LEARNING MADE EASY

VMware Software Special Edition

Self-Service Private Cloud

for dummies®

A Wiley Brand

Empowering teams with self-service provisioning

Accelerating IT delivery with self-service private cloud

Simplifying operations with VMware Cloud Foundation

Brought to you by

BROADCOM®

Alina Thylander
Taka Uenishi
Vincent Riccio

About Broadcom Inc.

Broadcom Inc., a Delaware corporation headquartered in Palo Alto, California, is a global infrastructure technology leader built on more than 60 years of innovation, collaboration, and engineering excellence.

With roots based in the rich technical heritage of AT&T/Bell Labs, Lucent, and Hewlett-Packard/Agilent, Broadcom focuses on technologies that connect our world. Through the combination of industry leaders Broadcom, LSI, Broadcom Corporation, Brocade, CA Technologies, Symantec's enterprise security business, and VMware, the company has the size, scope and engineering talent to lead the industry into the future.

Broadcom is focused on technology leadership and category-leading semiconductor and infrastructure software solutions. The company is a global leader in numerous product segments serving the world's most successful companies.

Broadcom combines global scale, engineering depth, broad product portfolio diversity, superior execution, and operational focus to deliver category-leading semiconductor and infrastructure software solutions so its customers can build and grow successful businesses in a constantly changing environment.

Self-Service Private Cloud

VMware Software Special Edition

by Alina Thylander, Taka Uenishi, and Vincent Riccio

for dummies®
A Wiley Brand

Self-Service Private Cloud For Dummies®, VMware Software Special Edition

Published by
John Wiley & Sons, Inc.
111 River St.
Hoboken, NJ 07030-5774
www.wiley.com

ISBN 978-1-394-34792-6 (pbk); ISBN 978-1-394-34793-3 (ePDF);
ISBN 978-1-394-34794-0 (ePub)

Publisher's Acknowledgments

Editor: Elizabeth Kuball
Acquisitions Editor: Traci Martin
Senior Managing Editor: Rev Mengle
Client Account Manager: Cynthia Tweed
Content Refinement Specialist: Umeshkumar Rajasekhar
Special Help: Brett McLaughlin, Lawrence Miller

Table of Contents

Introduction

The way organizations build and manage their IT infrastructure is evolving rapidly. Virtual data centers, often burdened by manual provisioning and heavy oversight, are giving way to a game-changing concept: the self-service private cloud (SSPC). SSPCs offer a fully automated, user-driven approach that eliminates slow, ticket-driven processes.

By harnessing the power of self-service automation and policy-driven governance, businesses can significantly lower operational overhead, enhance resource utilization, and speed up application delivery. This shift is about more than just easing IT workloads; it unlocks innovation. A well-architected SSPC guarantees scalability, security, and compliance, enabling teams to focus on building, not managing infrastructure. Whether you're a decision-maker or a technical leader (or both), embracing a modern SSPC is crucial for staying ahead in an increasingly competitive digital landscape.

About This Book

This book is your guide to understanding and building an SSPC and delivering a unified cloud consumption experience for application teams to manage all types of applications on a single platform. It breaks down the fundamental concepts and key capabilities needed to transition from a virtual data center to an automated cloud solution.

You'll gain insights into what defines a truly self-service cloud while learning to evaluate your organization's needs, select the appropriate cloud model, and design a scalable architecture. This book explores the implementation of self-service infrastructure as a service (IaaS), the creation of automated workflows, and the development of user-friendly service catalogs. Additionally, you'll discover how to operationalize your cloud effectively with workload management and cost optimization strategies, wrapping up with a list of benefits gained from adopting a self-service private cloud.

Foolish Assumptions

To make the most of this book, we assume

- You're familiar with vSphere and want to explore automation and a self-service approach to cloud.
- You want to get started with a private cloud and need a good product and jumping-off point.
- You're an IT decision-maker or technical leader who wants to transition to a self-service cloud model.

Icons Used in This Book

To help you navigate the content, we use the following icons:

REMEMBER

This icon points out important information you should commit to your nonvolatile memory or your noggin!

TIP

Tips are appreciated, but never expected, and I sure hope you'll appreciate these useful nuggets of information.

WARNING

These alerts point out the stuff your mother warned you about. (Well, probably not, but they do offer practical advice.)

Beyond the Book

If you're looking for more case studies, best practices, and technical documentation on self-service private clouds, visit `www.vmware.com/products/cloud-infrastructure/cloud-foundation-automation` for additional insights.

IN THIS CHAPTER

- Enabling self-service consumption with faster cloud access
- Automating infrastructure management to reduce IT overhead
- Selecting a scalable vendor for long-term flexibility

Chapter 1
Understanding Self-Service Private Clouds

A self-service private cloud (SSPC) offers organizations the flexibility of public cloud infrastructure with enhanced security and control. Unlike virtual data centers that require manual provisioning, an SSPC allows users to deploy and manage resources on demand. This chapter explores the key concepts, foundational layers, and essential vendor qualities for SSPCs.

Introducing VMware Cloud Foundation 9

VMware Cloud Foundation (VCF) Automation within VCF 9 enables IT to deliver an SSPC for application teams to build, run, and manage artificial intelligence (AI), Kubernetes (K8s), and virtual machine (VM)–based applications. The solution makes it easier to jump-start and scale a multitenant private cloud with infrastructure as a service (IaaS) resources available out-of-the-box (OOTB) to help bring applications to market faster while maintaining control with policy-based governance. The solution helps vSphere Infrastructure (VI) Admins evolve into Cloud Admins that can offer self-service consumption of infrastructure resources "as a service" to application teams. It helps IT to shift

from time-consuming, ticket-driven, multiteam, manual provisioning processes that are subject to error and rework to a more automated self-service environment.

Defining Self-Service Private Clouds

In an SSPC, instead of waiting for resources or navigating complex workflows, users deploy resources instantly through a self-service portal or application programming interface (API). This shift streamlines development, improves IT efficiency, and maintains security and compliance through automation and policy-driven governance.

Establishing core principles

An SSPC stands apart from virtual data center environments with several key principles:

- **On-demand infrastructure and automation:** Users deploy workloads autonomously, enhancing productivity while infrastructure resources adjust dynamically to eliminate repetitive manual tasks.
- **Policy-driven governance:** Security, compliance, and resource usage adhere to predefined policies, ensuring operational efficiency without added risk.
- **Unified cloud consumption experience:** A centralized interface allows users to manage resources across all application types, including AI, K8s (containers), and VMs.

Overcoming challenges

Virtual data center environments force IT teams to handle every provisioning request, leading to inefficiencies. Organizations are adopting SSPCs to:

- **Accelerate time to market:** Developers and IT teams deploy environments quickly, reducing bottlenecks in the development lifecycle.
- **Strengthen security and compliance:** Built-in policy enforcement ensures that all deployments meet security and regulatory requirements.

- **Optimize resourcing:** Automated scaling and governance prevent overprovisioning while reducing costs.

REMEMBER

Moving to an SSPC eliminates manual processes, lowers costs, and empowers teams with flexible cloud infrastructure while maintaining full control.

Identifying Key Layers of a Self-Service Private Cloud

An SSPC brings together multiple layers that work in sync to deliver automation, security, and scalability. Each layer plays a role in ensuring that infrastructure resources are provisioned efficiently, managed effectively, and secured appropriately. You should be familiar with the following layers:

- **Virtual infrastructure:** The foundational virtualization layer, encompassing compute, storage, and networking and designed for consistent and integrated operations across the stack
- **Built-in IaaS platform:** Core infrastructure services delivered OOTB, including compute, storage, networking, and data services, forming the base for higher-level capabilities
- **Extensible services platform:** The flexible extension layer, supporting both native and third-party services, allowing teams to integrate the tools and capabilities their applications require
- **Management and governance:** The centralized control layer, enabling tenancy, policy enforcement, and simplified operations for administrators through consistent tooling and automation
- **Cloud experience for end users and developers:** The consumption layer, providing a cloud-native experience with direct access to services like K8s, graphics processing units (GPUs) for AI workloads, data protection, and virtual private cloud (VPC) networking — all available through the platform

Choosing the Right Self-Service Private Cloud Vendor

Not all SSPC solutions are created equal. The right vendor provides more than just technology — it delivers a comprehensive approach to automation, security, and scalability that aligns with an organization's long-term cloud strategy and enables the organization to offer SSPC services, boosting developer productivity and user satisfaction. Choosing the right partner means evaluating how well a solution supports critical business, including the following:

- **Built-in automation:** A self-service SSPC solution automates provisioning, workload management, and policy enforcement, reducing IT overhead and enabling teams to focus on innovation.
- **Unified cloud experience:** An integrated platform for managing all workloads — AI, K8s, and VMs — through a single interface streamlines operations and enhances efficiency.
- **Security and compliance integration:** Governance should be inherent. Frictionless governance with embedded guardrails and features like role-based access control (RBAC), encryption, and policy-driven security help meet compliance requirements without slowing down self-service.
- **Scalability and flexibility:** As business needs evolve, a private cloud must adapt. A vendor that supports dynamic resource allocation, multitenancy, and a model that enables customers to run across both on-premises data centers and public cloud providers is essential.
- **Extensibility and ecosystem support:** The best solutions integrate with existing IT environments, including identity management, data protection, and third-party automation tools. Vendors that continually offer new extensible services further extend the value of the SSPC platform. Vendors that offer strong API support and broad ecosystem compatibility provide greater flexibility in cloud management.

With the right partner in place, organizations can unlock the full potential of an SSPC — accelerating innovation, improving efficiency, and maintaining control over their infrastructure.

IN THIS CHAPTER

- Assessing infrastructure, application, and business needs for cloud adoption
- Selecting a cloud model that balances control and scalability
- Automating resource provisioning to improve efficiency and governance

Chapter 2
Planning and Designing a Self-Service Private Cloud

Building a self-service private cloud (SSPC) starts with careful planning to ensure it meets business needs, supports diverse application and workload demands, and remains flexible. This chapter explores how to assess organizational requirements, choose the right cloud model, and leverage configuration and automation to create a scalable and efficient private cloud environment.

Assessing Your Organization's Requirements

An SSPC is only as effective as the planning behind it. Rushing into deployment without understanding your organization's needs can lead to inefficiencies, security gaps, and unexpected costs.

Identifying key stakeholders and use cases

An SSPC impacts multiple teams, and engaging the right stakeholders early ensures smoother adoption. Key groups to consider include the following:

- **IT and infrastructure teams:** Manage cloud stability, security, tenant management, and resource allocation.
- **Platform engineer teams:** Support developers and DevOps teams by building standardized infrastructure, automating repetitive tasks, and reducing manual overhead through the SSPC platform.
- **Developers and DevOps teams:** Require fast, on-demand access to computing resources to build, run, and manage applications.
- **Security and compliance teams:** Define access controls, encryption, and governance policies.
- **Finance and operations teams:** Oversee cloud costs, resource efficiency, and chargeback models.
- **Line-of-business teams:** Use self-service resources for analytics, artificial intelligence (AI), and application development.

Failing to align cloud capabilities with stakeholder needs can result in unnecessary restrictions or uncontrolled access, both of which hinder adoption.

REMEMBER

A successful SSPC balances autonomy with IT oversight, ensuring that users have the access they need without compromising security or efficiency.

Evaluating application and infrastructure requirements

Not all workloads require the same infrastructure, so understanding their needs is essential for resource allocation and performance optimization. Common workload categories include

- **Traditional enterprise applications:** Require stable virtual machines (VMs), storage, and predictable networking.

- **Cloud-native applications:** Depend on containers, Kubernetes (K8s), and API-driven orchestration.
- **AI and data analytics:** Demand high-performance computing, graphics processing units (GPUs), and large-scale data storage.
- **Edge computing:** Needs low-latency processing and hybrid cloud compatibility.

By analyzing workload patterns, organizations can prevent performance bottlenecks, avoid overprovisioning, and optimize cost-efficiency.

TIP

Not every workload benefits from self-service automation. Identify which workloads need direct user access versus controlled provisioning.

Defining automation and self-service expectations

Automation drives efficiency in an SSPC, but different organizations require different levels of control.

- **Fully automated self-service:** Allows users to provision, scale, and retire resources independently.
- **Controlled self-service:** Provides users with preapproved templates while maintaining IT oversight, as well as being able to apply changes to deployed resources with day 2 custom actions that enable users to make changes they need when they need it, instead of having to ask someone else and waiting for the changes to be made.
- **Limited self-service:** Requires IT approval for provisioning but still automates core processes.

The right approach ensures that automation enhances productivity without introducing security risks or unnecessary complexity.

Choosing the Best Cloud Operating Model for Your Organization

Enterprises should have a business strategy, an application strategy, and a cloud strategy with a cloud operating model that binds them together. A *cloud operating model* describes the people,

process, and technology changes required to evolve an organization to effectively manage its cloud.

TIP

Learn more about cloud operating models at `www.vmware.com/topics/cloud-operating-model`.

Evaluating private cloud deployment models

An SSPC can be implemented using different operating models, each with its own benefits and trade-offs:

- **Decentralized operations model:** This model prioritizes innovation and agility by giving individual workload teams maximum autonomy over their cloud resources. It's characterized by minimal central oversight and high complexity at the workload level, making it suitable for organizations that value speed and flexibility over standardization.
- **Centralized operations model:** Focused on control and standardization, this model maintains centralized management of cloud resources and policies. It provides low complexity for individual workloads but requires more centralized coordination, making it ideal for organizations with strong compliance requirements or limited cloud expertise.
- **Enterprise operations model:** This model emphasizes democratization of cloud services while maintaining governance and control. It typically involves medium complexity and provides a balance between innovation and standardization, supporting larger portfolios through centralized platform services and distributed execution.
- **Distributed operations model:** Designed for integration across complex, global operations, this model manages full portfolios across multiple regions and cloud providers. It requires sophisticated tooling and processes but provides the highest level of operational maturity and scale.
- **Hybrid operating model:** Many organizations adopt hybrid approaches that combine elements from different models based on specific use cases, compliance requirements, or organizational maturity. These models allow different workloads or business units to operate under different frameworks while maintaining overall consistency.

Although each model has distinct advantages, organizations must determine how well they align with workload characteristics, compliance needs, and IT skill sets.

Aligning cloud models with workload requirements

Different applications have unique demands for compute, storage, and networking. The cloud model should align with these requirements to ensure efficiency and performance:

- **Latency-sensitive applications:** Workloads requiring low-latency data access, such as real-time financial processing or industrial automation, perform best in private cloud environments where resources are physically close to users.
- **Cloud-native applications:** Applications built on microservices, containers, and K8s benefit from hyperconverged, where software-defined automation and dynamic scaling allow for flexible deployments.
- **Data-intensive applications:** AI, machine learning, and big data workloads require high-performance computing and scalable storage. Hyperconverged models often offer the best solutions by integrating GPU acceleration and distributed storage.
- **Regulated workloads:** Industries with strict compliance requirements, such as healthcare and finance, may need to keep workloads entirely on-premises within a traditional private cloud to maintain full data sovereignty.

Choosing a cloud model based on workload needs prevents performance issues, compliance risks, and unnecessary costs.

Managing cost and scalability

Beyond performance, organizations must evaluate how different cloud models affect cost structure and long-term scalability:

- **Traditional private clouds** require higher upfront capital expenditures (CapEx) for infrastructure investments but provide long-term cost predictability. Organizations that need full control over their infrastructure often favor this model despite the higher initial costs.

- **Hyperconverged infrastructure** reduces CapEx by consolidating resources into a streamlined architecture, making scaling easier, improving operational efficiency, and reducing IT labor costs.

In addition to cost structure, scalability is a key factor. Organizations should consider whether the cloud model supports elastic resource scaling, multiregion expansion, and workload portability to ensure future growth without disruptive migrations.

Creating Flexibility through Configuration and Automation

Automation plays a central role in making an SSPC possible. In a traditional virtual data center, IT teams must manually provision and maintain infrastructure, creating bottlenecks. With automation in an SSPC, infrastructure is deployed through preconfigured policies, reducing IT overhead while maintaining security and compliance.

REMEMBER

The strength of an SSPC isn't just in automation — it's in how automation and configuration work together to deliver control, scalability, and cloud operations.

Configuration ensures that flexibility doesn't lead to chaos. An SSPC enables organizations to define templates, blueprints, and policies that provide users with on-demand access to infrastructure without compromising governance. An SSPC built on automation and configuration delivers a cloud experience that is flexible, scalable, and easy to manage. Organizations gain the agility of public cloud, the control of private cloud, and the simplicity of a single platform.

IN THIS CHAPTER

- Establishing infrastructure to support automation and self-service
- Enforcing policies to maintain security and compliance
- Standardizing deployments with blueprints and catalogs

Chapter 3
Implementing Self-Service Private Cloud Capabilities

Building a self-service private cloud (SSPC) goes beyond selecting the right model — it requires a strong foundation, automated provisioning, policy guardrails, and structured self-service catalogs. This chapter explores how to set up key cloud components, configure self-service infrastructure as a service (IaaS), and create standardized blueprints to ensure efficient, scalable, and secure cloud operations.

Setting up Key Components

An SSPC relies on a well-architected foundation to achieve scalability, automation, and governance. Without solid infrastructure components, organizations risk inefficiencies, gaps in security, and poor usage of self-service capabilities.

Establishing core infrastructure

Before enabling self-service capabilities, organizations have to make sure their cloud infrastructure is designed for automation and scalability. This includes the following:

- **Compute:** A unified platform must support both traditional virtual machines (VMs) and modern containerized applications, enabling consistent operations, shared tooling, and reduced silos across environments.
- **Storage:** A software-defined storage strategy offers a dynamic, automated infrastructure tailored to your organization. It effectively supports block, file, and object storage, while hyperconverged solutions streamline management across storage, compute, and networking for enhanced performance.
- **Networking:** A software-defined networking (SDN) approach transforms network segmentation, traffic management, and security. It enables seamless connections with integrated load balancing, virtual private network (VPN) access, and automated firewall rules, while maximizing the benefits of isolated virtual private clouds (VPCs).


TIP

Learn more about VPCs at `www.vmware.com/topics/vpc`.

Enforcing identity and policy controls

An SSPC needs to provide users with access while maintaining security and compliance. Organizations should implement:

- **Identity and access management (IAM):** Role-based access control (RBAC) ensures that users and teams have appropriate permissions to provision and manage cloud resources.
- **Policy enforcement:** Automated guardrails prevent misconfigurations, unauthorized provisioning, and excessive resource consumption. Organizations can define quotas, security baselines, and compliance rules that apply across all their workloads.

Private cloud services

Organizations that require public cloud–like IaaS services can benefit greatly from utilizing private cloud services, such as VM, vSphere Kubernetes Service (VKS), Network, Volume, and VM Image. Application teams (for example, developers, DevOps, and platform engineering teams) can consume these services using a user interface (UI), command-line interface, or declarative Kubernetes (K8s)–style IaaS application programming interface (API) — for example, using the VKS service for deploying K8s clusters and the VM service for declaratively defining and provisioning VMs. Application teams can also leverage K8s manifests to provision VMs and VKS clusters, enabling agile development. A private cloud platform should also have the ability to provide additional extensible services, such as Private AI Services, Velero backup and recovery, Cert-Manager certificate management, Harbor image registry, Contour ingress controller, ExternalDNS, Secret Store, Argo CD, Istio Service Mesh, and data services, which enterprise IT admins can activate for easy consumption by application teams. These private cloud services can help accelerate developer velocity.

Preparing for automation and scalability

A scalable platform allows teams to move quickly without losing consistency, efficiency, or control. VMware Cloud Foundation (VCF) Automation supports this by managing not just infrastructure, but also the content used to define and deploy it — ensuring that images and blueprints are centrally maintained and version-controlled. Key enablers include the following:

- **Infrastructure-as-code (IaC):** Automating cloud deployment and configuration with IaC tools standardizes environments, reduces manual work, and helps to scale and apply DevOps best practices to infrastructure.
- **Workload lifecycle management:** Automated provisioning, scaling, and decommissioning of workloads prevents cloud sprawl and optimizes resource utilization. Beyond initial provisioning, automation extends to lifecycle management, allowing resources to scale dynamically based on workload demand. Policies can automatically decommission idle instances, reclaim unused storage, or migrate workloads based on performance thresholds.

- **Content library management:** Centralizing blueprints and VM images in a shared content library improves consistency across deployments, simplifies updates, and ensures that all teams use secure, approved configurations.

WARNING

Skipping any of these foundational steps can lead to performance issues, poor user consumption experience, security risks, and much higher operational costs.

Enforcing Identity and Policy Controls

An SSPC should enable users to provision and manage IaaS resources while maintaining governance and cost control. Proper configuration gives users the flexibility they need without creating operational risks or consuming unnecessary resources.

Defining resource provisioning policies

Organizations should establish clear policies to standardize how users request and consume cloud resources:

- **Compute quotas:** Limit VM sizes and resource allocations, including memory, based on team requirements.
- **Storage policies:** Enforce data retention, performance tiers, and automated backups to optimize storage usage.
- **Networking rules:** Preconfigure virtual networks, firewalls, and access control lists (ACLs) to maintain security and connectivity.

Automating infrastructure deployment

A well-configured self-service IaaS platform should leverage automation to streamline provisioning and enforce consistency. Key automation strategies include the following:

- **IaC:** Blueprints and scripts allow users to deploy repeatable, compliant infrastructure configurations.
- **Self-service provisioning portals:** User-friendly modern cloud interfaces enable teams to request, modify, and retire resources as needed.

- **Automated scaling:** Policies that dynamically adjust resources based on demand prevent overprovisioning and performance bottlenecks.

Ensuring security, compliance, and governance

Self-service access must be balanced with robust security controls to prevent misconfigurations and unauthorized usage. As environments grow more dynamic, manual processes no longer scale — organizations need to codify security and compliance requirements directly into cloud workflows.

By adopting a policy-as-code approach, teams can embed security, access control, and compliance checks into every stage of the infrastructure lifecycle, ensuring consistent enforcement without slowing down delivery. IaaS resource-based policy-as-code scales governance operations and helps to reduce the risk of human error and ensure that infrastructure resources adhere to organization requirements, thereby improving compliance.

- **RBAC** restricts provisioning capabilities based on user roles and team assignments.
- **Frictionless governance with embedded guardrails** allows admins to build custom policies for IaaS resources and eliminates the need for multiple external tools or add-ons to manage policies.
- **Logging and monitoring** tracks resource consumption, user activity, and security events to detect anomalies.
- **Compliance enforcement** uses policy-as-code frameworks to ensure workloads align with regulatory and internal governance requirements automatically.

Creating Self-Service Catalogs and Blueprints

An SSPC relies on catalogs and blueprints to provide users with fast, repeatable, and standardized infrastructure deployments. Instead of manually configuring VMs and K8s clusters,

networking, or security settings for every request, IT teams can define preapproved configurations that users deploy on demand. This improves operational efficiency, reduces misconfigurations, and ensures security policies are consistently enforced.

A self-service catalog acts as a repository of deployable infrastructure blueprints, allowing users to select and provision resources without requiring manual IT intervention. The catalog should include the following:

- **VM templates:** Preconfigured operating system (OS) images include security patches, performance optimizations, and required software packages.
- **Containerized workloads:** K8s clusters and containerized applications integrate with continuous integration/continuous delivery (CI/CD) pipelines, allowing developers to quickly deploy services that align with microservices architectures.
- **Graphics processing unit (GPU)–enabled deep learning VMs:** Users can easily select infrastructure for retrieval-augmented generation (RAG) workloads without IT intervention. This streamlining reduces IT dependencies, fosters agility, and enables quicker innovation by allowing data scientists to focus on their projects.
- **Application blueprints:** Multitier application stacks bundle compute, networking, and storage into a single deployable unit. This simplifies provisioning for complex workloads such as database-backed applications.
- **Security-hardened images:** Workloads are preconfigured with encryption, firewall settings, IAM, and compliance policies. By embedding security at the template level, organizations prevent misconfigurations and reduce exposure to vulnerabilities.

IN THIS CHAPTER

- Establishing repeatable deployment patterns to ensure scalability
- Managing workload lifecycles with automated scaling, optimization, and decommissioning
- Orchestrating infrastructure provisioning through automation
- Refining automation to maintain efficiency and governance

Chapter 4
Operationalizing Your Self-Service Private Cloud

A self-service private cloud (SSPC) is only effective if it runs efficiently and at scale. When the foundational components are in place, you'll need to focus on repeatability, lifecycle automation, provisioning orchestration, and ongoing optimization to ensure your organization's long-term success. Without structured automation, cloud environments can quickly become fragmented, leading to misconfigurations, wasted resources, and increased operational complexity.

Building Repeatable Deployments of Your Cloud

Ensuring consistent and repeatable cloud deployments is crucial to prevent configuration drift, security gaps, and operational inefficiencies. Without automation and standardization, IT teams

often rely on manual provisioning of workflows, causing delays and a lack of agility.

SSPC addresses these challenges by streamlining infrastructure delivery. It enables IT teams to deliver infrastructure as a service (IaaS) with speed, policy enforcement, and consistency across environments for virtual machines (VMs), containers, and artificial intelligence (AI) workloads. By providing a centralized, automated platform, admins can better align with business and compliance requirements.

With SSPC, platform engineers can prepare standardized infrastructure components for developers and DevOps teams and use GitOps to streamline infrastructure and application management. By leveraging Git as the source of truth for infrastructure and application configurations, GitOps enables platform teams to automate deployments, manage changes, and ensure consistency across the platform. Application teams can quickly provision and manage resources using a self-service interface, enhancing productivity and agility while adhering to IT guardrails. This brings the ease of public cloud–like services to internal operations.


TIP

Argo CD is a declarative, open source GitOps continuous delivery tool for Kubernetes (K8s). Argo CD can automatically control the cloud-native applications' deployment and manage their lifecycle. Argo CD is auditable, easy to understand, and easy to use. By running Argo CD on VMware Cloud Foundation (VCF), users can achieve the following benefits:

- Unified GitOps for all workloads from a single control plane — VMs, vSphere K8s Service (VKS) clusters, and vSphere Pods.
- Simplify operations with hands-free updates from Git to runtime.
- Define your entire stack as code that is reliable, repeatable, and fast.
- Built for VCF — integrated for seamless install, lifecycle management, and security.

Standardizing cloud deployments

Cloud deployments must be defined, version-controlled, and automated to ensure consistency and control. To achieve reliable and scalable deployments:

- **Use declarative IaC models.** Define infrastructure using formats like YAML Ain't Markup Language (YAML) to describe the desired state of compute, storage, and networking. VCF Automation can then interpret these definitions and build environments accordingly.
- **Implement standardized blueprints.** Create reusable blueprints for VMs, containers, and application stacks. Each blueprint should embed security configurations, performance optimizations, and connectivity rules by default.
- **Leverage parameterized configurations.** Avoid hardcoding environment-specific values. Use variables to adapt deployments for different teams, sites, or use cases without rewriting infrastructure logic.
- **Automate policy enforcement.** Integrate governance directly into deployment workflows by applying role-based access control (RBAC), tagging policies, and compliance validations as part of the provisioning process.

Maintaining deployment consistency across environments

An SSPC must support a variety of deployments all while maintaining uniform governance. To achieve this:

- **Enforce policy-driven provisioning.** Require that all deployed resources adhere to predefined security baselines, performance thresholds, and cost constraints.
- **Validate infrastructure before deployment.** Automate configuration testing and compliance checks to prevent misconfigurations from reaching production.
- **Use version-controlled deployments.** Store infrastructure definitions in a repository, ensuring that changes can be reviewed, audited, and rolled back when necessary.

Managing Workload Lifecycles through Automation

Effective cloud operations require more than just provisioning infrastructure. Workloads must be continuously monitored, optimized, and decommissioned to prevent resource sprawl, security risks, and unnecessary costs. Automation is an ideal solution for these problems.

Automating scaling and resource allocation

Cloud workloads should dynamically adjust based on real-time demand. Organizations can ensure scalability by:

- **Defining auto-scaling thresholds:** Configure rules that trigger scaling events based on central processing unit (CPU), memory, or network usage to maintain application responsiveness.
- **Using scheduled provisioning:** Automatically deploy workloads at predefined times, reducing idle capacity and ensuring that resources are available when needed.
- **Distributing workloads intelligently:** Automate placement policies to balance workloads, optimizing resource use and preventing localized congestion.

Optimizing performance and cost-efficiency

Running workloads efficiently requires continuous adjustments to prevent wasted resources and rising costs. Overprovisioned workloads lock up capacity, while underprovisioned ones cause performance issues. By tracking real-time metrics, organizations can dynamically adjust compute and storage allocations to match demand.

Performance optimization reduces waste by scaling workloads during peak usage and downsizing them when idle. Cost-efficiency improves through automated budget alerts, scheduled deployments, and workload placement strategies that shift non-critical tasks to lower-cost infrastructure.

REMEMBER

With automated monitoring and smart resource allocation, workloads stay performant and cost-effective without requiring constant manual oversight.

Automating decommissioning and cleanup

Unmanaged workloads can accumulate over time, leading to cloud sprawl and inefficiencies. Organizations should implement automated lifecycle management to:

- **Enforce workload expiration policies.** Automatically decommission temporary resources and enforce retention policies for unused applications.
- **Reclaim idle resources.** Identify and remove orphaned volumes, unused Internet Protocol (IP) addresses, and abandoned VMs to free up capacity.
- **Automate compliance cleanup.** Ensure that old logs, snapshots, and backups are archived or deleted according to regulatory policies.

Workload lifecycle automation ensures that resources remain optimized, responsive, and cost-efficient. Without automated scaling, monitoring, and cleanup, cloud environments can quickly become inefficient and difficult to manage.

Automating Infrastructure Provisioning and Management

With standardized deployments and workload automation in place, your next step is ensuring that your infrastructure remains agile, efficient, and policy-driven. Infrastructure provisioning must be responsive to demand, integrate with existing IT operations, and support real-time adjustments. Moving beyond initial automation, organizations must focus on orchestration, governance, and dynamic scaling to prevent resources sprawling and to maintain operational efficiency.

Advancing automation with orchestration

To build a fully orchestrated cloud infrastructure, automation must go beyond simple provisioning. As a general practice, you should:

- **Automate multistep provisioning workflows.** Ensure infrastructure components are provisioned in the correct order, avoiding configuration errors by automating dependencies like storage and network policies.
- **Utilize event-driven automation and streamline operational handoffs.** Respond to real-time triggers for actions (like workload migration) and automatically register provisioned VMs in a configuration management database (CMDB) for compliance reporting.
- **Standardize provisioning across environments.** Ensure consistent enforcement of performance, security, and compliance for deployments on-premises, in the cloud, or at the edge.

WARNING

Failing to orchestrate automation leads to inconsistent deployments, security gaps, and operational bottlenecks. A well-structured orchestration layer improves reliability and reduces effort across IT, security, and operations teams.

Enforcing continuous compliance and governance

Infrastructure automation must also align with security and operational policies. Instead of reacting to misconfigurations, your automations should

- **Embed security policies into provisioning.** Automate firewall rules, access controls, and encryption requirements at deployment instead of applying them manually after provisioning.
- **Enforce K8s governance.** Enforce policies, such as ensuring K8s clusters have only one control node for development and a minimum of three control nodes for production; enforcing minimum K8s versions for K8s clusters; and requiring K8s clusters to have a baseline pod security level.

- **Implement lifecycle rules for infrastructure resources.** Set expiration policies on temporary environments, and enforce automated cleanup processes to prevent resource sprawl.

TIP

Managing dependencies, enforcing policies, and dynamically adjusting resources ensures that infrastructure remains secure and efficient.

Monitoring and optimizing cloud resource utilization

After infrastructure is fully automated and workloads are self-service, real-time visibility and continuous optimization become essential to maintaining performance and cost-efficiency. Without active monitoring, even well-structured cloud environments can suffer from resource waste, unexpected cost overruns, or performance bottlenecks.

Effective monitoring starts with consolidated observability across compute, storage, and networking layers. Cloud platforms generate vast amounts of telemetry data, but organizations must focus on actionable insights instead of just collecting metrics.

WARNING

Collecting too many metrics without a clear optimization strategy leads to data overload rather than actionable insights.

Optimization requires more than adjusting individual workloads — it involves automating resource management based on dynamic demand patterns. Workloads should scale seamlessly when needed but also deallocate idle resources to avoid unnecessary costs. Automated policies can help rightsize VMs, enforce power schedules for nonproduction environments, and rebalance workloads to maximize infrastructure utilization.

Cost-efficiency is an ongoing process, too. Cloud usage patterns fluctuate, and a one-time optimization is almost never enough. Organizations should implement continuous evaluation cycles, using performance baselines to compare current and active consumption against historical trends. By integrating monitoring with provisioning automation, organizations can prevent overprovisioning before it happens, instead of fixing it later (often when costs have overrun or problems have occurred with the overall system performance).

Adopting an Automation-First Mindset

Operationalizing a SSPC goes beyond implementing automation tools; it requires integrating automation into the core design of infrastructure and policies. An automation-first mindset enhances efficiency, consistency, and scalability in cloud operations.

Organizations should embed automation from the start, creating repeatable processes for resource management while ensuring governance. Consistency across environments is key, as automation helps prevent configuration drift and maintain security across workloads. Infrastructure and workload automation should also be event-driven, adjusting to demand fluctuations and security or cost thresholds.

Automation is an ongoing effort, necessitating continuous monitoring and iterative improvements to stay aligned with business goals and technology changes.

TIP

Don't think about automation as just reducing manual effort — it's about building a resilient, scalable, and policy-driven cloud environment that evolves with your organization's needs.

IN THIS CHAPTER

- Isolating resources to prevent conflicts and improve security
- Managing virtual private cloud networks for controlled connectivity
- Automating governance to enforce compliance and security
- Reducing administrative overhead with policy-driven automation

Chapter 5
Managing Your Self-Service Private Cloud

A self-service private cloud (SSPC) has to provide more than just flexible access to resources — it must ensure that those resources are properly isolated, governed, and secured. Without a structured approach to segregation and management, cloud environments can become chaotic really quickly, leading to security vulnerabilities, inefficient resource use, and compliance challenges. Effective cloud management requires automated controls that allow teams to work independently while still maintaining centralized oversight — ideally through strong automation.

This chapter explores how to structure an SSPC for scalability and security. It covers isolating resources through organizational structures, managing virtual private cloud (VPC) networks for secure connectivity, and enforcing governance policies to automate security and compliance.

Isolating Resources and Infrastructure via Organizations

An SSPC has to support enterprise multitenant environments where different teams, departments, and even business units operate within the same infrastructure, as well as cloud service provider (CSP) multitenant environments where different clients operate within the same infrastructure. Without proper resource isolation, workloads can interfere with each other, your security risks increase, and operational control becomes more complicated. Organizations typically solve this by logically segregating infrastructure resources, ensuring that teams have access to only the resources they need while maintaining centralized governance.


TIP

Some organizations may have business- or industry-specific regulatory requirements to have complete isolation in their compute environments. For example, financial services firms often have both buy-side and sell-side business and typically implement strict internal controls like segregation of users, data, and infrastructure, to avoid potential conflicts of interest for risk management and regulatory compliance.

Structuring cloud resources with organizational boundaries

Segregating cloud resources requires a structured approach that allows organizations to apply policies, access controls, and quotas based on business needs, team ownership, or workload type. VMware Cloud Foundation (VCF) supports this through logical constructs such as organizations, projects, and namespaces, enabling organization admins to create resource envelopes similar to Amazon Web Services (AWS) member accounts. These boundaries make it easier to enforce security postures, manage resource utilization, and maintain operational control across diverse teams and applications.

The most common strategies include

- **Organization-based (or tenant-based) isolation:** Each tenant has no visibility or knowledge of any other tenants. This strategy is typically used when multiple separate entities share a common infrastructure — either with a service

provider model or with a centralized corporate IT department providing resources with self-service workload deployment and billing to different business units within the broader organization.

- **Project-based isolation:** Resources are grouped into projects, with each group or team having isolated compute, storage, and networking configurations. For example, large enterprises can divide infrastructure based on business functions, allowing departments like finance, engineering, or HR to operate independently while maintaining centralized oversight.
- **Namespace-based isolation:** Namespaces provide resource envelopes for applications and workloads, making it easy to apply governance, quota limits, and operational policies tailored to specific use cases or ownership groups. For example, development, testing, and production workloads can be completely isolated to prevent unauthorized changes from impacting live applications.

Enforcing resource quotas and limits

To avoid resource sprawl and overconsumption, cloud environments should enforce quotas and consumption limits at the organizational level. Best practices here include the following:

- **Define compute and storage quotas.** Restrict the number of virtual machines (VMs) and central processing unit (CPU) cores, as well as the amount of storage capacity, allocated to each team.
- **Implement budget controls.** Set spending limits and generate alerts when resource usage exceeds predefined thresholds — and add auto-shutdown to nonproduction workloads if this becomes a bigger issue.
- **Restrict privileged operations.** Prevent teams from modifying critical infrastructure settings that impact other tenants.

REMEMBER

Effective resource isolation balances team autonomy with centralized governance and ensures operational efficiency without compromising security or performance.

Managing Virtual Private Cloud Networks

An SSPC also has to provide network isolation and segmentation to ensure secure, efficient communication between workloads while still preventing unauthorized access. VPC networks allow organizations to logically separate traffic, enforce security policies, and provide granular control over how resources interact. Without proper VPC management, though, network congestion, misconfigurations, and security vulnerabilities can impact cloud operations.

Designing logical network segmentation

A well-structured VPC architecture ensures that cloud environments remain secure, scalable, and efficient. Organizations should design network segmentation based on workload type, security requirements, and operational needs. Common segmentation strategies that apply here include the following:

- **Environment-based segmentation:** Separates development, testing, and production environments to prevent accidental interference or unauthorized changes from impacting live systems
- **Application-tier separation:** Divides network traffic between front-end, middleware, and database layers, reducing lateral movement and improving security
- **Tenant-based isolation:** Ensures that different business units, departments, and teams operate within clearly distinct network spaces, preventing data leakage between these tenants

Each VPC should be configured with subnet structures, routing tables, and firewall rules that enforce traffic flow policies. Organizations should also implement network access control lists (ACLs) and security groups to filter inbound and outbound traffic based on workload sensitivity.

Implementing secure connectivity and traffic control

Beyond segmentation, network security controls must be in place to manage how workloads communicate within and outside the cloud environment. Think about the following key security approaches:

- **Defining micro-segmentation policies:** Restricts traffic flow between workloads to the minimum necessary, reducing exposure to lateral movement attacks
- **Enforcing encryption for in-transit data:** Ensures that all network traffic moving between cloud resources can be isolated
- **Configuring software-defined firewalls:** Implements per-tenant firewall policies that can adapt to changing network demands

A healthy combination of these security measures helps prevent unauthorized access, contain potential threats, and enforce compliance across cloud environments.

Automating network policy enforcement

Manually managing VPC networks at scale is inefficient and error prone. Organizations should automate network policy enforcement to maintain consistency and reduce administrative overhead. Best practices include the following:

- Applying role-based access control (RBAC) for network modifications to prevent unauthorized changes
- Using policy-driven automation to adjust routing and firewall rules based on workload requirements
- Monitoring traffic patterns with real-time analytics to detect anomalies and optimize network performance

REMEMBER

Even though some of these terms and ideas — like RBAC — show up a number of times in a number of chapters, they're still highly relevant. The same principles you'll use to set up a strong foundation still apply at the implementation and automation stages.

Implementing Governance and Policies

An SSPC provides agility, but without proper governance at scale, it can quickly lead to uncontrolled resource consumption, security vulnerabilities, and compliance risks. Effective governance ensures that users can access the resources they need without compromising security, efficiency, or operational stability. The challenge is balancing this control with flexibility — allowing teams to operate freely while keeping infrastructure secure, compliant, and optimized without overwhelming IT with oversight.

The most effective governance models are built on automation and extensibility. Instead of relying on manual approvals or checklists, organizations can define rules in code — known as policy-as-code (PaC) — and enforce them automatically as part of the infrastructure lifecycle. With this approach, governance becomes proactive, scalable, and aligned with how modern platforms are deployed. A well-governed private cloud enables IT teams to focus on delivering value, not chasing down misconfigurations or enforcing policy manually.

Establishing policy-driven governance

Governance in an SSPC should be embedded into every part of the provisioning process. By codifying policies and aligning them with infrastructure automation, teams can ensure compliance without slowing down delivery. PaC helps make these rules declarative, version-controlled, and testable — just like infrastructure-as-code (IaC).


TIP

VCF Automation leverages Kubernetes (K8s) validating admission policies to enforce infrastructure as a service (IaaS) resource policies. By using K8s's native policy engine, VCF Automation can ensure that VMs and vSphere K8s Service (VKS) clusters adhere to defined resource limits and consumption rules within namespace.

Native K8s multicluster management capabilities — including centralized VKS cluster fleet visibility, VKS policy management, data protection, and more — in VCF provide platform engineers with control over distributed K8s environments, while maintaining simplicity and operational consistency in a single, streamlined platform.

Key governance practices include the following:

- **Define RBACs.** Assign permissions based on roles, ensuring users have access only to the resources they need.
- **Enforce quota limits.** Automatically restrict compute, storage, and network usage per team, preventing overconsumption.
- **Implement tag-based policies.** Require tagging for cost tracking, security enforcement, and workload classification.
- **Use automated compliance checks.** Continuously scan infrastructure for misconfigurations or policy violations and trigger remediation actions.

By embedding policies directly into blueprints and provisioning workflows, and managing them using PaC, organizations ensure consistent enforcement across every environment — from dev to production.

WARNING

Even when implementing these best practices well, time and change will reveal inefficiencies or weaknesses. You should revisit your policies and automations regularly to ensure they stay useful and effective.

Automating security and compliance enforcement

Security and compliance must be dynamic, not static. In a self-service environment, enforcement should happen automatically the moment a workload is deployed or modified. With PaC, security policies become part of the deployment process — defined in version-controlled files, tested alongside infrastructure, and enforced automatically by the platform.

Examples of automated enforcement include

- **Enforcing encryption requirements:** Automate encryption for storage, backups, and in-transit data to meet compliance standards.
- **Auditing user activity and changes:** Track application programming interface (API) calls, configuration changes, and access attempts to detect anomalies.

- **Applying network segmentation dynamically:** Adjust firewall rules, virtual private network (VPN) access, and micro-segmentation policies based on workload sensitivity.
- **Triggering automated remediation:** Detect and correct security misconfigurations without requiring manual intervention.

PaC allows these controls to be applied uniformly, regardless of where workloads run or who deploys them — ensuring that security is never an afterthought.

Reducing administrative overhead with policy-based automation

Manual governance simply doesn't scale. Without automation, IT must constantly approve requests, validate configurations, and enforce policies — often after deployment. A PaC approach makes governance a built-in, not bolted-on, part of infrastructure delivery. It reduces friction while ensuring control. Helpful patterns include

- **Preconfigured approval workflows:** Use policy logic to automatically approve low-risk requests while routing sensitive changes for review.
- **Policy enforcement at deployment:** Validate configurations against governance rules during provisioning to prevent drift and noncompliance.

With governance and PaC in place, organizations can scale their SSPC confidently — knowing that automation enforces the rules and users have the freedom to operate within them.

IN THIS CHAPTER

- Evolving traditional IT roles to support cloud automation
- Automating AI, workload management, and governance
- Standardizing self-service catalogs for efficiency and control
- Optimizing content, security, and resource allocation

Chapter 6
Ten Benefits of Adopting VMware Cloud Foundation

A self-service private cloud (SSPC) delivers more than just automation — it transforms the private cloud consumption experience, enabling application teams to build, run, and manage all types of applications on a single platform, and more. This chapter outlines ten benefits of adopting VMware Cloud Foundation (VCF).

Advancing Your Career

Traditional vSphere Infrastructure (VI) admins play a critical role in managing virtualized environments, but as organizations shift toward an SSPC model, their responsibilities must evolve. Instead of focusing solely on virtual machine (VM) provisioning and infrastructure maintenance, cloud admins take a broader, automation-driven approach to managing resources, security,

and services. Embracing platforms like VCF allows VI admins to operate as cloud admins who:

- Orchestrate full-stack automation to streamline provisioning and reduce manual tasks.
- Enforce governance policies that ensure security, compliance, and efficiency.
- Optimize resource utilization to prevent waste and improve performance.

Instead of handling individual provisioning requests, cloud admins enable self-service, implement policy-driven automation, and ensure operational efficiency at scale. This shift reduces administrative overhead while increasing agility.


TIP

VCF Automation makes it easier for VI admins to jump-start and scale a multitenant private cloud. Guided workflows help VI admins understand what cloud concepts/governance constructs and configurations are required, enabling VI admins to become cloud admins.

Accelerating Time to Value

Broad infrastructure as a service (IaaS) adoption requires more than just making infrastructure available — it means removing friction from the request, approval, and provisioning process. A self-service catalog and self-service IaaS experience helps accomplish this by giving users fast access to either preapproved infrastructure blueprints or private cloud services. To successfully scale adoption:

- **Curate a self-service catalog.** Use standardized blueprints that can be consumed by users. Apply role-based access control (RBAC) to manage who can access specific catalog items to align with team roles and responsibilities.
- **Provide cloud services via a single, centralized consumption interface.** Access cloud services via a user interface, command-line interface, or declarative Kubernetes (K8s) IaaS application programming interface (API) for easy self-service consumption.

- **Automate provisioning workflows.** Eliminate manual steps and delays by automating approvals, configuration, and post-deployment tasks.

When users can choose what they need and deploy it themselves — without opening a ticket or waiting on manual approval — IaaS adoption grows quickly. This shift improves speed and consistency, and builds trust in IT services as fast, reliable, and developer-friendly.

Easily Implementing and Scaling Frictionless Governance

Without governance, an SSPC can quickly become unmanageable, leading to resource sprawl, security gaps, and unexpected costs. By embedding governance into cloud operations, organizations maintain efficiency while avoiding lots of unnecessary administrative overhead. Organizations benefit from governance by:

- Preventing resource overuse through automated quotas.
- Enforcing security policies across all deployments.
- Streamlining compliance by embedding regulatory requirements as part of the workflow request.


TIP

Admins can leverage catalog instance-centric policies (for example, leases, approvals, and day 2 actions) to control who can do what and how much, at the deployment level. Admins can also build custom policies for IaaS resources, using policy-as-code (PaC) or leverage predefined out-of-the-box (OOTB) policy templates to get started quickly and centrally implement resource policies, which manage namespace resources that deployments can consume, ensuring conformity and security of workloads.

Democratizing Infrastructure as Code

Manually configuring infrastructure leads to inconsistencies, errors, and delays. Infrastructure-as-code (IaC) shifts your cloud management to an automated, repeatable model where infrastructure is defined and managed through code. This approach streamlines operations while enabling faster, more consistent

deployments. The automated nature of IaC allows resources to scale up and down in a way that would be difficult to manage manually, especially in a containerized environment where microservices require infrastructure to be provisioned separately for each service.

However, its technical complexity, skill gaps, and steep learning curve (hurdles/barriers) can make it difficult for some organizations to adopt and take advantage of IaC. For users with little or no coding experience, VCF offers a unique low-code graphical user interface (GUI) that generates code on the side, so users can use the GUI or follow the code on the side to learn along the way.

With IaC in place, organizations accelerate provisioning, reduce overhead, and empower users with ready access to the infrastructure they need — while maintaining full control behind the scenes.

Building, Running, and Managing All Your Application Workloads on One Platform

An SSPC must support a mix of workloads, from legacy applications to cloud-native microservices and artificial intelligence (AI)–driven processing. Without a unified management approach, organizations struggle with siloed operations, inconsistent policies, and inefficient resource allocations. A well-designed SSPC brings:

- A single control plane for provisioning and managing all workload types.
- Consistent security and governance across VMs, containers, and AI models.
- Optimized resource allocation to ensure workloads get the right compute, storage, and networking resources.


TIP

For VI admins that are new to K8s, there's no need to worry. VCF has a built-in K8s runtime that can be managed via the user interface (UI). Simply follow the UI/guided workflows and set up the cloud environment to support K8s. Platform engineers can then easily create and manage K8s objects, using a low-code/no-code UI or K8s manifests.

Easily Getting Started on Private AI with Guided Workflows

AI workloads demand high-performance compute, dynamic scalability, and tight governance. Without automation, provisioning and managing the infrastructure needed to support them becomes a slow, error-prone process. An SSPC simplifies AI operations automating private AI service setup and by automating provisioning, scaling, and lifecycle management — reducing delays and enabling faster time to results.

Private cloud consumers like developers and data scientists benefit directly. With approved, reusable templates, they can quickly spin up graphics processing unit (GPU)–capable deep learning VMs, GPU-enabled K8s clusters, or even retrieval-augmented generation (RAG) workstations powered by pgvector and PostgreSQL — without opening a ticket or configuring infrastructure manually.

With governance policies in place, cloud admins can enforce resource quotas, track usage, and apply cost controls to ensure that high-performance infrastructure is used responsibly. Private AI automation allows teams to move fast while staying within operational and budgetary boundaries.

Increasing Efficiency with Seamless Orchestration and Extensibility

Orchestration in an SSPC ensures that every component, from compute to networking, operates as part of a seamless workflow. Orchestration eliminates the need for manual intervention by enabling complex, multistep deployments that respond dynamically to demand, security policies, and governance requirements. By integrating extensibility, your organization can

- Automate full-stack provisioning, from infrastructure to application services.
- Leverage event-driven automation to automate tasks and workflows triggered based on events occurring in the environment.

- Extend workflows with third-party integrations for security, monitoring, and compliance.

Supporting All Your Stakeholders with Multitenancy

An SSPC must support all your clients, teams, departments, and internal customers without compromising security, performance, or efficiency. Proper tenant management ensures that each group operates independently while still benefiting from shared infrastructure resources. By isolating tenants using organizations and virtual private clouds (VPCs), IT teams can enforce controls while enabling autonomy. This includes

- Assigning resource quotas to prevent overuse and maintain performance across tenants
- Applying security and governance policies at the tenant level to ensure compliance
- Enabling self-service access to approved resources while preserving isolation
- Tracking resource utilization per tenant to identify inefficiencies or spikes
- Implementing chargeback or showback models to increase accountability and promote responsible usage

With tenant-aware operations in place, cloud admins can monitor activity, optimize capacity, and ensure that each tenant gets the access they need — without impacting others.


TIP

VCF Automation provides a streamlined way to jump-start the creation of a multitenant cloud with quick-start wizards and guided workflows, so an admin who's new to multitenancy can quickly and easily get started setting up organizations (tenants) just like a cloud service provider (CSP).

Ensuring Consistency and Compliance with Content Management

An SSPC must also handle images and blueprints efficiently to ensure consistency across deployments. A centralized content library helps maintain version-controlled blueprints for repeatable, compliant deployments and lets you distribute approved images and configurations across environments. Automating content updates helps ensure that security and performance standards are enforced, reducing manual effort and the risk of outdated configurations.


TIP

With the new content management capabilities in VCF, managing and sharing standardized content across organizations and projects just got a whole lot easier, increasing efficiency and productivity by reducing time-consuming manual tasks and human intervention associated with managing content across different groups of users.

Reducing sprawl, Security Vulnerabilities, and Costs with Full Workload Lifecycle Management

Workloads in an SSPC must be efficiently managed from deployment to retirement. Without lifecycle management, organizations risk resource sprawl, security vulnerabilities, and escalating costs. A structured approach ensures that workloads are provisioned, monitored, optimized, and decommissioned according to business needs.

Effective workload lifecycle management starts with automated provisioning, ensuring that workloads are deployed consistently using predefined policies. As workloads run, performance monitoring and auto-scaling optimize resource consumption, preventing overprovisioning while maintaining availability. Finally, automated decommissioning removes unused workloads, freeing up resources and maintaining a clean cloud environment.

Unlock agility with self-service private cloud

Discover how to bring the speed, agility, and simplicity of the public cloud to your on-premises environment. *Self-Service Private Cloud For Dummies* shows IT leaders how to modernize infrastructure, automate operations, and empower teams through self-service access. Learn to streamline provisioning, enforce governance, and manage virtual machines, containers, and artificial intelligence workloads — all from a single platform. Unlock faster delivery, stronger control, and smarter cloud economics with self-service private cloud.

Inside...

- Define self-service private clouds
- Configure infrastructure as code
- Manage tenants and governance
- Scale resources with built-in guardrails
- Deploy virtual machines and clusters instantly
- Automate artificial intelligence and app workloads
- Track success with real metrics

Alina Thylander, **Taka Uenishi**, and **Vincent Riccio** are Product Marketing Engineers in Broadcom's VCF Division, focusing on infrastructure automation. Alina drives market adoption through core product marketing, Taka leads marketing efforts, and Vincent creates technical marketing content, all contributing to the division's success.

ISBN: 978-1-394-34792-6
Not For Resale

for dummies®
A Wiley Brand

WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley’s ebook EULA.