handling (SLAs, security, etc) can be applied to different traffic across the network.

As the demand for high-speed mobile data access has increased across a range of transport systems, new multi-operator connectivity will be required to enable passenger's access to data services but without each operator needing to deploy their own specific equipment. Likewise, rural areas could be served in a manner which provides connectivity via a single set of infrastructure where it is uneconomic for each operator to deploy their own. New business models and companies will emerge to fulfil this "Neutral Host" market; these companies will be built on a new range of open, lower cost networking solutions.

Key technologies which will enable the widespread roll-out of eMBB at moving speeds of up to 300km/h are:

C-RAN; Neutral host operators will be able to deploy dense connectivity in specific locations and scale the centralized processing using flexible BBU. Existing operators will use C-RAN to provide wide-scale roll-out of new 5G connectivity.

MEC; Pre-positioning content close to the edge, especially in mobile environments such as trains and planes, will be needed to keep the backhaul costs down. Likewise, rural areas where high-bandwidth connectivity may be expensive or prohibitive in terms of transit time (long circuits to very remote areas) will require MEC functionality.

CUPS; being able to separate traffic into different categories and treat according to SLA and/or ownership will enable defined classes of service.

3.6 Data plane performance

Virtualization and softwarization are key foundational elements in the construction of a new 5G core network and the benefits they can bring in terms of flexibility and time-to-market are clear.

In addition, the new demands on the 5G Core will require data rates in excess of 10Gbit/s across the network with minimal latency in some instances as discussed before. The goal of 5G is to decrease transit times and latency and all current virtualization solutions add latency to the throughput.

The ability to off-load traffic through the 5G software core components to ensure expedience of forwarding will therefore be important, and a number of new data optimization techniques are being developed.

Along with the well-known DPDK from Intel, other stack bypass solutions exist whereby traffic can avoid having to transit portions of the virtualization system. However, many of these optimization techniques are hard-tied to a particular virtual instance, be it a container or hyper-visor solution, and therefore impact the very flexibility which softwarization seeks to address.

Key technologies in 5G which will seek to address the impact on packet forwarding are:

- FPGA; Hardware based FPGAs can be used under control of the core elements to bypass and "switch" data in an optimal manner.
- SDN; the ability to program an entire network function chain such that optimal forwarding based on a traffic class and/or SLA is key to 5G.
- Control/User Plane Separation; Some traffic may not be concerned at high latencies (e.g. email), whereas others will be negatively affected (VR/AR). Multiple user-plans based on the traffic profile will exist in 5G with different technologies as appropriate for the traffic type.

3.7 Network efficiencies

3.7.1 Spectrum

Since the wide-scale installation of fiber backbones in the 2000s, while last-mile capacity has been problematic in some geographies, the wired Internet predominantly has infrastructure capable of delivering future traffic needs.

The same is not true in the raw commodity needed for mobile system – RF spectrum. This highly-valuable asset is tightly controlled and allocated and suffers from the dichotomy that the longer-reaching lower frequency spectrum is the one that provides the least bandwidth capacity whereas the higher-frequencies can provide the required bandwidth but do not have the required range.

Governments around the world are also keen to consolidate and limit spectrum and have already indicated that they expect 5G to be the vehicle by which some for-purpose legacy systems can be retired.

Governments around the world have been busy building out digital transmission systems for radio and television to replace analogue, usually not just to improve the range of services on-offer but also to gain from better spectral efficiencies in new modulation techniques. Switch-over to Digital Terrestrial Television is nearly complete in most EU countries saving substantial amounts of spectrum in the 800MHz band; a very useful band for future cellular solutions. Switch-over from broadcast AM and FM to digital systems has not progressed at the same rate, but take-up is increasing due to the rapidly falling cost of receivers; both Germany and the UK are now at over 50% of radio listening being on digital systems and both are considering switching-over.

3.7.2 Seamless mobility

5G will also need to tackle these challenges of how to make the service appear to be continuous while the connectivity underneath changes.

A good example which highlights not just the potential to consolidate spectrum more efficiently but also explains the challenges faced by 5G designers is the UK government's decision to replace the current standalone TETRA-based cellular system which is similar to 2.5G networks with an Emergency Service network on a future 5G solution as a VPN on one (or more) of the current MNOs. Key features required for emergency services include capabilities not in the 4G specification such as pre-emption (the ability to seize resources from other users), talk-groups both within a user-community and the ability to set-up new communities quickly, the role of a dispatcher and ultra-low call setup for Push-to-Talk (LTE has a Push-to-Talk over Cellular but the call setup is not fast enough for emergency services usage).

However, the current UK-wide TETRA network provides 100% geographical coverage unlike current 4G LTE. Providing national coverage instantly at switch-on has never happened with new mobile roll-outs so new techniques such as the ability to inter-operate with existing infrastructure will be required.

Additionally, the very notion of "national" coverage will challenge the concept of the competitive MNO – most operators will cover quite similar areas with similar revenue-potential equally well as they cannot afford to deploy to loss-making areas. Roaming between national operators while technically possible is not the-norm, is not seamless (an LTE re-attach is required which will cause all sessions to drop) and is a drain on battery resources (mobile devices in their "home" area currently do not search for alternate networks in order to preserve power).

Both of these examples, however, beg a similar question. In terms of national infrastructure spend, would it have been more efficient to have a single infrastructure solution such as 5G able to provide the variety of services?

The truth is that technology of the time when the infrastructure decisions were being made, mostly 2 or 3G, was not flexible enough to be able to deliver anything other than its primary service.

There is a lesson here for 5G – the solution needs to be flexible enough to be able to handle services which we haven't even considered yet.

3.8 Operational efficiencies

In the telecoms industry, across the space of about 30 years, we have seen several waves of technological change, but in many cases, the new technology has not fully replaced the original. For example, in 2017, a number of 2G networks still exist around the world in a fully operational and fully supported state.

Why is this? Firstly, the time involved in deploying national infrastructure is extended – most countries do not have full 4G coverage as of mid-2017 despite roll-outs having started many years previously.

Secondly, the funding cycles rely on new investment being driven by revenue derived from the infrastructure. In some aspects, there is a chicken-and-egg issue here – the argument may be made that if an entire network were to be put in place from day 1, new revenue would flow. However, experience of 3G shows that this is not always the case; the vast sums paid for spectrum in 3G have not yet been matched in new revenue. Also, macro-economic circumstances may slow the investment pay-back making the original network unprofitable. To some extent, 4G has suffered from a difficult macro-economic situation.

Thirdly, existing users cannot be disenfranchised simply because a newer technology exists; migrations between technologies need to be planned. This is especially the case with 2G which has found some niche uses in M2M and we find 2G equipment placed in remote and difficult-to-access locations such as ATMs and vending machines.

Finally cutting across from one technology to the other requires that the new technology provides service at-least-as-good as the one it is replacing in terms of coverage and services. Such technology transitions therefore happen over extended time-periods – an example of this is the moving from 405-line TV to 625-line TV in the UK which started in 1967 but was not completed until 1985 with the two systems running side-by-side. A similar move to digital terrestrial television was planned to take only 5 years, starting in 1998, but the Dot.com bubble of 2001 caused the eventual switch-over not to be completed until 2012 (Given, J. 2003).

Given current short funding and investment cycles and the uncertainty of the new revenue sources, 5G will need to offer more than simply a new, parallel set of infrastructure. At least in the early days, most of the revenue derived from 5G will come from savings in the operational state of existing technologies and introducing the ability to rapidly deploy new elements as use-cases are identified.

5G is more, therefore, than a new solution; it is about re-working existing solutions in a manner whereby cost-savings can be made. For example, moving today's 4G packet-core to a flexible platform which can be expanded to provide additional services and/or capacity is seen as a 5G facet – the current EPC solutions are difficult to expand and are built as a one-size-fits-all solution. Likewise, the ability to deploy new processing and/or storage solutions around the network is not currently possible and 5G will look to make this possible.

The architectural constraints of for-purpose hardware and software solutions inherent in 2G, 3G and 4G has not allowed optimum efficiency in the current deployments as it is inflexible. 5G will not only introduce new technologies in both packet-core and RF, but it will enable some of the existing solutions to be re-worked into more dynamic, more flexible deployments.

At the same time, presenting both new and existing elements as a single management domain will help reduce costs due to proliferation of OSS/BSS systems which has occurred as each new technology has brought with it its own management and operational model. 5G will look to address orchestration both within the new technologies and in the existing with the goal of reducing operational cost and complexity. Many studies into Management and Orchestration (MANO) are ongoing, such as the ETSI OSM group (OSM, 2017).

4. Telecom market realities

The telecom industry is being subjected to a number of positive and negative forces that is coming into effect in the timeframe of 5G and will be addressed in some cases by 5G and in some cases of changes outside 5G it-self but that can be seen as included in the network investment connected to 5G.

- Flat revenue from existing business segment with increased traffic in the networks leads to big investments in infrastructure with low or even negative margin. This is being addressed by seeking new segments of business and by lowering total cost per bit transported. The search for new business areas is one of the reasons things like extreme low latency is a requirement for 5G targeting among other things automotive and virtual reality applications.
- Over the top and cloud actors eating into premium revenues like voice and media distribution while leaving the operators with the cost of the transport at low revenue. This is being addressed by providing an execution environment close to the consumer to host the OTT service at a price as well as reduce the transport cost whenever possible. Also possible to leverage storage and security at the operator to bring part of the OTT services back.

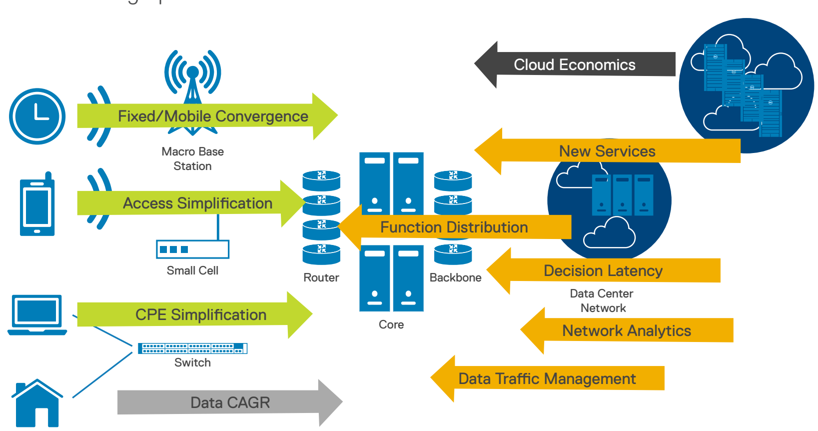

Figure 4: Shifts in the Mobile Architecture(s) and Value Chain

- The telecom networks is for historical reasons much more complex than what is common on the IT/Cloud side of business. In the past coming from a fixed telephony background this was one of the reasons why cellular networks had the stability and reliability that brought on the success of the past decades. Today the premium paid for the complexity is a threat to the operators existence and this is being addressed in the 5G timeframe by modernization of the infrastructure according to IT/Cloud praxis as well as with architectural modification from 3GPP going towards a more lightly coupled model.
- There is a convergence in several different domains that historically have been separate verticals that is happening in the 5G timeframe.
  - Convergence of fixed and mobile networks
  - Triple or quad play where data, voice, and media in different combinations is brought into the same networks
  - Convergence of infrastructure merging the principles from IT/Cloud with the network side. The cost point of the IT/Cloud side is an advantage but there are still functionality that is unique for the networks which needs to be handled.
  - Convergence of execution environments with resource pools spanning from the EDGE to the CORE and to the CLOUD.
- Automation, Management, Reporting and Analytics will become the value creation capabilities of the network to reduce service creation time dramatically from the classic deployment model of network function.

A significant driver of 5G definition is also the economic situation in the mobile industry. Prior to this generation of technology, revenue for the operators has been driven by two aspects; handset income and service-plan income.

Two elements have combined to make these sources of revenue less reliable. Firstly, smartphone penetration has reached saturation point in many markets – Kantar (2017) suggest that the double-digit growth globally ended in late 2015 with mature markets having reached full saturation several years before hand. While consumers may be tempted to move to a 5G-capable handset, this can no longer be assumed to be at a premium as was the case for the move to 3G (and to a lesser extent to 4G).

Secondly, ARPUs are dropping with a trend emerging towards $20. Statista (2017) show 2015 with a high of around $50/month, North America being the outlier in the developed world; figures such as $10/month in Portugal at the lowest end are more typical in developing nations. Driving this is the mass move away from operator-owned voice and messaging platforms to cheaper, often free-to-use, OTT services which are able to survive with much lower ARPUs. WhatsApp, for example, is estimated to be running at an annual ARPU of $0.06. Growth in the use of operator-

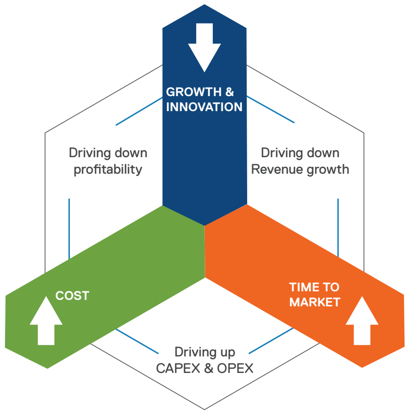

Figure 5: 5G Adoption Cycle: Drivers and Inhibitors

provided SMS applications has stalled; usage of OTT IM apps now exceeds that of the operator-provided solutions (GSMA Intelligence, 2017 p.34).

Likewise, the growth in the use of mobile video (GSMA Intelligence, 2017, p57) is stressing current networks – video is inherently bandwidth-hungry and delivering over a mobile infrastructure is costly. While operators have developed their own in-house video offerings to be able to capture additional revenue, the most popular video streaming sites remain the well-known OTT providers such as YouTube and Google.

Several commentators (CNBC, 2017) have noted that consolidation of providers is expected in the telecoms sector; tower and even network sharing is becoming more common across providers (Ericsson,2017). Additionally, the nature of the ownership of spectrum as it is held by one operator makes consolidation attractive when trying to roll-out new services as having more spectrum which may be re-allocated or re-farmed allows the movement and introduction of new services alongside existing ones. An example of this was the merger of Orange and T-Mobile in the UK where economies-of-scale allowed effective re-farming and the early introduction of 4G described in the paper "Assessing the case for in-country mobile consolidation", (2015. p.38)

Mergers do not happen at zero cost though; regulators have famously blocked mergers of existing operators such as between AT&T and T-Mobile in the US (WT Docket No 11-65: Staff Analysis and Findings, 2011).

The current technology does not help; services and transport are inextricably linked such that a handset is tied to a transport operator. Moving across operators or even between technologies is complex and not always possible. 5G will need to consider how the service and the last-mile connection can be separated seamlessly so that a given quality-of-service may be maintained as different connectivity options are used. This will have to be accomplished in a way which recognizes the value provided by the transport provider through revenue sharing. Pre-5G technologies do not expose the granularity required to be able to determine an outcome; the network delivers a single experience. Today's applications require a tighter association between transport capabilities and service-level; for example, an immersive AR/VR application requires specific bandwidth and latency so must be able to request these from the transport layer. Likewise, the transport layer must be represented to the application in a manner which describes its capabilities. In this way, several transport options may be presented to the application, allowing the application to select those which would enable the expected outcome to be delivered. 5G can therefore be seen not just as a new radio or core network technology, but as a way of intelligently utilizing exiting solutions as well.

Security is a major concern for two reasons:

- Current regulation on Lawful Intercept requires the ability to monitor an end-user without either them or other observers of the network being aware of the intercept
- Much Internet traffic is encrypted with the stated goal of the IAB (2014) to have all traffic encrypted.

Current mobile network technology has a simple solution to LI; 2.5G, 3G and 4G architectures are identical in that they have the concept of a central "anchor point." All traffic between every device can be seen at this point both within the mobile network and to the Internet. Therefore, intercepting traffic is straightforward – a tap is placed at the center. Likewise, intercepting at this point will cause no change in traffic flow and is therefore undetectable by anyone observing the network.

Even so, once intercepted, the predominance of encryption makes it very hard to be able to read the data flow; many times, security agencies are more concerned with the meta-data rather than the content and this isn't an issue. However, this is seen as a growing concern requiring, potentially, changes to IETF protocols to allow for some monitoring.

5G seeks to place services closer to users and therefore changes the underlying anchor-point-based architecture. Lawful Intercept hence becomes a distributed problem requiring compute and/or storage at the edge of the network. New techniques will be required in the core network to ensure that the non-observability of network traffic changes due to intercepts is maintained.

5. Telecom industry status

In previous "Gs" time, the telecommunications industry could be seen as a specific business defined by the special requirements around "carrier-grade" (99.999%) operation and large number of regulatory requirements. This industry was also bounded by the relatively few service providers and the even fewer providers of complex system solutions and components. This resulted in a situation where Network Equipment providers (NEPs) (Originally Nortel, Ericsson, Lucent, Alcatel and Siemens. Now, Huawei, Ericsson, Nokia, Alcatel Lucent, and Cisco) and operators created relationships in which the NEPs where extensions of the technology departments of the operators. This relationship was also amplified by operations outsourcing agreements and other contractual stipulations. In many instances, the existence of the operator and the NEP were inter-twined.

The relationship was such that the NEPs drove standards and development via 3GPP and then built equipment in a green field market that was not questioned by the operators since at the time it was more important to provide a working service and then build a market share at a premium in competition with other national operators. In an economic environment of rising customer ARPU, this relationship was in balance.

The solutions built were bespoke and proprietary with the exception of the 3GPP interfaces that enabled interoperability of the functional nodes in the network. In the same time the 5-9's requirements drove a verification, validation and acceptance methodology associated with a very high assurance, but at the cost of long lead times in deploying or upgrading network functions.

The technology development and innovation in cellular communications has been deliberate – iterative to ensure no major disruptions to carrier-class operations and timed to ensure that operators recoup some level of CAPEX outlay through service revenue. When compared with "Internet speed" deployment and availability of new services and features, deployment within the mobile industry has been slow and consequently, consumers have moved to other providers for their services and applications, diminishing the role of the cellular operator.to largely transport and connectivity. This move of revenue has challenged the ability of the operators to invest and likewise has had an impact on the NEPs.

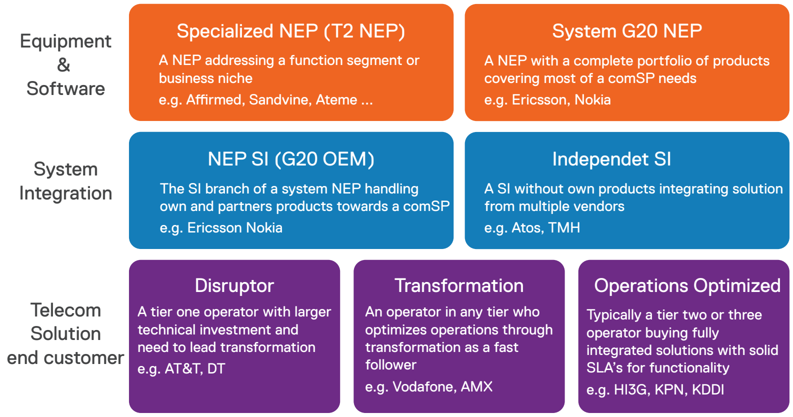

Figure 6: Shifting Roles Across the Mobile Supply Chain

Currently there are only four big NEPs left in the industry (Huawei, Ericsson, Nokia and Cisco), a number of potential challengers (ZTE, NEC and Samsung), and number of specialized NEP such as Affirmed, Mavenir, Metaswitch and others

As early as 2012, operators realized that a radical change was required and generated proposals around NVF to open-up the telecom infrastructure, citing the dominance of a handful of NEPs dominance as well as high costs by suggesting a separation of the execution environment from virtualized network functions running on top. This change required the cooperation of NEPs to provide such infrastructure agnostic network functions which was not in their best interest and development to this end has been slow. The result is that we now see virtualized verticals replacing the legacy verticals with little actual gain in the operational efficiency the operators badly require. Longer term, we will reach the full implementation of NFV as there are specialized NEPs prepared to produce the disaggregated network functions and challenge the incumbents.

In this scenario there is a need for a new actor who will build the **disaggregated telco data center** which can be either one of the big NEPs or one of the emerging system SI companies (e.g. TMH, Atos). The disaggregated telco data center is a key component for Dell EMC in 5G; investments in new infrastructure and software creates the momentum needed for a change of sourcing, development, system-integration and deployment principles to enable the disaggregation. The disaggregation also turns infrastructure from components in a telecom vertical into a full underlay independent infrastructure system that requires less telecom expertize.

Traditionally, infrastructure was sold to operators for telecom network functions as an **OEM through the NEPs**. The model has changed slightly and some equipment is now sourced directly but on strict specification from the providers of the network function software. The main reason for this lock in is due to the verification and certification processes pointed out previously. Every network function in every version is verified and certified on a specific set of infrastructure components. This process has contributed to the slow roll-out of new services.

Telecom procurement will evolve such that verification and certification will be based on scalability, orchestration and reporting. With such paradigm it would no longer be a question of how much can run on this hardware but rather what resources are needed for a function of this size. Thus, specification and procurement of infrastructure for a telecom environment would be a much more **like IT procurement**, building agnostic data centres as resource pools.

Abstraction of the infrastructure to enable certification of the solution, for example in a similar manner to VMWare whereby certification of the software is automatic on specific infrastructure where the VMWare installation is itself certified, enables rapid

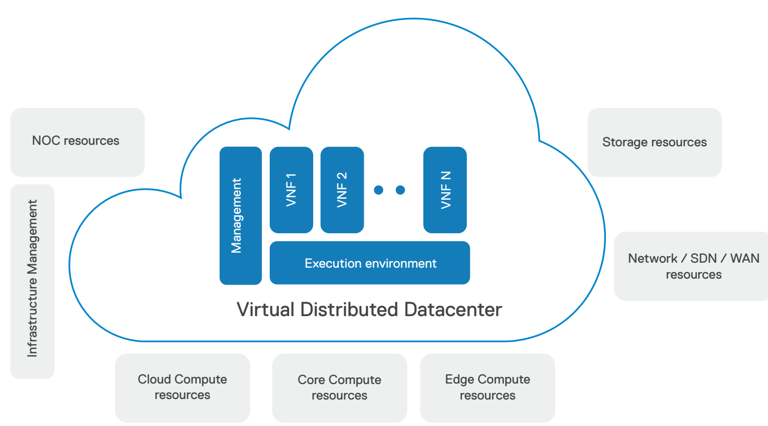

Figure 7: The Emergence of a New Mobile Cloud Blueprint

deployment on "known good" installations based on reference architectures rather than the current first-principles approach.

The telecom data center will therefore evolve from being a centrally located, for-purpose system into one encompassing compute in different physical locations with a single orchestration plane, thus creating one **homogenous execution environment** to place workloads in.

In this manner, operators will be able to deploy functions from data forwarding to MEC to C-RAN at on as-close to standard IT components as possible but with higher architectural specifications to handle the low latencies and near real-time handling of scheduling as required by the applications.

As this move to a mixture of locations including both central and on premise components evolves, the relationship with the end customer will change and new actors providing wholesale connectivity from a range of operators could emerge.

6. Technology advances

It can be seen that unlike previous mobile roll-outs, 5G is not a wholescale replacement of existing technology with something new – it is a way of evolving what is there and re-imagining the architecture so that maximum flexibility and operational efficiency may be made of current and future infrastructure.

Like in previous generation lifts but even more so for 5G there is a significant out of domain technology enabling the core technology of 5G. Areas in particular are in virtualization and automation/orchestration. There are also major changes in how networks are developed and deployed that comes from the agile methodologies in the broader software industry as well as use of micro services as a way to reach speed and flexibility.

The technology advances can be grouped into five major areas – the big 5 – comprising SDN, NFV, MEC, Air I/F and Automation. The latter area is rather wide and collects technologies for reduction of OPEX and TTM as well as flexible service creation.

6.1 Air interface

Probably the biggest difference in 5G than previous generations is the lack of a new air interface. Instead of mandating a new air interface as 2G, 3G and 4G have done, 5G uses the existing LTE air interface in existing spectrum but augments its capabilities by over-laying a new air interface above 6GHz (Qualcomm, 2017a).

In fact, this air interface is based on the same technology as that used in LTE, OFDM, but with several changes:

- 5G NR uses CP-OFDM which has narrower shoulders than LTE usage of OFDM meaning that better use can be made of the spectrum (less requirement for guard-bands)
- Single-Carrier versions of 5GNR exist (SC-OFDM and SC-FDMA) which are more suited to devices with limited battery life such as used in IoT in mMTC.

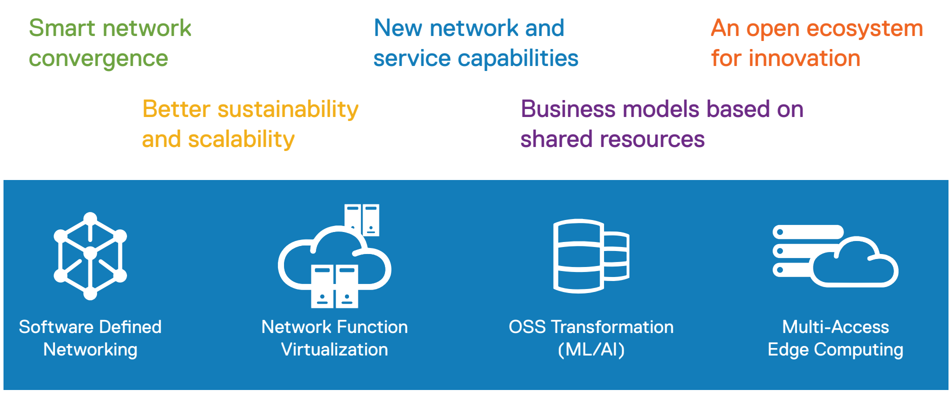

Figure 8: Technology Transformation Domains

- MIMO techniques are enhanced and aggregation can be accomplished across spectrum jumps allowing for massively enhanced data throughputs such as are appropriate for eMBB applications.
- Collaborative services will be able to be operated not only across spectrum used for 5GNR but also with LTE Advanced meaning that the 5G capabilities are additive to current services.
- Spectrum can be split between mMTC, eMBB and Mission-Critical as needed including in unlicensed space.

The use of higher frequencies will require a denser roll-out of base-stations; operators will need to be able to deploy systems rapidly and cheaply. New techniques in SDR will mean that many cell-sites will be “dumb” RRH type of deployments with the baseband processing backhauled to compute functionality located regionally and/or centrally.

C-RAN type of deployments may also no-longer be linked to a single operator as is the case with existing range extension such as DAS. So called “neutral host” solutions able to provide multiple operator access will appear, alleviating the difficulties associated with the in-building deployment of small-cells for LTE.

6.2 Packet core - general

Today’s 4G LTE Packet Core is built around the same design principles as that of the original 2.5G networks. That architecture was based on the voice central-office design with switching centers to provide transit service between endpoints and to/from the rest of the network.

While the end-station is mobile, the architecture is static, using a serious of network tunnels to create the illusion of mobility. This architecture has proved to be acceptable to deliver voice and text services, but is limited in its ability to provide the high bandwidths and low latency required for interactive video and the ultra-low latency required for IoT applications.

3GPP has published a new 5G Network Architecture in 23.501 (3gpp, 2017) – however, while this describes new interfaces and opens the way to potential network re-design, 3GPP does not envisage a whole-scale change in the operation of the packet core within the first deployments of 5G services. We can therefore expect to see the existing protocol stack of GTP and IP remain based on anchor-points for some time.

In the interim, two modifications to the current architecture will start to introduce 5G concepts to both the 4G LTE and 5GNR air interfaces in order to address the performance requirements of the new use-cases; virtualization (and by association, network slicing) and edge functionality.

6.2.1 Packet core – virtualization

A key technology which will be necessary to enable the flexibility and scalability required in the 5G-capable packet core is virtualization.

Firstly, the current 4G packet-core components are re-named and re-architected in 5G and will evolve as the 5G protocols are refined. Operators will need to be able to deploy these technologies rapidly, and the monolithic, for-purpose packet core components as deployed in 4G and before will not be appropriate. Orchestration and automation of deployment will also mean that the 5G core will be able to react to traffic utilization by scaling up/down components. As new use-cases emerge, the 5G core will be able to support these traffic types as operators will not have to install new equipment as was the case for 2, 3 and 4G.

Secondly, solving the latency issue for critical applications will require placement of services close to the user; having a virtualized environment able to host applications and software of use to the end-user will be key in 5G.

Third, 5G will be able to run network slices – independent instances of the core network for specific purposes (e.g. a closed-user group, specific traffic types or security requirements). Operators will be able to build a sliced network able to split by user and/or traffic type such that the best experience is given to each use-case rather than a one-size-fits-all as is the case today.

Being able to provide flexible virtualization features located at all levels of the network but with a single operational view and coupled with automation features which understand traffic flows, network constraints, etc and take autonomous action to maintain service against defined SLAs is key to delivering a virtualized 5G core network.

6.2.2 Packet core – edge services

Just as with the wired Internet where the proliferation of CDNs close to the subscribers has improved video delivery, the wireless Internet requires that services are placed close to the user where optimal data experience is required.

Delays of over 50ms are not uncommon on today's 4G LTE EPC, unacceptably high for the control loop timings of IoT applications or some video such as AR/VR.

Multi-access Edge Computing (Dahmen-Lhuissier, S., 2017) considers how content and applications may be moved to the edge of the network, often at the cell site or pre-aggregation site in a mobile network, in order to provide lower latencies than would be possible where services are centrally located.

MEC instances can also be placed into network slices whereby a user community accessing the same content source or application can be re-directed to use a local version located on compute close to the edge.

Likewise, elements of the packet-core, currently a fixed architecture, can also be relocated or co-located on edge compute functions removing latency from the session.

The key to a successful MEC solution is that the orchestration and operation should be transparent to not only the end user who need make no change to their application, but also to the operator; the network orchestration will build and deploy appropriate services to edge devices.

MEC may be deployed on existing 4G LTE networks as well as being a core component of 5G.

Common to both virtualization and slicing of the packet core and the ability to deploy edge services is the need for a universal compute solution available at every point on the network where appropriate functionality may be deployed.

The benefit of having such a platform is that new, hitherto undefined applications may also be hosted on this platform; experience of the rapid growth of the Internet has shown us that trying to second-guess what applications and services will appear in the future is futile and that having the ability to be able to adapt to whatever the market demands is key.

An example of a possible future hosted application is the move towards centralized (or partially centralized) baseband processing for RF known as C-RAN. With the developments in RRH technology, baseband processing is becoming less of a cell-site feature and more a compute function. As 5G will see a proliferation of cell sites, baseband processing could become a very important distributed compute requirement.

6.3 MEC

As stated before the edge is not really a place but rather the ability to move selected workloads closer to the clients and with that reduce latency as well as cutting transport cost for services when possible.

MEC in itself is not enabled by any unique technology shift and depend on a combination of SDN, Automation and equipment suitable to sit in a site far out in the aggregation network. MEC is more driven by a number of use cases that has not existed before the 5G launch with the exception of CDN (content distribution networks) like Akamai who has placed caching equipment on distributed sites for the same reasons that we are now looking at MEC in a wider perspective.

6.4 SDN

SDN is a key for many other advances like for orchestration to be efficient the network configuration must be under automation or API control as well. In the same way with a software defined transport layer with richer function sets that what has been available in the past, there is a possibility to move low touch network functions into the transport network. The creation of distributed virtual cloud environments require transparent networking between locations or private network structure to be set up per application or tenant in the network which would not be possible under the existing manual paradigms.

Advances in switch and router hardware as well as development of SDN controllers has accelerated in the past years and is expected to accelerate further in the years to come.

A consequence of bringing SDN into main stream telecom is that the organizational setup in the industry with separate network departments will have to go from running the network operations into running the equipment and handling interfaces and policies.

Technology of special interest in the SDN area is domain specific languages to set up service chains with processing in the switching infrastructure or in of load accelerators placed in banks. With the switching infrastructure fast approaching a switching fabric with function placement all the way from the L2 switch through accelerators and on to a smart NIC including a virtual switch component in the processing domain

6.5 Automation/orchestration/ reporting/analytics

This area is both the most complex due to its broad nature as well as the multitude of technologies need to make it work. Generally speaking this is the area where agile service creation is enabled and with that the area where a lot of the value in the 5G network will be created.

Virtualization in various forms including hypervisors and containers have already been brought into the telecom domain with NFV, and that in turn requires software to manage the new resource types. Here are the cloud stacks like OpenStack and VMWare represented that fans out from managing virtual resources to also manage workloads and services.

Going the other way there is technology emerging for bare metal management, BMaaS, and new initiatives in infrastructure management like RedFish in the RCA sphere that makes for higher utilization and more flexibility.

For automation to work there is a need for policy based decision systems and information gathering for the decision models. Current advances in analytics and big data fits well into the area of network analytics in 5G that will enable the whole chain from reporting giving the state of the network to automatic problem detection and solving further to predictive analysis of the networks. These functions are today deployed in other industries and fits well into the requirement picture for 5G.

Another area from the IT industry that will be fundamental for 5G is catalogue based service creation both for application and service management.

Finally automated security analysis and enforcement is becoming far more important with the increased flexibility and openness of the 5G networks. These technologies are being used elsewhere in enterprise IT but have not been seen as needed for the closed and static operations of telecom.

The way applications are built in other industries with µServices and agile deployment will drive the use of software frameworks like Cloud Foundry even though they are not yet ready to take on the full scope of telecom software. In the extension it will also lead to some parts of the telecom software stack is being deployed as a service from the cloud utilizing technology similar to Amazon AWS and Salesforce.

7. Dell EMC's position in 5G

It is increasingly obvious that the cellular telecommunications industry is at a point where the future 5G network will require levels of flexibility not currently possible from monolithic solutions. Orchestration across multiple domains and the ability to consolidate and re-architect today's services to react quickly to changing network conditions and for network capabilities to adapt to service requirements, while enabling rapid deployment of new functionality is the panacea that will enable operators to seek out new revenue streams.

New actors from the non-traditional Telco space continue to appear and expect to be able to deploy their services into the 5G environment in the same manner in which they deploy into the cloud.

New customer relationships will be formed where neither the traditional NEP nor the MNO holds the contract with the end-user. This will require more openness in terms of access to the infrastructure than has been the case up to now.

Operators will no longer have the ability to run long-term test and validation cycles given the increased set of services and the myriad of combinations of solution. Instead, operators will expect the infrastructure components to be tested as a set of Reference Architectures and with the ability for the components to monitor and correct for deviation from those references.

OSS/BSS must be simplified, work across multiple component providers and have a degree of autonomy not currently found. Additionally, it must understand and react not just on infrastructure KPIs but on end-user experiences and SLAs.

As the industry moves to Software Defined infrastructure in both RAN and network core, sensible placement of functionality between software and hardware components based on cost per component rather than software vs hardware for pure ideological reasons will prevail; in some instances, it will be more efficient to use components such as FPGAs for data processing – for example, in high-speed packet forwarding, encryption off-load and C-RAN processing – but in others, a software-based solution may be more advantageous. The key is open access to each of the components and orchestration across them such that the infrastructure makes the choice based on overall efficiency, not by the component chosen by a specific manufacturer.

In short, tomorrow's 5G infrastructure requirements with its emphasis on multi-actor, distributed, work-flow-based deployment in an efficient and agile manner looks a lot like today's large-scale Enterprise IT solutions and DellTechnologies is therefore uniquely positioned to provide solutions at SP-scale.

While elements of the 3GPP 5G solution are still under discussion, the 5G Systems Architecture 23.501 is defined and it is possible to map the major components to Dell infrastructure.

It is also interesting to note that a distributed bus-based approach to the common core functions has been taken with 5G – this again is in common with distributed processing and message-bus techniques prevalent in Enterprise IT environments. Each of these has an interface defined by the prefix "N" (e.g. Nnnsf, Nausf) described as a "services-based interface" more akin to a data model than a traditional protocol.

Each of these elements exists per-slice, under control of the orchestration. There therefore has to be a compute environment provided capable of spinning up/down each of the functional elements as required.

Mapping each of these components to their resource requirement makes for interesting reading.

Each of these elements exists per-slice, under control of the orchestration. There therefore has to be a compute environment provided capable of spinning up/down each of the functional elements as required.

An initial judgement can also be taken on the requirements for these components and summarized in Table 1:

	Compute Intensive	Database System	Message Bus	Data Forwarding	Latency Critical
NSSF		✓	✓		
NEF			✓		
NRF			✓		
PCF			✓		
UDM	✓	✓	✓		
AF					
AUSF	✓	✓	✓		
SMF	✓	✓	✓		✓
AMF		✓	✓		
UPF	✓			✓	✓
(R)AN	✓			✓	✓

Figure 9: Mobile Blueprint and the Dell EMC Opportunity

Compute intensive functions, especially in RAN and UPF, lend themselves very well to hardware off-load techniques such as SmartNIC and/or FPGA – developments are already being seen in these areas.

Orchestration and Service Assurance techniques found in Enterprise IT will also become more important – management of multiple networks is currently challenging for operators with items in their own domain. Managing multiple instances across domains and locations will be even more onerous unless work-flow methods are used to control elements. Dell's experience in service assurance in large-scale Enterprise IT is highly relevant.

The increased network and service agility will require tools like Service assurance suite and Real time analytics as well as orchestration tools like the VmWare stack.

8. Conclusions

As an infrastructure supplier Dell EMC has a limited role in driving and defining the functions and services that makes up 5G, but has a significant role and opportunity in driving the evolution of the execution environments and the implementations of the functions and services. For this to be possible and to have credibility in the telecom industry a high level of expertise is needed around the telecom systems of previous generation and the emerging 5G systems. This should be combined with identified assets like the Dell EMC strong position in enterprise systems and modifications to the product portfolio, both hardware and software, to cater for the needs emerging around 5G.

5G represents a huge wave of investments starting 2019 and peaking around 2022-2025 for building a new communications structure on top of and to some degree replacing the current 2G/3G/4G systems. This combined with the rapid grown in subscribers, data volumes and services makes the investments needed even larger. A majority of the network functions in the 5G timeframe is made up of PROCESSING, STORAGE and NETWORKING and can be translated into Dell EMC products and a clearly addressable market.

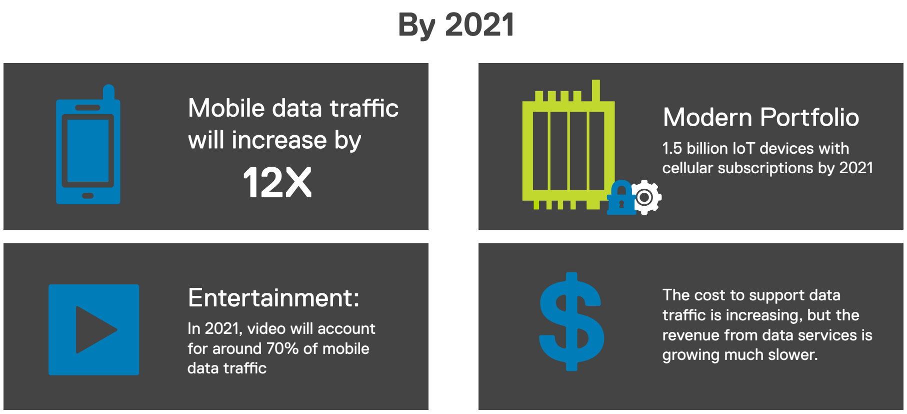

Figure 10: The CSP Challenge of Diminishing Returns

Software to address the OPEX saving and Service Creation needs of the telecom industry is a critical part of an infrastructure portfolio in the 5G timeframe. Software is of the types ORCHESTRATION, MANAGEMENT, REPORTING and ANALYTICS. In selecting software to develop or source from 3PP it is vital to look both at up to what level it makes sense to invest (infrastructure vs. application management), and if a segment is needed to be internally addressed for the company to be perceived as a primary partner to the Operator or if it is just a component needed in the infrastructure.

For the hardware infrastructure components there are gaps in the portfolio especially when addressing MEC and C-RAN that is deployed in sites with telecom specific requirements that needs to be fulfilled. Shallow depth and limited power/cooling still combined with ECC as server components mandatory. Well-developed remote management capabilities will also be a requirement due to the large amount of remote sites with limited access

When addressing MEC it should be noted that the EDGE is not a place and that there will be no specific MEC products deployed. Instead MEC is the moving of workloads to the edge of the systems close to the user for reduced latency and cost of transport. This view again drives the point that apart from hardware that can sit in an edge location the majority of MEC is around distributed processing environments and management placing workloads in a desired place.

For C-RAN the requirements are similar to MEC, but also add fulfilment and assurance mechanisms for near real-time processing and extreme low latency towards the radio heads. Functionally, this can be done in most cases with existing Dell EMC infrastructure components but an additional layer of API's and software is needed besides the building form factors suitable for centralized radio sites.

To tie together the distributed execution environment there will be a need for SDN/SD-WAN solutions connected to high touch service capable L2 switch elements. The open networking solutions and the continuation of those with for example Barefoot based technology can be a good fit, but as it is telecom and now becomes distributed auditing and security overlays are needed as well as API layers for consumption of the network resources from the telecom network functions.

At the other end of the spectrum there is a clear place for centralized cloud in parts of the telecom network which cannot be fulfilled by the standard IT cloud lacking geographic restriction of data and security mechanisms needed by telecom.

The trend for telecom under NFV has been to do everything in standard X86 processing components and while this is still true for a majority of functions, there are parts in the gateways and transport that are better done in service capable switches and FPGA/GPU components. By blending X86 with accelerated functions under a software control layer for easy consumption the COTS infrastructure can be positioned for a higher value.

IOT as a central part of the 5G transformation is an opportunity but it has to be recognized that only a subpart of the IOT revenue will be in the cellular networks and with the telecom operators. There will be aggregation gateways connected to both cellular and fixed networking as well as devices directly connected with cellular and other direct radio access forms like LORA and WIFI. In this complex landscape the value creation will be done in the service creation software bundling the access with security, functionality and backend services like storage and big data.

Summary is that strong partnerships with selected system software providers bringing Dell EMC components into solutions ranging from hardware to complex software is the key to harness 5G.

9. Call To Action

Dell EMC involvement in 5G is already in-process, with teams from SP Solutions, PowerEdge Advanced Engineering, and Strategy and Operations already involved in various phases of theoretical and applied engineering, industry thought leadership, and customer and partner trials.

In September 2017, Dell EMC provided its first outbound messaging on 5G through a combination of a SDxCentral Webinar and 5G blog (co-authored with Intel). Dell EMC has follow-on demonstrations planned for SDN/NFV World Congress (with partners Wind River and AltioStar), and

At present, Dell EMC is engaged with a number of CoSP, specifically:

- AT&T: Dell EMC and AT&T have been collaborating on introducing container services to the network edge for Control and User Plane separation (CUPS) and a foundation for Cloud Radio Access Networks (C-RAN)
- Verizon: Dell EMC and Verizon continue to evolve their virtualized network infrastructure towards the edge for the deployment of IoT services
- Sprint: Sprint has been developing Clean CUPS Core for Packet Optimization (C3PO) on Dell EMC infrastructure, and has announced their reference solution on Dell EMC DSS9000
- Telstra: Dell EMC and Telstra have engaged on transformation of their central office infrastructure to support virtualized network functions at the network edge
- Korea Telecom: Dell EMC is providing infrastructure for their initial pilots of 5G networks through Nokia, planned for launch at the 2018 Olympics
- Telefonica: Dell EMC is in discussions with Telefonica for the inclusion of Dell EMC technologies in 5TONIC, a laboratory of excellence focusing on 5G technologies

For Dell EMC to continue to grow and be successful in winning the 5G transition, the following **Call to Action** is recommended:

1. Define and consolidate a single 5G strategy across Dell EMC
2. Continue to engage with customers and partners in collaborating on 5G use-cases, leveraging deep technical industry expertise to drive engagements
3. Increase external messaging related to Dell EMC's role in 5G, with focus on industry-in versus product-out, including working with customers and partners to take joint innovations forward through outbound social media, leveraging Dell EMC Social Sphere.

A. References

3GPP. (2017). www.3gpp.org - /ftp/specs/archive/23_series/23.501/. [online] Available at: http://www.3gpp.org/ftp/specs/archive/23_series/23.501

Aijaz, A., Dohler, M., Aghvami, A., Friderikos, V. and Frodigh, M. (2017). Realizing the Tactile Internet: Haptic Communications over Next Generation 5G Cellular Networks. IEEE Wireless Communications, 24(2), pp.82-89.

Assessing the case for in-country mobile consolidation. (2015). [ebook] London: GSMA. Available at: http://www.gsma.com/publicpolicy/wp-content/uploads/2015/02/Assessing-the-case-for-in-country-mobile-consolidation-report.pdf

Cisco. (2017). Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2016–2021 White Paper. [online] Cisco. Available at: https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html

CNBC. (2017). JP Morgan sees US telecom sector consolidation, T-Mobile deal. [online] Available at: http://www.cnbc.com/2017/01/23/jp-morgan-sees-us-telecom-sector-consolidation-t-mobile-deal.html

Dahmen-Lhuissier, S. (2017). Multi-access Edge Computing. [online] ETSI. Available at: http://www.etsi.org/technologies-clusters/technologies/multi-access-edge-computing

Statista. (2017). Mobile ARPU by country 2015 | Statista. [online] Statista. Available at: https://www.statista.com/statistics/203642/forecast-for-the-global-average-revenue-per-mobile-user-in-2015-by-region/

Ericsson. (2017). MBNL network consolidation project completed by Ericsson. [online] Available at: https://www.ericsson.com/en/press-releases/2010/11/mbnl-network-consolidation-project-completed-by-ericsson

ETSI. (2017). Harmonised Standards for IMT-2000. [online] Available at: http://www.etsi.org/technologies-clusters/technologies/mobile/imt-2000

Given, J. (2003). Turning off the Television: Broadcasting's Uncertain Future. University of New South Wales Press (UNSW).

IAB. (2017). IAB Statement on Internet Confidentiality | Internet Architecture Board. [online] Available at: https://www.iab.org/2014/11/14/iab-statement-on-internet-confidentiality/

IEEE Spectrum: Technology, Engineering, and Science News. (2017). Popular Internet of Things Forecast of 50 Billion Devices by 2020 Is Outdated. [online] Available at: http://spectrum.ieee.org/tech-talk/telecom/internet/popular-internet-of-things-forecast-of-50-billion-devices-by-2020-is-outdated

GSMA Intelligence, G. (2017). GSMA Intelligence — Research — Global Mobile Trends. [online] Gsmaintelligence.com. Available at: https://www.gsmaintelligence.com/research/?file=357f1541c77358e61787fac35259dc92&download.p34

ITU. (2017a). ITU towards "IMT for 2020 and beyond". [online] Available at: http://www.itu.int/en/ITU-R/study-groups/rsg5/rwp5d/imt-2020/Pages/default.aspx

ITU. (2017b). Focus Group on IMT-2020. [online] Available at: http://www.itu.int/en/ITU-T/focusgroups/imt-2020/Pages/default.aspx

Kantar. (2017). Double Digit Smartphone Market Growth is over - Global site - Kantar Worldpanel. [online] Available at: https://www.kantarworldpanel.com/global/News/Double-Digit-Smartphone-Market-Growth-is-over

Mobile Future. (2016). The Rise of Mobile: 11.6 Billion Mobile-Connected Devices By 2020. [online] Mobilefuture.org. Available at: http://mobilefuture.org/the-rise-of-mobile-11-6-billion-mobile-connected-devices-by-2020/

DELL EMC

OSM. (2017). OSM. [online] Available at: http://osm.etsi.org

Qualcomm. (2014). The Evolution of Mobile Technologies. [online] Qualcomm. Available at: http://www.qualcomm.com/media/documents/files/the-evolution-of-mobile-technologies-1g-to-2g-to-3g-to-4g-lte.pdf

Qualcomm. (2017a). Making 5G NR a reality | Qualcomm. [online] Qualcomm. Available at: https://www.qualcomm.com/documents/making-5g-nr-reality

Qualcomm. (2017b). VR and AR are pushing the limits of connectivity, but 5G is coming to our rescue | Qualcomm. [online] Qualcomm. Available at: https://www.qualcomm.com/news/onq/2017/02/01/vr-and-ar-are-pushing-limits-connectivity-5g-our-rescue

Storno. (2017). UK Public Mobile Radiophone Service. [online] Available at: http://www.storno.co.uk/radiophone.htm

TelecomTV. (2017). ITU agrees on key 5G performance requirements for IMT-2020. [online] Available at: http://www.telecomtv.com/articles/5g/itu-agrees-on-key-5g-performance-requirements-for-imt-2020-14401/

Titcomb, J. (2017). Mobile web usage overtakes desktop for first time. [online] The Telegraph. Available at: http://www.telegraph.co.uk/technology/2016/11/01/mobile-web-usage-overtakes-desktop-for-first-time/

WT Docket No 11-65: Staff Analysis and Findings. (2011). [ebook] Washington DC: FCC. Available at: http://hraunfoss.fcc.gov/edocs_public/attachmatch/DA-11-1955A2.pdf

B. Abbreviations

3GPP	Third Generation Partnership
AMPS	Advanced Mobile Phone System
AR	Augmented Reality
ARPU	Average Revenue Per User
BBU	Baseband Unit
CDMA	Code Division Multiple Access
CP-OFDM	Cyclic Prefix Orthogonal Frequency Division Multiplexing
COTS	Common Off The Shelf
C-RAN	Cloud Radio Access Network
DAS	Distributed Antenna System
eMBB	Enhanced Mobile Broadband
EPC	Evolved Packet Core
ETSI	European Telecommunications Standards Institute
EVDO	Evolution Data Optimized
FPGA	Field-Programmable Gate Array
GPRS	General Packet Radio Service
GPU	Graphics Processing Unit
GSM	Global System for Mobile Communications
IM	Instant Messaging

IoT	Internet of Things
ITU	International Telecommunications Union
LI	Lawful Intercept
LTE	Long Term Evolution
M2M	Machine to Machine
MANO	Management and Orchestration
MEC	Multi-Access Edge
mMTC	Massive Machine Type Communication
MNO	Mobile Network Operator
NEP	Network Equipment Provider
NMT	Nordic Mobile Telephone
OEM	Original Equipment Manufacturer
OFDM	Orthogonal Frequency Division Multiplexing
OTT	Over the Top
PSTN	Public Switched Telephone Network
RAN	Radio Access Network
RF	Radio Frequency
RRH	Remote Radio Head
SC-FDMA	Single Carrier Frequency Division Multiple Access
SC-OFDM	Single Carrier Orthogonal Frequency Division Multiplexing
SDR	Software Defined Radio
SLA	Service Level Agreement
SMS	Short Message Service
TACS	Total Access Communications System
TCP/IP	Transmission Control Protocol/Internet Protocol
TD-SCDMA	Time Division Synchronous Code Division Multiple Access
TETRA	Terrestrial Trunked Radio
UMTS	Universal Mobile Telecommunications Service
URLLC	Ultra Reliable Low Latency Communications
VoLTE	Voice over LTE
VPN	Virtual Private Network
VR	Virtual Reality
W-CDA	Wideband Code Division Multiple Access

All images courtesy of Intel.

DELL EMC