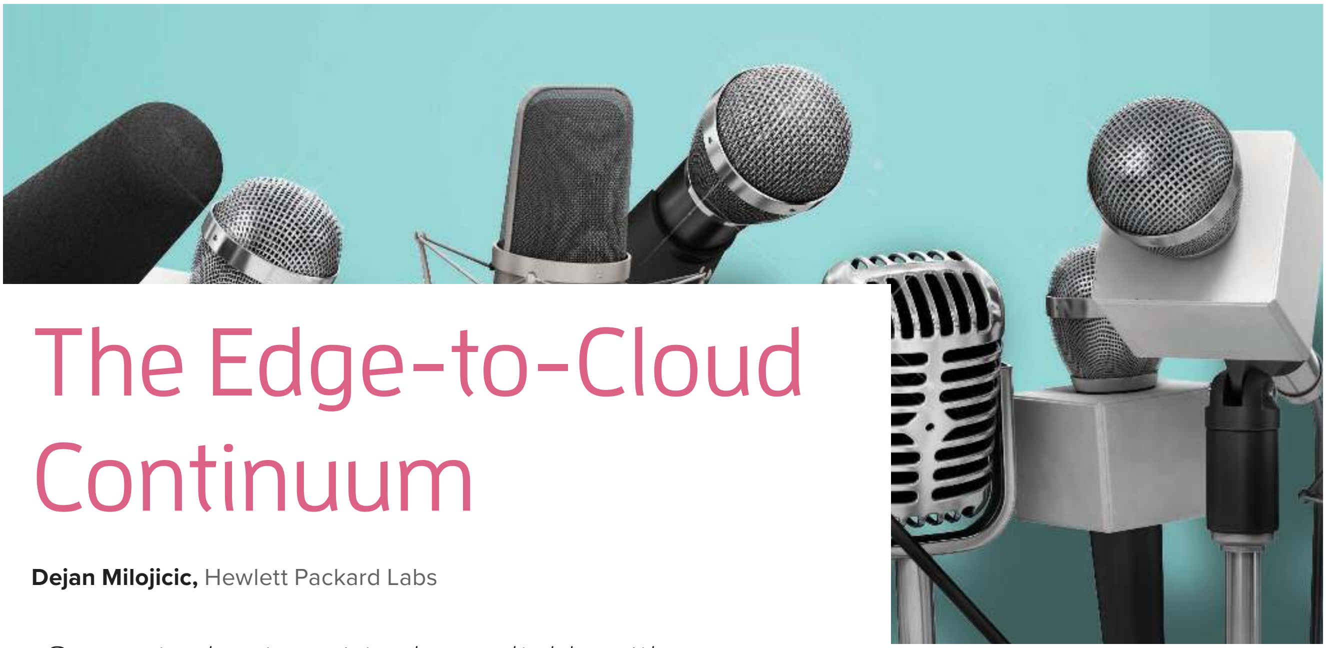

The Edge-to-Cloud Continuum

Dejan Milojicic, Hewlett Packard Labs

Computer hosts a virtual roundtable with three experts to discuss the opportunities and obstacles regarding edge-to-cloud technology.

In *Computer*, virtual roundtables (VRTs) are virtual panels. We ask a series of questions, typically roughly a dozen, to a group of experts via email to get their perspective on an important technology that is relevant to businesses and humanity. One difference between VRTs and real face-to-face panels is offline versus real-time interaction. A VRT is different from a face-to-face-panel, where answers from one participant can immediately affect the answers of others. In a VRT, each participant contributes his or her own input independently; however, we allow the panelists to eventually see input from the other contributors and evolve their thinking. Therefore, the dynamics are somewhat slower and perhaps more thorough. In other VRTs, editors kept the process entirely separate and independent until publication.

In this VRT, our topic is related to edge-to-cloud technology, including critical technological, business, and humanitarian aspects that involve a connection from the source of data at the edge to the ultimate information and processing in the cloud and everything in between. For the edge-to-cloud VRT, we invited three experts to respond to 12 questions. Their written responses may have undergone minor edits; however, as editors, we attempted to keep their words as verbatim as possible. The three experts are Tom Bradicich (Hewlett Packard Enterprise), Adam Drobot (OpenTechWorks), and Ada Gavrilovska (Georgia Institute of Technology). (See "Roundtable Panelists.") It is important to note that the opinions of the experts are their own, with no input from the article editor. We hope the readers who are interested in technologies from the edge to the cloud will find these questions and responses insightful.

COMPUTER: The pendulum from centralized to decentralized has been moving back and forth throughout the history of computing, with mainframes, grids, and clouds on one side and clusters, peer-to-peer networking, and the edge on the other. Where do you see immediate future of computing? What are some of the lessons from the past that we should keep in mind?

TOM BRADICICH: Technology companies must keep their eye on two dynamics. The first is the demand side

Digital Object Identifier 10.1109/MC.2020.3007297
Date of current version: 21 October 2020

 0018-9162/20©2020IEEE

ROUNDTABLE PANELISTS

Tom Bradicich is a vice president, a Hewlett Packard Fellow, and the global head of the edge and Internet of Things (IoT) labs at Hewlett Packard Enterprise. He serves on the board of directors of Aspen Technology and the advisory board of the University of Florida, and he is an advisor to Womenincloud.com and the chair of the IoT Community industry. Contact him at tom.bradicich@hpe.com.

Adam Drobot is a technologist and manager with more than 40 years of experience in industry, government, and academia. He spends his time on start-ups, strategic consulting, and industrial associations. He is a Senior Member of IEEE and a fellow of the American Physical Society. Contact him at adam.drobot@opentechworks.com.

Ada Gavrilovska is an associate professor in the School of Computer Science, Georgia Institute of Technology. She currently serves as an associate editor of *IEEE Transactions on Cloud Computing* and a program cochair of the USENIX Annual Technical Conference. Contact her at ada@cc.gatech.edu.

of changing customer sentiments and needs, and the second concerns supply-side changes in technology capability, cost, and new applications. This is a fine balance. Frequently, technology vendors can share new ideas about which customers, who are wholly preoccupied with their businesses, are unaware. With distributed clouds and edge compute, I've personally seen this where we (supply side) introduced a new product category to accommodate pent-up customer demand to process and act on data right at its source. And now we're building the capability to cloud manage these remote infrastructures. This is profound, as not long ago the marketplace was limited to the paradigm of "send everything to a remote cloud." See my article at https://www.synnexcorp.com/stellr/blogs/7-reasons-to-compute-at-the-edge/.

ADAM DROBOT: A significant trend and precursor to what we have today was the emergence of client-server systems during the early 1980s. I remember an important article by John Markoff in *The New York Times* from 1991 called "The Attack of the Killer Micros." It recognized the domination and explosive growth of PCs and workstations and capped a major shift to decentralization in the patterns of how computing is done. Despite this, centralized computing systems never went out of style. Underlying the changes have been shifts in technology and, just as important, new business models. There are several ideas that are important here.

1. Computing, as used in practice, is hierarchical, heterogenous, and distributed. The demand for cycles is still exponentiating. It is a continuum that runs from specialized embedded devices to highly capable, standards-based individual terminals; on-premise data centers; the edge (gateways, edge computing, fog, and mist); the larger data centers in the cloud; and then to hyperscale data centers and high-performance computing (HPC). The advances in technology are happening up and down the hierarchy, but it's worthwhile to chop the computing space into three buckets. All three are growing but have very different dynamics. The three would be specialized computing, where the computing fabric is highly customized and often proprietary—the embedded world; computing that relies on standards and volume—general-purpose infrastructure; and specialized facilities focused on extremes in performance, where custom processors and architectures dominate—HPC. There is a trend to migrate applications, wherever possible, to general-purpose computing infrastructure and, at the same time, create software frameworks and platforms that exploit the complete continuum.
2. Computing technology continues to evolve, and exponential improvements are happening across multiple axes, not just Moore's law. The big deals are virtualization and compartmentalization, enabling multipurpose, multitenant computing. Another direction is running on bare iron, using

large arrays of low-power, low-cost processors. Miniaturization and energy efficiency continue apace, driving the deployment of general-purpose computing on mobile platforms. The growth of multicore processors is also dramatically increasing compute power up and down the hierarchy. The proliferation of various accelerators, whether they are graphics processing units, tensor processing units, or artificial intelligence (AI) and machine learning (ML) engines, is enabling powerful new applications. While there is doubtlessly more on the technology front, it's worthwhile to mention that software, whether in applications or operating systems, is also on an exponential curve in terms of what it can do. The lesson here is that, as much as has been accomplished, we are nowhere near done with technology. In the wings, but longer term, is the promise of quantum computing.

3. The business landscape has also changed. To use an analogy, during the past century, we saw the emergence of utilities and common infrastructure across the globe. These provided water, gas, electricity, and roads. In addition, common services, such as ambulances, firefighting, waste removal, sewage treatment, and many others, became part of the ordinary. Life as we know it without that infrastructure would not be possible. Computing has reached a point where it is now well on the way to becoming a ubiquitous and well-established infrastructure. The fact that someone can start a new business or add a business line and have access to powerful data storage, processing, and software capabilities without having to integrate the capabilities from scratch has lowered the threshold for what it takes to be successful and create incredible opportunities for innovation. Along with this has come the open source movement that redefines, in many ways, how businesses cooperate and compete at the same time—the underpinnings for software as a service, infrastructure as a service, and platforms as a service.

ADA GAVRILOVSKA: It's curious to reflect back at the time when cloud computing was emerging and to remember many conference keynotes promising, by 2020, all that will be needed will be stored in just a dozen megadata centers. This, of course, did not consider the extent of the growth of data, in large part accelerated precisely because of cloud computing, and the renaissance of AI that keeps finding new ways to make use of all these data. True, there is tremendous compute power today consolidated in just a dozen megadata centers, but they are far from adequate to address the demand of modern workloads. With computing permeating every aspect of industry and society, the richness of workloads is such that there is sufficient demand for many of these paradigms to coexist and complement each other. When I think about the cloud and the edge today, there are fundamental reasons why they must coexist: the speed of light and data movement energy will limit what cloud-only solutions can deliver. The law of big numbers will continue to make megadata centers a more cost-efficient solution for workloads that are not limited by latency and the backhaul bandwidth.

***COMPUTER*:** What are the key market forces that drive this continuum of computing? For example, how does delivery as a service impact businesses that are moving to operational expenditure (OPEX) and away from capital expenditure (CAPEX) and so forth?

BRADICICH: It's clear today that as a service (aaS) affords the value propositions of lowering CAPEX, agility of skills, and obsolescence protection. This has been proven in the cloud and the data center. Today, we can impute these benefits on infrastructures at the edge and on remote distributed clouds. However, the challenges of remote and edge services delivery differ greatly from those in data centers, such as low bandwidth, limited on-site IT skills, hostile environments, and connectivity reliability. In addition, the edge is more complex, as it includes operational technologies (OTs), such as control systems and industrial protocols, that are not seen in the data center. Comprehending both IT and OT systems in aaS offerings provides great customer value and differentiation.

DROBOT: The most important dynamic is digitization and the newly emergent infrastructure services and technologies that support it. For simplicity, this consists of manufacturing, goods and products, services, and processes. In each of these, enterprises and their customers are instrumenting what they do, generating more data than ever, and then extracting value by solving the hard problems that computers are great at. The first area affected in a fundamental way is the visibility into the way we conduct business, the consequences of how we design and build things, how our customers use the goods and services we provide, how the goods and services perform, and finally the experience that our customers have. What the data then enable us to do is improve efficiency in manufacturing and service delivery, avoid disruptions, optimize the design of goods and services, find better approaches to maintenance and repair, dramatically improve the customer experience, innovate with new functionality and capabilities, and restructure whole ecosystems with better business models. In some ways, what is attractive for enterprises is that they can take on more complex challenges and deliver value individualized to the customer while hiding that complexity. The result is better options where CAPEX is deployed; new tradeoffs between CAPEX and OPEX; optimizing CAPEX and, at the same time, improving the efficiency and effectiveness of OPEX spending; and most of all, innovation in new products and services that serve customers better, contribute to the customer experience, and raise the quality of life on a societal scale.

GAVRILOVSKA: There are a number of intertwined market forces driving these developments. From the providers' perspective, I think the main questions are around the business models needed to sustain infrastructure investments. Many current examples of enterprise adoption are largely driven by managing operation costs related to data movement, network expenses, and reliability. Emerging applications, such as augmented reality (AR) and virtual reality (VR), need edge computing to meet their core functional requirements. I think that the real transformation will come by moving beyond point solutions to a place where edge computing is available through models similar to current cloud computing offerings. In recent months, we have seen announcements of partnerships among mobile network and cloud providers. If such partnerships lead to the cloudification of mobile networks, this will be a major accelerator for reshaping the next phase of the edge–cloud continuum.

***COMPUTER*:** What are the key market verticals/industries that are suitable for this approach? Oil and gas? The Industrial Internet of Things (IoT)? Cybersecurity? What else?

BRADICICH: We live world of perpetual connectivity, pervasive computing, and precision control. These three Cs are driving the edge-to-cloud continuum in virtually all industries and even our homes. Verticals, such as industrial manufacturing and oil and gas, can benefit greatly as we unleash insights and act on the data in remote edges in real time (for example, plant floors and drill rigs). This results in higher production and cost efficiencies, which accelerate the time to revenue and reduce expenses.

Emerging applications, such as augmented reality and virtual reality, need edge computing to meet their core functional requirements.

DROBOT: For me, this is a short answer. I can't think of a vertical that will not be touched or many horizontal services that will not go down this path (such as security). There is a significant gradation in where the verticals are in their progression to digitization and exploitation of the computing continuum. Eventually they will all get there. What it means is that we have only scratched the surface. In areas such as industrial manufacturing, any company that has not internalized this is at risk. In areas such as retailing, for companies that did not jump on the bandwagon, the carnage is well under way. In what we may have thought of as technological backwaters, such as small-scale manufacturing, it is an amazing revolution. It's important to understand that adoption will roll out for the next 20–30 years. It will come in waves of increasing capability in the underlying technologies and yet-undiscovered applications in the way it is used. What will be deployed every 10 years will likely eclipse the sum total of what was deployed in all previous decades.

GAVRILOVSKA: The use cases that are reaping benefits from an edge–cloud continuum are so diverse; this clearly points to the fact that there isn't a single market vertical that will drive this space. If anything, what we see are efforts toward the consolidation of the underpinning technologies, including some forms of standardization, to make it possible to cross-pollenate the technical advances across market domains to further bootstrap the ecosystem and to make infrastructure and service deployments broadly available and commoditization economically viable.

***COMPUTER*:** What are the key programming paradigms for edge-to-cloud technology? Services computing? Do we need new middleware, or will existing versions suffice?

BRADICICH: The portability of applications and middleware is key to leveraging the many compute domains in the edge-to-cloud continuum. There's value in executing workloads at the edge and in the cloud and in the many places in between. It's clear that the emergence of containers helps here. However, there's currently customer demand for "stack identicality," whereby the same legacy software stacks running in data centers can run at remote edges and in distributed clouds. This affords the customer benefits of having identical IT skills, leveraging volume software licensing agreements, and having one software part number to deploy and manage. This also benefits the supply side, as independent software vendors don't have to develop pared-down "edge versions" of their software.

DROBOT: This is where the battle is today. To take full advantage of what's possible, there is a steep learning curve. We do have things that work and deliver value, and hence a good case for a return on investment is already established but far short of what it could be. The issues are the modularity of software and its reuse, composability, and orchestration. At a higher level, it is the coalescence of IT and OT concerns. Platforms and frameworks are part of the answer, but they tend to be too inflexible. I would say that we are searching for something that's a long-lived set of middleware, protocols, and new tools above the distributed computer operating system level that enable the rapid formation and dissolution of computing networks at will. There is also a tight interweaving of other technologies with computing: storage, communications, sensors, actuators, and human interfaces. Specifying application programming interfaces does not address what's necessary. As a guess, I would venture that most applications being built today grapple with compute needs that involve a sense of time and distance and that need the full hierarchy of computing. The requirements and tradeoffs are complicated, and the tools to do those tradeoffs in any automated fashion have a way to go. This is a key question and exposes a challenge in the way we think about computing.

GAVRILOVSKA: With decades of distributed systems, particularly given the current proliferation of cloud computing, there is a plethora of mature and sophisticated technologies addressing the requirements of different workloads, including batch analytics, interactive, streaming data, and so on. We should look to leverage these, not just because of prior investments in these technologies but because of the existing applications and developer base. To support the continued rapid development of the edge ecosystem and accelerate its economic viability, it will be critical to seamlessly deliver benefits to existing workloads and programming frameworks. To achieve this, at the system level, we will need technologies that will open up flexibility in realizing cloud- versus edge-based deployments for existing code bases, with as little developer effort and retraining as possible.

At the same time, there are many inherent differences between the edge and the cloud. Current cloud-based programming frameworks lack flexibility to express geolocality, deal with the resource diversity that's present across edge and edge–cloud infrastructure elements, deal with the significant asymmetry and variability in connectivity, and so forth. A major assumption around which cloud computing frameworks are built is that of scaling to practically limitless resources, which, by definition, does not hold at the edge. This clearly points to gaps in current programming frameworks. Moving forward, we need both new programming tools and systems support for the automated refactoring of existing applications and data management frameworks so they can operate in edge-based deployments as well as new edge-native programming abstractions and middleware. These new technologies will need to enable edge applications that behave correctly while also realizing the cost and performance benefits that edge computing is promising to deliver.

COMPUTER: Is there a role for virtualization on both ends of the spectrum, and how do most of the recent approaches, such as containers and functions as a service, play the edge-to-cloud continuum?

BRADICICH: Yes (see my previous answer on stack identicality from the edge to the cloud and in between).

DROBOT: Absolutely. The ability to aggregate at the edge and in the cloud and use a compute engine to serve multiple purposes is at the heart of digitization; that means the disaggregation of functionality and virtualization when containers provide a good mechanism for isolation and preventing function interference. Moving what used to be embedded software to general-purpose computing has a tremendous benefit. That does not mean it's easy, and there are hard problems to solve, such as the control of latency and jitter and the delivery of predictable and repeatable behavior. For many verticals, computing services that go from the edge to the cloud in public/hybrid deployments are very much on the rise but often require integration with the management of noncomputing assets. That exacerbates the resources that have to go into reintegration and retesting every time a part of a solution or application is altered.

GAVRILOVSKA: Virtualization is a powerful technique for masking any diversity and dynamism in the available resources from upper levels of the stack. When considering the edge computing landscape, we see a great diversity of technologies coming together across the full stack, from application frameworks and services to hardware-level designs. At the same time, we can expect support for multitenancy to be unavoidable at the edge. Look, for instance, at Chick-fil-A. The fast-food chain has generated a lot of attention with its early adoption of IoT and edge computing in its restaurants, bringing in modern technologies, such as Kubernetes and Docker containers, to support a diverse mix of workloads for anything from long-term data analytics to real-time food preparation. To support the mix of workloads, enable workloads to be easily ported to different infrastructure

A consensus is emerging around the adoption of container-based technologies and orchestrators, such as Kubernetes.

elements, and provide for online adaptation to variability in the demand for different services and resources, virtualization will inevitably play a role at the edge.

Looking at recent developments, despite the many verticals and contexts targeted by edge computing, a consensus is emerging around the adoption of container-based technologies and orchestrators, such as Kubernetes. This is evident from a number of different open source proposals backed by mobile network operators, enterprise technology vendors, and even the Linux Foundation. There is, indeed, an opportunity to extract additional efficiencies and performance through the shared use of runtime-level functionality via higher-level function-as-a-service models. One important aspect to keep in mind, though, is the stateful nature of many edge functions, which makes a case for edge-native solutions that will accommodate such workloads in combination with the demand for high responsiveness and the efficient use of backhaul bandwidth.

COMPUTER: How does AI benefit from deployment from the edge to the cloud, and what are the new challenges, such as model training at the edge, elastic models, distributed learning, and so on?

BRADICICH: Consider some first principles: predictions are more accurate with large data sets and samples, and the data at the edge are more voluminous than all other types of big data combined. Hence, the predictive dimension of AI will be highly efficacious when applied at the edge and at remote data centers. To mine insights and act in real time, ML (which is very compute intensive) must also be done remotely at the sources. This, in turn, requires enterprise-class compute and HPC at the edges to avoid the thrashing and latency of recompiling ML models back in the remote cloud.

DROBOT: When someone says *AI* today, I have no idea what he or she means, so for the purpose of answering this question, I am assuming that we are talking about some form of ML. A typical example is to use the results of an AI model that is close to an endpoint for control, which means a fast timescale and close to the edge. If it's in a car, that means it has to be on the vehicle itself. If it's something like a telecommunications radio access network, then it can be at an aggregation point where the distance between a tower and the compute engine is determined by latency and jitter constraints. The management system for the applications is likely to be at an intermediate point (a fog installation) that collects, analyzes, and routes data; handles less-time-sensitive functions; and provides updates to the control unit. Finally, the heavy-duty processing for ML is likely to need access to computing at a larger level and massive storage, and it is much more likely to be done in the cloud, if not a hyperscale facility. I would associate this with processes on a much longer timescale. As mentioned in the answer to a previous question, what is needed is a holistic approach for how such applications are written, the development of middleware that automates and speeds up the development and update cycle, and the ability to manage other noncomputing but necessary assets.

GAVRILOVSKA: One ingredient in AI is data. If we look at where data are coming from, it's from sensors, devices, the contexts where data are used, and so on. Growth predictions for data volumes generated at the edge are staggering, and from a cost perspective, we understand that blindly moving all this data to cloud-based ML frameworks is not sustainable. It is equally ineffective to blindly discard parts of this data in random manners. The proliferation of edge computing infrastructure opens up opportunities to operate on and process more data, despite any cost-, performance-, and privacy-related data movement constraints. For instance, the edge provides for distributed deployments of ML frameworks that facilitate the localization of where AI functionality is executed and that create contextualized models tailored for the operational environment. There is a lot of momentum in the research community toward revisiting the distributed ML framework originally built for data centers, including for geographically distributed data centers, and making them edge ready.

There are a number of challenges here to fully realize the opportunity for localized learning and low-latency inference serving that can be derived from the edge. Some are related to the limited resource footprint at the edge. As a result, the software–hardware codesign for AI and the integration of accelerators will be important. Issues that affect how data, models, and processing are shared across AI clients and applications are similarly important. All of this will arguably lead to greater diversity in the capabilities of the edge element that must be adequately dealt with at the level of the distributed orchestrators controlling edge deployments. Another set of challenges is related to dealing with different types of biases that may arise due to contextualized use of AI. This will require new mechanisms for managing the model accuracy, fairness, and other important metrics.

COMPUTER: Cybersecurity, trust, and privacy: how do they work end to end?

BRADICICH: Security must be built into all the three *C*s: connectivity, compute, and controls. This is a challenge as the edge-to-cloud continuum opens both additional and unfamiliar points of vulnerability. Now, each of the billions of "things" (the *T* in the IoT) presents another new entrance point to the network. At Hewlett Packard Enterprise, we're architecting a combination of proven and new secure attestation of endpoints and anomaly detection using AI. We're also developing two means of reducing vulnerability. First, we're designing direct cloud-to-endpoint connectivity, which can completely eliminate vulnerable controllers and switches, and second, we're building proven enterprise-class IT security right into OT systems, which is an industry first. Allow me to suggest reading my article where I further discuss the three *C*s, the four-stage IoT/edge architecture, and the seven reasons to compute at the edge. It is available at https://iiot-world.com/connected-industry/the-intelligent-edge-what-it-is-what-its-not-and-why-its-useful/.

DROBOT: They work after a fashion but need improvement. It's unlikely that we will achieve 100% cybersecurity; that what we provide in privacy will suit everyone; and that the level of trust will be sufficient to overcome the comfort level of "server huggers," skeptics, and Luddites. It is important to keep raising the bar and just as important to anticipate what new vulnerabilities we are creating and put considerable effort into mitigating them. The hard part is that this is not just a technical issue; we have operators and users in the mix, and the bundle of cybersecurity, trust, and privacy depends on their behavior. Education and the better design of how users (mere mortals) interact with the computing environment are a must. There are a few technologies, such as homomorphic computing, on the horizon, and if they come to fruition, they may help tremendously. Last, it's important to be prepared for things to go wrong and recover gracefully, much along the lines of the National Institute of Standards and Technology security framework but with the actionable details that matter.

GAVRILOVSKA: Cybersecurity will continue to be a hard problem. In one way, edge computing massively increases the attack surface through the addition of new software stacks and hardware elements, each with their own vulnerabilities. A straightforward application of cloud-based solutions will have scalability challenges, or worse, it will lead to overheads that may obviate the expected performance benefits of operating at the edge. Leveraging hardware-level, security-centric acceleration, such as for secure enclaves and different cryptographic operations, will be critical for building edge-native solutions. There will likely be a need for new, richer trust models, as edge computing brings new stakeholders in the end-to-end pipeline, in the form of edge infrastructure and service providers.

At the same time, the emergence of edge computing will create new opportunities. The edge is closer to devices and endpoints where many attacks originate; for instance, there is an increased trend toward distributed denial-of-service attacks for IoT bots. The edge offers a landing point for deploying new detection, prevention, and containment strategies. Similarly, in terms of privacy, edge computing provides for localized data analytics without requiring data to be first aggregated in centralized cloud locations. As more edge computing solutions and offerings start becoming available, the scope of these challenges and opportunities will start becoming clearer, and I see ahead tremendous opportunities for innovation in this space.

***COMPUTER*:** Power has been a critical resource in large cloud data centers, but aggregate power quantities, and even individual edge device power capacities, are also limiting resources. Is there a holistic view of power from the edge to the cloud?

BRADICICH: Edge-to-cloud continuums can both reduce energy and better distribute the energy demand by, for example, eliminating duplicate IT resources, such as storage, by decimating data in real time and reducing network infrastructures by computing at the edge. In addition, we're pioneering the convergence of OT and IT onto the same systems chassis. Metaphorically, much as the smartphone reduces energy by physically combining a cell phone, camera, GPS, and music player, the new product category, converged edge systems, affords the same energy-saving benefits at the edge.

DROBOT: For the larger cloud, fog (intermediate computing facilities), and edge operators who have to pay the bills for power, the understanding is very good. They make their decisions on where to locate facilities by paying a lot of attention to the cost and usage of power (including multiple sources for redundancy) and to the cost and power for connectivity. In the enterprise space, with a few exceptions, the situation is much more complex and less well engineered. Where the greatest improvements are needed is for the smaller players who don't necessarily have the bench strength to deal with the issues.

GAVRILOVSKA: I don't think we're there yet. Data centers have made major advances in their power usage effectiveness. Power efficiency and power proportionality are a first-class design principle across components, including CPUs, memories, network equipment and entire data centers, and certainly for end-user devices. But having a holistic view requires a fine-grained accountability of power, and such capabilities are still missing given the workload multitenancy that exists across end devices, edge and cloud servers, and the networks that interconnect them. Achieving this is made more complex due to microservice architectures. As more computing starts happening outside data centers, answering this question will become more critical.

***COMPUTER*:** What importance and impact do 5G technologies have for edge-to-cloud deployment?

BRADICICH: The promise of 5G includes improved bandwidth, an increased number of access points, and higher reliability. These attributes will promote the deployment of remote and distributed edge clouds. These, in turn, enable more intelligence at the edge and within the continuum to execute compute- and network-intensive AR and AI in real time. There's a "spectrum of insight" from the edge to the cloud, whereby some insight is derived at the edge, some back in the cloud, and some best revealed in between within the continuum. 5G gives us more options in this regard.

DROBOT: Communications are crucial for cloud deployments and designing solutions that are viable in many of today's settings. The communication infrastructure consists of many technologies, and 5G is only one of them. The first priority, both for business reasons and to satisfy technical requirements, is to ensure that the communications infrastructure has capacity and coverage. Within that is the next-level requirement that the infrastructure must support mobile and nomadic applications; that's where 5G comes in. It brings three major improvements: a reduction in latency and jitter so that control-oriented applications are possible; a significant increase in bandwidth, enabling data to flow at the volumes and velocity demanded by emerging applications; and finally, a high degree of automation, permitting quality-of-service performance beyond the best-effort delivery on the existing Internet. In turn, this makes the IoT; the Industrial Internet; and technologies such as the digital twin, VR/AR, the exploitation of AI, and digital additive manufacturing a reality. It creates a growing market for edge-to-cloud technologies with solutions that are attractive in manufacturing, the design of products, services, and processes.

GAVRILOVSKA: Edge computing and 5G are inextricable. The use of commodity compute infrastructure in the edges of the network is a key ingredient in building cost-effective solutions for delivering the latencies and bandwidths promised by 5G; in that sense, edge computing is an enabler for 5G. At the same time, the applications that we expect will be transformed by 5G, including AR/VR, automation, smart cities and infrastructure, ultrahigh-definition video, and so forth, are precisely the prime drivers for computing at the edge since many of their end-to-end requirements cannot be practically addressed with purely cloud-based service deployments. Yes, there are many use cases for edge computing that are not tied to 5G, particularly in the context of industrial automation and certain enterprise solutions, but I see the broader availability of 5G and the workloads it will enable as an instrumental driver for the commoditization of edge computing.

COMPUTER: How is the convergence of IT and OT happening, and is edge to cloud enabling OT in new ways?

BRADICICH: It's instructive to portion IT/OT convergence into three categories: 1) process convergence, breaking silos by promoting line-of-business cooperation with the chief information officer on OT systems connected to IT networks, such as virus, security, and network support; 2) software–data convergence, where remote data and analytics results flow directly into corporate IT systems, affecting corporate operations, customer relationship management, and budgeting; and 3) physical convergence, integrating OT functions (controls systems, data acquisition, and industrial networks) into the same box as IT functions (compute, storage, and management). The customer benefits are great, including the seven benefits of convergence: less space, a faster deployment time, a lower cost, less energy, fewer systems to buy and manage, less error-prone cabling, and cross-functional new applications. See my blog at https://community.hpe.com/t5/iot-at-the-edge/7-ways-industries-benefit-from-ot-and-it-convergence/ba-p/6975867#.Xsu_vGhKg2w.

DROBOT: It is. Edge-to-cloud technology is enabling the convergence of IT and OT, but there is another important angle here. In the end, one needs IT and OT that are strongly bonded together without creating undue burdens on the frictional costs of development, deployment, and operations. The manner in which applications and software are developed in IT organizations and some digital companies rarely achieves the high levels of reliability and the stringent requirements where human safety and other aspects of criticality are concerned. You would not want your high-speed train operating from the cloud if it was hackable, if the connectivity could fail, or if you were trying to perfect your software from failures during beta testing. It may be that OT actually forces us to confront some of the inherent habits in IT and forces us to do a better job. There is likely to be an increase in the formality of the processes used for deploying edge-to-cloud technology, and the trick is to implement that formality without losing the flexibility prevalent in today's development and operations practices.

GAVRILOVSKA: I view OTs as early adopters and demonstrators of the benefits of the IoT and edge computing. In many industry verticals, industrial automation, inventory management, and predictive maintenance have been transformed through the years by automation enabled by sensors, the online processing of sensor data, and real-time analytics using on-premise edge infrastructure. These types of previously purpose-built systems are being generalized in the form of the current edge computing trends. But there are still important ways in which OTs are impacted by current developments. One way is a function of economics: open source or broadly available solutions, market competition, and access to a large skilled workforce base pave way for the deeper penetration of the edge and the IoT across many aspects of OT. Another way is through advances in reliable communications technologies, making it possible to "cut the cord" and, at the same time, increase the demand for real-time monitoring and controls. It is also important to point out AI and AR/VR. These technologies present a demand for data and low latency; they are considered important drivers for edge computing and have also had an undeniable impact on OT.

***COMPUTER*:** What are earlier unforeseen applications enabled by the edge-to-cloud continuum?

BRADICICH: First, it must be understood that enterprise-class, no-compromise IT infrastructures can be unleashed from the data center and placed at distributed edges and edge data centers. When this is fully embraced, it will shatter the predominant point of view that all data must be sent to, and processed in, centralized clouds. This annihilates a mutually exclusive object of compute-intensive versus low-latency workloads. Specifically, applications, such as video analytics and AI/ML (which are data and compute intensive) can now be applied at edge locations, resulting in a lower-latency response and action. Remember that with the three *A*s of the intelligent edge—acquire, analyze, and act—it's the acting that causes the business benefit. Acquiring and analyzing are simple means to the end of acting on the data.

DROBOT: Historically, many of the ideas we are discussing for applications that employ the cloud continuum have existed for quite some time. The big difference between what existed in the past and what exists now is that the necessary infrastructures are more mature, and the burden does not fall on a single entity to design, build up, and integrate the capability from scratch for a single purpose and operate it on its own. That used to take national means. Examples to cite would be the National Air Control System, the Sound Surveillance System array for hunting submarines during the Cold War, the NASA exploration probes, the distributed National Oceanic and Atmospheric Administration radar facilities for weather prediction, and large academic research networks that tied major facilities to computing centers.

In the commercial world, it was the back-end systems that ushered in the rise of Amazon and the retailing revolution and the emergence of the shared economy (from Airbnb to Zipcar). In the aerospace industry, it included the development of the last generation of aircraft by the two largest manufactures, Boeing and Airbus, each using hundreds of contractors around the world. While many of the examples did not use the public cloud, they relied on the wide proliferation and penetration of the Internet and the succession of crossed thresholds that made many of the ideas affordable and practical for all. It was hard to do when a typical connection was a T1 line or a 300-baud terminal connection.

GAVRILOVSKA: I am probably the most surprised by the diversity of industry segments where there has been a development of solutions based on edge computing. The examples range from traffic management to fast-food restaurants and oil and gas. That each of these use cases can benefit from edge computing is obvious. That different companies from very diverse industry domains were able to see a sufficiently big potential to make the technology investments themselves; that, to me, was not expected.

***COMPUTER*:** Could an edge-to-cloud application help humanity in preventing viruses? An example might include globally deployed sensors connected to a decentralized cloud service that could trace the spread of pandemics.

BRADICICH: The edge is where the action is. There are many benefits of capable IT and OT infrastructures with remote visualization that are deployed for this activity. Decentralized clouds and edge deployments offer three *M*s: monitor, maintain, and monetize. The limits come only from one's imagination. What would you do if you could constantly monitor customer behavior, your products, your patients, and your remote facilities? Concerning health care, underserved and geographically dispersed populations that may not have access can be included in both monitoring and health-maintenance initiatives. Concerning monetizing, such telemedicine services lower costs and increase patience reach, benefitting the health-care providers, too. See my article at https://www.cioapplications.com/cxoinsights/the-3-m-s-and-3-c-s-of-the-edge-creating-new-insights-and-new-business-nid-1019.html.

DROBOT: It is already happening. There are innumerable organizations and companies that track medical data around the globe to predict the spread and progression of diseases. An example is a platform called Cloudbreak. Others include the installation at the Johns Hopkins Coronavirus Resources Center, the World Health Organization, our own Centers for Disease Control and Prevention, and almost all the major medical research institutions in virology. Having said that, cloud computing is no magic fix. The underlying science and understanding of how viral diseases impact human beings and reservoir hosts are not resolved, and computing may play a greater role in helping there.

What's amusing is that people, when they don't feel well, will often tweet or post what they are feeling and what they perceive their symptoms to be on social media. Natural language processing with data from crawling the web is surprisingly effective at picking up the patterns. It is great at ringing the alarm bells but not quite as good at figuring out how to deal with a disease. The medical research, clinical, and public health community, on the other hand, is very engaged in applying the computing infrastructure to everything from understanding the structure of specific viruses to contact tracing. One of the successful deployments, running off the cloud, is at the airports in Taiwan: infrared and hyperspectral imagers were posted at entry points and originally developed to contain severe acute respiratory syndrome. The system identifies passengers with elevated temperatures. This question actually raises an important point: the engagement between the end users and the computing services providers must be much closer; it's how knowledge of what's possible meets the knowledge of what's doable and practical in a specific context.

GAVRILOVSKA: Edge computing can, indeed, be leveraged for some type of privacy-preserving or trusted, decentralized contact monitoring and tracing applications, which could prevent a spread of the virus of the form we're witnessing now. In fact, it may already be used for this purpose in some parts of the world. There are also current demonstrations of the use of edge computing, with compute elements placed on lampposts and in drones, for virus detection via the dynamic generation of heat maps based on infrared cameras, and we know that detection, along with tracing, is key to controlling the spread of the virus.

COVID-19 also exposed many gaps in connectivity, and connectivity is a problem that the edge–cloud continuum is designed to address by trading (it is much cheaper and more easily deployable) compute resources to reduce the demand on connectivity. While we still ultimately need other ways to build immunity to eliminate the dangers of the virus, edge computing can solve many related problems. During shelter-in-place times, connectivity is not just about advanced automation or infotainment; it's about basic services, such as health care, education, and livelihoods. We all look forward to a future when we will be able to be back at work, but some of the changes that we are witnessing now will likely remain. ■

DEJAN MILOJICIC is a distinguished technologist at Hewlett Packard Labs. Contact him at dejan.milojicic@hpe.com.

Digital Object Identifier 10.1109/MC.2020.3027329