GSMA Intelligence

ANALYSIS

Understanding 5G: Perspectives on future technological advancements in mobile

December 2014

Contents

Executive summary

5G offers enormous potential for both consumers and industry

As well as the prospect of being considerably faster than existing technologies, 5G holds the promise of applications with high social and economic value, leading to a 'hyper-connected society' in which mobile will play an ever more important role in people's lives.

The GSMA will work for its members and with its partners to shape 5G

As the association representing the mobile industry, the GSMA will play a significant role in shaping the strategic, commercial and regulatory development of the 5G ecosystem. This will include areas such as the definition of roaming and interconnect in 5G, and the identification and alignment of suitable spectrum bands. Once a stable definition of 5G is reached, the GSMA will work with its members to identify and develop commercially viable 5G applications. This paper focuses on 5G as it has developed so far, and the areas of technological innovation needed to deliver the 5G vision.

There are currently two definitions of 5G

Discussion around 5G falls broadly into two schools of thought: a service-led view which sees 5G as a consolidation of 2G, 3G, 4G, Wi-fi and other innovations providing far greater coverage and always-on reliability; and a second view driven by a step change in data speed and order of magnitude reduction in end-to-end latency. However, these definitions are often discussed together, resulting in sometimes contradictory requirements.

Sub-1ms latency and >1 Gbps bandwidth require a true generational shift

Some of the requirements identified for 5G can be enabled by 4G or other networks. The technical requirements that necessitate a true generational shift are sub-1ms latency and >1 Gbps downlink speed, and only services that demand at least one of these would be considered 5G use cases under both definitions.

Achieving sub-1ms latency is a hugely exciting challenge that will define 5G

Delivering 1ms latency over a large scale network will be challenging, and we may see this condition relaxed. If this were to happen, some of the potential 5G services identified may no longer be possible and the second view of 5G would become less clear. This paper looks at some of the challenges that must be overcome to deliver 1ms latency.

At the same time 4G will continue to grow and evolve

Technologies such as NFV/SDN and HetNets are already being deployed by operators and will continue to enable the move towards the hyper-connected society alongside developments in 5G. Considerable potential also remains for increasing 4G adoption in many countries, and we expect 4G network infrastructure to account for much of the $1.7 trillion the world's mobile operators will invest between now and 2020. Operators will continue to focus on generating a return on investment from their 4G (and 3G) networks by developing new services and tariffing models that make most efficient use of them.

Introduction

Objectives of this report

The purpose of this report is to take a step toward clarifying what '5G' really means in the technological sense, by: reducing 5G to its fundamental core (including acknowledging what it is arguably not); expanding on some of the use case scenarios that 5G might enable; and discussing conceivable implications for operators in terms of network infrastructure and commercial opportunities. This can only be achieved by framing the discussion around 5G in a broader context alongside existing network technologies and those currently in development.

In summary, there are three key questions that this report will ask:

1. What is (and what isn't) 5G?
2. What are the real 5G use cases?
3. What are the implications of 5G for mobile operators?

Notes on terminology

GSMA Intelligence's definition of 4G includes the following network technologies: LTE, TD-LTE, AXGP, WiMAX, LTE-A, TD-LTE-A, LTE with VoLTE and WiMAX 2.

Due to the commonality of operator definitions classifying LTE and TD-LTE as 4G technologies, we follow this convention. This differs from the ITU's strict definition of transitional versus true 4G. Also, where we use the term 'LTE' in this document it incorporates all LTE variants (LTE, TD-LTE, AXGP, LTE-A and TD-LTE-A). Finally, for simplicity we do not consider WiMAX in this analysis, so where the term '4G' is used it incorporates all LTE variants but not WiMAX (a transitional 4G technology) or WiMAX 2 (a true 4G technology). Therefore for the purpose of this report the terms '4G' and 'LTE' are interchangeable.

What is 5G?

Evolution beyond mobile internet

From analogue through to LTE, each generation of mobile technology has been motivated by the need to meet a requirement identified between that technology and its predecessor (see Table 1). For example, the transition from 2G to 3G was expected to enable mobile internet on consumer devices, but whilst it did add data connectivity, it was not until 3.5G that a giant leap in terms of consumer experience occurred, as the combination of mobile broadband networks and smartphones brought about a significantly enhanced mobile internet experience which has eventually led to the app-centric interface we see today. From email and social media through music and video streaming to controlling your home appliances from anywhere in the world, mobile broadband has brought enormous benefits and has fundamentally changed the lives of many people through services provided both by operators and third party players.

Generation	Primary services	Key differentiator	Weakness (addressed by subsequent generation)
1G	Analogue phone calls	Mobility	Poor spectral efficiency, major security issues
2G	Digital phone calls and messaging	Secure, mass adoption	Limited data rates – difficult to support demand for internet/e-mail
3G	Phone calls, messaging, data	Better internet experience	Real performance failed to match hype, failure of WAP for internet access
3.5G	Phone calls, messaging, broadband data	Broadband internet, applications	Tied to legacy, mobile specific architecture and protocols
4G	All-IP services (including voice, messaging)	Faster broadband internet, lower latency	?

Table 1: Evolution of technology generations in terms of services and performance
Source: GSMA Intelligence

More recently, the transition from 3.5G to 4G services has offered users access to considerably faster data speeds and lower latency rates, and therefore the way that people access and use the internet on mobile devices continues to change dramatically. Across the world operators are typically reporting that 4G customers consume around double the monthly amount of data of non-4G users, and in some cases three times as much. An increased level of video streaming by customers on 4G networks is often cited by operators as a major contributing factor to this.

The Internet of Things (IoT) has also been discussed as a key differentiator for 4G, but in reality the challenge of providing low power, low frequency networks to meet the demand for widespread M2M deployment is not specific to 4G or indeed 5G. As Table 1 suggests, it is currently unclear what the opportunity or ‘weakness’ that 5G should address is.

Two views of 5G exist today:

View 1 – The hyper-connected vision: In this view of 5G, mobile operators would create a blend of pre-existing technologies covering 2G, 3G, 4G, Wi-fi and others to allow higher coverage and availability, and higher network density in terms of cells and devices, with the key differentiator being greater connectivity as an enabler for Machine-to-Machine (M2M) services and the Internet of Things (IoT). This vision may include a new radio technology to enable low power, low throughput field devices with long duty cycles of ten years or more.

View 2 – Next-generation radio access technology: This is more of the traditional 'generation-defining' view, with specific targets for data rates and latency being identified, such that new radio interfaces can be assessed against such criteria. This in turn makes for a clear demarcation between a technology that meets the criteria for 5G, and another which does not.

Both of these approaches are important for the progression of the industry, but they are distinct sets of requirements associated with specific new services. However, the two views described are regularly taken as a single set and hence requirements from both the hyper-connected view and the next-generation radio access technology view are grouped together. This problem is compounded when additional requirements are also included that are broader and independent of technology generation.

5G technology requirements

As a result of this blending of requirements, many of the industry initiatives that have progressed with work on 5G (see Appendix A) identify a set of eight requirements:

- 1-10Gbps connections to end points in the field (i.e. not theoretical maximum)
- 1 millisecond end-to-end round trip delay (latency)
- 1000x bandwidth per unit area
- 10-100x number of connected devices
- (Perception of) 99.999% availability
- (Perception of) 100% coverage
- 90% reduction in network energy usage
- Up to ten year battery life for low power, machine-type devices

Because these requirements are specified from different perspectives, they do not make an entirely coherent list – it is difficult to conceive of a new technology that could meet all of these conditions simultaneously.

Equally, whilst these eight requirements are often presented as a single list, no use case, service or application has been identifed that requires all eight performance attributes across an entire network simultaneously. Indeed some of the requirements are not linked to use cases or services, but are instead aspirational statements of how networks should be built, independent of service or technology – no use case needs a network to be significantly cheaper, but every operator would like to pay less to build and run their network. It is more likely that various combinations of a subset of the overall list of requirements will be supported 'when and where it matters'.

Finally, while important in their own right, six of these requirements are not generation-defining attributes. These are considered below:

Perceived 99.999% availability and 100% geographical coverage:

These are not use case drivers, nor technical issues, but economic and business case decisions. 99.999% availability and 100% coverage are achievable using any existing technology, and could be achieved by any network operator. Operators decide where to place cells based on the cost to prepare the site to establish a cell to cover a specific area balanced against the benefit of the cell providing coverage for a specific geographic area. This in turn makes certain cell sites and coverage areas - such as rural areas and indoor coverage - the subject of difficult business decisions.

Whilst a new generation of mobile network technology may shift the values that go in to the business model that determines cell viability, achieving 100% coverage and 99.999% availability will remain a business decision rather than a technical objective. Conversely, if 100% coverage and 99.999% availability were to be a 5G 'qualifying criteria', no network would achieve 5G status until such time as 100% coverage and 99.999% availability were achieved.

Connection density (1000x bandwidth per unit area, 10-100x number of connections):

These essentially amount to 'cumulative' requirements i.e. requirements to be met by networks that include 5G as an incremental technology, but also require continued support of pre-existing generations of network technology. The support of 10-100 times the number of connections is dependent upon a range of technologies working together, including 2G, 3G, 4G, Wi-fi, Bluetooth and other complementary technologies. The addition of 5G on top of this ecosystem should not be seen as an end solution, but just one additional piece of a wider evolution to enable connectivity of machines. The Internet of Things (IoT) has already begun to gain significant momentum, independent of the arrival of 5G.

Similarly, the requirement for 1,000 times bandwidth per unit area is not dependent upon 5G, but is the cumulative effect of more devices connecting with higher bandwidths for longer durations. Whilst a 5G network may well add a new impetus to progression in this area, the rollout of LTE is already having a transformational effect on the amount of bandwidth being consumed within any specific area, and this will increase over the period until the advent of 5G. The expansion of Wi-fi and integration of Wi-fi networks with cellular will also be key in supporting greater data density rates.

Meeting both of these requirements will have significant implications for OPEX on backhaul and power, since each cell or hotspot must be powered and all of the additional traffic being generated must be backhauled.

Reduction in network energy usage and improving battery life:

The reduction of power consumption by networks and devices is fundamentally important to the economic and ecological sustainability of the industry. A general industry principle for minimising power usage in network and terminal equipment should pervade all generations of technology, and is recognised as an ecological goal as well as having a significant positive impact on the OPEX associated with running a network. At present it is not clear how a new generation of technology with higher bandwidths being deployed as an overlay (rather than a replacement) on top of all pre-existing network equipment could result in a net reduction in power consumption.

Some use cases for M2M require the connected device in the field to lie dormant for extended periods of time. It is important that innovation in how these devices are powered and the leanness of the signaling they use when becoming active and connected is pursued. However, this requirement is juxtaposed with 5G headline requirements on data rate – what is required for mass sensor networks is very occasional connectivity with minimal throughput and signaling load. Work to develop such technology predates the current 5G requirements and is already being pursued in Standards bodies.

These six requirements should be and are being pursued by the industry today using a range of techniques (some of which are covered later in the paper) but these amount to evolutions of existing network technology and topology or opportunities enabled by changing hardware characteristics and capabilities. These will in turn open business opportunities for operators and third parties. However, none of these business opportunities exist today – they are constrained by limitations greatly governed by economics, and much of these six requirements are motivated by improving the economic viability of those opportunities, rather than filling technological gaps that explicitly prohibit these opportunities, regardless of the amount they might cost to enable.

Thus in the strictest terms of measurable network deliverables which could enable revolutionary new use case scenarios, the potential attributes that would be unique to 5G are limited to **sub-1ms latency** and **>1 Gbps downlink speed**.

Potential 5G use cases

Imagining the mobile services of the next decade

As with each preceding generation, the rate of adoption of 5G and the ability of operators to monetise it will be a direct function of the new and unique use cases it unlocks. Thus the key questions around 5G for operators are essentially:

a. What could users do on a network which meets the 5G requirements listed above that is not currently possible on an already existing network?
b. How could these potential services be profitable?

Figure 1 illustrates the latency and bandwidth/data rate requirements of the various use cases which have been discussed in the context of 5G to date. These potential 5G use cases and their associated network requirements are described below.

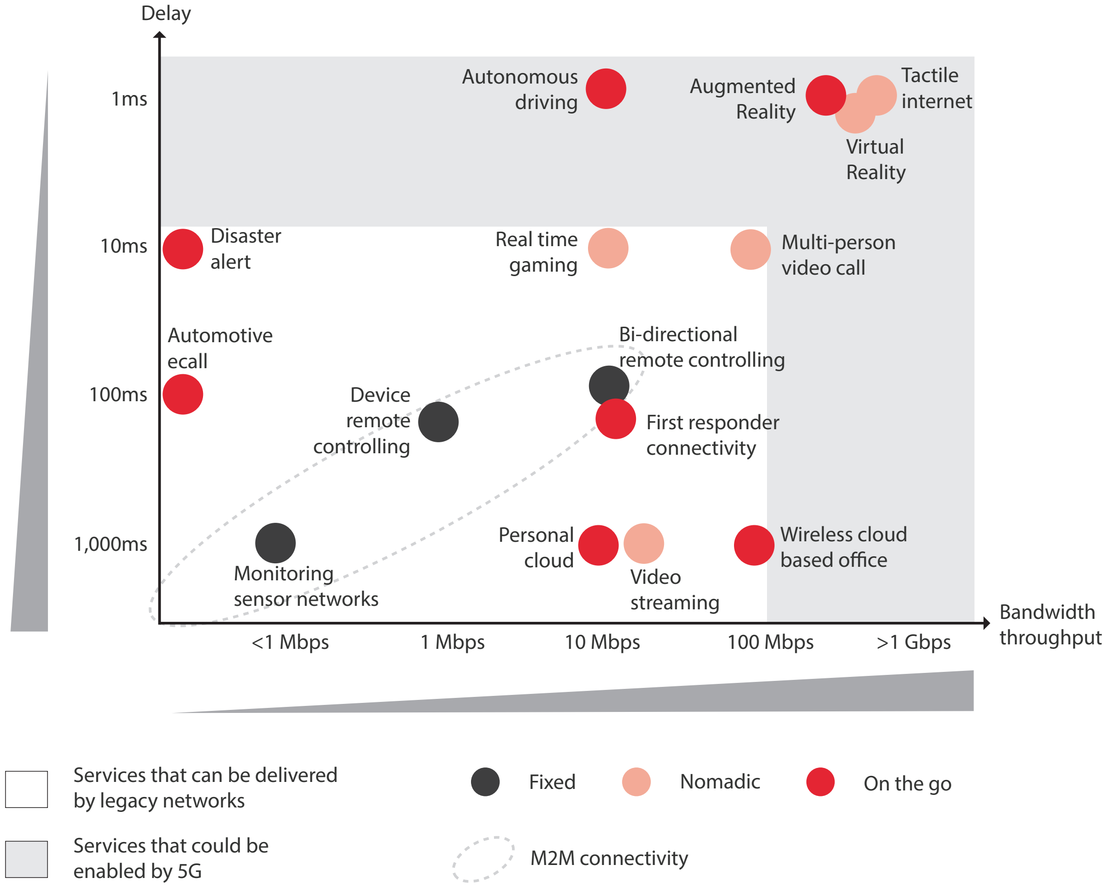

Figure 1: Bandwidth and latency requirements of potential 5G use cases

Source: GSMA Intelligence

Virtual Reality/Augmented Reality/Immersive or Tactile Internet

These technologies have a number of potential use cases in both entertainment (e.g. gaming) and also more practical scenarios such as manufacturing or medicine, and could extend to many wearable technologies. For example, an operation could be performed by a robot that is remotely controlled by a surgeon on the other side of the world. This type of application would require both high bandwidth and low latency beyond the capabilities of LTE, and therefore has the potential to be a key business model for 5G networks.

However, it should be pointed out that VR/AR systems are very much in their infancy and their development will be largely dependent on advances in a host of other technologies such as motion sensors and heads up display (HUD). It remains to be seen whether these applications could become profitable businesses for operators in the future.

Autonomous driving/Connected cars

Enabling vehicles to communicate with the outside world could result in considerably more efficient and safer use of existing road infrastructure. If all of the vehicles on a road were connected to a network incorporating a traffic management system, they could potentially travel at much higher speeds and within greater proximity of each other without risk of accident - with fully-autonomous cars further reducing the potential for human error.

While such systems would not require high bandwidth, providing data with a command-response time close to zero would be crucial for their safe operation, and thus such applications clearly require the 1 millisecond delay time provided in the 5G specification. In addition a fully 'driverless' car would need to be driverless in all geographies, and hence would require full road network coverage with 100% reliability to be a viable proposition.

Wireless cloud-based office/Multi-person videoconferencing

High bandwidth data networks have the potential to make the concept of a wireless cloud office a reality, with vast amounts of data storage capacity sufficient to make such systems ubiquitous. However, these applications are already in existence and their requirements are being met by existing 4G networks. While demand for cloud services will only increase, as now they will not require particularly low latencies and therefore can continue to be provided by current technologies or those already in development. While multi-person video calling - another potential business application - has a requirement for lower latency, this can likely be met by existing 4G technology.

Machine-to-machine connectivity (M2M)

M2M is already used in a vast range of applications but the possibilities for its usage are almost endless, and our forecasts predict that the number of cellular M2M connections worldwide will grow from 250 million this year to between 1 billion and 2 billion by 2020, dependent on the extent to which the industry and its regulators are able to establish the necessary frameworks to fully take advantage of the cellular M2M opportunity.

Typical M2M applications can be found in 'connected home' systems (e.g. smart meters, smart thermostats, smoke detectors), vehicle telemetric systems (a field which overlaps with Connected cars above), consumer electronics and healthcare monitoring. Yet the vast majority of M2M systems transmit very low levels of data and the data transmitted is seldom time-critical. Many currently operate on 2G networks or can be integrated with the IP Multimedia Subsystem (IMS) – so at present the business case for M2M that can be attached to 5G is not immediately obvious.

A true requirement for a generational shift?

Thus many of the services that have been put forward as potential 'killer apps' for 5G do not require a generational shift in technology, and could be provided via existing network technologies. Only applications that require at least one of the key 5G technical requirements – sub-1ms latency and >1 Gbps downlink speed – can be considered true next generational business cases.

Of these two requirements, reducing latency to sub-1ms levels may provide the greatest technical challenge (see page 12). Meanwhile, as discussed in more detail in Appendix B, operators are already making a considerable amount of progress in increasing the data speeds of their existing networks by adopting LTE-A technologies (see Figure 2). While it is important to note that although many of the use cases and services discussed in this section do not strictly require 5G, they could offer an enhanced user experience on a 5G network. However this amounts to an incremental benefit that is more difficult to market than a genuine new service, and not a core component of any 5G business case.

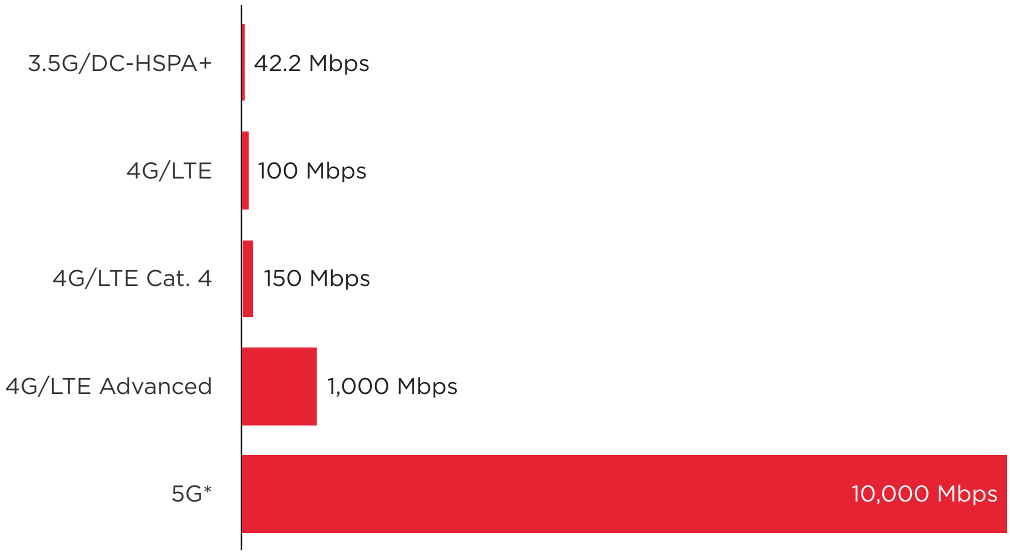

Figure 2: Maximum theoretical downlink speed by technology generation, Mbps
(*10 Gbps is the minimum theoretical upper limit speed specified for 5G)
Source: GSMA Intelligence

The implications of 5G for mobile operators

The progress from initial 3G networks to mobile broadband technology has transformed industry and society by enabling an unprecedented level of innovation. If 5G becomes a true generational shift in network technology, we can expect an even greater level of transformation. There are varying implications of providing an increased level of connectivity or developing a new radio access network (RAN) to deliver a step change in per connection performance, or a combination of the two. This means that the final design of a 5G network could be any one of a range of options with differing radio interfaces, network topologies and business capabilities.

While a shift to 5G would be hugely impactful, the industry will need to overcome a series of challenges if these benefits are to be realised, particularly in terms of spectrum and network topology.

5G spectrum and coverage implications

While there are a number of spectrum bands which could potentially be used in meeting some of the 5G requirements identified to date, there is currently a substantial focus on higher frequency radio spectrum. As discussed in Appendix A, operators, vendors and academia are combining efforts to explore technical solutions for 5G that could use frequencies above 6GHz and reportedly as high as 300 GHz. However, higher frequency bands offer smaller cell radiuses and so achieving widespread coverage using a traditional network topology model would be challenging.

It is widely accepted that 'beam-forming' - the focussing of the radio interface into a beam which will be usable over greater distances – is an important part of any radio interface definition that would use 6GHz or higher spectrum bands. This however means that the beam must be directed at the end user device that is being connected. Since the service being offered is still differentiated from fixed line connections on the basis of ‘mobility’, the beam itself will have to track the device. This is innovation that could make 5G an expensive technology to deploy on large scale, since each cell may have to support several hundred individual beams at any one time and track the end users that are connected via these beams in three dimensional space.

High-order MIMO (Multi-Input, Multi-Output) is another method for increasing bandwidth that is often discussed. This is where an array of antennae is installed in a device and multiple radio connections are established between a device and a cell. However, high-order MIMO can have issues with radio interference, so technology is required to help mitigate this problem. This tends to focus on a need for the radio network to adjust its beam to take into account the specific orientation of the antenna at any given time.

All of this is incremental research and development over and above that currently being conducted for 4G. The use of bands higher than 6GHz will likely require operators to invest in an entirely new RAN since it will have fundamentally different masthead requirements. Given the level of infrastructure required to achieve the desired network topology, operators may be forced to rethink their existing business models. New technology is rarely a cheap option, and the nature of the new technology that is required in the radio network makes it very power-intensive, hence counter to the stated requirement for significant reduction in overall network power consumption.

That said, vendors are researching ways to include beam forming and MIMO technology in mobile devices. As a result, the process of identifying and aligning internationally around common bands for 5G will have a clear dependency on the technology that can be identified to overcome band usage in high frequencies for wide area coverage.

Can 1 millisecond latency be achieved?

Achieving the sub-1ms latency rate identified as a technical requirement for 5G necessitates a new way of thinking about how networks are structured, and will likely prove to be a significant undertaking in terms of technological development and investment in infrastructure.

Despite the inevitable advances in processor speeds and network latency between now and 2020, the speeds at which signals can travel through the air and light can travel along a fibre are governed by fundamental laws of physics. Subsequently services requiring a delay time of less than 1 millisecond must have all of their content served from a physical position very close to the user’s device. Industry estimates suggest that this distance may be less than 1 kilometre, which means that any service requiring such a low latency will have to be served using content located very close to the customer, possibly at the base of every cell, including the many small cells that are predicted to be fundamental to meeting densification requirements. This will likely require a substantial uplift in CAPEX spent on infrastructure for content distribution and servers.

If any service requiring 1 millisecond delay also has a need for interconnection between one operator and another, this interconnectivity must also occur within 1 kilometre of the customers. This could well be the case in a service such as social networking content pushed into augmented reality. Today, inter-operator interconnect points are relatively sparse, but to support a 5G service with 1 millisecond delay, there would likely need to be interconnection at every base station, thus impacting the topological structure of the core network. Roaming customers would need to have visited network contextual roaming capabilities, and have content relevant to their applications available directly from the visited network, posing challenges for the existing roaming model.

In the most extreme case, it would make sense for a single network infrastructure to be implemented, which would be utilised by all operators. This would mean all customers could be served by a single content source, with all interaction and interconnect with localised context also being served from that point at the base station. This would also imply that only one radio network would be built, and then shared by all operators.

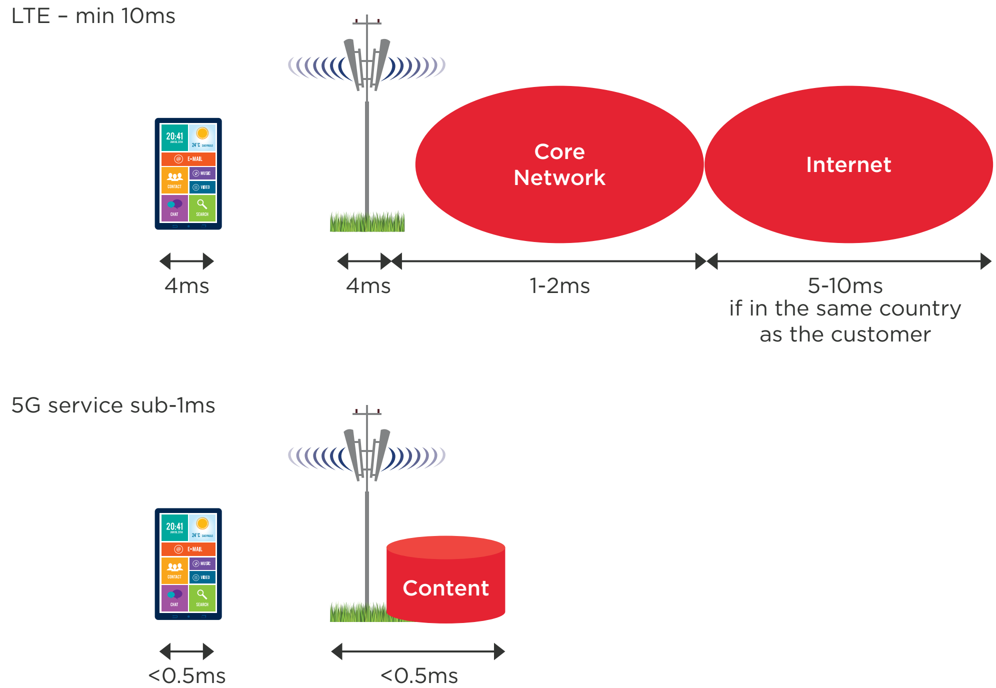

Figure 3: Latency performance for LTE compared to latency requirement for 5G
Source: GSMA

Such a model would considerably reduce CAPEX in the network build (since rather than say four operators building four parallel networks, only a single network would be built) but would require unprecedented levels of co-operation between operators. It would also impact the nature of inter-operator competition, shifting focus to services rather than data rate and coverage differentiation. It would also make spectrum auctions somewhat irrelevant, since only one radio network being built would mean there would only be one bidder and one license per market.

Once this is all realised, it is likely that requirements for sub-1ms delay will be relaxed or possibly removed entirely from 5G, rather than industry committing to the massive upheaval and resource acquisition that would be implied. If this were to happen, it may draw into question the viability of coupling services such as augmented and virtual reality, immersive internet and autonomous driving with mobility. However, if those services were removed from the expected service set, the justification for the technological view of 5G would also become questionable.

Continuing development of network technologies: what 5G isn't

To further enhance the mobile broadband experience for customers, operators are continuing to develop their 4G networks through the deployment of LTE-Advanced technologies. Many are also deploying technologies such as network function virtualisation (NFV), software defined networks (SDN), heterogeneous networks (HetNets) and low power, low throughput (LPLT) networks. These allow different network upgrade paths and expansion of coverage through integration of broader wireless technologies, as well as potentially having a positive effect in the total cost of ownership of the network.

The term 5G is sometimes used to encapsulate these technologies. However, it is important to clarify that these technological advancements are continuing independently of 5G. While these are areas that will have significant impact on the mobile industry over the coming years, explicitly including them under the term 5G has the potential to adversely affect progress in the industry between now and the realisation of 5G as a commercial service.

A summary of these technologies follows:

Network Function Virtualisation (NFV) and Software Defined Networks (SDN)

NFV is a network architecture concept that enables the separation of hardware from software or 'function', and has become a reality for the mobile industry due to the increased performance of 'common, off-the-shelf' (COTS) IT platforms. SDN is an extension of NFV wherein software can perform dynamic reconfiguration of an operator's network topology to adjust to load and demand, e.g. by directing additional network capacity to where it is needed to maintain the quality of customer experience at peak data consumption times. A number of operators have built or are building part or all of their LTE networks using NFV and SDN as the basis.

These technologies in combination can potentially reduce operator CAPEX as they offer a cheaper and simpler network architecture that is easier to upgrade, while OPEX is also reduced through power savings as network capacity is only provided when and where it is needed. However, shifting from existing structures to IT-based soft functions will bring new complexities for operators in terms of network provisioning and management, as well as requiring a new skill set within operator staff.

Heterogeneous Networks (HetNets)

HetNet refers to the provision of a cellular network through a combination of different cell types (e.g. macro, pico or femto cells) and different access technologies (i.e. 2G, 3G, 4G, Wi-fi). By integrating a number of diverse technologies depending on the topology of the coverage area, operators can potentially provide a more consistent customer experience compared to what could be achieved with a homogenous network.

Small cell deployments are a key feature of the HetNet approach as they allow considerable flexibility as to where they are positioned, however, the use of more cells brings implications in terms of power supply and backhaul, especially when they are located in remote areas. Wi-fi can also play a significant role in HetNets, both in terms of data offload and roaming.

HetNet technology has typically been developed in relation to data networking, but recently voice has been brought under the scope as well, not least because of support for Wi-fi calling being available in Apple's iPhone 6 which was released in September 2014.

Conclusions: enabling innovation through industry-wide collaboration

The many initiatives and discussions on 5G going on around the world by governments, vendors, operators and academia demonstrate the continuing ethos of collaboration and innovation across the industry. In these debates we must ensure that we continue to co-ordinate with aligned goals to maintain momentum in completing the definition of 5G.

The key 5G considerations at this stage are:

When 5G arrives will be determined by what 5G turns out to be

As discussed earlier, there are currently two differing views of what 5G is. The first view makes its implementation somewhat intangible – 5G will become a commercial reality when sufficient industry voices say so, but this will be something that is difficult to measure by any recognisable metric. The second approach is more concrete in that it has a distinct set of technical objectives, meaning that when a service is launched that meets those objectives it will count as the advent of 5G.

As the requirements identified for 5G are a combination of both visions, in some cases the requirement set is self-contradictory – for example, it would not be possible to have a new RAN with beam forming and meet a requirement for power reduction, because beam forming uses a lot more power than today's RAN. As a result, there must be an established answer to the question of what 5G is before there can be an answer to the question of when it will arrive.

The case for a new RAN should be based on its potential to improve mobile networks

The principal challenge in the 5G specification is the sub-1ms latency requirement, which is governed by fundamental laws of physics. If, as discussed above, this challenge proves too much and the requirements for sub-1ms delay are removed from 5G, the need for a new RAN would be questioned. Whether a new air interface is necessary is arguably more of a question of whether one can be invented that significantly improves mobile networks, rather than on a race to the arbitrary deadline of 2020.

This raises the question of where the industry should go next. Without a new air interface, the '5G' label makes less sense, as the industry would need to shift to the evolutionary view of 5G - with the new networks building on LTE and Wi-fi by adding new functionalities and architecture.

5G should not distract from more immediate technological developments

Technologies such as multiple-carrier LTE-A, NFV/SDN, HetNets and LPLT networks will form an important part of the evolution of mobile networks. Each has the potential to offer tangible benefits to operators within the next few years, and so the industry should not risk losing focus on the potential benefits of these technologies in the short and medium term. Also, the term '5G' should always be associated with the definition of new radio technology. Everything else is the net result of other forms of innovation.

LTE remains very important and will continue to evolve

There remains considerable potential for future LTE growth, which still only accounts for 5% of the world's mobile connections. LTE penetration as a percentage of connections is already as high as 69% in South Korea, 46% in Japan and 40% in the US, but LTE penetration in the developing world stands at just 2%. Hence there is still a substantial opportunity for operators to generate returns on their investment in LTE networks.

LTE technology will also continue to develop, with operators already making a considerable amount of progress in increasing the data speeds of their existing networks by adopting multiple-carrier LTE-A technologies. Therefore, while there remain monetisation and interconnect issues around LTE, these advancements will enable operators to offer many of the services that have been put forward in the context of 5G long before 5G becomes a commercial reality.

The industry should make full use of governmental interest and resources

As detailed in Appendix A, there is a considerable level of governmental interest worldwide in the subject of 5G, not to mention a substantial amount of funding available for research and development in the field. It is important that the industry leverages this and effectively channels the focus and resource into something meaningful for both operators and their customers. This should be implemented in a coordinated framework to avoid a fragmented vision of 5G for different parts of the world.

5G is an opportunity to develop a more sustainable operator investment model

If previous generations of mobile technology have taught us anything, it is it that, as with each preceding generation, 5G will unlock value in ways we cannot and will not anticipate. Services that were initially expected to have a negligible impact became hugely popular (e.g. SMS), while those expected to be the 'next big thing' have been slow to gain traction (e.g. video calling). Through the development of 5G, we as an industry can expect a paradigm shift in the way that all of the stakeholders in the mobile ecosystem play their role. Regulators especially can use this as an opportunity to create healthier environments that stimulate continuing investment in next generation technology.

Some of the business cases that have worked well for 3G and 4G technologies may not be the right ones for 5G. By actively conceiving and exploring 5G business cases at an earlier stage, operators will have greater potential to shape the new paradigm.

The GSMA will continue to work with its members to shape the future of 5G

Whichever form 5G eventually takes, the GSMA, as the association representing the mobile industry, looks forward to contributing to the development of a 5G ecosystem through collaboration and thought leadership. The GSMA's focus is on:

• working with its operator members to identify and develop commercially viable 5G applications
• collaborating in the work being undertaken in terms of research, development and definition of 5G technologies by industry groups such as 3GPP, NGMN and ITU-R, and contributing to the various working groups in these areas
• identifying requirements around roaming and interconnect
• driving the development of the regulatory framework for 5G by identifying suitable spectrum bands for its operation, and working with governments around the world to ensure international alignment within those bands
• creating a forum for relevant parties to discuss 5G through e.g. GSMA boards and committees, industry workshops, Mobile World Congress etc.

The successful shift to next generation networks can only be achieved through strong industry-wide collaboration. The GSMA will continue communicating through subsequent papers to influence the strategic direction of 5G development, as the business case and technical requirements for 5G become clearer. In order to realise the immense opportunity that 5G represents for the industry, the GSMA will do all it can to ensure that the next generation of telecommunications deliver innovation and consumer benefits in an economically viable way.

Appendix A: Current 5G industry activity

Since 2012, a number of initiatives have been established to define and develop 5G and there have also been a considerable number of statements from interested parties such as governments and infrastructure vendors. Having fallen behind Eastern Asia and North America in terms of mobile technological advancement due to a relatively slow rollout and adoption of 4G networks, European governments are particularly keen to get ahead of the curve in the 5G space and there have been a number of announcements from Neelie Kroes, European Commission (EC) Vice President for Digital Agenda, on the subject going back to Mobile World Congress 2013. The governments of Japan, South Korea and China have also been particularly active in driving the 5G agenda.

Meanwhile, vendors such as Ericsson, Huawei, NSN and Samsung all began research and development towards 5G in 2013, and this year mobile operators have also begun making announcements regarding their own 5G laboratory trials. A summary of the key parties, milestones and targets is below.

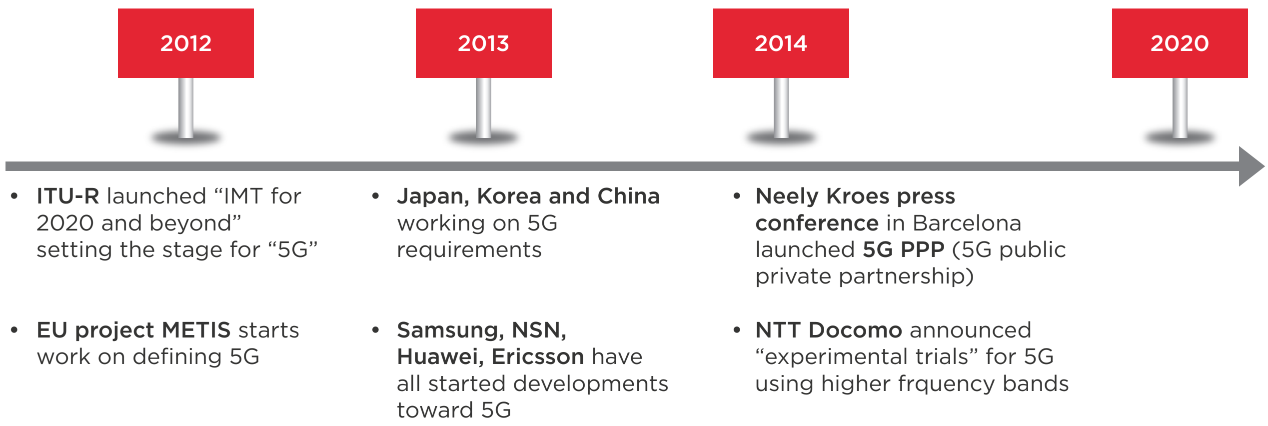

Figure 4: Timeline of key events in 5G developments

Source: GSMA Intelligence

ITU Radiocommunication Sector (ITU-R)

The ITU-R plays a key role in in the global management of the radio-frequency spectrum and satellite orbits as well as being the body that defined the criteria for previous generations of technology – the IMT-2000 family of technologies correlates directly to 3G, whilst the intent was that IMT-Advanced technologies would be 4G. However, for 4G the relationship between ITU-R IMT definitions and specific 'G's became broken. IMT-Advanced only identifies two technologies as meeting the criteria laid out by ITU-R for 4G – LTE-A and WiMAX2. Operators and equipment vendors blurred this definition by marketing LTE, WiMAX and even HSPA+ as '4G'. LTE and WiMAX are in fact included in the IMT-2000 technology group, and so, if the association between ITU-R IMT groups and 'generations' were to be maintained, LTE and WiMAX would be 3rd Generation, rather than 4th.

In early 2012, ITU-R began a programme to develop "IMT-2020" (International Mobile Telecommunications 2020), setting the stage for the 5G research activities that have since emerged across the world. In 2015, the organisation plans to finalise its "Vision" of the 5G mobile broadband connected society. This view of the horizon for the future of mobile technology will be key in setting the agenda for the World Radiocommunication Conference (WRC) 2015, where discussions regarding additional spectrum will take place in support of the future growth of the industry.

NGMN Alliance

The NGMN (Next Generation Mobile Networks) Alliance is a forum made up of 24 mobile operators and various other mobile industry ecosystem companies including network and handset vendors, and research institutes. NGMN began working on identifying requirements for 5G standards in Q4 2013 and plans to present a white paper detailing end-to-end requirements for 5G at its industry conference in March 2015. The paper is intended to support the standardisation and subsequent availability of 5G from 2020.

The NGMN Alliance has positioned itself as the lead organisation driving the 5G agenda, although is yet to make any public statement on what the requirements it defines might be.

European Commission

The EC's 5G research activities began in November 2012 with the co-funding of METIS (Mobile and wireless communications Enablers for the Twenty-twenty (2020) Information Society), a consortium of 29 partners spanning vendors, operators, the automotive industry and academia focused on the next generation of mobile and wireless communications systems for year 2020. A year later METIS published its five key 5G scenarios, 12 test cases and seven Key Performance Indicators for 5G, associated with technical requirements. The project is due to release its final report in April 2015.

In December 2013, the EC went further and announced a joint 5G research and innovation project with the private sector - The 5G Infrastructure Public Private Partnership (5G PPP) - with collective funding of €4.2 billion, of which €700 million will come from the commission itself, reflecting its desire to seize the initiative in 5G development. 5G PPP will facilitate research into solutions, architectures, technologies and standards for 5G infrastructure, and aims to ensure that at least 20% of 5G standards essential patents (SEP) are developed and owned by European organisations, while ensuring that European vendors retain at least 35% of global market share in the supply of future network infrastructure.

National governments

Outside of Europe, the majority of 5G research appears to be confined to Eastern Asia, with China, Japan and South Korea all working independently on defining 5G requirements. China's 5G initiative, named 'IMT-2020', is a combination of three government agencies and has established eight working groups with the aim of promoting the development of 5G technologies in the country.

Meanwhile, Japan's '2020 and Beyond Ad Hoc' (20B AH) group was established by the Association of Radio Industries and Businesses (ARIB) in September 2013 to study the concept, function and architecture of mobile communications systems going into the next decade, as well as the services and applications those systems could offer. The country has set an ambitious target of having commercial 5G services available in time for the 2020 Olympic Games in Tokyo.

Equally ambitious 5G targets have been set in South Korea. The country's '5G Forum' group's website states that "The 5G technology is expected to be commercialised by 2020 with 1,000-time speed of current LTE data transfer." The Korean Ministry of Education, Science and Technology has allocated $1.6 billion of funding to the project.

Individual operators and vendors

More ambitious yet is South Korean market leader, SK Telecom, who announced last July that it had signed an agreement with Ericsson to develop 5G technology in time to demonstrate a network at the 2018 Winter Olympics in Pyeongchang. Earlier that month, the vendor had already demonstrated a 5Gbps data throughput speed in laboratory trials, in the 15 GHz frequency band.

Again the majority of research and innovation at the operator and vendor level is taking place in Eastern Asia, with both Huawei and Samsung having reportedly achieved latencies of less than 5 milliseconds in laboratory trials. Japan's NTT DoCoMo has also begun conducting extensive "experimental trials" of potential 5G technologies across multiple frequency bands. The operator has partnered with various vendors to test technologies in a number of spectrum bands, including Alcatel-Lucent (3-6 GHz), Fujitsu (3-6 GHz), NEC (5 GHz), Ericsson (15 GHz), Samsung (28 GHz) and Nokia (70 GHz).

Appendix B: 4G opportunities and challenges

LTE is still in the early stages of its lifecycle

Historically, cellular technologies have adhered to an approximate 20-year cycle from launch to peak penetration, with around ten years between the launch of each new technology (see Figure 5). The first commercial LTE networks went live in 2009 and based on historical precedent we would not expect the technology to reach a peak level of connections until around 2030.

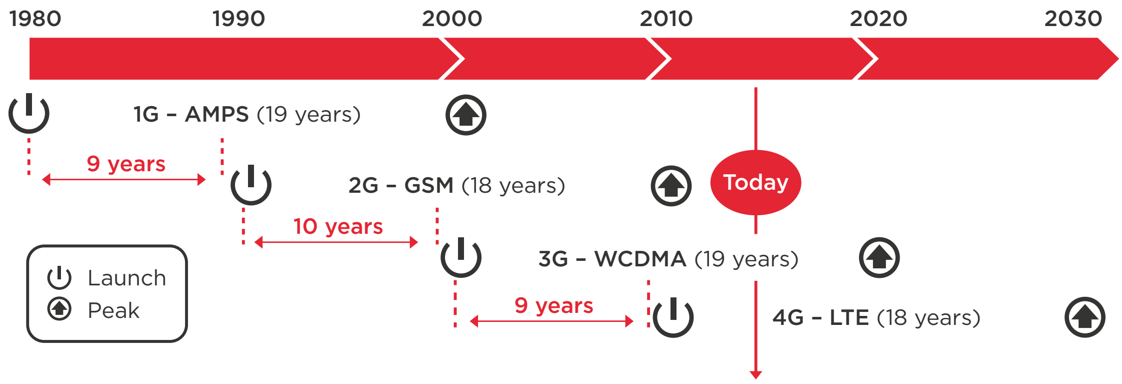

Figure 5: Evolution of mobile technology by generation, 1980 onwards

Source: GSMA Intelligence

In reality, the adoption of LTE is proceeding at a faster rate than its predecessor technologies (see Figure 6), yet we still do not expect LTE connections to peak until well into the next decade. The technology is still at an early stage in its lifecycle, with networks currently confined to just 110 of the world's 237 mobile markets. Hence LTE still represents a considerable growth opportunity for the industry – at present, only around a third of the world's mobile operators (293) have live LTE networks. Assuming all known future network launches go ahead as planned, 158 countries will soon have at least one LTE operator – yet this still leaves one third of the world's mobile markets as untapped territory for LTE services.

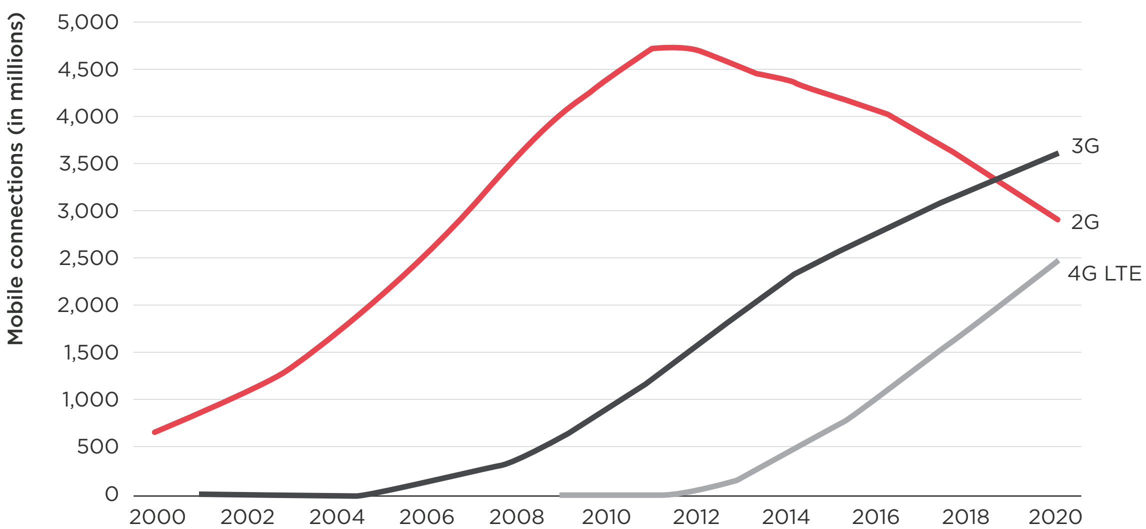

Figure 6: Total cellular connections, global, by technology generation

Source: GSMA Intelligence

In terms of coverage, we expect LTE networks to reach 26% of the world's population by the end of 2014 (see Figure 7), although the technology will account for only 6% of global connections at that time, illustrating the considerable growth potential for LTE even in regions with widespread networks such as the Americas, Europe and Oceania, where we expect coverage to increase from around three in five people on average in 2014 to more than four in five by 2020.

An expected proliferation of launches will also bring coverage to more than two thirds of the population of Asia by that time, while almost one in five Africans will also be covered by LTE networks, more than double the current proportion. Thus globally we expect that LTE coverage will rise from a quarter now to more than 60% of the world's population by 2020 – meaning that 4.9 billion people will potentially have access to the technology.

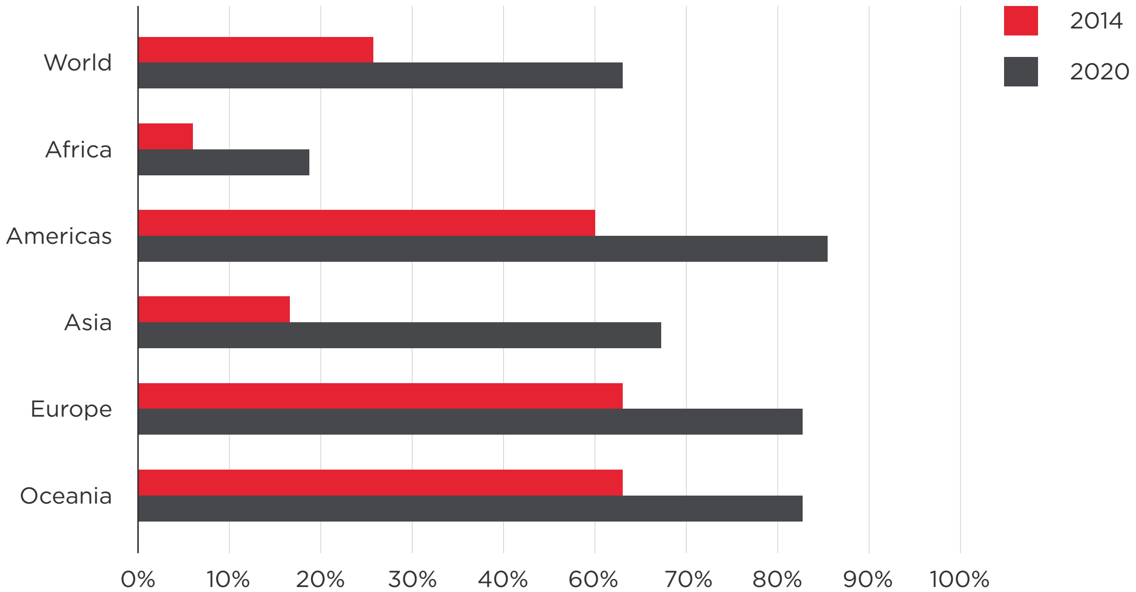

Figure 7: LTE network coverage forecasts, as a % of population by region

Source: GSMA Intelligence

Opportunities for further evolution of LTE

As a technology, LTE continues to develop. Operators are already making a considerable amount of progress in increasing the data speeds of their existing networks by adopting dual-carrier LTE-A technologies, which can achieve theoretical downlink speeds of up to 300 Mbps. As of October 2014, some 22 operators had already launched LTE-A, and we are aware of firm commitments by a further 47 to implement the technology. All in all 15 countries across the world now have live LTE-A networks, and this figure will increase to 35 assuming that all currently planned networks successfully make it to the commercial launch stage.

LTE-A should be able to meet mobile broadband demand (in terms of speed) for several years to come and will provide operators with increasing opportunities to develop attractive and profitable 4G services. In addition, 3GPP is also working on optimising congestion control for more efficient use of M2M on LTE networks.

LTE is driving increasing CAPEX levels

Given the potential that remains for increasing 4G adoption in many countries, we expect to see considerable further investment in the technology. We forecast that the world's mobile operators will invest $1.7 trillion in network infrastructure over the period 2014-2020 (see Figure 8), much of which will be in 4G networks. This outlay is a considerable uplift on the estimated $878 billion invested over the period 2009-2013 and underlines the industry's commitment to meeting the exponentially increasing demand for mobile broadband services as well as connecting 'the next billion people' to the internet.

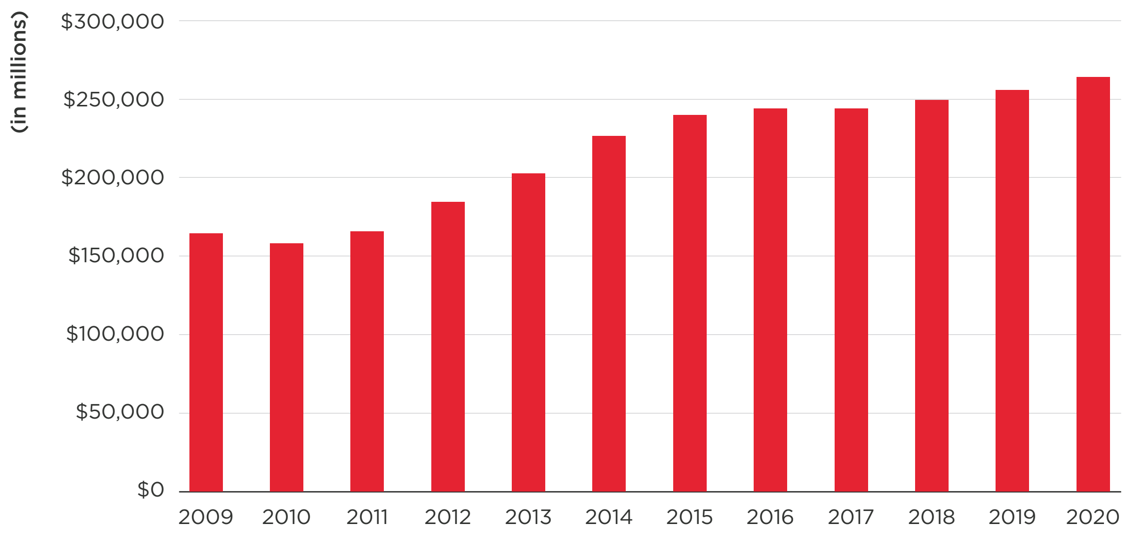

Figure 8: Total mobile operator CAPEX forecast, annual, in million US$

Source: GSMA Intelligence

The monetisation challenge that remains for LTE

The rapid migration towards LTE in the world's most advanced mobile markets has driven a surge in data usage, with 4G users typically consuming twice as much data per month as other users. However, while the introduction of LTE has led to an uplift in ARPU in some instances, the impact on revenue varies widely depending on the market. For example, in South Korea KT reported an LTE ARPU of KRW 44,300 ($41.91) in Q2 2014, 31.8% greater than their blended ARPU for the same period. Operators in the US are seeing similar trends with Verizon Wireless – the largest LTE operator globally with 53.7 million 4G connections in Q2 2014 – announcing that its Q2 2014 ARPA (average revenue per account) was up 4.7% on a year earlier to $159.73 (based on an average of 2.8 connections per account).

However, in regions such as Europe, the migration towards LTE is at a significantly earlier stage and while they have reported similar trends in terms of data consumption, mobile operators in these regions are not yet seeing the same positive impact on revenue from LTE as witnessed in 'digital pioneer' markets such as South Korea, the US and Japan. In many cases, European operators are pricing 4G at the same price as 3G from the outset, while those that initially charged a premium for 4G are having to re-evaluate in the face of strong competition. Hence, the most significant challenge around LTE for many operators remains the monetisation of the networks that they have invested heavily in. While operators are transitioning their tariff structures to become increasingly data-centric, the continued decline in voice and SMS revenues is in many cases yet to be offset by corresponding increases in data revenues.

Thus, operators must seek to manage the change in usage patterns and pricing to minimise any cannibalisation of voice and SMS revenues and to ensure that margins are protected and future investment in LTE and other network technologies remains viable.

LTE interconnect and roaming issues

Interconnect is another area where LTE still has significant challenges to overcome. The GSMA has made progress in this area through the definition of IP eXchange (IPX), a technology with an all-IP core that can provide an improved interconnection which enhances the richness and quality of LTE data roaming. IPX allows the data roaming experience to be managed from end-to-end and manipulated in real time, which is useful for providing services requiring particular attributes e.g. high bandwidth, low latency.

However, the wider adoption of voice over LTE (VoLTE) has been constrained by the absence of a standard IP-based interconnect technology for voice, largely due to operator concerns about being unable to effectively manage and bill VoLTE traffic in the same way as traditional voice calls. Reaching an agreed technical standard for VoLTE interconnect is crucial, as voice services must provide a consistent experience for customers over any network, anywhere in the world. Hence the GSMA will continue to work towards the goal of delivering a seamless VoLTE interconnect and roaming service for consumers while protecting the commercial interests of operators.

Authors

Dan Warren Senior Director, GSMA Technology
Calum Dewar Lead Analyst, GSMA Intelligence

Other GSMA contributors to this report:

Hyunmi Yang Chief Strategy Officer
Javier Albares Head of Corporate Strategy
Elisa Balestra Corporate Strategy Manager
Scott Burcher Senior Analyst, GSMA Intelligence
Matthew Bloxham Senior Director - Head of Policy Research, Government & Regulatory Affairs
Wladimir Bocquet Senior Director - Spectrum Policy, Government & Regulatory Affairs

About GSMA Intelligence

GSMA Intelligence is the definitive source of global mobile operator data, analysis and forecasts; and a publisher of authoritative industry reports and research.

Our data covers every operator group, network and MVNO in every country worldwide — from Afghanistan to Zimbabwe. It is the most accurate and complete set of industry metrics available, comprising tens of millions of individual data points, updated daily.

GSMA Intelligence is relied on by leading operators, vendors, regulators, financial institutions and third-party industry players, to support strategic decision-making and long-term investment planning. The data is used as an industry reference point and is frequently cited by the media and by the industry itself.

For more information, visit gsmaintelligence.com/about/

gsmaintelligence.com • info@gsmaintelligence.com • @GSMAi

GSMA Intelligence, The Walbrook Building, 25 Walbrook, London EC4N 8AF