Digital-first organizations are looking to edge computing as a key differentiator for creating business value. As companies distribute infrastructure and workloads outside of the datacenter, an automation strategy is essential for consistency and scalability.

Edge Growth Drives Need for Automating the Last Mile

April 2023

Written by: Dave McCarthy, Research Vice President, Cloud and Edge Infrastructure Services, and Jevin Jensen, Research Vice President, Infrastructure and Operations

Introduction

Edge solutions will drive the next wave of innovation as organizations execute digital transformation initiatives. Whether to automate operations, deliver rich customer experiences, or introduce new business models, edge solutions hold many benefits for enterprises. In addition, CIOs want to take advantage of cloud-native application design and software-defined infrastructure with the freedom to deploy anywhere.

This evolution is part of an expanded definition of hybrid cloud. The concept has evolved from a duality of on-premises and public cloud resources to include edge locations such as factories, hospitals, and retail stores. This flexibility in workload placement is essential to achieving the right balance of performance and security.

AT A GLANCE

KEY STATS

IDC forecasts the overall market for edge hardware, software, and services will reach $317.4 billion by 2026.

» 70% of organizations look to edge solutions as a key element of business transformation.

» There was 50% annual growth in edge production transactions from 2021 to 2022.

The move by organizations toward diversifying their footprint is significant because they realize that much of their differentiation and ability to create competitive advantage occurs in these edge locations. For manufacturing, edge computing combined with artificial intelligence (AI) has the potential to optimize processes, resulting in increased yield and reduced defects. Healthcare providers can automate the analysis of MRI scans to diagnose a patient and recommend a specialized treatment plan more accurately. Retailers can better understand customer behavior and provide personalized promotions and product recommendations, while using AI to detect shoplifting at self-checkout stations.

As edge environments grow, there is an increased requirement for consistent management and interoperability to reduce complexity. Adoption of enterprise open source technologies and, ultimately, an open hybrid cloud helps minimize vendor lock-in, facilitates standards-based integration, and leverages the open source community to accelerate the development of new capabilities.

Regardless of the use case, one thing holds true: The distributed and heterogeneous nature of these systems requires a common automation platform that makes it easier to deploy new sites and manage logistics of mass configuration while maintaining policy-based security that extends from edge to cloud.

Defining the Edge

In the simplest of terms, edge brings infrastructure and applications closer to the producers and consumers of data. As data becomes more distributed, organizations need to capture, process, and utilize this data in real time near its point of creation.

While the concept of distributed infrastructure has been around for many years, interest in modern edge deployments has emerged more recently as a massive industry trend. According to IDC forecasts, the overall market for edge hardware, software, and services is on track to reach $317.4 billion by 2026, with a CAGR of 14.8%.

The importance of edge has transcended IT departments and caught the attention of the C-suite. IDC survey data shows that 70% of enterprises view edge computing as a key element of business transformation including the implementation of new business models. This recognition is causing a shift in IT investments. Over the next two years, 74% of enterprises will increase their spend on edge solutions by an average of 37%. Despite this growth, it is important to note that edge is a complement to — not a replacement for — cloud infrastructure.

Edge has the potential to address limitations inherent in centralized architectures, including:

» **Security/compliance.** Whether due to government regulation or corporate governance, there can be restrictions to where data may reside. As jurisdictions continue to pursue data sovereignty legislation, businesses are challenged with compliance.

» **Latency.** This refers to the amount of time between a request and a response. Whether introduced by the network itself or the number of hops between endpoint and server, latency represents a delay that can negatively affect real-time applications.

» **Cost.** The proliferation of Internet of Things (IoT) and other connected equipment has led to an increase of data generated in remote locations. This data can be expensive to transmit and store in a core datacenter or the cloud, especially if it is needed only for short-term analysis.

» **Resiliency.** As organizations increase their dependence on technology, they need the ability to maintain operations in situations where connectivity to the cloud is not guaranteed.

Edge locations consist of both service provider and customer-owned facilities. In many cases, organizations utilize a mix of locations to achieve their business objectives.

Challenges of the Edge

Despite the benefits that edge computing can provide, organizations are challenged in how best to implement, manage, and secure these solutions. Some of the challenges can be attributed to the fact that many edge locations lack local IT staff, emphasizing remote management. These locations may also lack the same physical security and standardization found in a datacenter, requiring a different approach.

While these might seem like common IT concerns, edge solutions often introduce industry-specific elements. This means potentially integrating with OT hardware, software, and network protocols. IT/OT convergence also affects how different teams within an organization interact and define policies and procedures. This problem is compounded by organizations' growing dependence on these critical systems.

Edge computing also brings a new context to scalability. Most IT organizations are familiar with how to scale infrastructure in a dense datacenter environment, and a large organization might have several of those locations. Compare that with the challenges of managing hundreds or thousands of sites that contain edge workloads.
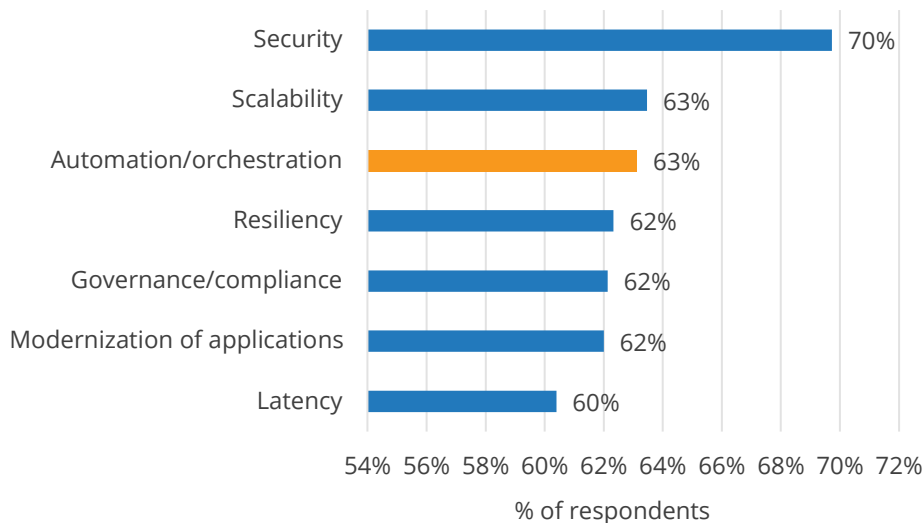
However, the edge requires a different type of scalability due to the number and variety of remote locations where the technology is deployed. A key challenge for IT is to design and operate effectively in these remote environments in a way that is consistent with existing datacenter approaches. As a result, IT often relies on bespoke technology that requires a significant amount of integration and manual human oversight. This approach adds time and cost to these projects, and the lack of standardization increases the possibility of security risks.

The Importance of Automation

IDC defines IT infrastructure automation as the ability to automatically provision, deploy, and manage IT resources across the private datacenter and public cloud environments. This functionality includes day-to-day operations of servers, storage, networks, security, containers, and other critical IT infrastructure. In addition, configuration management, cost transparency, reporting analytics, and AI are essential elements for enterprises looking to build genuinely autonomous operations. IDC research found that IT automation is one of the top priorities for these leaders in the next two years (see Figure 1).

FIGURE 1: *IT Automation Is a Top Priority for Edge Investments*

Q *How important is the following to edge infrastructure investments over the next two years?*

Category	%
Security	70%
Scalability	63%
Automation/orchestration	63%
Resiliency	62%
Governance/compliance	62%
Modernization of applications	62%
Latency	60%

% of respondents

n = 1,500

Source: IDC's EdgeView Survey, January 2022

Seasoned IT leaders want to deliver business value through cost controls and new strategic projects in the post-pandemic economy, which features rising inflation and a tight labor market. In addition, business models are changing more rapidly than ever as organizations move to a digital-first mindset. Operations teams play a crucial role in delivering this value

through IT resiliency. However, customer experience and revenue can decline when resiliency service-level objectives (SLOs) are missed. IT automation can be a way to provide both quick returns on investments and improved resiliency for the enterprise. For example, IDC found that 59% of large enterprises (with revenue greater than $2 billion) reported a return on investment of fewer than 12 months for IT automation projects.

Automation can be even more urgent within an edge environment. The edge can bring new challenges of limited or no IT staffing in remote locations such as branches, stores, warehouses, or plant floors. These remote edge facilities are often more focused on OT. The OT found in edge facilities has historically concentrated on hardware devices such as those needed to print, pack, and ship physical goods. These environments often lack the resources and controls found in IT-focused large cloud datacenters. Modern datacenters typically include redundant cooling, multiple power grids with generators, 24 x 7 physical security, expert onsite staffing, reliable operation processes, and resilient networks. IT focuses more on enabling applications and data resiliency.

In digital-first organizations, IDC sees a convergence of IT and OT. As more organizations begin this integration, it will quickly become a standard to maintain competitiveness. OT must adopt the data governance and integrity processes of its IT peers. This data foundation will continue to be developed and enhance the infrastructure in the cloud and at the edge. IT must support OT's nimble and more advanced operational analytics capabilities and quick decision making while improving upon zero trust cybersecurity for remote devices. OT needs the same resiliency that organizations achieve in the datacenter. Resiliency may require extending tools and trusted automation solutions to achieve a similar SLO at the edge. Advancing these goals will require a mix of in-house expertise and external services as well as the adoption of new architectures that deliver flexibility and technology resilience at scale.

Considering Red Hat Ansible Automation Platform for Edge Automation

Red Hat Ansible Automation Platform 2 is a fundamental building block for edge infrastructure. It is one element of Red Hat's portfolio of edge solutions, including Red Hat OpenShift for container support and Red Hat Enterprise Linux for the operating system and high-availability clustering.

Ansible Automation Platform 2 provides an enterprise framework for building and operating IT automation at scale, enabling users across an organization to create, share, and manage automation — from development and operations to security and network teams. IT managers apply guidelines on using automation within individual groups, and automation creators can write tasks that use existing knowledge. In addition, Ansible Automation Platform 2 provides a secure and stable foundation for deploying end-to-end automation.

Adopting containerization at the edge simplifies capacity scaling, increases resiliency, and improves consistency. Ansible Automation Platform 2 uses containerization to package, distribute, and execute automation across environments. This capability, coupled with the platform's resilient distributed architecture, enables organizations to bring IT services rapidly, consistently, and securely to the edge. Ansible Automation Platform 2 supports multiple installation targets, including Red Hat OpenShift. This capability enables organizations to drive automation using their preexisting container platform while reducing the complexity of running containers at scale.

Another new capability being introduced to the Ansible Automation Platform is Event-Driven Ansible. Event-Driven Ansible is an important innovation with the potential to help many organizations expand their use of automation. Organizations can quickly automate common tasks based on real-time events or triggers by using already familiar Ansible Rulebooks that are similar to Ansible Playbooks but use conditional if-then rules. Creating Ansible Rulebooks will be

familiar to Ansible users because both use the YAML language. Enterprises are looking to improve their "day 2" IT operations by improving consistency and accuracy, reducing the mean time to resolution (MTTR) and improving efficiency. Event-Driven Ansible triggers an automation rulebook with prebuilt desired responses when certain conditions occur. This allows incident resolution to occur very rapidly and can help to deliver system-centric responses to changing conditions across the environment. The use cases are flexible, such as applying a patch to an edge device when it comes online; rerouting network traffic when an outage is detected on an endpoint or sensor; and turning on the lights when a badge is swiped in an office building. The possibilities are varied and extensive.

Allowing IT operations to be more proactive via event-driven automation improves the digital customer experience and reduces middle-of-the-night calls to IT. This proactive response is part of IDC's vision for autonomous operations that supports a self-driving digital business. Organizations running Red Hat Ansible Automation Platform 2 can now implement an event-driven automation process, as shown in Figure 2.

FIGURE 2: *Red Hat Event-Driven Ansible*

The Ansible Automation Platform 2 Portfolio
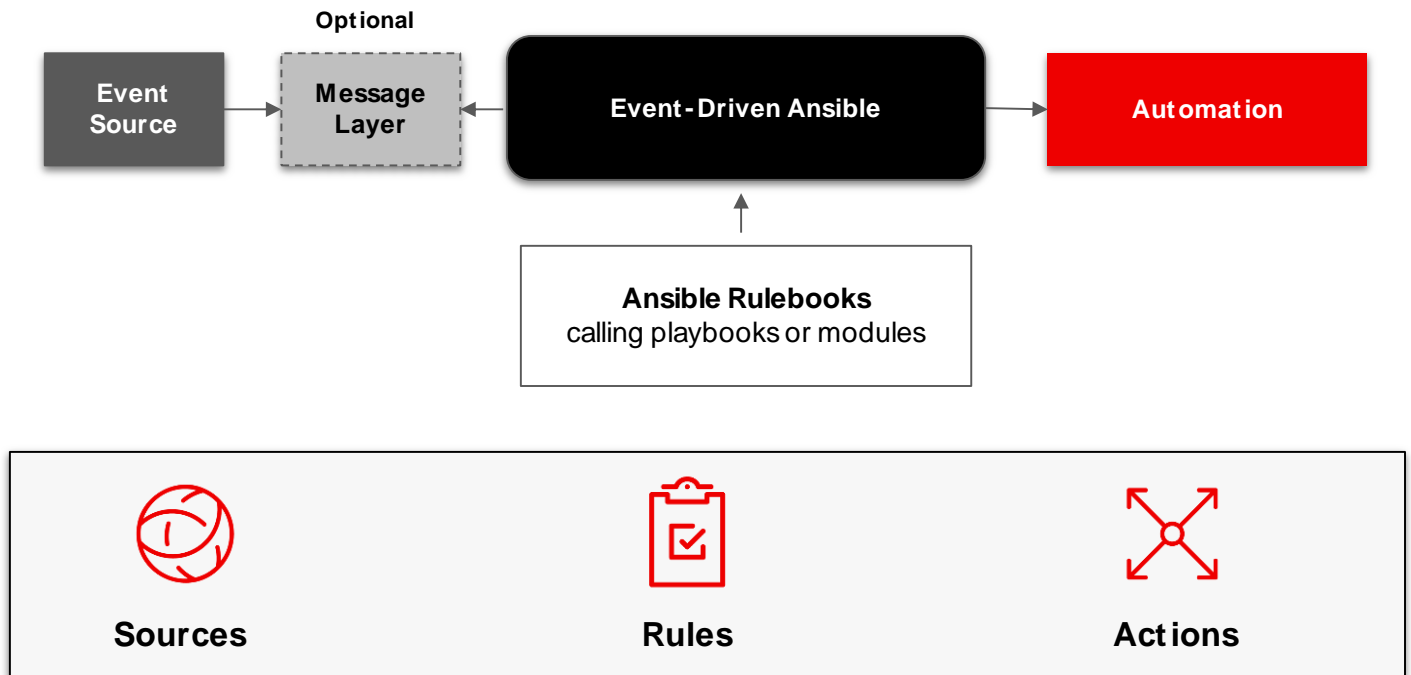
Source: Red Hat, 2023

Ansible Moves to the Cloud

Red Hat has dramatically increased the options for running Ansible Automation Platform 2 in the past 12 months. New and existing customers can now run their Ansible infrastructure in public clouds including Google Cloud, AWS, and Microsoft Azure. Red Hat's 24 x 7 support remains available for any public cloud installation as it does today for on-premises customers.

Ansible Automation Platform is available via the Google Cloud and AWS marketplaces and installs in minutes into the customer's cloud environment with consolidated billing, and Ansible purchases count toward any spending agreement targets. The environment is self-managed by the end customer with updates and maintenance of the Ansible Automation Platform being the responsibility of the customer. On Microsoft Azure, the process is identical with availability on the Azure Marketplace and billing integrated to current Azure spending and invoices. Red Hat goes a step further with Azure by offering the option to fully manage the platform, freeing customers to focus on what and where to automate.

Moving to a public hyperscaler for an automation solution allows customers to drive automation more quickly throughout their business. Quicker implementation, flexible procurement, and proven configurations mean IT teams are spending more time automating tasks than managing their environment. IDC expects adoption of cloud hosting of Ansible Automation Platform to grow rapidly. The edge is one area that may benefit from cloud hosting of automation platforms. Edge applications are already outside the core datacenter by definition, and managing them from the cloud reduces complexity.

There is more to manage in today's complex edge environment than just containers. As a result, customers demand an open platform and wide-ranging support. Red Hat supports other operating systems, including Windows Servers, and multiple hypervisors, with over 50 modules for VMware alone. Also, at a minimum, automation platforms need to support additional infrastructure elements, such as network routers, switches, firewalls, and load balancers. Ansible supports popular hardware products from Cisco, F5, Juniper, Arista, Palo Alto, NetApp, Dell, and others. In addition, Red Hat and its contributors developed Certified Content Collections modules for prebuilt playbooks and automation code to help enterprises rapidly implement a new automation platform solution.

Automation mesh, a feature introduced in Ansible Automation Platform 2, provides a simple, flexible, and reliable way to scale automation of large inventories across diverse network topologies, platforms, and regions. Mesh uses a secure overlay network, and distributed worker nodes run automation closer to the endpoints that need it, improving resilience to latency and intermittent connection disruptions.

Challenges

The complexity of modern IT infrastructure is often needed to enable a successful digital business. With the rapid growth in edge transactions, IT operations must prepare to manage and optimize this critical area. The definition of the edge itself means enterprises must manage dozens or hundreds of locations, each with unique applications. In addition, the dependencies between development, operations, and security teams and technologies make it challenging to deliver digital transformation on the edge in a consistent manner. Stretched operations teams must meet a growing list of project implementations in a tight labor market where job openings can go unfilled for many months. Using automation to assist understaffed operations teams is a way to keep them providing mission-critical support even without fully staffed positions. Automation at scale is the logical choice to deal with these challenges.

As the need to address these complexities grows, it is imperative to provide a single platform capable of supporting multiple teams. These teams include DevOps and SRE professionals, cloud and network architects, and remote OT groups. Collaboration among these teams is key to delivering business value. Enterprises should ensure they get full capabilities from Red Hat Ansible Automation Platform 2 by accessing the platform not just for automation and security but also for collaboration, compliance, and analytics. By leveraging a platform solution across multiple teams, enterprises

can achieve a faster ROI when deploying new edge solutions. For these reasons, automation at the edge is a crucial investment choice for IT executives even in a time of tightening IT budgets.

Conclusion

IDC projects that more than 750 million new business applications will be created by 2025. Many of these will be on the edge. Therefore, IDC recommends that DevOps teams consider container deployments to establish edge solutions rapidly and with resiliency. As these applications grow in scope and complexity, developers and the IT operations and infrastructure teams that support them will need to manage and compress application delivery cycles effectively.

Accordingly, advanced capabilities and tools that foster developer productivity are also rapidly gaining acceptance in the marketplace. Ansible Automation Platform 2 addresses these challenges by providing an array of tools needed to implement enterprisewide automation at scale. In addition, Red Hat added new tools to assist developers in being a part of the solution by allowing the easy creation of automation content.

As enterprises ramp up their digital transformation projects, keeping operations involved throughout the journey is vital. Delivering the total business value of digital transformation projects requires a first-class customer experience. Neither internal nor external customers will tolerate slow or unreliable applications. Customers tell IDC that they have realized a 58% reduction in downtime after implementing an automation solution. Downtime reduction coupled with cost and efficiency savings of IT automation of the edge is piquing the interest of many organizations.

IDC recommends that DevOps teams consider container deployments to establish edge solutions rapidly and with resiliency.

About the Analysts
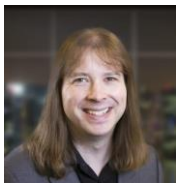
Dave McCarthy, *Research Vice President, Cloud and Edge Infrastructure Services*

Dave McCarthy is a Vice President within IDC's worldwide infrastructure practice, where he leads a team of analysts covering shared (public) cloud, dedicated (private) cloud, and edge strategies. Benefitting both technology suppliers and IT decision makers, Dave's insights delve into how hybrid and multicloud platforms provide the foundation for next-generation workloads, enabling organizations to innovate faster, automate operations, and achieve digital resiliency.

Jevin Jensen, *Research Vice President, Infrastructure and Operations*

Jevin Jensen is Research Vice President, Intelligent CloudOps Markets service at IDC. He covers infrastructure as code/GitOps infrastructure automation, cloud cost transparency, DevOps, hybrid/public/multicloud management platforms, and edge management.

MESSAGE FROM THE SPONSOR

About Red Hat Ansible Automation Platform

Red Hat Ansible Automation Platform is a single, flexible platform that includes all of the tooling needed for building, deploying, and managing end-to-end automation at scale. Learn more about Red Hat Ansible Automation.

» [The business value of Red Hat Ansible Automation Platform](#)

» [IDC Quick Take: Red Hat Ansible Automation Platform 2 Is Re-Architected for Speed and Ease of Use](#)

» [The IT automation imperative: What executives need to know](#)

IDC Custom Solutions

The content in this paper was adapted from existing IDC research published on www.idc.com.