Enabling Mobile Augmented and Virtual Reality with 5G Networks

Written by,

Alisha Seam (alisha.seam@att.com) Sr. Member of Technical Staff | Palo Alto Foundry

Amy Poll (aj264h@att.com) Lead Outreach Manager | Ecosystem & Innovation

Remound Wright (rw084c@att.com) Sr. Product Development Engineer | Palo Alto Foundry

Dr. Julius Mueller (jmu@att.com) Sr. Member of Technical Staff | Palo Alto Foundry

Faraz Hoodbhoy (faraz.hoodbhoy@att.com) Director of Outreach | Ecosystem & Innovation

With special thanks to our friends at the DIRECTV Digital Innovation Lab, AT&T Corporate Strategy, and AT&T Labs:

Jarrell Pair (jp002m) and Matthew Rasmussen (mr181d) | AT&T Entertainment Group – Product Concepting

Tony Driscoll (ad420r), Suzanne Farhoud (sfo48y), and Sean Nowling (sn415d) | AT&T Corporate Strategy

Dr. John Oetting (jo1742), Dr. Bo Han (cn5542), and Claude Noshpitz (cn5542) | AT&T Labs

January 2017

Executive Summary

Though the underlying technology and concepts defining Augmented Reality (AR) and Virtual Reality (VR) have been in development for decades, advancements in graphics, computing, sensors, and networking have finally placed this emerging industry on the cusp of mass-market adoption. As the supporting hardware and infrastructure continue to develop, immersive systems have the potential to inspire a multi-billion-dollar industry, and even be as game-changing as the advent of the PC. However, despite unprecedented development and investment in the past few years, the AR/VR industry is still facing the substantial difficulty of deploying high quality systems with the mobility and robust user-experience necessary pull this technology from the pages of science fiction into our mainstream reality.

The challenges of delivering the wealth of rich and compelling experiences that will give this new medium an enduring reason to exist are decidedly nontrivial. Immersive and interactive graphical applications require extraordinary levels of computing power relative to traditional forms of media and interaction. Even the most superior mobile AR/VR hardware currently available is suffering from limited battery life, overheating, and insufficient computing capabilities – challenges that we believe will persist throughout the next decade. Mobile-driven AR/VR engineers are eager to harness the power of cloud computing and alleviate the strain placed on end user devices by offloading computationally-intensive tasks to powerful remote servers. However, they have thus far found network bandwidth and latency constraints to be prohibitive.

Enabling true ubiquitous connectivity and ultra-low-latency mobile access to remote computing power is essential for AR/VR technology to achieve horizontal market penetration. We believe that the immersive media industry is unlikely to realize its full potential – beyond niche applications – on today's networking and computing infrastructure. Consequently, this challenge presents a significant opportunity for network operators as we move towards the 5G era: not only to enable the growth of a potentially disruptive technology industry, but also to share in its success.

Classification and Industry Applications of Immersive Systems

Though the underlying technology and concepts defining immersive media have been in development for decades, advancements in graphics, computing, sensors, and networking have finally placed this industry on the cusp of mass-market adoption.[1] In fact, experts claim that the progress of the last 3 years outpaces that of the 20 years before.[2] With sufficient ecosystem investment and development, we believe that these immersive systems are poised to become ubiquitous and have implications far beyond the video gaming and military applications where their journey began. In the years ahead, we could see considerable adoption within the fields of healthcare, retail, and entertainment. As technology advances, immersive media has the potential to inspire a multi-billion-dollar industry, and even be as game changing as the advent of the PC.[3]

The classification of immersive systems is still the subject of debate, and a clearer picture will emerge as the industry matures. Immersive experiences exist along the spectrum of what Paul Milgram first described as the 'Reality-Virtuality Continuum'[4], and we employ a rudimentary taxonomy to distinguish between systems according to varying levels of immersion and methods of composing real and virtual objects.[‡] We occasionally reference these terms, but for the purposes of this paper we generally omit this granularity and divide the systems merely into two broader categories:

- Augmented Reality (AR): transparent displays with digital overlays upon the physical world (including MR)
  Industry initially driven by enterprise and utility-oriented applications
- Virtual Reality (VR): user vision occluded and experiences confined to a synthetic environment
  Industry initially driven by consumer and entertainment-oriented applications

Furthermore, given that this discussion will be primarily centered on technology development and supporting infrastructure rather than content or end-user experiences, we often refer to AR/VR as a single unit. AR and VR largely operate upon the same graphical software engines and are driven by the same hardware processing components. Most importantly, they both encounter similar obstacles in the face of existing networking and computing paradigms.

Head-Mounted Displays: The Evolution of Personal Computing

Just as smartphones placed the world at our fingertips, Head-Mounted Displays (HMDs) will deliver it right in front of our eyes. Following the evolution of computing interfaces through history – from command line, to desktop operating systems, to touchscreen interfaces – we have transitioned from a world where one would need to be properly trained in command line programming to one where one can use a smartphone or tablet without any prior experience at all.

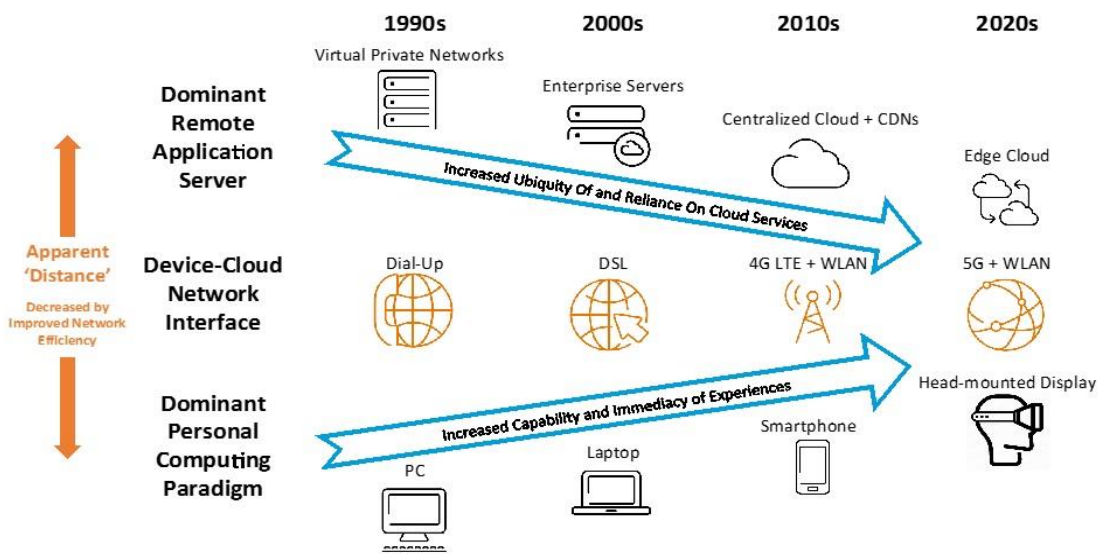
The addressable market and scope of applications for computing has broadened significantly as the user interface became more akin to natural gestures and interactions.[3] This medium also gives us a wider field of view, such that our window into this virtual world is no longer limited by the dimensions of a display screen. Fundamentally, by blending the realms of the virtual and the physical, HMDs present an ever more intuitive way to leverage the power of a computational source. This ease of use, along with the wide spectrum of potential applications, will create the potential for Augmented and Virtual reality to emerge from vertical-specific use cases into a broader computing platform.[3]

[‡] It is important to note that this taxonomy deviates significantly from the canonical definitions that were presented by Milgram and that have been traditionally employed by the research community - in which Augmented Reality (AR) and Augmented Virtuality (AV) both fall under the broader scope of Mixed Reality (MR), which encompasses the entire spectrum between the extremes of Real and Virtual environments. These terms have recently been heavily adapted by industry in an effort by companies to differentiate their offerings to consumers and developers. Thus, rather than a strict set of definitions, the taxonomy we offer here is an attempt to bring a non-rigorous understanding and order to the systems currently being presented in today's immersive media market.

The IT-Networking Convergence Has Unlocked the Power of the Cloud

As technology has evolved and converged, the flow of information has become streamlined, driving compute and storage off personal devices while simultaneously enabling those devices to become more powerful and intuitive with each subsequent generation. However, computing revolutions are not a byproduct of the device technology alone – the entire ecosystem and infrastructure must be developed to support these seismic shifts.

In the 1990s, telecommunications companies (telcos), which previously offered primarily dedicated point-to-point data circuits, began to offer Virtual Private Network (VPN) services. Rather than building out physical infrastructure to allow more users to have their own connections, telcos were now able to optimize resource utilization toward overall bandwidth usage efficiency and could offer the same quality of service at a fraction of the cost. In these early stages, the term "cloud" was used to represent the computing space between the provider and the end user. As computers became ubiquitous, engineers and technologists explored ways to make large-scale cloud computing power available to more users through time-sharing. They experimented with algorithms to optimize the infrastructure, platforms, and applications to prioritize computing resources and increase efficiency for end users. By the 2000s, cloud-driven XaaS providers began to deliver enterprise-level applications to end users via the internet.[5]



Historical perspective and future projections of the IT-Networking convergence

Over time, the IT-Networking ecosystem has become increasingly adept and dynamic; cloud data centers have become more powerful and geographically distributed, and networks have become dramatically more efficient in order to support them. As a result, remote application servers have become easier to access, bringing more information, media content, utility applications, and realistic forms of communication directly to the consumer. To prepare for the next generation of ultra-low-latency and compute-intensive technology such as the lightweight, affordable mobile HMDs that will enable pervasive AR/VR adoption, cloud providers have already begun to invest in "edge clouds" with a distributed presence that positions their cloud servers geographically closer to the end user.

Technology Adoption is Driven by Useful and Unique Experiences

Whether for consumer use or enterprise use, both AR and VR technology face the challenge of delivering a high enough value proposition to add another device to the current portfolio of offerings, which already consists of desktops, notebooks, tablets, and smartphones.[3] Previous iterations of Head-Mounted Displays were unable to gain sufficient traction due to limited battery life, privacy concerns, social stigma, and most importantly, an inability to deliver sufficiently complex and compelling experiences.[6] The best early adopters of a product are not tech-enthusiasts, but rather the people who "needed the early versions of the product *yesterday* and will rapidly start getting benefit from using the device right now".[7] A see-though display offers a convenient, hands-free interface, but this form-factor feature alone is not sufficient to drive a personal computing revolution.

Not all immersive content is created equal, and the power of this new medium is not implicit. The novelty era of branded VR stunts is rapidly drawing to a close, and consumers are beginning to demand more than the low-quality, 'casual AR/VR' experiences offered by the first wave of affordable HMDs and content.[2] We posit that to address new classes of problems and provide a verifiable step-change capability set, HMDs will need to enable complex, dynamic graphical interfaces and intuitive, interactive experiences that truly are not achievable with a smartphone today.

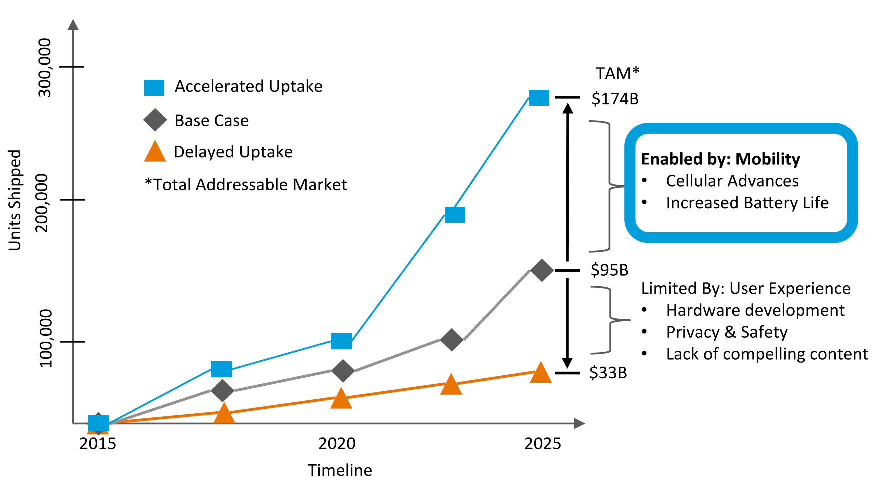
True Mobility of AR/VR Systems: The Key to Widespread Market Adoption

Though maintaining the quality of experience is essential, we believe that the AR/VR industry will not be able to realize its true disruptive potential until Head-Mounted Displays can be driven without dependence on the instantaneous access to the content and power provided by a tethered PC. In today's market, consumers are unwilling to be restricted in their ability to access and share content. AR/VR will be fundamentally limited by the experiences it can deliver, and we believe that mobility is key to horizontal market penetration.

We anticipate different levels of desired mobility for predominantly-AR vs. predominantly-VR use cases. In VR systems, the user's vision is obstructed. Thus, VR is best suited for use cases in entertainment, training, etc. and most likely will not be heavily utilized outdoors due to safety, social, and practical considerations. Despite these constraints, we believe that consumers will not want to be physically tethered to a PC or even restricted to a dedicated 10x10 space to use VR headsets. We believe that users will desire the moderate freedom and flexibility of "nomadic mobility," in which connectivity is wireless but not ubiquitous (i.e. limited to disjoined hotspots).[8,9]

In contrast, we believe that AR must be 'truly mobile' to capture its full market potential. Without this degree of agility, AR will not be viable for field operations or the personal computing use cases that have made smartphones so ubiquitous and powerful today. Mobility is necessary for users to meaningfully interact with the surrounding environment and receive information anytime, anywhere, and on-demand. While the smartphone is an impressive piece of hardware, it is the ubiquitous connectivity that makes it truly powerful.[10]

Critical Factors Driving Mass Adoption



Enabled Markets

	Level of Mobility	Markets Captured
Accelerated Uptake	Truly Mobile (Cellular)	• Retail • Engineering & Field Operations • AR Entertainment • Military/Public Service & Safety • Real estate
Base Case	Nomadic (Wi-Fi)	• Education • Healthcare • Video Entertainment • Live event streaming
Delayed Uptake	Fixed (Ethernet)	• Video gaming

AR/VR forecast scenarios, enabling characteristics, and captured markets [3,11]

The Computational Demands of AR/VR Systems

As discussed, we believe that Head-Mounted Displays have the potential to revolutionize personal computing and that AR/VR could become a pervasive new content medium – but only if the underlying technology can deliver sufficiently compelling experiences. However, the computational challenges of delivering this highly immersive, highly interactive content with a satisfying user experience are not trivial.[12]

> *"AR/VR is all about tricking the brain to believe that it's looking at a real world. Any discrepancy that disturbs this illusion leads to negative experiences (such as nausea). And compute is the key component to creating a virtual world that looks the same as the real world."*
> *– Rishi Ranjan, Founder of Gridraster Inc.*

Achieving Immersion Requires Steadfast Adherence to Strict Neurological Response Metrics

AR/VR functions are measured against a stricter set of neurological functions than traditional passive media. Immersive systems typically require a head-mounted display (HMD) that smoothly moves the display screen in tandem with the user's head. The goal of this medium is 'true immersion' — creating an illusion so real that it convinces the human brain, the world's finest computer. The experience is therefore fundamentally rooted in the human audio-visual system, and requires an entirely new approach to the display problem.[13]

The benchmark for determining if technologies are adequately 'responsive' has always been relative to the processing abilities of the human body (reaction time). In the past, the metric used to evaluate content refresh speeds was the reaction time for a human brain to process a screen, which is approximately 80ms.[14] However, for AR/VR, the industry is driving the conversation around the Vestibulo-Ocular Reflex (VOR) – the neurological process by which the brain coordinates eye and head movements in order to stabilize images on the retina – which is critical to synchronizing virtual and real objects to create a coherent view. The entire VOR process takes the brain 7ms, a more than 10x reduction over screen-to-brain propagation.[15]

Though standards are still being determined (especially for AR), the current VR industry best practices dictate a maximum motion-to-photon latency of 20ms. The general consensus is that target response latency for both AR and VR systems is easily 3 to 10 times lower than for today's standard non-HMD visualization.[16]

AR/VR Visual Requirements Push the Limits of Hardware Display Functions

This frame latency requirement translates into increased demands on the processors that generate the graphical displays. In some traditional graphics workloads, low frame latency can be essential to high-quality rendering, but in many cases, high resolution or visual effects take priority. For example, most movies operate at 24 frames per second, and a game is considered 'playable' at 30 frames per second (fps).[17] In contrast, low frame latency is absolutely necessary for true immersion and thus an essential performance metric for AR/VR. High quality and highly dynamic experiences are generally recommended to be produced at 100fps or above. Lower frame rates will not only break the immersive experience, but also lead to user disorientation and nausea (known as simulator sickness). The frame rate requirement places further stress on the screen refresh rate (measured in Hz), because the display capability effectively places a cap on the number of frames shown to the user.

In addition, high quality screen resolution is essential to immersive AR/VR systems. Screen resolution has been improving over the past 20 or 30 years but development has been driven almost exclusively for two-dimensional displays, which have significantly more relaxed requirements than AR/VR.[18] In Head-Mounted Displays, pixelation is much easier to perceive given that the screen sits mere inches from the user's face. The actual number of pixels rendered per second (with an acceptable frame latency) for today's higher-quality VR systems is roughly four times higher than that of the average PC display (1080P resolution), and twice that of a high-end PC display.[12, 18]

3D Physics Calculations, Tracking, and Scene Rendering Must Occur in Real Time

Beyond the display functions, AR/VR systems require substantially more real-time calculations to properly render 3D environments and respond to user inputs. These systems can only deliver a truly immersive experience if they are able to respond to input sensors. Besides head motion tracking, the input sensors are used to track 6 Degrees of Freedom (6DOF) 3D position, hand motions, controllers, etc. AR systems further require environmental image processing and spatial tracking to seamlessly integrate a user's real and virtual environments.

For each frame, the processors must then run extensive computational functions based on these inputs to calculate how every single object would react and appear to the user according to the laws of physics. Lighting and shading on each object must be layered on top of the object's positioning as well. This is the only way to trick a human brain into perceiving virtual objects as though they were real.

The Requirement for Powerful PCs to Derive Performance from GPUs

GPU development has been a critical factor driving AR/VR industry growth. PCs provide all types of power: CPU processing, memory access speeds, RAM storage, etc. but for AR/VR, the GPU is the most significant differentiating upgrade.[19] The GPU is the single largest generator of performance data points in 3D interactive content benchmarks because the graphics rendering operations cannot be performed nearly as efficiently by the CPU alone.[20] A high-powered GPU helps to keep frame rates smooth while rendering more realistic graphics to maintain 'immersion'.

Unfortunately, to accommodate the GPUs and achieve their desired performance, high quality AR/VR headsets are currently physically tethered to very powerful and expensive PCs. These requirements create larger adoption barriers due to high price points and limited mobility, which detracts from the overall experience. Even as companies begin to release higher-end wireless headsets in 2017, we anticipate that the quality of the user experience will still be highly dependent on proximity to the PC that drives them.

The Mobile-PC GPU Capability Gap

The challenges facing current mobile GPUs are two-fold:

1. Current mobile GPUs are significantly weaker than even their consumer-grade PC counterparts. As a rough back-of-the-envelope comparison, we compare the system on chip (SoC) GPU on board one high-end smartphone on the market today against the GPU that drives a premium PC-connected full-immersion experience. One of most defining metrics of GPU performance is its FLOP count (floating-point operations per second). The mobile GPU can handle ~500 GFLOPS, whereas minimum requirement for the PC-tethered experience is ~3500 GFLOPS and most in the industry even recommend upgrading to the next model for GPU headroom with 6000 GFLOPS. Thus, even high-end VR-capable mobile devices are lagging the PC requirements by **7X**.

2. The compute-demand of PC GPUs is expected to increase exponentially as users demand higher quality experiences. Currently, the PC-tethered headset runs at 2K resolution per eye at 90fps, but experts forecast that 16K/eye at 120fps may be the standard for true immersion – a 10X display upgrade. Some are even driving the conversation around 24K/eye at 144fps.18[, 21] Thus we estimate that, at a minimum, the ability to render these 'truly immersive experiences' without an external computing source would require a **70x improvement** over the capabilities of current high-end mobile devices.

For the past few decades, the tech industry has pointed to Moore's law (the trend that processing power doubles every 18 months, primarily dependent number of transistors per square inch) as an indicator of future hardware capabilities.[22] Relying solely on this trend as a benchmark for the future, it would take more than **9 years** to achieve a 70X increase in computing performance.[23]

Additional Hardware Limitations: Overheating and Battery Life

Moore's Law is rapidly reaching its physical limits as transistor sizes rapidly approach atomic scale, bringing with them a whole host of power and current leakage obstacles. Hardware engineers are aggressively pursuing alternative modalities of sustaining the processing improvement trend, but many are still in R&D phases, and passive cooling and sustaining battery life are proving to be substantial physics barriers. The intense compute demand of mobile AR/VR applications results in overheating, which further depletes an already limited battery life on mobile devices.[24] Unfortunately, there is no Moore's Law equivalent for battery technology, and designing a safe, small, low-cost power source to keep pace with aggressive computing demands remains an uphill battle. Until then, the stress that AR/VR places on SoC-driven devices means that only low usage time is possible, causing a significant barrier to providing the high quality, rich, and immersive experiences that are necessary to drive adoption of this new content medium.

Network Requirements of Cloud-Assisted Mobile AR/VR Systems

The limitations of mobile phones and similar SoC-driven systems naturally motivate the desire to offload computationally intensive tasks to a powerful remote server. Cloud computing and caching have been central to the revolutionary capabilities of modern mobile devices and the whole ecosystem of supporting enterprises. Further, as the computing demands of emerging applications grow, so do the capabilities of the cloud.

Mobile-driven AR/VR engineers are eager to harness the power of cloud computing by designing 'thin-client' applications, which require minimal resources on the end user device and offload computationally intensive tasks to powerful cloud servers.[25, 26, 27] However, they have found the lack of available network bandwidth, and more importantly the induced latency, to be prohibitive.

Bandwidth Limitations and the Need for Increased Uplink

Current mobile networks are not equipped to address the increased bandwidth requirements of AR/VR applications. This industry has only just begun to explore the possibilities of mass-market applications, merely touching the tip of the iceberg of interactivity and immersion with experiences such as 360-video. However, even for 360-video, industry best practices dictate a minimum of 20-30Mbps for a satisfactory streaming experience, comparable to bitrates required for the higher end resolutions (4K) of traditional 2D video streams.[28] Even the higher quality mobile 360 videos do not deliver anywhere near the dynamic range and resolution that 360 cameras are capable of capturing [29], and as mobile display technology advances, the requested bitrate is only expected to increase. Furthermore, today's highly dynamic and interactive applications are primarily downloaded and hosted on the client (device) side. Streaming complex, interactive AR/VR 3D graphics will require unprecedented data rates.

Another noteworthy deviation from the delivery of traditional, static media content is that AR/VR streaming, like applications such as system control and IoT analytics – is poised to require increased bandwidth allocation for both downlink *and uplink* data streams. Existing networks designate significantly more spectrum to downlink data delivery.[30] This asymmetry may need to be reevaluated as new network traffic patterns emerge.

Techniques such as view-adaptive streaming are being designed to alleviate bandwidth requirements by strategically delivering content in response to user inputs such as head-position. However, the tradeoff is that the client device then requires continuous interaction with the hosting application server – for which a guaranteed low latency and stable connection are critical.

Network Paths to the Centralized Cloud: The Latency Challenge

Responsiveness of web-based applications is measured by the end-to-end network latency (time to travel across nodes to the server then back to the end user) plus the application latency (delay induced by processing, buffering, display, etc.). An interaction or service is considered 'real-time' when this total response time is significantly faster than the time constants of the application, such that the delay incurred by communication and computing appear negligible to the end user.[31] Analyses by major web service companies have demonstrated that even minor apparent lags translate into user dissatisfaction and application abandonment.[32, 33, 34] Responsiveness becomes an even more important factor when attempting to maintain the immediacy and realism of immersive media.

The applications that are dominant today (asynchronous or one-way-delivery applications like instant messaging, web page loading, etc.) can tolerate 100+ms of latency, and are well supported by the combination of existing network infrastructure, content delivery networks (CDN) and centralized cloud computing. This architecture is optimized for the delivery of static, predefined content with minimal levels of dynamic interaction.[14] However, we are now on the cusp of a new set of dynamic, hyper-local applications for which 'real-time' interaction takes on a whole new meaning because the demands of responsiveness have shifted by an order of magnitude. As discussed previously, interactive AR/VR content responsiveness is measured against the Vestibulo-Ocular Reflex (7ms), whereas static content is measured against screen-to-brain propagation (80ms). The constrained neurological metrics translate into a proportional constraint on application responsiveness, placing unprecedented demands on network latency.[14]

5G Adoption Will Be Driven by Next-Generation Use Cases

Despite unprecedented progress and investment in the past few years, the Augmented and Virtual Reality technology industry is still facing the substantial challenge of deploying high quality systems with the mobility and robust user-experience necessary to drive mass-market adoption. As discussed, we believe that these immersive systems are unlikely to realize their full potential – beyond the video gaming industry, scattered novelty experiences, and niche enterprise applications – if built solely upon existing networking and computing infrastructure.

Network Demands of Emerging Applications

As with each preceding generation, the rate of adoption of 5G will be a direct function of the new and unique use cases it unlocks.[35] The evolution from 3G networks to mobile broadband technology (3.5G+) transformed industries by delivering a significantly enhanced mobile internet experience, eventually leading to the app-centric interface we know today. From email and social media to music and video streaming to the control of home appliances from anywhere in the world, mobile broadband fundamentally changed the lives of many people through services provided both by operators and third party players.[35]

Given the network demands of emerging technology industries such as AR/VR, there is much excitement and anticipation around the arrival of 5G network capabilities.[36, 37, 38, 39, 40] 5G specifications reference everything from extreme mobility (e.g. high speed trains) to dramatic increases in energy efficiency and network reliability, leading to a hyper-connected society in which mobile technology will play an ever more important role in our lives.[41] In the strictest terms of network deliverables, 5G is being driven around expectations for 1 millisecond end-to-end (E2E) latency and 1-10 Gbps downlink speeds to end points in the field (i.e., not just a theoretical maximum).[14] However, despite the hype and anticipation, the specific means to achieving these metrics remains largely undefined. 5G is not characterized by a single radio access technology upgrade. Rather, it is a full portfolio of access and connectivity solutions addressing the aggressive demands of mobile communication beyond 2020.[42] Thus, delivering the high social and economic value that 5G promises necessitates a critical analysis of network architecture and technology deployment strategy in context of the applications it hopes to support.

If implemented effectively, 5G has the potential to be yet another true generational shift in network technology and enable an even greater level of disruptive innovation. Operators must therefore critically assess the network requirements of emerging technologies and identify those that truly require a step increase in network capabilities. The projected latency and bandwidth/data rate requirements of the various use cases that have been discussed in the context of 5G are depicted below:

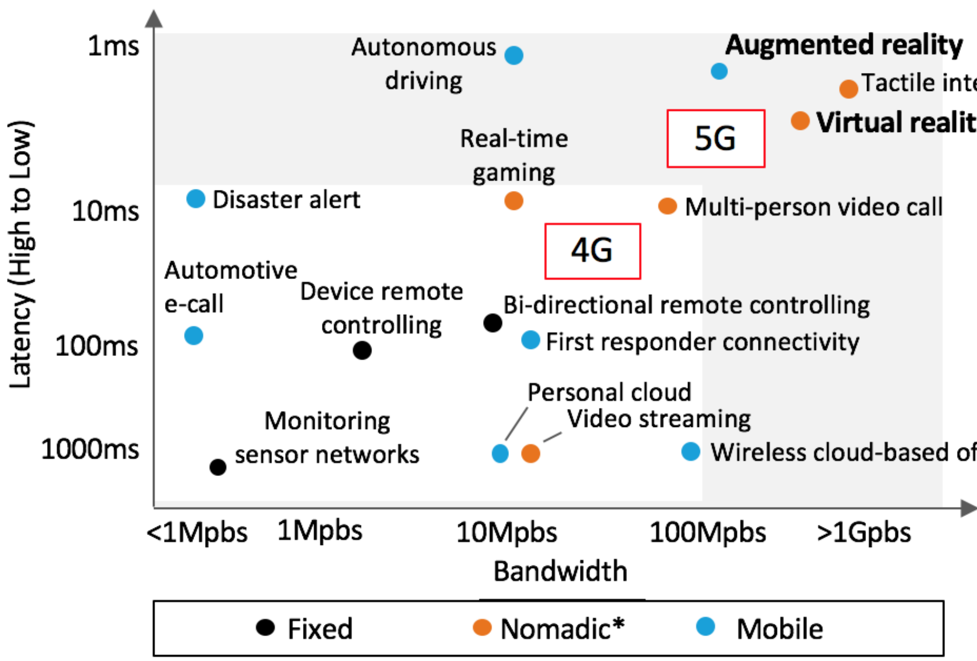


Figure 1: Projected network demands of emerging **technologies**[35, 43]
***Nomadic: Connectivity is wireless but not ubiquitous - limited to disjoined hotspots**

As indicated by the required metrics, several of the services that enterprises claim as potential 'killer apps' for 5G do not necessarily require a generational shift in technology and could in fact be provided within the technical realm of 4G networks. Many are driven by increased coverage, connection density, and energy efficiency – efforts that improve the economic viability of these services, rather than filling technological gaps that explicitly prohibit them. Thus, GSMA Intelligence argues that only the use cases with network requirements on the order of 1ms latency and >1Gpbs downlink speeds (i.e. within the shaded region of the graph) – such as autonomous driving, tactile internet, and augmented and virtual reality – have the "generation-defining attributes" that will most effectively drive 5G adoption and monetization.[35]

To prepare for the next generation of ultra-low-latency and compute-intensive technology such as the lightweight, affordable mobile HMDs that will enable pervasive AR/VR adoption, cloud providers have already begun to invest in "edge clouds" with a distributed presence that positions their cloud servers geographically closer to the end user. 5G networks will need to keep pace with these infrastructure demands as well, not only by continuing to increase the capacity and efficiency of network functions, but also by integrating computing resources directly into the communication network.

Designing for the Greatest Challenge Yields the Greatest Opportunity

The Augmented and Virtual Reality (AR/VR) industries present a significant challenge to operators because they are poised to stress both network latency and bandwidth capabilities simultaneously. Additionally, Goldman Sachs is projecting a possible AR/VR adoption curve similar to that of smartphones.[3] It is therefore critical to design 5G architecture in anticipation of these requirements and guard against network congestion. However, within this challenge also lies a significant opportunity.

As infrastructure becomes increasingly software-defined and application-agnostic, the risk of investing in industries with indeterminate futures is heavily mitigated. Technology evolution has given rise to a 'digital convergence': the merging of telecommunications, computing and broadcasting into a single digital bit-stream, blurring the technology challenges and delivery architecture across these industries.[44] Many of the challenges facing the AR/VR industry translate broadly across emerging applications such as system control (i.e. driverless cars), IoT and sensor analytics, machine learning, etc. Though initial capital investment may be high, the infrastructure deployed to support AR/VR is relatively 'future-proof' because application-specific platform services can be dynamically adapted as each of these industries develop throughout the coming decade. Designing a network capable of addressing the use cases with the most extreme latency, bandwidth, and computing requirements has the power to enable an entire class of emerging industries and give rise to a truly 'next-generation' technology ecosystem.

References

[1] Financial Times. "Virtual Reality Gets its Mass Market Headset on". https://www.ft.com/content/f8087e6e-8c66-11e6-8aa5-f79f5696c731
[2] MEC Global. "Spotlight on VR-AR-360". Issued January 2016.
[3] Goldman Sachs. Virtual and Augmented Reality Industry Report. 13 January 2016.
[4] Paul Milgram, Haruo Takemura et al. "Augmented Reality: A class of displays on the reality-virtuality continuum". SPIE 1994.
[5] "The History of Cloud computing". http://www.eci.com/cloudforum/cloud-computing-history.html. Accessed 30 Nov 2016.
[6] Ian Altman. "Why Google Glass failed and why Apple Watch could too". Forbes.com. 28 April 2015.
[7] Dan Kaplan. "The real reason Google Glass failed". https://threadling.com/product-marketing-google-glass/. Accessed 17 Dec 2016.
[8] Shatzkamer, Kevin. "The IP of Mobility". Cisco.com. 15 June 2009.
[9] UBS Global Research. "Where will 5G fall on hype versus reality spectrum?". 16 November 2016.
[10] Laugesen, John and Yufei Yuan. "What Factors Contributed to the Success of Apple's iPhone?". McMaster University. 24 June 2010.
[11] Goldman Sachs Forecast Revision "Augmented Reality starts to liven up the AR/VR market" issued July 2016
[12] D. Kanter, "Graphics Processing Requirements for Enabling Immersive VR," AMD White Paper, July 2015.
[13] D. Kanter, "Graphics Processing Requirements for Enabling Immersive VR," AMD White Paper, July 2015.
[14] Weldon, Marcus K. "The Future X Network: A Bell Labs Perspective". 2016.
[15] Weldon, Marcus K. "The Future X Network: A Bell Labs Perspective". 2016.
[16] D. Kanter, "Graphics Processing Requirements for Enabling Immersive VR," AMD White Paper, July 2015.
[17] Steam User Forums. http://forums.steampower.com. Accessed 4 December 2016.
[18] "Steam Hardware & Software Survey". http://store.steampowered.com/hwsurvey. Accessed on December 2016.
[19] Nordquist, Brett. "The Rise of the GPU". 21 August 2016.
[20] Interview with Solutions Architecture Team at Nvidia Corporation. 3 November 2016.
[21] Jarrell Pair Interview with Qualcomm Engineering on 360 Video Requirements. 5 December 2016.
[22] Investopedia. "Moore's Law". http://www.investopedia.com/terms/m/mooreslaw.asp.
[23] Interview with Rishi Ranjan, CEO and Founder of Gridraster. 28 Nov 2016.
[24] Hamid Sarbazi Azad. "Advances in GPU Research and Practice". Morgan Kauffman Copyright. 2016.
[25] Interview with Rishi Ranjan, CEO and Founder of Gridraster. 28 Nov 2016.
[26] Daniel Wagner and Dieter Schmalstieg. "Making Augmented Reality Practical on Mobile Phones". Parts 1 and 2. Graz University of Technology.
[27] Interview with Benjamin Cooley, Founder of WayGate (VRTV). 8 November 2016.
[28] Interview with Jarrell Pair regarding Facebook Developer Conference (F8) 2015. March 2015.
[29] Nokia. "Ozo Specifications". https://ozo.nokia.com/ozo_en/nokia-ozo-specs/. Accessed on 19 December 2016. Samsung. "Gear 360". http://www.samsung.com/global/galaxy/gear-360/. Accessed on 19 December 2016. Cabral, Brian K. "Introducing Facebook Surround 360: An open, high-quality 3D-360 video capture system". 12 April 2016.
[30] Liopiros, Kosas. "Asymmetry and the impending (US) spectrum crisis". Financial Times.com. 28 May 2013.
[31] Daniel Wagner and Dieter Schmalstieg. "Making Augmented Reality Practical on Mobile Phones". Parts 1 and 2. Graz University of Technology.
[32] Khan, Farhan. "The Cost of Latency". Digital Realty.com. 10 March 2015.
[33] Shalom, Nati. "Amazon found every 100ms of latency cost them 1% in sales". GigaSpaces.com. 13 August 2008.
[34] Eaton, Kit. "How One Second Could Cost Amazon $1.6 Billion In Sales". Fastcompany.com. 15 March 2012.
[35] GSMA Intelligence. "Understanding 5G: Perspectives on future technological advancements in mobile". December 2014.
[36] Delaware North. "The Future of Sports". Futureof.org. http://futureof.org/sports-2016/. Accessed on 14 December 2016.
[37] Next Generation Mobile Networks. "5G White Paper". 17 February 2015.
[38] Fettweis, Gerhard P. "A 5G Wireless Communications Vision". Microwave Journal. 14 December 2012.
[39] DMC R&D Center, Samsung Electronics Co., Ltd. "5G Vision". February 2015.
[40] National Instruments. "5G: The Internet for Everyone and Everything". Accessed on 14 December 2016.
[41] Next Generation Mobile Networks. "5G White Paper". 17 February 2015.
[42] Ericsson White Paper. "5G Radio Access". http://www.ericsson.com/res/docs/whitepapers/wp-5g.pdf. April 2016.
[43] UBS Global Research. "Where will 5G fall on hype versus reality spectrum?" 16 November 2016.
[44] Blackman, Colin R. "Convergence between telecommunications and other media: How should regulation adapt?. 1998.