Edge Computing: A Building Block for Pervasive Computing

Abstract

Today, hyperscale cloud providers offer ubiquitous computing to support different types of applications and systems. The resulting collection, forwarding, and processing of massive quantities of data allows companies to operate more intelligently, act swiftly, and make informed decisions. The drawback: sending all the data to the cloud can lead to latency - a delay prior to the actual transfer of data. Moreover, the cloud's reach is currently limited in remote areas, requiring computing power to be moved from distant datacenters to the edge of networks.

This paper discusses how organization can leverage edge computing to achieve sub-second responses that are crucial to maximizing application potential, user experience, quality of service, and ultimately, the success of the digitized infrastructure.

Edge Computing: The New Frontier in Data Processing

Edge computing pushes computing resources physically closer to devices by deploying either a device that performs the computation itself, a computing node that acts as a miniature cloud, or a combination of the two. Such core-computing node-edge architecture helps reduce latency, minimize data threats, improve bandwidth, and decrease costs, while increasing privacy-policy compliance and reliability (see Figure 1).

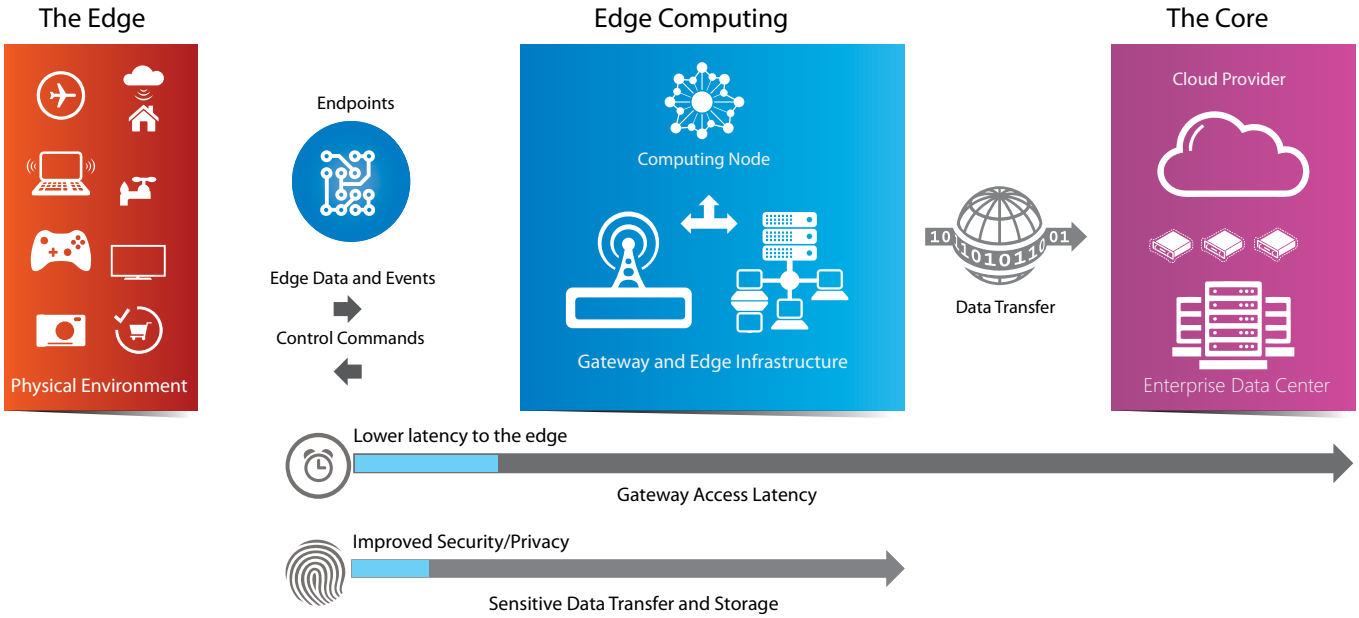

Figure 1: A high-level view of the edge computing architecture

Given the enormous amount of data collected by edge devices, pre-processing it, including sampling and blanking, right at the edge device or at a computing node, can greatly reduce the volume of data to be transferred. Not only does pre-processing help cut down on energy and bandwidth costs, it also ensures privacy and security by redacting sensitive and personal identifiable information (PII) from data records and reducing the application attack surface.

Unlocking Value from Edge Computing: Some Use Cases

Edge computing can significantly improve application performance and user experience by moving both computing and data closer to the user in order to offer dynamic and customized content. With intelligence pushed to the edge of the network across a spectrum of edge devices, edge computing enables interesting use cases across industries, some of which are:

IoT gateways for industrial automation

With edge computing, connected devices consume less backhaul bandwidth by processing most of that data at the edge in an IoT gateway close to the source, rather than in the cloud. This enables organizations to take advantage of telemetric data exposed by industrial devices that can otherwise be lost over time. By collecting and rationalizing this data at the edge, industrial IoT gateways improve the efficiency and effectiveness of local decision making, enterprise mobility, secure data exchange, remote monitoring, data consolidation, and remote system control. For instance, FogHorn is a company that develops edge intelligence software for industrial IoT (IIoT). Several companies across sectors ranging from oil and gas to smart buildings and manufacturing have used FogHorn's intelligence platform to gather real-time equipment insights and reduce processing and storage costs by 100 to 1000x.[1]

Content delivery in the entertainment and media industry

By caching content — be it a web page, video, or audio file — at the edge, end users can enjoy a better experience. A massive improvement over traditional web servers, edge caching brings down latency to milliseconds. The content delivery network (CDN) market continues to be dominated by a few players who have global cache networks. With edge computing, an organization can spin its own custom micro-cache at the edge of the network with more flexibility and customization than if it were to use a CDN service provider. Verizon, for example, uses edge to store video content in physical locations that are closer to end users, reducing the amount of traffic on its network by up to 20%.[2]

Voice and video recognition for telecom companies

As voice assistants such as Cortana, Alexa, Google Assistant, and Siri gain prominence, and video calling using data services grows rapidly, hauling all voice and video content back to the cloud is becoming costly for telecom companies. At the same time, users expect low-latency responsiveness for a seamless experience without taking a hit on their data plans. Edge computing enables the execution of machine learning (ML) models, such as those used for voice and video analysis, close to end users and their devices. This means using edge computing, a megabyte of voice recording can be converted into just a few bytes of text without the request making a roundtrip to the cloud. EdgeConneX is a US-based company that offers edge infrastructure solutions through well-located data centers to help mobile networks bring services, data, content, and applications closer to the edge for better streaming and speed.[3]

Cloud storage for consumer and enterprise data

As more and more people as well as businesses have started putting their data on the cloud, service providers recognize how crucial it is to ensure availability and reliability of the service, round the clock. A storage gateway at the edge can act as a read-write cache, which can be tapped into in the event of the cloud storage service being unavailable during outages or unplanned downtime. Moreover, on a slow speed network link, this gives the impression of a fast local storage array, enabling an always-on and accessible user experience. For instance, American company Dropbox, Inc., a leading provider of file hosting services, has built an extensive network of points of presence (PoPs) around the world that it calls as Edge – the network already has some half a billion registered users creating terabits of traffic.[4]

Augmented reality systems in the construction and manufacturing sectors

Augmented reality (AR) and virtual reality (VR) offer several benefits in terms of time and cost reduction. However, the biggest hurdle in comprehensively deploying mobile AR is the lack of desired computational and graphical performance in current devices. By implementing special hardware at the edge, users can offload demanding AR algorithms from mobile AR headsets to the edge to enable its application for use cases that are bandwidth-hungry and latency-sensitive. The result:

ability to create high-quality, photo-realistic renders, and immersive, AR and VR experiences for architecture, product design, and manufacturing. Microsoft's second gen HoloLens will bring together AR, cloud computing, edge computing, and distributed computing to enable edge computing applications to provide more compelling overall experiences.[5]

Edge Computing in Conjunction with the Cloud Will See Increased Adoption

Gartner estimates that IoT endpoints will reach an installed base of 25.1 billion units by 2021.[6] During the next five years, the proportion of edge applications based on new deployments of complete stacks of technology is also expected to rise dramatically.

Much of the current attention on edge computing comes from the need for IoT systems to deliver disconnected or distributed capabilities in the digital world. Information technology (IT) and operational technology (OT) are converging in numerous sectors and industries, including healthcare, transportation, defense, energy, aviation, manufacturing, mining, oil and gas, natural resources, telecommunications, and utilities. These industries generate large quantities of data that can be more efficiently filtered, analyzed, and acted on, locally.

Edge computing, along with cloud computing, is adding a new dimension to the multi-data-center, multi-cloud model for tackling the challenges of scale, resiliency, and security across industries. While no single application is likely to drive the emergence of a general-purpose edge computing platform, emerging use cases taken together will create a platform ecosystem around the edge, one that can handle heterogeneous workloads while enabling new possibilities. Clearly, companies that embrace edge computing will be well-positioned for a more distributed future by enabling solutions that are better designed for latency sensitive workloads where every millisecond matters.

References

1. FogHorn, Real-time Edge Intelligence for Industrial IoT, accessed June 13, 2019, https://www.foghorn.io/
2. Light Reading, Verizon Gets Serious About Edge Computing (January 2019), accessed June 13, 2019, https://www.lightreading.com/the-edge/verizon-gets-serious-about-edge-computing/d/d-id/749185
3. EdgeConneX, About EdgeConneX, accessed June 13, 2019, https://www.edgeconnex.com/company/about-edgeconnex/
4. Dropbox, Dropbox traffic infrastructure: Edge network (October 2018), accessed June 13, 2019, https://blogs.dropbox.com/tech/2018/10/dropbox-traffic-infrastructure-edge-network/
5. Microsoft, New HoloLens 2 gives Microsoft the edge in the next generation of computing (February 2019), accessed June 13, 2019, https://news.microsoft.com/innovation-stories/hololens-2/
6. Gartner, Gartner Identifies Top 10 Strategic IoT Technologies and Trends (November 2018), accessed June 13, 2019, https://www.gartner.com/en/newsroom/press-releases/2018-11-07-gartner-identifies-top-10-strategic-iot-technologies-and-trends

About The Author

Sanjay Tupe

Sanjay Tupe is a delivery partner in the Software Services group of TCS' HiTech business unit. He has over 18 years of experience and has played several key roles in product engineering and cloud infrastructure provisioning. Sanjay specializes in managing the delivery of cloud infrastructure manageability platforms, cloud analytics and insights, accessibility and security assessments for enterprise products, LOB applications, and cloud workloads. He has a Post Graduate Degree in Solid Mechanics and Design from the Indian Institute of Technology, Kanpur, India.

Contact

Visit the HiTech page on www.tcs.com

Email: HiTech.Marketing@tcs.com

Blog: HiTech Bytes

About Tata Consultancy Services Ltd (TCS)

Tata Consultancy Services is an IT services, consulting and business solutions organization that delivers real results to global business, ensuring a level of certainty no other firm can match. TCS offers a consulting-led, integrated portfolio of IT and IT-enabled, infrastructure, engineering and assurance services. This is delivered through its unique Global Network Delivery Model™, recognized as the benchmark of excellence in software development. A part of the Tata Group, India's largest industrial conglomerate, TCS has a global footprint and is listed on the National Stock Exchange and Bombay Stock Exchange in India.

For more information, visit us at www.tcs.com

TCS Design Services I M I 06 I 19