

the  
**GORILLA**  
**GUIDE**<sup>®</sup> to...



# Navigating Edge Computing

The Risks and Challenges of Edge Computing

**SEAN MCMANUS**

# Navigating Edge Computing

---

By Sean McManus

## TABLE OF CONTENTS

---

Introduction.....	4
What Is Edge Computing?.....	4
Key Drivers for Edge Computing .....	8
Current State of Edge Computing.....	10
Delivery Models for Edge Computing.....	11
Risks and Challenges of Edge Computing.....	15
How To Prepare for the Changes at the Edge.....	18
The Future of the Edge.....	19
Baby Steps.....	21

Copyright © 2022 by ActualTech Media

All rights reserved. This book or any portion thereof may not be reproduced or used in any manner whatsoever without the express written permission of the publisher except for the use of brief quotations in a book review. Printed in the United States of America.

### ACTUALTECH MEDIA

6650 Rivers Ave Ste 105 #22489 | North Charleston, SC 29406-4829  
[www.actualtechmedia.com](http://www.actualtechmedia.com)

# Publisher's Acknowledgements

---

## EDITORIAL DIRECTOR

Keith Ward

## DIRECTOR OF CONTENT DELIVERY

Wendy Hernandez

## CREATIVE DIRECTOR

Olivia Thomson

## SENIOR DIRECTOR OF CONTENT

Katie Mohr

## PARTNER AND VP OF CONTENT

James Green

---

## ABOUT THE AUTHOR

**Sean McManus** is a writer specializing in technology. He helps some of the world's biggest technology companies to communicate clearly. He also writes introductory programming books, including *Mission Python* and *Scratch Programming in Easy Steps*. Visit his website at [sean.co.uk](http://sean.co.uk) for free chapters.

# Introduction

---

Welcome to this Gorilla Guide To...® Navigating Edge Computing, Foundation Edition. Edge continues to gain traction in the enterprise, as more options become available for processing data closer to where it's generated and required.

Telcos, hyperscale cloud service providers, and colocation data centers will make up a complex ecosystem of edge service providers. This book, intended for IT leaders and business owners, will outline the opportunities and the risks of edge computing moving forward.

Although edge computing has been used in sectors such as retail for about a decade, it has now moved mainstream. The arrival of private 5G will make the Internet of Things (IoT) much more scalable. At the same time, organizations are looking for more automated ways to manage their businesses, in the pandemic and beyond.

We start our journey with an overview of edge computing, which can mean different things to different people.

## What Is Edge Computing?

---

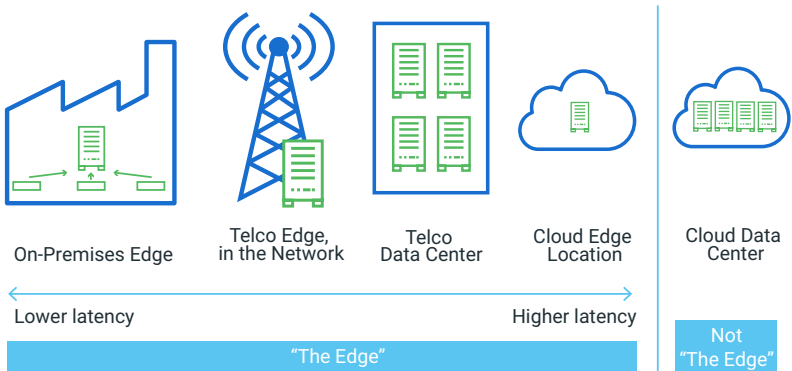
In recent years, data has increasingly been processed in a centralized cloud. Edge computing puts processing power closer to where data is generated and used—at the edge of the network—and avoiding a round trip to the cloud.

There’s no universal definition of where “the edge” is (see **Figure 1**). For ultra-low latency, an application could run on-premises. Alternatively, it could reside within the telco network in a location such as a central office, or in a data center a short hop away. When cloud service providers (CSPs) talk about “the edge,” their definition is often only as precise as a particular city.

Keeping data at the edge of the network solves several challenges inherent in the cloud. These include:

- **Latency issues**, especially as 5G dramatically cuts the “over the air” latency and the backhaul to the cloud risks becoming the new bottleneck.
- **Bandwidth availability**, where large volumes of data must be processed, for example for video analytics.

### Where is the Edge?



**Figure 1:** Edge locations range from on-premises to cloud service providers’ data centers in large urban areas

- **Network availability issues**, for example for mobile devices that might travel out of network coverage, or for locations where connectivity would be expensive or impractical.
- **Data sovereignty rules**, which sometimes require data to be stored locally and not in the cloud.
- **Cloud storage and bandwidth costs**, by processing locally and sending only the essential data to the cloud.

At the same time, edge computing can make a huge amount of computing power available to connected devices. Wearables can offload their storage and compute to the edge server, for example, so they can be lightweight and small.

Many early edge computing use cases are video-based, because it's easy to set up cameras, and the maturity of deep learning frameworks makes it easy to analyze video feeds. For example, computer vision can be used for quality control on assembly lines, and perimeter surveillance applications that stop a machine if someone goes too close to it.

---

**Edge computing can make a huge amount of computing power available to connected devices.**

---

Industrial applications based on instrumenting assembly lines have also been successful, because they help to avoid the high costs of downtime and rework. These applications

can be used for predictive maintenance (fixing things just before they break), and process control, for things like ensuring that products are assembled in the right order.



### **EDGE COMPUTING USE CASE: RAIL INSPECTION**

Duos Technologies is using edge computing to speed up train inspections in the United States. The aim is to check for defects, while impacting the rail service as little as possible.

Usually, inspections are carried out manually. Using the new technology, trains can drive at up to 60 mph through a portal kitted out with bright lights and cameras, which capture images of the train.

There are nine cameras, each generating 1GB of data per rail car. For the average train of 120 cars, the total data captured is approximately 1TB. What's more, each portal might scan up to 40 trains a day.

Data needs to be processed quickly to ensure the train is inspected before it reaches its destination. Sending that much data to the cloud, often from locations with poor network coverage, would be impractical, so Duos Technologies uses edge computing. Data is processed in a small building next to the portal, using artificial intelligence software.

Find out more in Network World's article: "[4 Essential Edge-Computing Use Cases.](#)"

# Key Drivers for Edge Computing



Edge computing is now entering the mainstream, [according to Gartner Inc.](#), as organizations start to adopt new business applications, using the IoT in particular.

Money will be diverted from the cloud to new edge vendors, [according to Forrester Research](#), shaving five points off public cloud growth. This will build on the growing momentum of the edge, which shows no signs of slowing. And the arrival of private 5G makes IoT applications more scalable.

## THE IMPACT OF COVID-19

One of the drivers of growth at the edge will be the COVID-19 pandemic, which has made business as usual impossible for many companies. [IDC predicts](#) that COVID-19 response will be the main accelerator for 80% of edge-driven investments and business model changes.

Creating safer environments for staff and customers is a high priority. Edge computing can be used for thermal analysis, automated checking of masks, video-based contactless access control, proximity analysis to enforce social distancing, and people counting for capacity management.

Due to the pandemic, simple human interactions continue to be risky, so many businesses will be looking at how they can automate more processes. As the rail inspection

example earlier shows, computer vision at the edge can be used to automate human processes such as inspection and quality control. Automation can also help to drive down costs, at a time when many companies have suffered huge losses as a result of the pandemic.

Companies need to reduce their reliance on specific individuals, as far as possible. People may be unavailable at short notice, as a result of illness or the need to self-isolate to protect others. Edge computing can enable augmented reality (AR) to guide assembly and repair without the need for on-site experts.

Working from home has become the new normal for many people. Hosting virtual desktop and business application software at the edge can make applications more responsive, and employees more productive.

## **PRIVATE 5G**

5G will enable many more devices to be connected, for lower-latency and higher-bandwidth data sharing. The combination of the edge and 5G will ultimately enable new entertainment use cases (such as high-definition streaming and gaming on mobile devices), and is happening now.

The opportunity now is in using private 5G to connect devices, gather more data than before, and process it locally for near-real-time feedback. Private 5G networks can be created for specific applications, businesses, or sites, and are good for connecting IoT devices.

# Current State of Edge Computing

---

Edge computing is still relatively immature, although that's changing. Those using the edge recently [told Forrester](#) that the biggest benefit was collecting data from mobile and other devices.

It's not new, however—for more than a decade, there have been successful edge computing implementations. In retail, point-of-sale systems may combine local stock information and sales processing with the cloud for aggregated reporting and updates. In industry, machines can be instrumented to collect data and works-in-progress can be tagged to help with production flow and maintenance.

---

**Working from home has become the new normal for many people. Hosting virtual desktop and business application software at the edge can make applications more responsive, and employees more productive.**

---

However, these applications have typically been developed or customized for each customer by systems integrators (SIs) and independent software vendors (ISVs). There are no

standards for edge infrastructure (beyond the obvious communications standards), so it's hard to integrate solutions from different vendors.

## **Delivery Models for Edge Computing**

---

The market for hosting applications at the edge is relatively new, but has the potential to grow fast. Few organizations will see strategic value in owning the edge platform, so they will buy into “Edge as a Service.” That will also prove to be cheaper, and less demanding of skills, than building it themselves.

Many different types of providers are vying for a place in the new market for edge infrastructure. That includes CSPs, telecommunications companies, networking companies, and content delivery networks (CDNs).

Colocation data centers will continue to play a part, too. Aggregators can make it easier to find data centers closer to your desired location.

---

**The market for hosting applications at the edge is relatively new, but has the potential to grow fast.**

---

## PUSHING CLOUD INFRASTRUCTURE TO THE EDGE

The hyperscale CSPs are pushing their infrastructure closer to large population centers, to enable their customers to use the same familiar architecture at the edge that they already use in the public cloud.

[AWS Local Zones](#), for example, promise “single-digit milli-second latency.” They’re available now in Los Angeles, and are previewing in Boston, Houston, and Miami.

Similarly, [Azure Edge Zones](#) bring the Microsoft Azure software stack to specific named cities, a long way from Azure’s regional data centers. Azure Edge Zones will be available in New York, Los Angeles, and Miami.

Both Amazon Web Services (AWS) and Azure offer appliances with their cloud software stack built-in, which can be placed on-premises for extremely low latency. Azure Private Edge Zones are compact devices that provide access to compute and storage, and can also host networking workloads. [AWS Outposts](#) are managed appliances that bring the AWS stack to virtually any data center, as long as it can connect to an AWS region.

## WORKING WITH TELCOS

The lowest latency is only possible by working with telcos, who can host infrastructure within the network or provide a fast path to an edge data center. Some telcos see value in offering edge services, but it’s challenging to build an edge

infrastructure stack, with no proven return on investment for them yet. For both CSPs and operators, cooperation makes sense.

For Azure, Private Edge Zone partners include several leading telcos. Microsoft has also partnered with AT&T to offer Azure hosting within the operator's data center. The [Azure Edge Zone with Carrier](#) infrastructure offers latency of less than 10 milliseconds from mobile device to application.

Meanwhile, [AWS is partnering with Verizon](#) to offer [AWS Wavelength](#), which hosts the AWS cloud within the operator's network.

[Google has opened up its edge network to telcos](#) to host their applications, too, with AT&T being the first on board.

## CONTENT DELIVERY NETWORKS

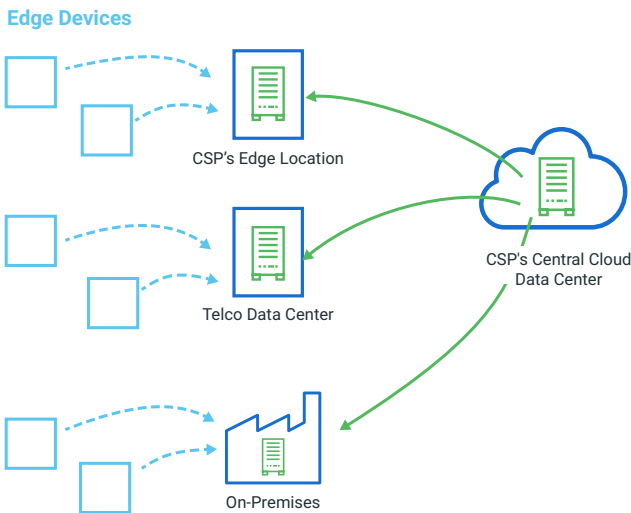
The edge is also a desirable target for CDNs that already have a rich network of locations for hosting content. Akamai now offers its [Intelligent Edge Platform](#), [Fastly offers a serverless compute environment at the edge](#), and Limelight Networks has an [Edge Compute platform](#), too.

Steve Miller-Jones, vice president of edge strategy and solution architecture at Limelight, [told TechRepublic](#): “Customers are integrating edge capability into their deployment pipeline to deploy their own processes and workflows. They don't have to do any distribution; the code just runs where their requests show up.”

## MANAGED APPLICATIONS

Managed applications will continue to be important: for example, where SIs/ISVs build a retail or industrial application for a particular company. In some cases, they may be able to use the new architecture options. However, Gartner doesn't expect public cloud solutions for the edge (such as AWS and Azure) to achieve more than a 30% share of edge deployments by 2025.

Gartner contrasts “cloud-out” and “edge-in” approaches, which define the infrastructure used. With edge-in deployments, the edge application is developed independently of the cloud it will connect to, and will likely use its own identity management and programming stacks. With cloud-out deployments, the public cloud stack is used at the edge (see **Figure 2**).



**Figure 2:** The cloud-out approach sees the CSP's servers move to edge locations, including on-premises, the telco data center, and the CSP's own edge data center

---

**Gartner doesn't expect public cloud solutions for the edge (such as AWS and Azure) to achieve more than a 30% share of edge deployments by 2025.**

---

The cloud-out approach might not offer the flexibility that SIs and ISVs need as they develop new edge solutions, and expand on existing ones. Gartner notes that many of them are using their own proprietary code or working with opensource software such as the [LF Edge stack](#).

## **Risks and Challenges of Edge Computing**

---

There are a number of challenges for the IT organization, and the company as a whole, to overcome as they launch applications at the edge.

### **ORGANIZATIONAL CHALLENGES**

Some culture clash may be inevitable: Edge deployments are typically led by the business owners of the applications or by the operational technology (OT) teams. For these teams, who may be working with systems that support machinery or infrastructure, the priorities are predictability and reliability. This is different from the IT team, which may adopt more cloud-like thinking, based on redundancy, scalability, and agility.

For successful edge implementations, IT and OT will need to work together. Edge computing will increasingly see operational technology touching IT systems and the cloud. Without close cooperation, there's the risk of the OT team building up silos and duplicate infrastructures that mirror what IT already has in place.

## MANAGEMENT CHALLENGES

The distributed nature of the edge means that the IT organization will need new capabilities to orchestrate, manage, and monitor applications across the distributed infrastructure. At least some of that infrastructure is likely to be owned by a service provider, so transparency and service-level agreements (SLAs) will be key.

IT organizations will need to evolve beyond monitoring the health of the hardware and software in the data center, and start to proactively manage the end-to-end Quality of Experience (QoE) offered by the application. This approach is commonly used by telcos. One method uses software probes that simulate and measure the performance of real application workloads in the distributed network.

## SECURITY CHALLENGES

Introducing new devices and compute capability at the edge creates new security challenges.

Security at the edge is complex. There may be different technology stacks to protect, in distributed locations outside

the data center. At the same time, security solutions at the edge are untested and immature, and standards are not yet available.



**Edge computing expands the attack surface, because there are more devices to attack, there's more data, and that data is widely distributed.**

Additionally, there are new network connections to protect, including from sensors to the edge processing location. Edge locations may not have data center-grade physical security, being in locations such as street cabinets.

The blast radius, i.e., the ability of an attack to damage adjacent systems and spread, may also increase with IT and OT becoming more tightly integrated. In particular, OT that has been isolated and offline in the past may now become connected, with the risk that it could be attacked from the IT side.

IT and OT teams need to work together and ensure there's a holistic approach to security. There needs to be routine vulnerability assessments and patching, anomaly detection (to identify data tampering), and zero-trust authentication of devices and users. Critical infrastructure must remain isolated from edge assets.

## RISK OF VENDOR LOCK-IN

While cloud-out solutions from the likes of AWS, Azure, and Google Cloud Platform (GCP) promise consistency with the public cloud, they also increase the risk of lock-in because they're not compatible with each other. They also have poor support for legacy infrastructure, making it harder to integrate with existing technology.

## SCALING UP

For many organizations, it will be difficult to go from a single proof-of-concept to a fully scaled production solution, because orchestration solutions and edge locations are relatively immature. Some organizations will nevertheless find immense value in edge solutions that don't require multiple sites, such as implementations that cover a single factory. Over the course of the year, the number of hosted edge locations will likely increase.

# How To Prepare for the Changes at the Edge



Organizations need to bring together business owners, OT teams, and IT teams to co-develop the edge strategy. The strategy will benefit from their pooled experience, but this will also help to avoid duplication and incompatibilities across their domains.

---

**Throughout the lifecycle of the project—the initial spark of an idea, through solution procurement, to the production operation—security must be front of mind.**

---

Start by identifying edge projects with a relatively quick return on investment to the business. There will be a lot of progress, and better technologies are quickly becoming available. Organizations may prefer to defer large projects until they have gained experience with smaller projects, allowing time for the available tools and solutions to mature.

Organizations should attempt to find overlap in edge projects, to enable experience and solution components to be shared as much as possible. SIs and ISVs should be involved in this discussion: The aim is to evolve away from bespoke implementations and toward more reusable solutions. Open source frameworks such as LF Edge may help with this.

Throughout the lifecycle of the project—the initial spark of an idea, through solution procurement, to the production operation—security must be front of mind.

## **The Future of the Edge**

---

Nobody is predicting the end of the public cloud. Edge computing will be complementary to the cloud, balancing low-latency local processing at the edge with the cloud for bulk storage and analysis.



**Many use cases will remain a better fit for the public cloud, and even some edge applications might only need data capture locally.** Data capture at the edge doesn't necessarily imply that it must be processed locally, if time isn't a factor. AWS understands this well: [AWS Snowmobile](#) enables up to 100PB of data to be collected in a 45-foot container and physically shipped to AWS for upload. At the low end, [AWS Snowcone](#) stores 8TB of data. Data can be processed locally, or the unit can be shipped to AWS for upload to the cloud.

The lack of even de facto standards for edge computing will mean that more than 85% of enterprises will deploy multiple incompatible edge in the near term, [according to Gartner](#).

Standardization will take time. This is still an evolving market. There will be a number of technology stacks emerging for the edge before there is a shake-out and consensus builds around architectures.

As public 5G takes off, service providers may expand their investments in edge infrastructure to target consumer-facing applications such as AR, virtual reality, and gaming. This will broaden the options available for enterprises, too.

---

**IT and OT teams need to work together and ensure there's a holistic approach to security.**

---

# Baby Steps

---

As you've seen in this Gorilla Guide, the rise of edge computing is highly disruptive. It will enable new applications by cutting latency. It will take compute and storage resources out of the centralized cloud and the on-premises data center, and push them into new edge locations. Only by using edge computing will it be possible to fully benefit from the massive connectivity and high bandwidth of 5G, because 5G does nothing to solve the latency of backhaul to the cloud.

---

**Standardization will take time. This is still an evolving market. There will be a number of technology stacks emerging for the edge before there is a shake-out and consensus builds around architectures.**

---

However, edge technologies are still in their infancy. Organizations are faced with a range of hosting options and will face challenges in orchestration, management, and security.

An increased focus on automation and the rise of private 5G makes it a good time for organizations to start developing their first edge deployments. They should focus on small projects that build their skills and deliver a quick return on investment, so they can pivot later as technologies mature.

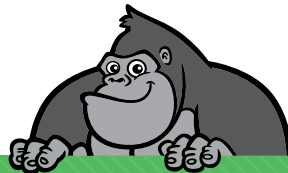
# About ActualTech Media



ActualTech Media is a B2B tech marketing company that connects enterprise IT vendors with IT buyers through innovative lead generation programs and compelling custom content services.

ActualTech Media's team speaks to the enterprise IT audience because we've been the enterprise IT audience.

Our leadership team is stacked with former CIOs, IT managers, architects, subject matter experts and marketing professionals that help our clients spend less time explaining what their technology does and more time creating strategies that drive results.



If you're an IT marketer and you'd like your own custom Gorilla Guide® title for your company, please visit <https://www.gorilla.guide/custom-solutions/>