White Paper

Improving the Private Cloud Experience with VMware Cloud Foundation 5.2

Sponsored by: VMware by Broadcom

Dave McCarthy
September 2024

IDC OPINION

The shift to cloud services has had a significant impact on organizations. The ability to deploy resources on demand and scale quickly has increased the pace of innovation. This unlocks new capabilities faster than ever, allowing businesses to optimize their operations and develop new products and services.

Topping the priority list of new capabilities is artificial intelligence (AI). While AI technologies have been available for some time, generative AI (GenAI) has grabbed the attention of CEOs and other business leaders. IDC expects organizations will spend $346.6 billion to implement GenAI from 2024 to 2027, and GenAI implementation will increase total AI implementation spending by $199 billion over the next four years.

The agility of cloud and rapid deployment plans for AI are a perfect match. For this reason, IDC predicts that by 2025, 70% of enterprises will form strategic ties to cloud providers for GenAI platforms, developer tools, and infrastructure, requiring new corporate controls for data and cost governance.

It is important to note that cloud technologies are no longer limited to hyperscale service providers. Instead, the cloud has evolved into an operating model that can be deployed anywhere to meet specific workloads' various performance and security needs. This is driving interest in platforms that create a consistent infrastructure layer across private cloud, hybrid cloud, and multicloud environments.

Private AI running on private clouds has also emerged as a preferred method of protecting corporate and customer data in AI models. Private AI is an architectural approach that balances business gains from AI with the organization's practical privacy and compliance requirements. However, building and operating cloud infrastructure

that spans multiple locations and service providers while simultaneously supporting traditional and modern applications is challenging and complex.

VMware Cloud Foundation (VCF) is a full-stack private cloud platform that simplifies diverse cloud environments. It enables organizations to accelerate developer productivity and embrace cloud-native and AI technologies to deliver apps and services to the market faster. With the release of VCF 5.2, VMware continues to enhance its ability to provide a comprehensive framework for modern infrastructure, ensuring that developers have a seamless cloud experience and that organizations maintain robust security and resilience.

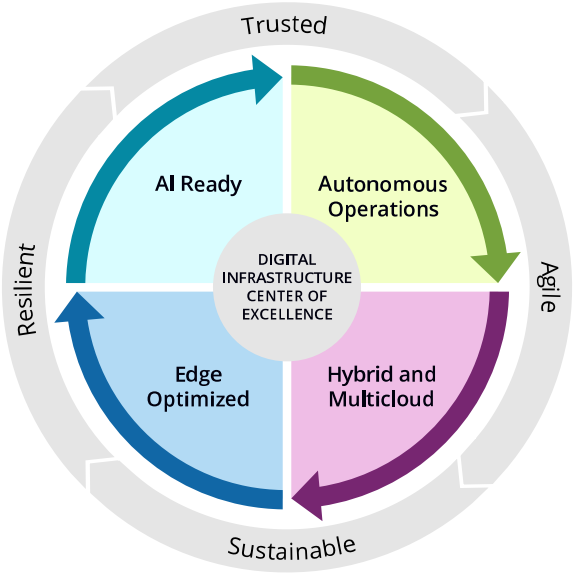
SITUATION OVERVIEW

Digital infrastructure provides mission-critical underpinnings for agile businesses, including emerging workloads and use cases powered by high-performance, data-intensive AI technologies; automated DevOps; and data science toolchains.

IDC's Future of Digital Infrastructure framework provides enterprise decision-makers with a view of how emerging innovations and disruptions across the digital infrastructure landscape will offer important opportunities for competitive advantage in the coming years (see Figure 1). The framework highlights the vital role that operating models, sourcing strategies, and governance processes play in ensuring that budgets, policies, and outcomes are aligned across the business.

FIGURE 1

IDC's Future of Digital Infrastructure Framework



Source: IDC, 2024

The framework focuses on five key pillars:

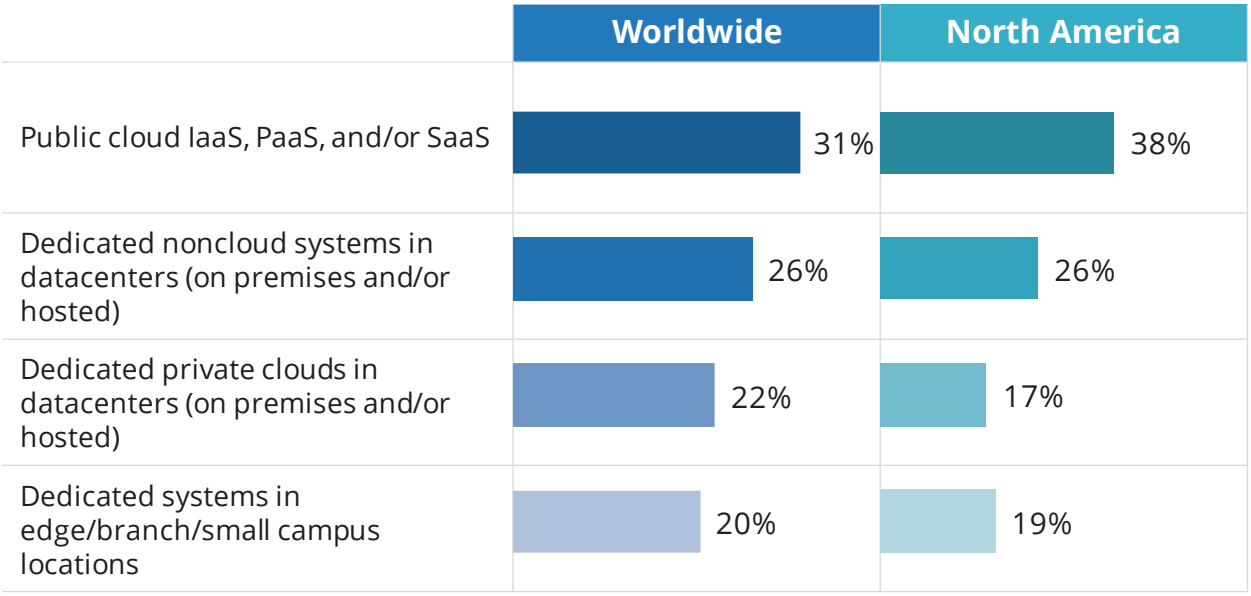
- **AI-ready infrastructure** technologies, products, and cloud services are optimized for the scale, performance, cost, sustainability, and interoperability requirements of emerging AI and other high-performance, data-intensive workloads.
- **Edge-optimized architectures** anticipate and accommodate the increasingly distributed nature of enterprise computing and data management required by the convergence of IT and operating technology and increased demands for widely distributed network connectivity and location-independent workloads.
- **Hybrid and multicloud** interoperability recognizes that applications and data are deployed depending on the specific needs of the individual workloads as well as the interactions required when multiple applications and data repositories need to interconnect; interoperable environments with a consistent infrastructure layer benefit the entire enterprise by enabling workload portability and modernization as well as supporting seamless data and process links.
- **Autonomous operations** take full advantage of AI, observability, and automation to enable organizations to manage, scale, and secure infrastructure consistently across datacenters, colocation and hosting sites, mobile and edge locations, and public cloud infrastructure-as-a-service (IaaS) and software-as-a-service (SaaS) platforms.
- **Digital infrastructure centers of excellence** provide collaborative governance and strategic coordination across IT, cloud, line-of-business, DevOps, and data science teams; these teams must promote tech debt avoidance, interoperability, and coordinated engagement with strategic vendors and ecosystem partners.

The framework's pillars come together when considering how organizations plan to deploy GenAI (see Figure 2).

FIGURE 2

GenAI Workloads and Data Deployment Over the Next 18 Months

Q. What is your best estimate of where your organization will deploy product GenAI workloads and data over the next 18 months?



n = 881 for worldwide, n = 361 for North America

Source: IDC's *Future Enterprise Resiliency and Spending Survey, Wave 1,* January 2024

Some 68% of IDC survey respondents indicate their preferred deployment model for GenAI will be single-tenant, dedicated infrastructure. This typically involves some element of on-premises private cloud combined with public cloud resources.

There are several reasons that on-premises infrastructure still exists. The most common are:

- **Latency:** The distance between an endpoint and a remote datacenter introduces network latency, prohibitive to real-time applications.
- **Cost:** As more data is created outside of datacenter environments, transferring and centrally storing this information can be expensive.
- **Data sovereignty:** Whether due to government regulations, industry standards, or corporate governance, more organizations must keep data local.
- **Business continuity:** Mission-critical applications must continue operating when the cloud or the internet is unavailable.

Given the current state of digital infrastructure–enabled innovation, it has never been more important for organizations to understand, anticipate, and exploit emerging

digital infrastructure technologies and operating models. Similarly, unified collaborative governance around infrastructure investments, architectures, strategic vendors, and tech debt mitigation is more important than ever.

VMWARE CLOUD FOUNDATION 5.2

VMware Cloud Foundation is a full-stack private cloud platform that supports digital transformation initiatives by enabling organizations to accelerate developer productivity. VCF embraces cloud-native and AI technologies to deliver apps and services to market faster.

Offering a consistent infrastructure layer and user experience for private cloud deployments, VCF streamlines resource management, accelerates innovation, and improves operational efficiency by reducing organizational silos.

The VCF platform is designed to address three strategic areas:

- **Modernized infrastructure:** VCF plays a crucial role in transforming traditional IT infrastructures into a more agile and adaptable private cloud environment. This transformation is key to enabling organizations to deploy core private cloud use cases that deliver strategic outcomes, lower the total cost of ownership, and increase productivity. VCF provides organizations with a uniform infrastructure layer that results in a consistent cloud operating model across cloud endpoints, combined with automation and orchestration to standardize and simplify the entire infrastructure life cycle, including day 0 deployment, day 1 provisioning, and day 2 operations, including patches, updates, compliance, troubleshooting, diagnostics, and logging.
- **Cloud consumption experience:** To keep continuous development pipelines running at peak efficiencies for modern workloads and private AI, it's critical to ensure that developers have frictionless access to application code, infrastructure services, runtime environments, system tools, libraries, and registries. VCF includes an embedded, upstream-compliant Kubernetes runtime via Tanzu Kubernetes Grid (TKG), combined with a cloud consumption interface and a set of infrastructure and automation services, including infrastructure as code, software configuration management, and infrastructure pipelines.
- **Security and resilience:** VCF provides a consistent, secure platform that can extend the security architecture with intrusion detection and recovery, addressing key challenges such as ransomware threats, disaster scenarios, and advanced security architectures.

As a complete cloud platform, VCF provides a set of software-defined services for compute, storage, network, container, and cloud management:

- **VMware Cloud Foundation Infrastructure Stack:** Incorporates VMware's industry-leading infrastructure solutions — vSphere for compute, vSAN software-defined storage, and NSX software-defined networking — focusing on the installation, operations, and life-cycle management of the entire stack and streamlining the deployment, configuration, and updating processes
- **Management and orchestration:** Delivers comprehensive operations, automation, and analysis across the full-stack infrastructure platform
- **Hybrid cloud extension:** Enables seamless migration, workload rebalancing, and disaster recovery across the private, edge, and public clouds

VCF delivers a unified and automated platform, simplifying the deployment of a fully integrated IaaS stack. Enabling consistent, secure, and agile operations across private and public clouds, VCF ensures that organizations can flexibly scale their infrastructure to meet evolving business needs. It consolidates disparate functions into a single, integrated platform, significantly reducing the complexity and overhead typically associated with private cloud deployments.

NEW IN VCF 5.2

VCF 5.2 introduces new features that map to the three strategic areas previously mentioned.

Modernized Infrastructure

- **VCF import for vSphere and vSAN:** Reduces downtime and simplifies the process of onboarding existing vSphere and vSAN environments into VCF
- **VMware Cloud Foundation Edge:** Addresses the need to extend cloud capabilities to edge locations, ensuring consistent operations and infrastructure management across distributed environments
- **vSAN Max:** Supports disaggregated storage that provides petabyte-scale centralized storage for VCF environments
- **Offline depot:** Enables updates in isolated and/or air-gapped environments for dark sites

Cloud Consumption Experience

- **Independent TKG service:** Empowers developers to manage containerized workloads with great flexibility, accelerating development cycles and enhancing productivity
- **Simplified adoption of virtual networks:** Streamlines process for adopting virtual networks that reduces complexity to increase adoption of network segmentation and management

- **Cloud admin launchpad/dashboard:** Includes intuitive interfaces for managing resources and automating tasks, such as VCF Automation for DevOps to streamline workflows
- **Improved GPU monitoring for private AI:** Incorporates VCF Operations that show a dashboard summary of GPU-equipped clusters for additional metrics and telemetry

Security and Resilience

- **Live patching and flexible upgrades:** Minimize operational disruptions through live and flexible patching to ensure systems remain secure without impacting productivity by providing greater control over schedules and deployments
- **Dual DPU support:** Enhances computational efficiency and workload management to make high-performance data processing more accessible and reliable
- **Edge redundancy and load balancing:** Include additional capabilities to ensure uninterrupted service availability
- **Network detection and response:** Include complementary network security technologies that seek to automatically monitor, detect, analyze, and respond to sophisticated cyberthreats

These new features help customers improve resource utilization, enhance developer speed, and provide faster time to value for the business.

CHALLENGES/OPPORTUNITIES

Broadcom's acquisition of VMware has raised questions about the future of the product portfolio. Industry observers have been concerned about whether Broadcom will continue to invest in new features and maintain the partnerships critical to ensuring compatibility in multivendor environments.

VCF 5.2 is the first major release since the acquisition, representing a new unified product development strategy. By simultaneously releasing all products under a single cadence, customers will benefit from increased coordination of feature sets, allowing for quicker adoption across the organization.

CONCLUSION

The rise of cloud computing has revolutionized how organizations operate. By enabling on-demand resource allocation and rapid scaling, the cloud has accelerated innovation and spurred the development of new products and services. At the forefront of these new capabilities is AI, particularly generative AI, which has captured the attention of

business leaders worldwide. Massive investments are being made in AI technologies, and the cloud's agility makes it the ideal platform for deploying these solutions.

Many organizations are forming strategic partnerships with cloud providers to capitalize on the potential of GenAI. However, the complexity of managing cloud infrastructure across multiple locations and service providers is a significant challenge. In addition, the sensitive nature of data in AI models has led to a growing interest in private AI solutions.

VMware Cloud Foundation offers a solution to these challenges by providing a unified platform for managing diverse cloud environments. VCF simplifies infrastructure management, accelerates developer productivity, and enhances security and resilience. With the latest release, VCF 5.2, VMware continues strengthening its position as a comprehensive framework for modern infrastructure, empowering organizations to harness the power of cloud and AI while safeguarding their data.

MESSAGE FROM THE SPONSOR

About VMware Cloud Foundation

VMware Cloud Foundation (VCF) is the industry's first private cloud platform to deliver public cloud scale and agility with private cloud security, resilience and performance, and low overall total cost of ownership. VCF supports customers' digital innovation with faster infrastructure modernization, a unified cloud experience, and better cyber resiliency and platform security and compliance.

VMware Cloud Foundation delivers a consistent, unified cloud experience across any environment - whether it's a customer managed data center, edge location or any cloud endpoint with license portability. VCF empowers customers to deliver any application by deploying both traditional or container apps with self-service catalogs with complete automation.

ABOUT IDC

International Data Corporation (IDC) is the premier global provider of market intelligence, advisory services, and events for the information technology, telecommunications, and consumer technology markets. With more than 1,300 analysts worldwide, IDC offers global, regional, and local expertise on technology, IT benchmarking and sourcing, and industry opportunities and trends in over 110 countries. IDC's analysis and insight helps IT professionals, business executives, and the investment community to make fact-based technology decisions and to achieve their key business objectives. Founded in 1964, IDC is a wholly owned subsidiary of International Data Group (IDG, Inc.).

Global Headquarters

140 Kendrick Street
Building B
Needham, MA 02494
USA
508.872.8200
Twitter: @IDC
blogs.idc.com
www.idc.com