JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Edge Computing with 5g Enabled Wireless Networks

*** Ashwini K.L. Rao, Assistant Professor, Dept. of Computer Science, Reva University, Bangaluru.**

**** Samyukta.D. Kumta, Assistant Professor, Dept. of Computer Science, Reva University, Bangaluru.**

Abstract

This paper attempts to study 5g **enabled Edge Compute** Platform that helps to build and run applications and services elastically. Build and deploy applications with unparalleled scale, reliability, and security. 5G is having a positive impact on edge adoption in the enterprise as it will make it possible to install lower power but faster computational power at the edge. 5G and edge are complementary. 5G will connect the next wave of smart devices, resulting in exponential growth of data at the edge. The vision is that 5G will connect the next wave of smart devices, resulting in exponential growth of data at the edge.

The 5G mobile network is the latest global wireless standard developed by 3GPP. While 4G networks provide connectivity to most current cell phones, 5G can enable connectivity for an expanded set of devices, from machines to vehicles. 5G has been designed to meet many challenging requirements, including: Significantly higher throughput (up to 20 Gbps peak data rates based on IMT-2020 requirements) in mobile broadband, such as 4K videos streaming or virtual reality (VR). Ultra-low latency (less than 1 millisecond) for real-time communications, such as surveillance drone maneuvering. Massive network capacity, with plans to exceed the number of connected devices per unit area by over 100x compared with 4G/LTE technologies. A more uniform user experience and more users, i.e. higher reliability in crowded areas such as stadium and sports events. Improved energy efficiency, including reducing power consumption and the delivery of cellular Internet of Things (IoT) devices that can last for more than 10 years. The new 5G core network follows a Service-Based Architecture (SBA), which allows for the adoption of cloud-native technologies and open source software into telecom networks and brings in more collaboration, innovation, and openness. Based on this vision, the International Telecommunication Union (ITU) has envisioned several use cases and defined them under three main categories in the "IMT Vision for 2020 and beyond" (the categories are shown in figure 1). Most of us might focus on the enhanced Mobile Broadband (eMBB) use case, but the biggest disruption of 5G will be by massive Machine Type Communications (mMTC) and Ultra-Reliable and Low Latency Communications (URLLC) type use cases such as autonomous vehicles, smart city, and industrial automation.

Keywords: Edge computing, Mobile edge computing, Ultra low latency, 5G, Cloud architecture, Data sovereignty,

Introduction

It’s hard to get the full picture of 5G without understanding edge computing. Edge computing is computing that takes place at or near the physical location of either the user or the source of the data, which results in lower latency and saves bandwidth.

Effectively, edge computing is a concept that enables services to be hosted close to the service consumers (e.g., subscribers, machines, and devices). Being closer provides benefits, including significantly reducing end-to-end latency, decreasing load and saving cost on the backhaul transport network and enhancing data privacy (i.e., private 5G use case). The edge computing paradigm is the only way to realize the latest 5G URLLC use case, and it also addresses many challenging mobile network issues in 5G technologies.

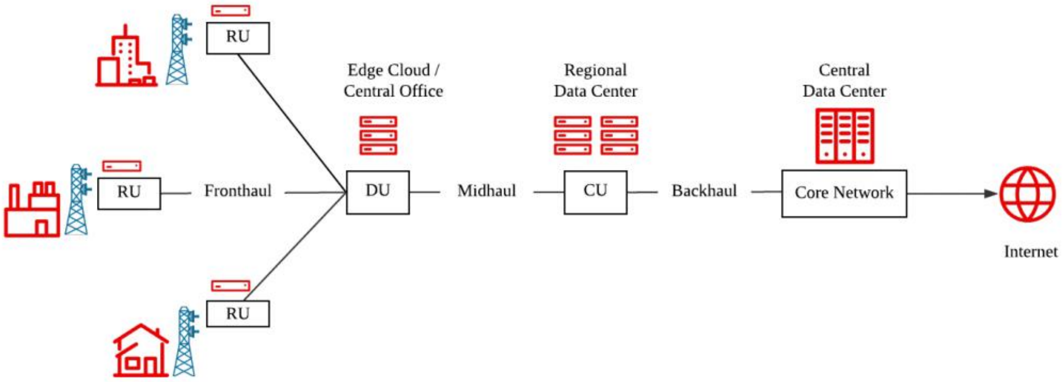
Edge computing is the key to achieving the promise of many new 5G use cases, including virtual and augmented reality (VR/AR), IoT, Industrial IoT, autonomous cars and drones, real-time multiplayer gaming, and recently the deployment of open Radio Access Network (open RAN). I will elaborate on the emerging use cases in this section.

Mobile broadband service is the main use case of 3G/4G, and remains the first 5G commercial use case for the public. 5G eMBB mobile broadband service promises 10~100x broadband speed improvement over 4G through the 5G New Radio (a.k.a. NR), especially in higher frequency bands (i.e., mmWave spectrum allocated in above 24GHz) and denser small cell sites.

To evolve this massive radio network toward an open, intelligent, virtualized (or containerized), and fully interoperable RAN, open RAN standards and deployment have been piloted by several mobile operators. Open RAN is possibly the first prevailing edge computing use case in 5G mobile network.

As the following sample open RAN architecture describes, the modern 5G radio access network has been split and de-coupled into the Central Unit (CU), Distributed Unit (DU), and Radio Unit (RU) through open interfaces like fronthaul, midhaul, and backhaul.

This architecture disaggregates RAN hardware (HW) and software (SW) through the vendor-neutral and general-purpose processor platforms (i.e., Intel x86-64 or ARM). It also allows multi-vendor interoperability and enables a diverse ecosystem for mobile operators to choose best-of-breed options for their 5G deployments.



To date, we have seen several successful open RAN CU/DU deployments on top of the edge cloud since 2020. This underlying edge cloud is based on commercial-off-the-shelf (COTS) servers with hardware acceleration components like FPGA, eASIC, or GPU. Besides, even the RU in the remote cell sites has gradually shifted to a white box hardware built on an open and disaggregated architecture.

These edge and far edge deployments typically run on a massive scale. For example, there are about 400,000 cell sites in the United States, so it is quite common to see thousands of compute nodes in each deployment. As the telco industry continues its shift from proprietary to open based technology, open RAN is expected to be quite popular in the near future.

In recent customer projects, Red Hat saw several open RAN implementations on top of Red Hat edge cloud solutions, including a Virtual Network Functions (VNF) based implementation on Red Hat OpenStack Platform and OpenStack Distributed Compute Node (DCN, for centralizing the control plane and optimizing the cost in edge sites), as well as a Containerized Network Functions (CNF) based implementation on Red Hat OpenShift. For the incoming low latency and real-time use cases, such as VR/AR, autonomous and connected vehicles, drone control, telesurgery and medical robotic instruments, and real-time multiplayer online gaming, the edge computing paradigm is the most effective way to ensure the low latency and also minimize the backhaul bandwidth and cost.

As mobile operators can’t build all the use cases themselves, they will need to attract partners and developers to build an optimized ecosystem on top of their edge cloud. Independent software vendors (ISV) and enterprise developers demand a platform that can help abstract the network complexity, including making use of the mobile network capabilities (e.g., user location, SIM-based authentication, prepaid/postpaid payment mechanism, etc.) in a more secure way.

Objective:

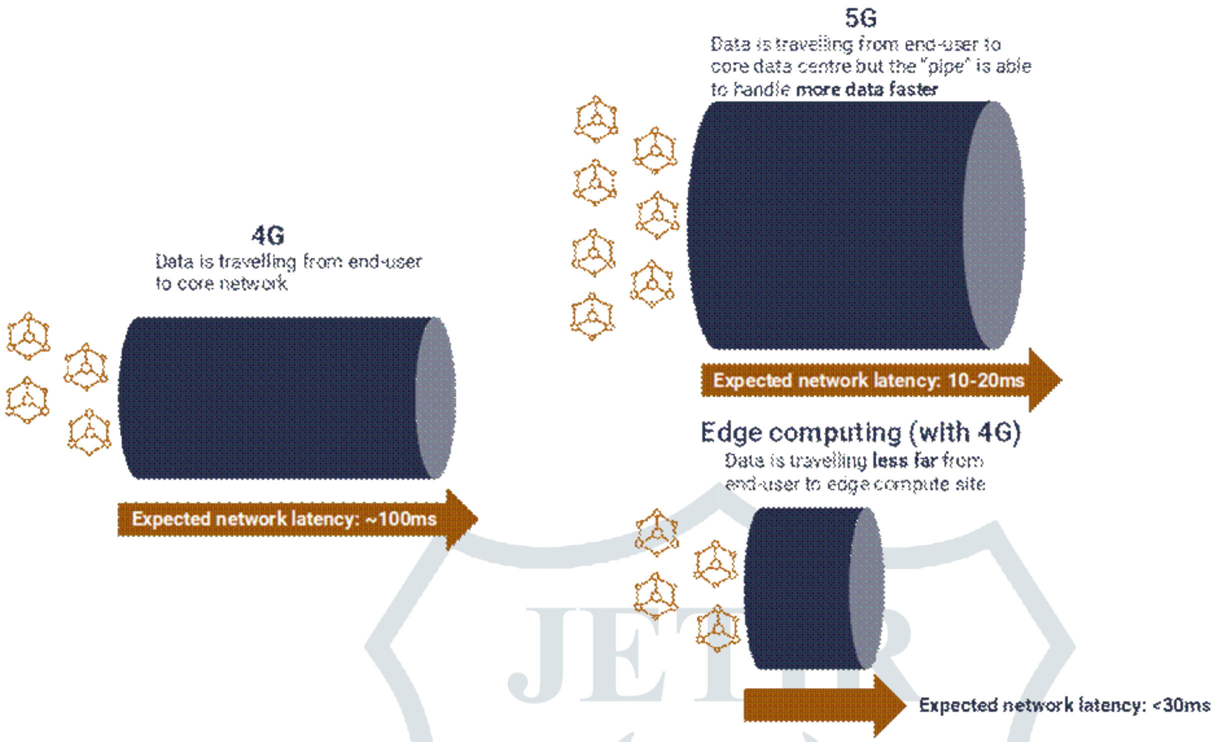
This paper intends to explore and analyze how **Edge computing** brings computation and data storage closer to where data is created by people, places and things. **Combined with 5G**, creates opportunities to enhance digital experiences, improve performance, support data security and enable continuous operations in every industry.

5G and edge computing

5G and edge computing are two inextricably linked technologies: they are both poised to significantly improve the performance of applications and enable huge amounts of data to be processed in real-time. 5G increases speeds by up to ten times that of 4G, whereas mobile edge computing reduces latency by bringing compute capabilities into the network, closer to the end user. It is inherent to 5G standards as it is the only way to meet the latency targets that have been set (1ms network latency). While telecoms operators have reported that 5G in the lab can deliver network speeds that are more than twenty times faster than LTE1, this will not reflect the experience of the average user. There are still major unknowns in how 5G will achieve these speeds – ultra-low latency standards will only be revealed in 3GPP’s Release 16 later this year. We feel it is likely 5G will rely upon edge computing to reach the targets that are being set.

1. The gradual approach operators are taking to deploy 5G – the 5G go slow cycle – will mean coverage of “full 5G” will be insufficient to cultivate an ecosystem of new applications. However, edge could seed a 5G market even before widespread coverage.

Figure 1: Edge computing will reduce network latency by processing data closer to the end-user



What is the 5G go slow cycle?

Operators in different markets have diverse approaches towards 5G. research shows operators in the US, China, Korea and Japan are chasing 5G most enthusiastically and back that enthusiasm with real investment.

This is characterised by T-Mobile's CTO Neville Ray: "*We are all in on 5G. Every dollar we spend is a 5G dollar, and our agreement with Nokia underscores the kind of investment we're making to bring customers a mobile, nationwide 5G network.*"

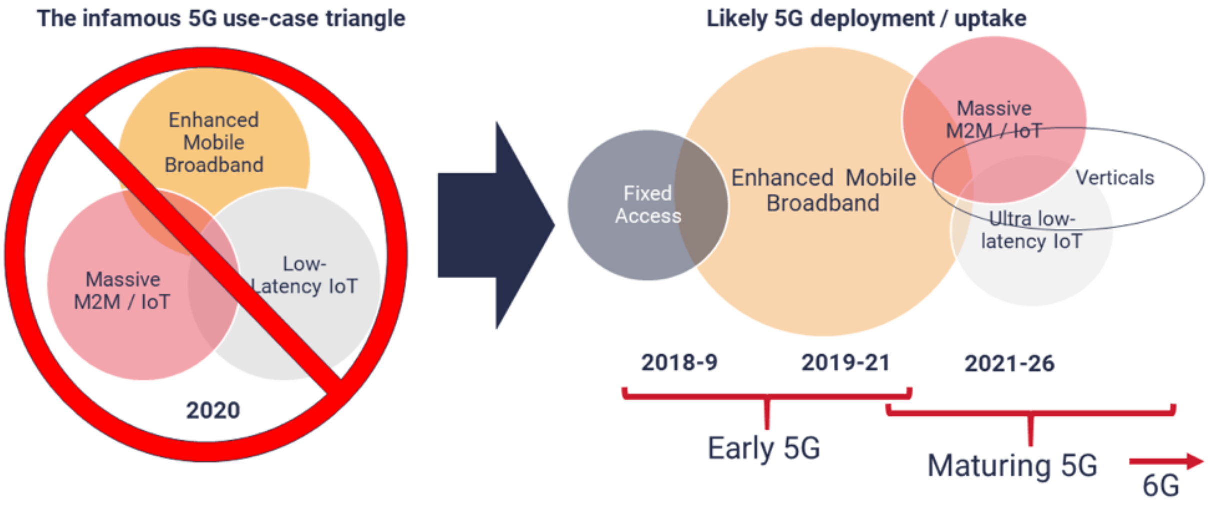
Initially, operators in Europe, such as Orange, Telefonica, Veon, Deutsche Telekom, BT, Telenor and Vodafone, did not hold back from taking part in the 5G jamboree but were more muted in their approach. Most importantly, they continue to invest heavily in their LTE networks (for example in LTE-A Pro) in their pursuit of a "glide path" to 5G. More recently, some have announced early 5G deployments in major cities, such as BT (EE) and Vodafone.

BT's Consumer Division CEO, Marc Allera, talks wisely about learning lessons from the past

"3G technology was overhyped, the price was overhyped and the consumer was underwhelmed by the experience. That said, we learnt some valuable lessons, and opened people's eyes to doing more on their mobile. 3G taught us a lesson in financial prudence while 4G taught us how to deploy rapidly at scale."

Operators in South East Asia, the Middle East and Europe are (privately) cautious about the prospects for rapid roll-out of broad coverage. They think 5G offers compelling cost advantages over LTE for enhanced mobile broadband (eMBB) in denser urban locations and potentially fixed wireless access (FWA) in certain suburban ones (where fibre has not been widely adopted). We set this out in our recent report: 5G: 'Just another G' – yet a catalyst of change.

Figure 2: Many 5G benefits will not been delivered until 2025 and beyond >



The timeline for 5G standards, development and deployment is staggered. Not all use cases will be commercially available at the same time. This creates a predicament over how to expand beyond these 5G islands. Without wider coverage, how can operators catalyse ecosystems to invest in new services (low-latency, immersive or Massive MIMO)?

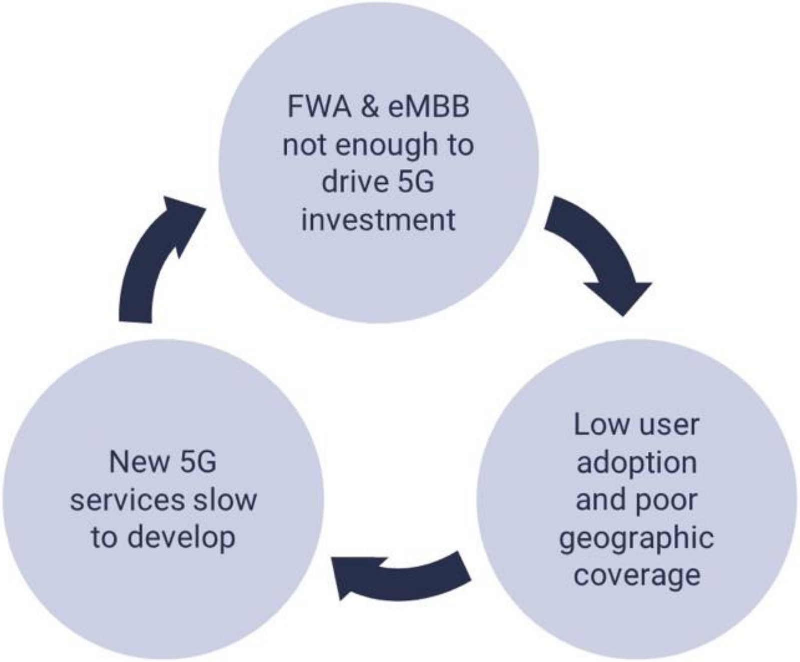
Plus, much of the promise of 5G rests on being able to enable new use cases that are not just a step up from LTE mobile broadband, for example those that are characterised as ultra-low latency and massive M2M. Enterprises are already demanding these capabilities, yet the telecoms industry is still trying to agree standards for these later releases of 5G.

The role of 5G in the manufacturing industry – $740Bn opportunity by 2030

At the same time, without coverage of 5G in these new use case domains and end-users with 5G devices, application owners will not be incentivised to develop applications that use the technology.

These questions, particularly when considering the massive CAPEX investment that 5G requires, have only be exacerbated by the COVID-19 pandemic. At a critical time, 5G spectrum auctions in countries such as Spain and Austria have been delayed because of health and financial concern. Operators are trying to navigate uncertainty in how our global economy will recover from the virus, and many will not accelerate 5G investment and innovation until there’s more certainty on how their business will be affected.

Figure 3: The 5G go slow cycle has only be exacerbated by COVID-19



Short term: Mobile edge computing is a key technology towards 5G

Somewhat counter-intuitively, 4G and edge compute can help accelerate 5G coverage and device adoption rather than delay it. Astute operators can mix commercialised edge compute and 4G to create “good enough” capabilities for 5G services to fall back onto. And these will overcome limitations in 5G geographic or device coverage. While 4G + edge compute won’t be as high performance or cost efficient as 5G, they will be good enough to provide the basis for securing greater demand for services.

Examples of 5G edge computing applications

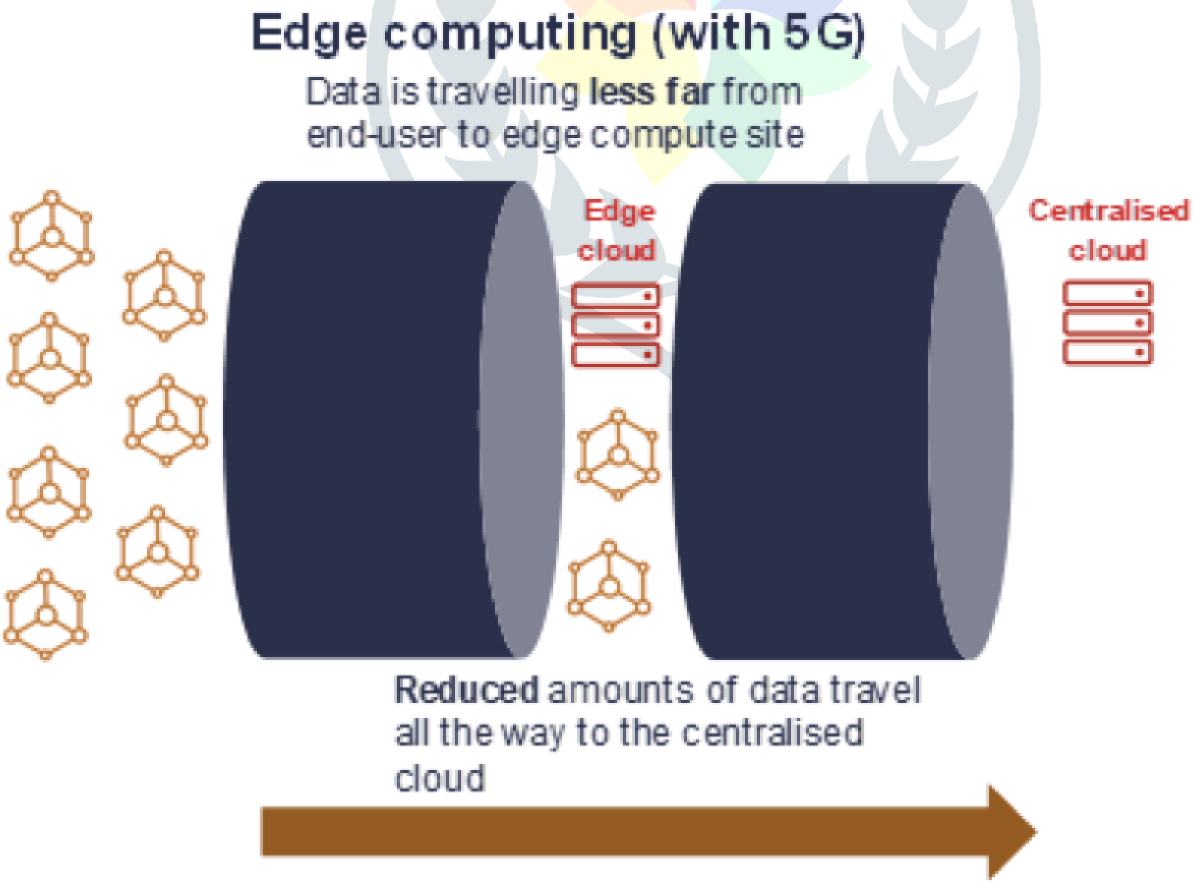
-
  - A gaming company that wants to create a distinct gamer experience needs high-end GPUs mainly found in 5G phones. The company may find its addressable market is limited because there are few existing devices. By using edge compute, it could address a broader base of 4G devices, at least initially, and succeed ahead of wider device adoption. More information on cloud gaming can be found in our recent report Cloud gaming: What’s the telco play?
  - A maintenance company wants to introduce lightweight AR headsets to its teams in the field to provide them with relevant information in real-time. The company needs to be sure that these could be used acceptably outside 5G coverage areas. The “5G” experience might be enough to convince the company of AR’s potential, but scaled adoption will require coverage. More information on this particular use case and how edge computing can impact it can be found in our recent report What edge developers want from telcos now.
  - A manufacturer wants to use 5G IoT sensors to monitor the machine tools in its plant in real-time. The manufacturer needs to standardise processes across other plants and scale this technology, however these

plants may be in remote areas without 5G coverage. Edge computing would enable real-time processing of data using devices on 4G networks, which could then move to a 5G network in the long term.

Short term: Mobile edge computing is a key technology towards 5G

In the long term, the question will not be 5G or edge computing, but how to combine both technologies to deliver use cases. This is for two key reasons:

1.
    1. To achieve ultra-low latency, necessary for further out use cases like autonomous drones or remote telesurgery, the combination of 5G and edge computing will be necessary. This means both a bigger, faster pipe in conjunction with a shorter distance for the data to travel.
    2. Edge can enable operators to change their backhaul business models. For data-heavy applications, such as those requiring high-definition video or extensive data analysis, even with 5G, sending data constantly back to the cloud will be expensive and deteriorate the customer experience. Instead, data could be filtered out, with the full stream travelling only as far as a local edge site, before being analysed, rationalised, and only what is necessary streamed and stored in the centralised cloud. If operators decouple access and backhaul connectivity pricing, offering reduced backhaul for streaming to the edge rather than the centralised cloud, they can incentivise application developers to use edge computing sites, rather than on-premise or on-device workarounds. For more information, see Decouple from cloud connectivity to succeed in edge compute.



Conclusion

In summary, 5G needs edge computing to drive demand for its services. Today, there are only nascent markets for the types of applications 5G enables: augmented reality, mass IoT, robotics, AUVs/drones, etc. Edge computing can provide developers an environment to create the 5G applications that do not exist today even without "full 5G" being available yet.

References

1. Hamilton, Eric (27 December 2018). "What is Edge Computing: The Network Edge Explained". cloudwards.net. Retrieved 2019-05-14.
2. Gartner. "Gartner Trend Insights report 2018" (PDF). Gartner. Archived (PDF) from the original on 2020-12-18. Retrieved 2021-05-26.
3. "Globally Distributed Content Delivery, by J. Dilley, B. Maggs, J. Parikh, H. Prokop, R. Sitaraman and B. Weihl, IEEE Internet Computing, Volume 6, Issue 5, November 2002" (PDF). Archived (PDF) from the original on 2017-08-09. Retrieved 2019-10-25.
4. Nygren., E.; Sitaraman R. K.; Sun, J. (2010). "The Akamai Network: A Platform for High-Performance Internet Applications" (PDF). ACM SIGOPS Operating Systems Review. 44 (3): 2–19. doi:10.1145/1842733.1842736. S2CID 207181702. Archived (PDF) from the original on September 13, 2012. Retrieved November 19, 2012. See Section 6.2: Distributing Applications to the Edge
5. Davis, A.; Parikh, J.; Weihl, W. (2004). "EdgeComputing: Extending Enterprise Applications to the Edge of the Internet". 13th International World Wide Web Conference. doi:10.1145/1013367.1013397. S2CID 578337.
6. Gartner. "2021 Strategic Roadmap for Edge Computing". www.gartner.com. Archived from the original on 2021-03-30. Retrieved 2021-07-11.
7. "IEEE DAC 2014 Keynote: Mobile Computing Opportunities, Challenges and Technology Drivers". Archived from the original on 2020-07-30. Retrieved 2019-03-25.
8. MIT MTL Seminar: Trends, Opportunities and Challenges Driving Architecture and Design of Next Generation Mobile Computing and IoT Devices
9. "What is fog and edge computing?". Capgemini Worldwide. 2017-03-02. Retrieved 2021-07-06.
10. "ETSI - ETSI Blog - What is Edge?". etsi.org. Retrieved 2019-02-19.
11. "CloudHide: Towards Latency Hiding Techniques for Thin-client Cloud Gaming". ResearchGate. Retrieved 2019-04-12.
12. Anand, B.; Edwin, A. J. Hao (January 2014). "Gamelets — Multiplayer mobile games with distributed micro-clouds". 2014 Seventh International Conference on Mobile Computing and Ubiquitous Networking (ICMU): 14–20. doi:10.1109/ICMU.2014.6799051. ISBN 978-1-4799-2231-4. S2CID 10374389.
13. Patrizio, Andy (2018-12-03). "IDC: Expect 175 zettabytes of data worldwide by 2025". Network World. Retrieved 2021-07-09.
14. "What We Do and How We Got Here". Gartner. Retrieved 2021-12-21.

15. Ivkovic, Jovan (2016-07-11). The Methods and Procedures for Accelerating Operations and Queries in Large Database Systems and Data Warehouse (Big Data Systems) (PDF). National Repository of Dissertations in Serbia (Doctoral thesis) (in Serbian and American English).
16. Shi, Weisong; Cao, Jie; Zhang, Quan; Li, Youhuizi; Xu, Lanyu (October 2016). "Edge Computing: Vision and Challenges". IEEE Internet of Things Journal. 3 (5): 637–646. doi:10.1109/JIOT.2016.2579198. S2CID 4237186.
17. Merenda, Massimo; Porcaro, Carlo; Iero, Demetrio (29 April 2020). "Edge Machine Learning for AI-Enabled IoT Devices: A Review". Sensors. 20 (9): 2533. doi:10.3390/s20092533. PMC 7273223. PMID 32365645.
18. "IoT management". Retrieved 2020-04-08.
19. Garcia Lopez, Pedro; Montresor, Alberto; Epema, Dick; Datta, Anwitaman; Higashino, Teruo; Iamnitchi, Adriana; Barcellos, Marinho; Felber, Pascal; Riviere, Etienne (30 September 2015). "Edge-centric Computing". ACM SIGCOMM Computer Communication Review. 45 (5): 37–42. doi:10.1145/2831347.2831354.
20. 3 Advantages of Edge Computing. Aron Brand. Medium.com. Sep 20, 2019
21. "Vehicle-Mounted Active Denial System (V-MADS)". Global Security. Archived from the original on 5 March 2008. Retrieved 2 March 2008.
22. "DVIDS – News – New Marine Corps non-lethal weapon heats things up". DVIDS. Retrieved 1 November 2014.
23. "Effects on the human body: Extremely low frequency RF | Radio Frequency | Radio Spectrum". Scribd. Retrieved 8 March 2021.
24. Kinsler, P. (2010). "Optical pulse propagation with minimal approximations". Phys. Rev. A. 81 (1): 013819. arXiv:0810.5689. Bibcode:2010PhRvA..81a3819K.