Reference Design: VMware Cloud Foundation on Lenovo ThinkAgile VX

Last update: 26 December 2023
Version 1.5
Configuration Reference Number: CLDVW03XX12

Describes reference design for VMware Cloud Foundation on Lenovo ThinkAgile VX appliances

Includes VMware Tanzu Kubernetes platform for modern applications development and VMware Private AI Foundation framework for AI/ML workloads

Includes details about hybrid cloud connectivity to Amazon Web Services and Microsoft Azure

Contains Lenovo XClarity integrators for VMware SDDC products

Chandrakandh Mouleeswaran

Cristian Ghetau

LENOVO PRESS

Table of Contents

1 Introduction

This document describes the reference design of VMware Cloud Foundation (VCF) on Lenovo® ThinkAgile VX servers. VCF on ThinkAgile VX is a way to implement a hybrid cloud solution as a rack based integrated system. This solution is built using ThinkAgile VX hardware from Lenovo, VMware Software Defined Data Center (SDDC) software capabilities and Lenovo XClarity integrators. These three major components come together to give the customers a turnkey hybrid cloud solution with tight integration for ease of management. It provides customers a hyperconverged infrastructure (HCI) solution with automated life cycle management (LCM) capabilities. This document also covers the different components required for implementing an on-premises VMware Cloud Foundation appliance along with a description of various ThinkAgile VX servers available from Lenovo for the customer to pick the right sized solution for their business needs.

The intended audience of this document are IT professionals, technical architects, sales engineers, and consultants to assist in planning, designing, and implementing SDDC products. General understanding of server virtualization, cloud services and VMware software is expected to get the most out of the paper.

This reference design covers the following VMware products:

- vSphere 8.0U2 which provides compute virtualization
- vSAN 8.0 U2, which provides software defined storage (SDS)
- VMware Cloud Foundation 4.2.1 which automates the entire SDDC system lifecycle and simplifies software operations.
- NSX-T Data Center 3.1.3 which provides network virtualization and security by using software defined networking (SDN) and supports private, public, and multi-clouds.
- vRealize Suite 8.5.1, which provides cloud management capabilities for private, public and hybrid clouds with support for multiple hypervisors
- Tanzu Kubernetes Grid 1.3 which provides a container platform to run Kubernetes 1.20.4 in vSphere to build and deploy modern applications leveraging support from the opensource ecosystem.
- VMware HCX 4.2 which provides infrastructure abstraction and management allowing multi-cloud connectivity and hybrid workflows for Enterprise & Provider Clouds

This document provides an overview of the business problem that is addressed by VCF and embedded SDDC products and the business value that is provided by the SDDC products and Lenovo ThinkAgile VX certified nodes for hybrid cloud and modern applications deployment. A description of customer requirements is followed by an architectural overview of the solution and a description of the logical components. The operational model describes the architecture for deploying into small to medium enterprises. Performance and sizing information is provided with the best practices and networking considerations for implementing SDDC products.

See also the Reference Architecture for VMware vCloud Suite (lenovopress.com/lp0660) which uses network shared storage instead of VMware vSAN.

2 Business problem and business value

This chapter provides a summary of the business problems that this reference design is intended to help address, and the value that this solution can provide.

2.1 Business problem

With rising costs and complexity, it is becoming increasingly harder to manage IT infrastructure in a data center to address private cloud, hybrid cloud and container workloads. As it changes over time, the infrastructure becomes more fragile and more difficult to know the impacts of making changes. Overlaid on the infrastructure issues are the business demands to both reduce costs and at the same time provide a platform to develop more flexible polyglot applications that can meet the business and end-user demands for agility, stability, performance, availability, and easier upgradability.

2.2 Business value

VMware Cloud Foundation (VCF) is a hybrid cloud platform to deploy VMware SDDC for private cloud based on the VMware Validated Design and to integrate with public clouds running VMware SDDC clouds. It provides software defined services for compute, storage, networking, and cloud management to run different workloads. It simplifies installation, upgrade and patch management of SDDC components through lifecycle management either through online or offline.

VCF built on ThinkAgile VX hardware and embedded with VMware SDDC provides all the hardware and software needed for building an enterprise infrastructure platform to support virtualized and containerized workloads that is flexible, easy to manage and easy to change for future needs. By virtualizing compute, storage and networking, SDDC is less dependent on physical hardware. Together with the addition of policy driven configuration, lifecycle management and on demand provisioning, SDDC makes it easier to manage, extend and upgrade the underlying infrastructure to address monolith and microservices architectures. The Lenovo ThinkAgile VX certified nodes and appliances solution for VMware SDDC provides businesses with an affordable, interoperable, and reliable industry-leading cloud solution to manage all of their virtualized and containerized workloads.

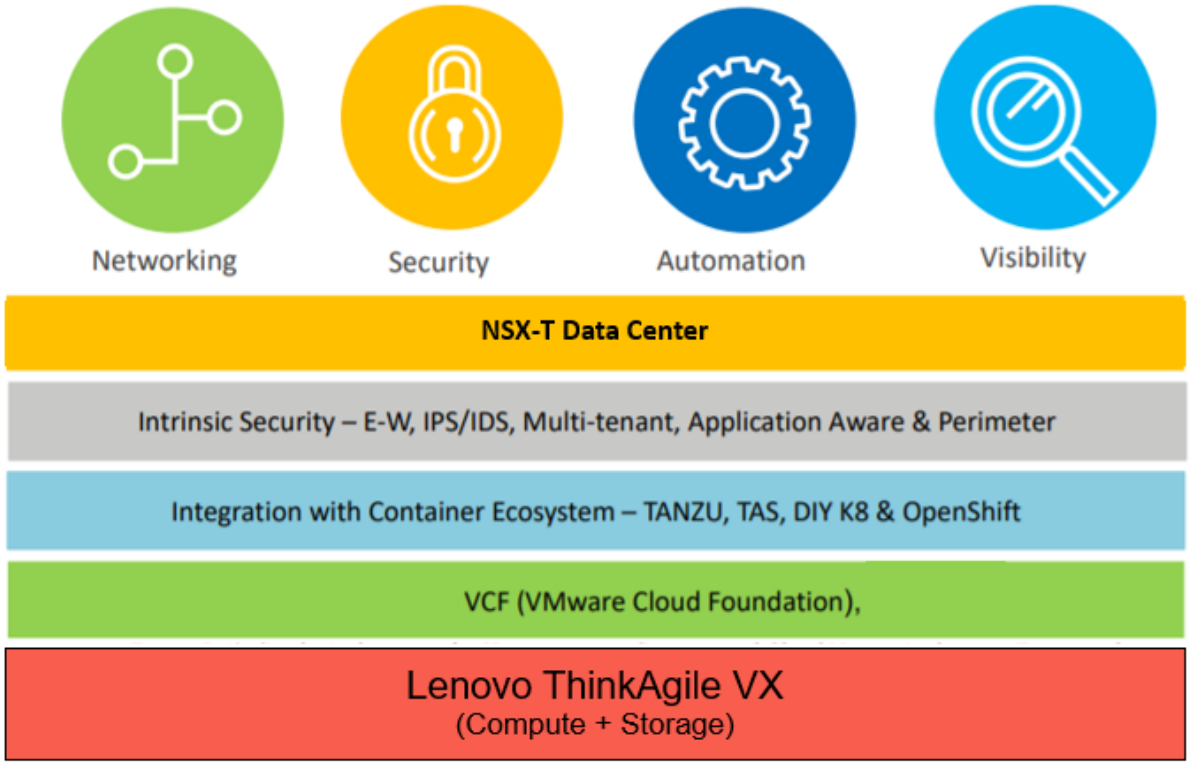

Figure 1: Lenovo ThinkAgile VX for VMware SDDC

3 Requirements

This chapter describes the functional and non-functional requirements for this reference design.

3.1 Functional requirements

The following section describes the functional requirements that are needed for typical multi cloud deployments. Figure 2 shows a simplified use-case model for hybrid cloud deployments.

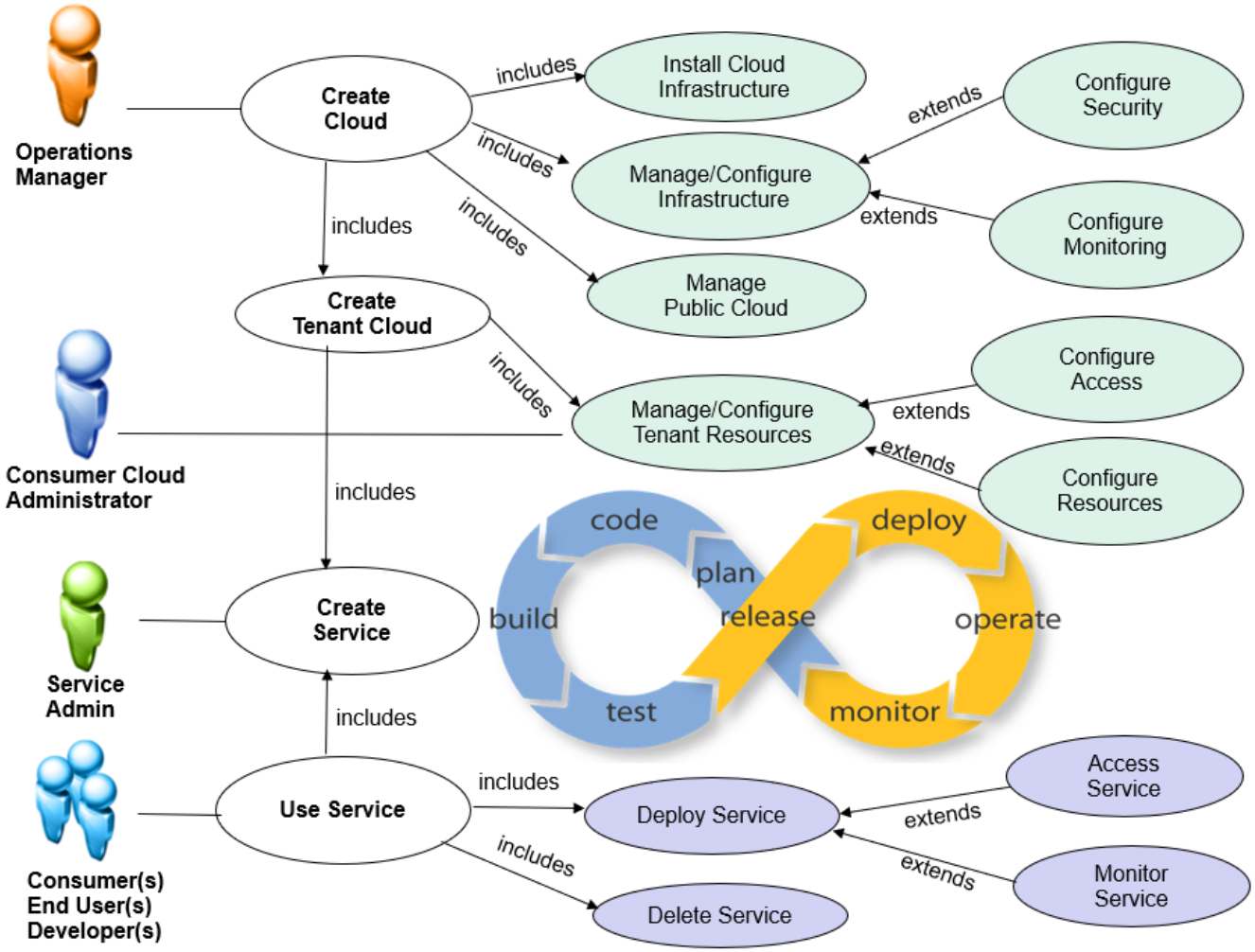

Figure 2: Use case model

Table 1 lists the functional requirements.

Table 1: Functional requirements

Requirement name	Description
Virtualization	Solution supports compute, storage, and network virtualization
Containerization	Mange and deploy containers on the virtualized infrastructure
Monitoring, event and capacity management	Monitors the health of the cloud infrastructure, collection and management of exception events, and capacity planning
Self-service automation	Solution provides on boarding, provisioning, and management of services and VMs from a service catalog
Approval and workflow	Provides the capability to approve, modify, deny, and delegate service requests
Cloud administration	Provides capabilities to administer a cloud environment, such as adding storage or computational resources in the cloud pool or defining new segregated networks
Image management	Provides capabilities to create VMs and containers, establish version control, search for and compare images, and delete images from the virtual images templates repositories
Service management	Provides capabilities to create services, establish version control, search for services, and delete services from the service templates catalog repositories
Access and authorization Controls	Provides the capabilities to create users and groups and to establish authorization to certain features in the cloud, such as tenant cloud administration, service developer, and user service requester
Virtual Machine Migration	Migrate applications, virtual machine and templates between private and public clouds.
Migrate Security Policies	Migrate network and security policies such as firewall rules to public cloud and vice versa,
Network Extension	Retain virtual machines network properties (L2 and L3) across clouds.
Catalog Management	Maintain common catalog for templates across clouds.
Hybrid Cloud Integration	Supports connectivity to seamlessly migrate and manage workloads in multi cloud environments.
Opensource ecosystem	Supports integration and flexibility to leverage open-source software in the platform.
DevSecOps	An advanced approach to security that simplifies and automates container operations across multi-clouds.

3.2 Non-functional requirements

Table 2 lists the non-functional requirements that are needed for typical cloud deployments.

Table 2: Non-functional requirements

Requirement name	Description
Backup/Recovery	Solution support for integrated backup
Ease of installation	Reduced complexity for solution deployment
Ease of management/operations	Simple management of infrastructure and cloud software
Supportability	Available vendor support

Requirement name	Description
Scalability	Solution components scale with increase in number of concurrent users, VMs/services provisioned per minute or per hour
Flexibility	Solution supports variable deployment methodologies
Security	Solution provides ways to secure customer data
Reliability, availability, and serviceability (RAS)	High availability and resiliency of cloud management and managed infrastructure

4 Architectural overview

This chapter gives an architectural overview of SDDC products. Figure 3 gives an overview of how those products are deployed into management, edge and compute and additional compute clusters and seamlessly integrated with different public clouds. This separation of function into these clusters allows for scaling in larger environments.

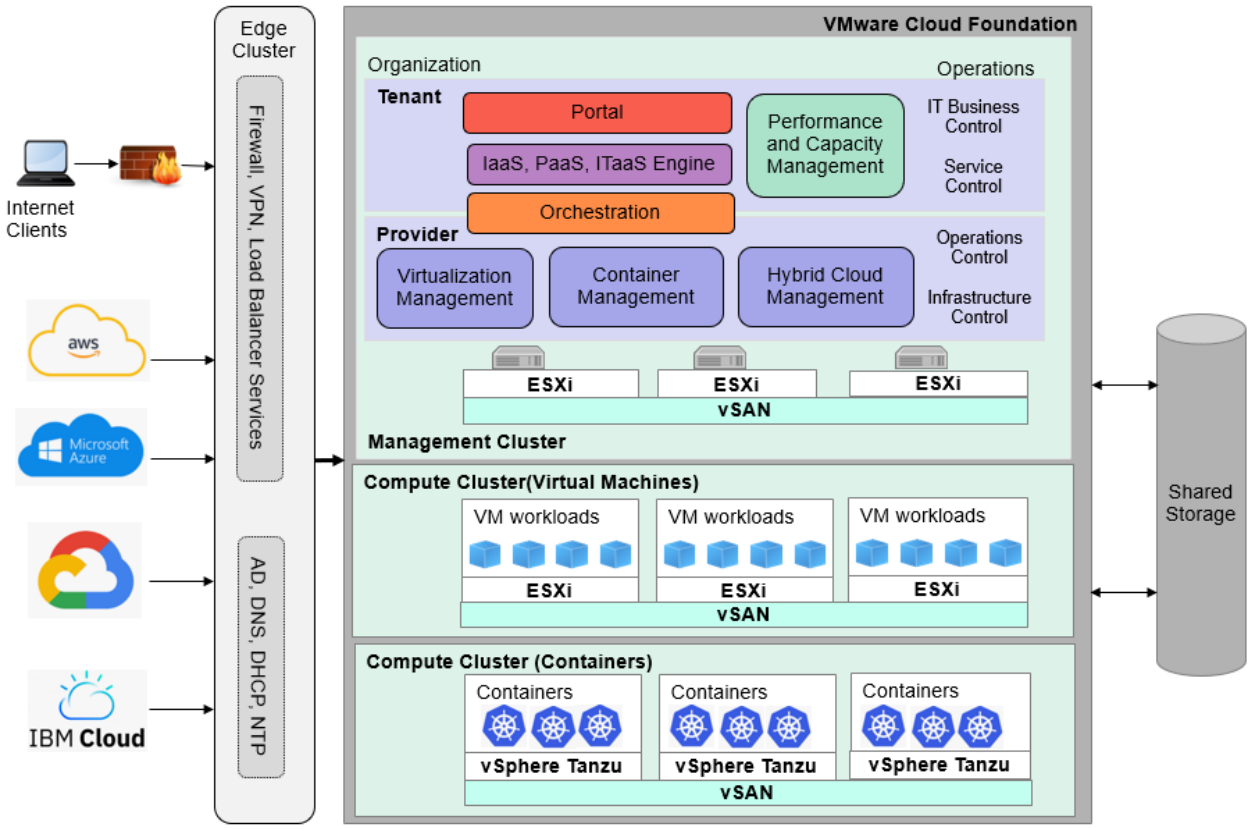

Figure 3: Conceptual design of a SDDC environment

The management cluster runs the components required to support SDDC and is used for management of virtualization and container platforms, public cloud management, monitoring, and infrastructure services. A management cluster provides resource isolation which helps these services to operate at their best possible performance level.

Dedicated edge cluster required for large environments and for small medium deployments, the edge services can coexist in either management or compute clusters. Edge provides protected capacity by which internal data center networks connect via gateways to external networks. Networking edge services and network traffic management occur in this cluster and all external facing network connectivity ends in this cluster. The shared edge and compute cluster also supports the delivery of all other (non-edge) customer workloads and there can be one or more compute clusters, depending on the customer environment. Multiple compute clusters can be for different organizations or tenants, different workload types, or to spread the load in a large enterprise.

5 Component model

This chapter describes the component model for VMware SDDC and optionally extending it into public clouds with hybrid cloud connections. Lastly the HyTrust suite of software is described which provides additional security protection features.

5.1 VMware SDDC Components

Figure 4 shows an overview of the major components of the VMware SDDC.

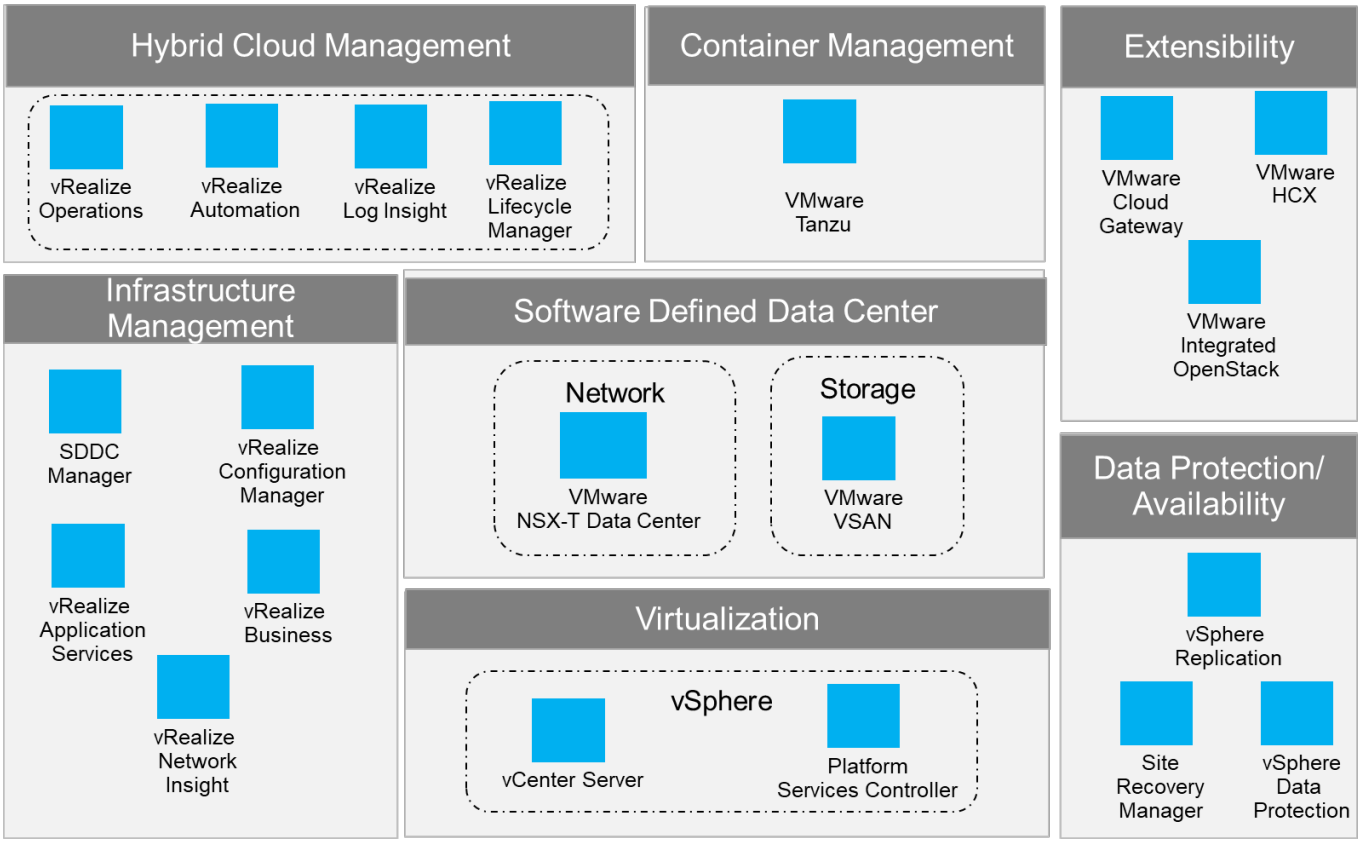

Figure 4: SDDC components

The VMware SDDC features the following components:

ESXi hypervisor	Provides bare-metal virtualization of servers so you can consolidate your applications on less hardware.
vCenter Server	Provides a centralized platform for managing vSphere environments and includes vSphere replication and vSphere data protection.
Platform Services Controller (PSC)	Provides a set of common infrastructure services that encompasses single sign-on (SSO), licensing, and a certificate authority (CA).

VMware Cloud Foundation (VCF)	Suite of components to deploy and manage your software-defined data center (SDDC)
SDDC Manager	Provides management interface to VCF. It performs deployment of ESXi, vRealize Suite, NSX-T and lifecycle management operations.
Cloud Builder	Used to deploy and configure the first cluster of the management domain and transfer inventory and control to SDDC Manager
vRealize Suite Lifecycle Manager	Provides deployment options such as install, configure, import, and upgrade vRealize Suite environments and perform drift analysis and view the health of those environments
vRealize Automation	Provides a self-service, policy-enabled IT and application services catalog for deploying and provisioning of business-relevant cloud services across private and public clouds, physical infrastructure, hypervisors, and public cloud providers.
vRealize Operations	Provides a set of components for automation of operations including infrastructure health, configurations and compliance, application discovery, and monitoring of hardware and software.
• vRealize Operations Manager	Provides comprehensive visibility and insights into the performance, capacity and health of your infrastructure.
• vRealize Configuration Manager	Provides automation of configuration and compliance management across your virtual, physical, and cloud environments, which assesses them for operational and security compliance.
• vRealize Infrastructure Navigator	Provides automated discovery of application services, visualizes relationships, and maps dependencies of applications on virtualized compute, storage, and network resources.
• vRealize Hyperic	Provides monitoring of operating systems, middleware, and applications that are running in physical, virtual, and cloud environments.
vRealize Business for Cloud	Provides transparency and control over the costs and quality of IT services that are critical for private (vCloud Suite) or hybrid cloud (vRealize Suite) success.
vRealize Log Insight	Provides analytics capabilities to unstructured data and log management, which gives operational intelligence and deep, enterprise-wide visibility across all tiers of the IT infrastructure and applications. Standard for vRealize Suite.
vRealize Network Insight	Provides end-to-end management and helps you gain visibility for NSX, VMware SD-WAN, VMware Cloud on AWS, Tanzu Kubernetes Grid

vCenter Site Recovery Manager (SRM)	Provides disaster recovery capability with which you can perform automated orchestration and non-disruptive testing for virtualized applications by using ESXi hypervisor only. SRM is standard for vCloud Suite and optional for vRealize Suite.
NSX-T Datacenter	NSX provides network virtualization and functions and is part VMware's vision of the SDDC. Refer "VMware NSX-T Data Center"
VMware Hybrid Cloud Extension (HCX)	Provides Hybrid cloud connectivity between on premise VMware cloud and public cloud SDDC
Tanzu	Kubernetes based container platform for vSphere and supports development of modern applications.

The SDDC products also have dependencies on the following external components:

Identity source	Identity sources (Active Directory, OpenLDAP, or Local OS) or similar is required to implement and operate the vCloud Suite or vRealize Suite infrastructure.
DNS	DNS must be configured for connectivity between vCenter Server, Active Directory, ESXi hosts, and the VMs
DHCP/TFTP	PXE boot is required for vSphere Auto Deploy functionality.
Time synchronization	Accurate time keeping and time synchronization is critical for a healthy infrastructure. All components (including ESXi hosts, vCenter Server, the SAN, physical network infrastructure, and VM guest operating systems) must have accurate time keeping.
Microsoft SQL Server database	Many of the SDDC components come with embedded PostgreSQL database or they can use Microsoft SQL Server as external database depending on the component and the intended environment.

Other software components such as Lenovo XClarity Administrator are not shown. As well as providing management of Lenovo hardware, XClarity Administrator also has plugins for VMware vCenter, VMware vRealize Orchestrator, and VMware vRealize Log Insight which are further described in "Systems management for Lenovo servers" on page 34.

5.2 VMware vSAN

VMware vSAN is a Software Defined Storage (SDS) solution embedded in the ESXi hypervisor and provides flexible configurations with mix of SSD, NVMe and HDDs. VMware vSAN All Flash pools flash devices for caching and capacity tiers and vSAN Hybrid uses flash for cache and magnetic disks for capacity across three or more 10 GbE connected servers into a single shared datastore that is resilient and simple to manage.

VMware vSAN can be scaled to 64 servers, with each server supporting up to five disk groups, with each disk group consisting of a one solid-state drives (SSDs) or NVMe drives for cache and up to seven SSDs or hard disk drives (HDDs) for capacity. Performance and capacity can be easily increased by adding components, such as disks, disk groups, flash devices, or servers.

The flash cache is used to accelerate reads and writes. Frequently read data is kept in read cache; writes are coalesced in cache and destaged to disk efficiently, which greatly improves application performance. vSAN All Flash uses cache for write back cache only and reads happens through capacity drives.

VMware vSAN manages data in the form of flexible data containers that are called *objects*. The following types of objects for VMs are available:

- VM Home
- VM swap (`.vswp`)
- VMDK (`.vmdk`)
- Snapshots (`.vmsn`)

Internally, VM objects are split into multiple components that are based on performance and availability requirements that are defined in the VM storage profile. These components are distributed across multiple hosts in a cluster to tolerate simultaneous failures and meet performance requirements. VMware vSAN uses a distributed RAID architecture to distribute data across the cluster. Components are distributed with the use of the following two storage policies:

- Number of stripes per object. It uses RAID 0 method.
- Number of failures to tolerate. It uses either RAID-1 or RAID-5/6 method. RAID-5/6 is currently supported for an all flash configuration only.

VMware vSAN uses the Storage Policy-based Management (SPBM) function in vSphere to enable policy driven VM provisioning, and uses vSphere APIs for Storage Awareness (VASA) to make available vSAN storage capabilities to vCenter. This approach means that storage resources are dynamically provisioned based on requested policy, and not pre-allocated as with many traditional storage solutions. Storage services are precisely aligned to VM boundaries; change the policy, and vSAN implements the changes for the selected VMs. Table 3 lists the vSAN storage policies.

Table 3: vSAN storage policies

Storage Policy	Description	Default	Maximum
Failure Tolerance Method	Defines a method used to tolerate failures. RAID-1 uses mirroring and RAID 5/6 uses parity blocks (erasure encoding) to provide space efficiency. RAID-5/6 is supported only for All Flash configurations. RAID 5 requires minimum 4 hosts and RAID 6 requires minimum 6 hosts. When RAID 5/6 is chosen, RAID 5 is used when FTT=1 and RAID 6 is used when FTT=2.	RAID-1	N/A

Primary level of failures to tolerate	Defines the number of host, disk, or network failures a VM object can tolerate. For *n* failures tolerated, n+1 copies of the VM object are created and 2n+1 hosts with storage are required. For example with a FTT=1, RAID-1 uses 2x the storage and RAID-5/6 uses 1.33x the storage. When FTT=2, RAID-1 uses 3x the storage and RAID-5/6 uses 1.5x the storage.	1	3
Secondary level of failures to tolerate	Works only for stretched clusters and defines the number of disk or host failures a storage object can tolerate for each of the sites. A storage object with the primary level of failures “m” and secondary level of failures “n” can tolerate “n” host or disk failures in addition to “m” site failures. Supported values are 0 to 3 depending on the fault tolerance method (erasure coding can tolerate up to 2 failures). For each of the sites the number of required hosts in order to tolerate “n” failures is “2n+1” for mirroring and 4 or 6 for erasure coding(failures would be 1 or 2 respectively)	0	3
Number of disk stripes per object	The number of HDDs across which each replica of a VM object is striped. A value higher than 1 might result in better performance, but can result in higher use of resources.	1	12
Object space reservation	Percentage of the logical size of the object that should be reserved (or thick provisioned) during VM creation. The rest of the storage object is thin provisioned. If your disk is thick provisioned, 100% is reserved automatically. When deduplication and compression is enabled, this should be set to either 0% (do not apply) or 100%.	0%	100%
Flash read cache reservation	SSD capacity reserved as read cache for the VM object. Specified as a percentage of the logical size of the object. Should be used only to address read performance issues. Reserved flash capacity cannot be used by other objects. Unreserved flash is shared fairly among all objects.	0%	100%
Force provisioning	If the option is set to Yes, the object is provisioned, even if the storage policy cannot be satisfied by the data store. Use this parameter in bootstrapping scenarios and during an outage when standard provisioning is no longer possible. The default of No is acceptable for most production environments.	No	N/A
IOPS limit for object	Defines IOPS limit for a disk and assumes a default block size of 32 KB. Read, write and cache operations are all considered equivalent. When the IOPS exceeds the limit, then IO is throttled.	0	User Defined

Disable object checksum	Detects corruption caused by hardware/software components including memory, drives, etc. during the read or write operations. Object checksums carry a small disk IO, memory and compute overhead and can be disabled on a per object basis.	No	Yes
Data locality	Specify the data location. Either the preferred fault domain or Non-preferred fault domain in a stretched cluster, or set to Host local to pin the VMs folder and VMDKs to the host it was created on. This policy is only valid for objects with the primary level of failures to tolerate = 0. Default value: None	None	N/A

5.3 VMware vSAN Data Persistence Platform

VMware vSAN Data Persistence Platform(DPp) supports adding third party k8 operators, storage classes and vCenter services to provide S3 compatible scalable object storage for containers and integrated with VMware Cloud Foundation with VMware Tanzu. vSAN DPp enables running stateful services and provides multitenancy and it is supported on both vSAN Shared Nothing Architecture (SNA) and vSAN Direct configurations. It is integrated with complete lifecycle management and maintenance operations, and it provides seamless experience to run third party Kubernetes services. It enables to run various data science and AI/ML workloads on vSAN at scale without compromising performance.

5.4 VMware NSX-T Data Center

VMware NSX-T™ Data Center is an SDN solution that allows the creation of overlay networks with the same capabilities that are available in the physical network. Clients can build multi-tier application networks and implement micro-segmentation to mitigate against threats that penetrate through the perimeter firewall. VMware NSX can be used with VMware vSphere hypervisor and also with several other hypervisors.

When deployed, VMware NSX-T is a collection of virtual machines that work collectively to support the overlay network. These components are distributed across multiple hosts or clusters and can tolerate simultaneous failures while providing optimal performance. Table 4 lists the NSX-T components.

Table 4: NSX-T Components

Component	Description
NSX Manager	management plane for the NSX-T Data Center and provides configuration and orchestration of logical switching and routing, edge services, security services and distributed firewall.
NSX Policy Manager	Provides policy-based access to NSX-T Data center services.
Cloud Service Manager	Manages all public cloud NSX-T environment communications.
NSX Controller	Distributed state management system that controls virtual networks and overlay transport tunnels.

Component	Description
Transport Node	The ESXi hosts are transport nodes and the communication happens through one or more VTEP endpoints on the hosts.
Virtual Tunnel Endpoint (VTEP)	VMkernel interface that is created by the NSX-T manager during the initial preparation of the ESXi Host to participate in the overlay network.
Edge Services Gateway	The Edge Services Gateway gives you access to all NSX Edge services, such as firewall, NAT, DHCP, VPN, load balancing, and high availability. Each Edge Services Gateway can be configured for single or multiple services and have a total of 10 uplink and internal network interfaces. The internal interfaces connect to secured port groups and act as the gateway for all protected virtual machines in the port group.
Logical Switch	Provides a representation of Layer 2 switched connectivity across many hosts with Layer 3 IP reachability between them. It used to isolate tenants from each other.
Distributed Router	East-West routing and it is handled by transport nodes.
Service Router	Edge nodes serve stateful centralized services NAT, DHCP server, VPN, Gateway Firewall, Bridging, Service Interface, Metadata Proxy for OpenStack. Provides north-south routing.
Physical Router	A physical router that is logically connected to each ESXi host in the data center.
Two-Tier routing	Multi-tier routing can be design using DR, SR and physical routers across gateways
Virtual Routing Forwarding (VRF)	virtualization method that consists of creating multiple logical routing instances within a physical routing appliance. It provides a complete control plane isolation between routing instances.
Distributed Firewall (DFW)	provides stateful protection of the workload at the vNIC level and enforcement occurs in the hypervisor kernel, helping deliver micro-segmentation
Load Balancer	Provides Layer 4 and Layer 7 load balancing features

Table 5 lists the various logical networks in which these components are deployed.

Table 5: NSX Component Logical Networks

Logical Network	NSX Component/Service
Management Plane	NSX Manager, Policy Manager, Cloud Service Manager
Control Plane	NSX Controllers
Data Plane	NSX VIBs, NSX Edge, NSX Firewall, NSX Logical (Distributed) Router, Transport Zones

Figure 5 shows the standard set of icons that are defined by VMware to represent the various NSX-T components.

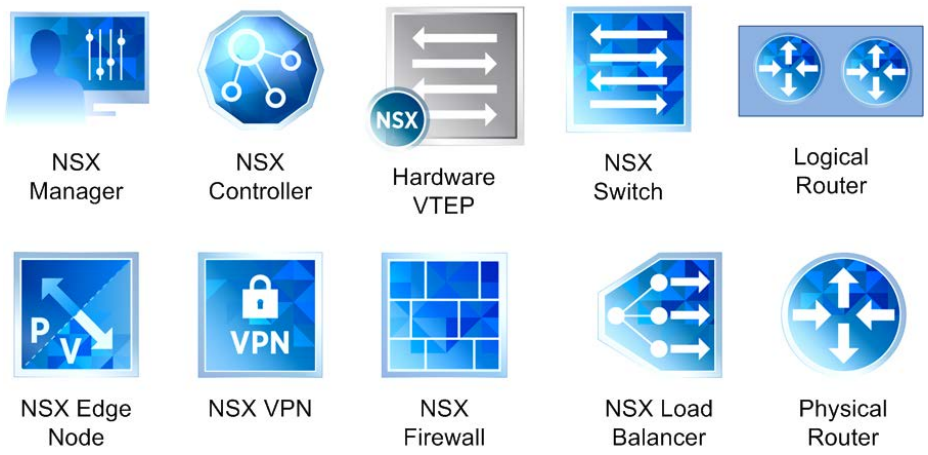

Figure 5: NSX-T Standardized Icons

5.5 Hybrid Clouds

VMware SDDC can run either in on-premises or on any other public clouds such as Amazon Web Services (AWS), Microsoft Azure, IBM Cloud and Google Cloud Platform. VMware vRealize can manage workloads across clouds and workloads can be seamlessly provisioned and migrated across different SDDC environments.

5.5.1 VMware Hybrid Cloud Extension (HCX)

Enables on-premises SDDC workloads to migrate and rebalance to different public clouds running VMware Cloud. The migration can be done live or batch or scheduled and vRealize Network Insight helps to monitor the migration. NSX Hybrid Connect can be used to migrate virtual machines between two on-premises VMware SDDC cloud. HCX supports various features for proxy and WAN optimization to improve throughput and do migration at scale.

Table 6: VMware Hybrid Cloud Extension support

Source Cloud	Components	Target Cloud
On Premise VCF	HCX	On Premise VCF
On Premise VCF	HCX, NSX Hybrid Connect	VMware Cloud on Amazon Web Services (AWS)
On Premise VCF	VMware NSX® Advanced or Enterprise through IBM Cloud	IBM Cloud for VMware Solutions
On Premise VCF	HCX Connector, HCX Cloud Manager Appliance	Google Cloud VMware Engine,
On Premise VCF	VMware HCX Connector, Azure VMware Solution HCX Cloud Manager	Azure VMware Solution

5.6 VMware Tanzu Kubernetes Platform

A Tanzu Kubernetes Cluster is a full distribution of the open-source Kubernetes container orchestration platform that is built, signed, and supported by VMware. vSphere with Tanzu offers a VM Service functionality that enables DevOps engineers to deploy and run VMs, in addition to containers, in a common, shared Kubernetes environment. By using vSphere with Tanzu the vSphere Administrator can turn a vSphere cluster to a platform for running Kubernetes workloads in dedicated resource pools. The vSphere administrator can manage and monitor vSphere Pods, VMs, and Tanzu Kubernetes clusters by using the vSphere Client.

5.6.1 vSphere with Tanzu

Both, containers and VMs, share the same vSphere Namespace resources and can be managed through a single vSphere with Tanzu interface. The VM Service addresses the needs of DevOps teams that use Kubernetes but have existing VM-based workloads that cannot be easily containerized. It also helps users reduce the overhead of managing a non-Kubernetes platform alongside a container platform. When running containers and VMs on a Kubernetes platform, DevOps teams can consolidate their workload footprint to just one platform. Figure 6**Error! Reference source not found.** below shows the virtualization and container components in vSphere with Tanzu architecture.

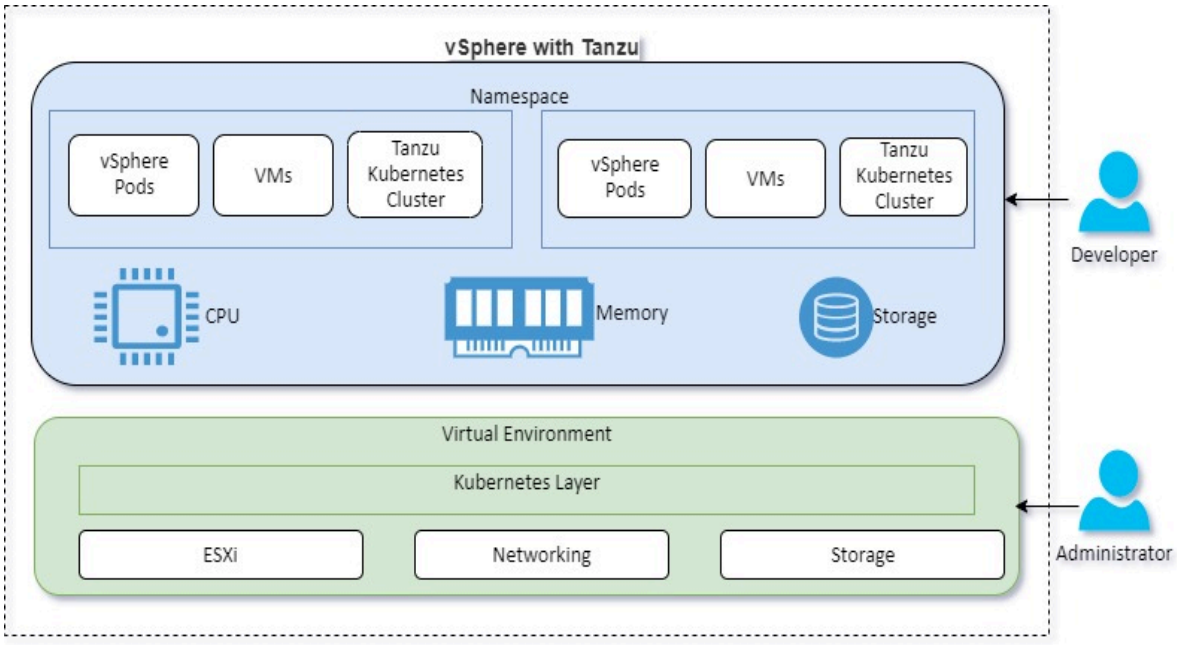

Figure 6: vSphere with Tanzu

The Table 7 shows the list of core components in Tanzu

Table 7: vSphere Tanzu Components

Component	Description
Supervisor Cluster	A cluster that is enabled for vSphere with Tanzu is called a Supervisor Cluster. It runs on top of an SDDC layer that consists of ESXi for compute, NSX-T Data Center or vSphere networking, and vSAN or another shared storage solution
Kubernetes control plane VM	Three Kubernetes control plane VMs in total are created on the hosts that are part of the Supervisor Cluster and three control plane VMs are load balanced
Tanzu Kubernetes Grid Service	Kubernetes control plane runs directly on the hypervisor layer
vSphere Pod	A vSphere Pod is a VM with a small footprint that runs one or more Linux containers. It is equivalent to Kubernetes pod. vSphere Pods are Open Container Initiative (OCI) compatible and can run containers from any operating system as long as these containers are also OCI compatible
vSphere Namespaces	Provides shared resource pools to run containers to isolate applications and tenants. The vSphere administrator can set limits for CPU, memory, storage, as well as the number of Kubernetes objects that can run within the namespace. The vSphere administrator can set limits for CPU, memory, storage, as well as the number of Kubernetes objects that can run within the namespace
Tanzu Kubernetes Cluster	Kubernetes clusters created by Tanzu Kubernetes Grid Service to run workloads
Spherelet	An additional process called Spherelet is created on each host. It is a kubelet that is ported natively to ESXi and allows the ESXi host to become part of the Kubernetes cluster
Container Runtime Executive (CRX)	CRX includes a para-virtualized Linux kernel that works together with the hypervisor. CRX uses the same hardware virtualization techniques as VMs and it has a VM boundary around it. A direct boot technique is used, which allows the Linux guest of CRX to initiate the main init process without passing through kernel initialization. This allows vSphere Pods to boot nearly as fast as containers
Tanzu CLI	A command line interface to access and run commands to manage Kubernetes cluster and containers.

5.6.2 Tanzu Kubernetes Shared Services

The Tanzu Kubernetes platform leverages SDDC components and many opensource components to manage and provide operational services for containers running on Kubernetes cluster. The plugins can be chosen based on the compatibility preferences to use across multi cloud environments. The VCF platform hides all complexities in configuring these components and provides GUI based one click deployment for Tanzu.

Table 8 shows the list of shared services components supported on Tanzu Kubernetes platform.

Table 8: Tanzu Kubernetes Core Services

Services	Description
Infrastructure platform	vSphere 6.7U3, vSphere 7.x, vSphere 8.x, VMware Cloud on AWS, Azure VMware Solution
Cluster Lifecycle Management	Core Cluster API (v0.3.14), Cluster API Provider vSphere (v0.7.6)
Kubernetes node OS distributed with TKG	Photon OS 3, Ubuntu 20.04
Bring your own image	Photon OS 3, Red Hat Enterprise Linux 7, Ubuntu 18.04, Ubuntu 20.04
Container runtime	Containerd (v1.4.3)
Container networking	Antrea (v0.11.3), Calico (v3.11.3)
Container registry	Harbor (v2.1.3)
Ingress	NSX Advanced Load Balancer Enterprise (v20.1.3), Contour (v1.12.0)
Load Balancing	NSX Advanced Load Balancer Essentials, HA Proxy
Storage	vSphere Container Storage Interface (v2.1.0) and vSphere Cloud Native Storage
Authentication	LDAP or OIDC via Pinniped (v0.4.1) and Dex
Observability and Monitoring	Fluent Bit (v1.6.9), Prometheus (v2.18.1), Grafana (v7.3.5), Tanzu Mission Control*
Backup and migration	Velero (v1.5.3)
Service Mesh	VMware Tanzu Service Mesh*
Policy and Management	Tanzu Mission Control*
Image Build	Tanzu Build Service*
Data Flow	Spring Cloud Data Flow*
Database	Tanzu Data Service*
Image Catalog	Tanzu Application Catalog*
API Gateway	Spring Cloud Gateway*

*These services are part of VMware Tanzu Advanced Edition which is not included in VCF editions.

5.6.3 Tanzu Mission Control

Tanzu Mission Control(TMC) provides centralized management and operations for multi cloud Kubernetes deployments which enables developers to work seamlessly across different environment without compromising security and governance. Tanzu Mission Control also well integrated with Tanzu observability and Service Mesh

and provides cluster lifecycle management, data protection, policy management and centralized authentication and authorization capabilities which enables operators and infrastructure teams to manage efficiently. Tanzu Mission Control is offered through VMware Cloud Services which provides also Tanzu Application Services and Data services as subscriptions.

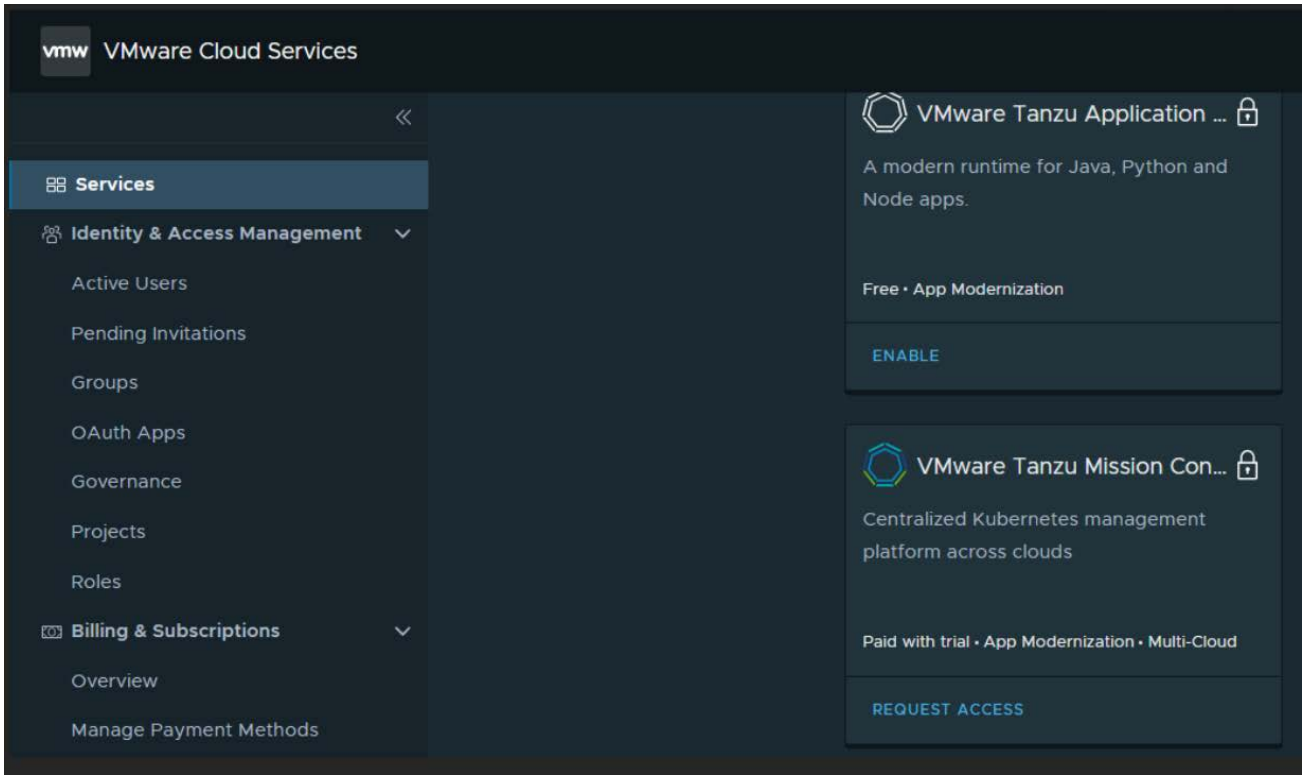

Figure 7: VMware Cloud Services and Tanzu Mission Control

The Tanzu Mission Control Console provides methods to integrate the vSphere with Tanzu Supervisor Cluster which allows us to create, attach and manage Kubernetes clusters and organize them into logical groups for easier management of their resources, such as apps, users and security.

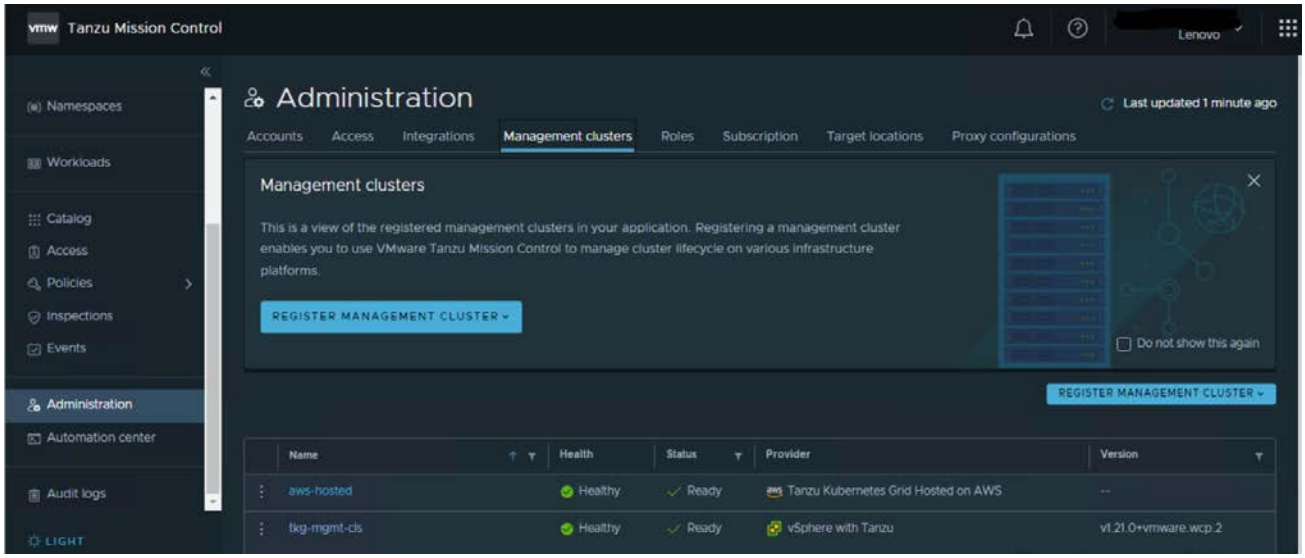

Figure 8: Tanzu Mission Control

Tanzu Mission Control offers several Kubernetes cluster management capabilities, such as:

- Cluster Lifecycle Management – Kubernetes clusters from different providers can be attached, created, resized or deleted
- Cluster Diagnostics – the health and resource usage of clusters can be viewed from a centralized point (TMC Console)
- Cluster Inspection – preconfigured inspections can be run against the clusters using the opensource Sonobuoy to ensure the consistency across the clusters
- Data Protection – clusters data resources can be backed-up and restored using Velero
- Access Control – TMC allows the use of federated identity management and granular role-based access control
- Policy Management – Kubernetes cluster and other organizational objects can be managed using different policy types, such as: access policy, image registry policy, network policy, quota policy etc.
- Events and Audit logs – TMC provides different levels of audit logging, enabling log entries of that describe what was done, who performed the action and when and where the action was performed

Through the TMC Console the Kubernetes resources can be organized and viewed in two different ways, allowing the operations administrators to maintain control over the clusters and providing self-serve access to namespaces for different teams:

- Cluster Groups – provides an Infrastructure View by organizing the clusters into groups that better fits the business needs
- Workspaces – provides an Application View by organizing the managed namespaces into logical groups that better aligns with the development projects

TMC integrates Tanzu Observability Services which provides full-stack Kubernetes monitoring and smart out-of-the-box alerting via Tanzu Observability by Wavefront, it also offers support for monitoring and logging with open-source Prometheus and Grafana services.

Tanzu Observability by Wavefront is a high-performance streaming analytics platform that supports observability for metrics, counters, histograms and traces/spans, that can scale to very high data ingestion rates and query loads. The information collected is visualized in dashboards and charts from which alerts can be directly created. Using the Wavefront query language (WQL) the desired information can be extracted by combining a series of customizable filters and functions.

5.6.4 Tanzu Service Mesh

Tanzu Mission Control integrates Tanzu Service Mesh(TSM) which provides enterprise grade service mesh which solves the challenges associated with a distributed microservices application across multiple clusters and clouds by expanding the services outside of a Kubernetes cluster. Tanzu Service Mesh supports many Kubernetes platforms, providing reliable control and security for resources across multiple clusters. Tanzu Service Mesh supports Amazon EKS, Amazon EKS Anywhere (EKS-A), Azure Kubernetes Service(AKS), Anthos google kubernetes engine (GKE) and Redhat OpenShift.

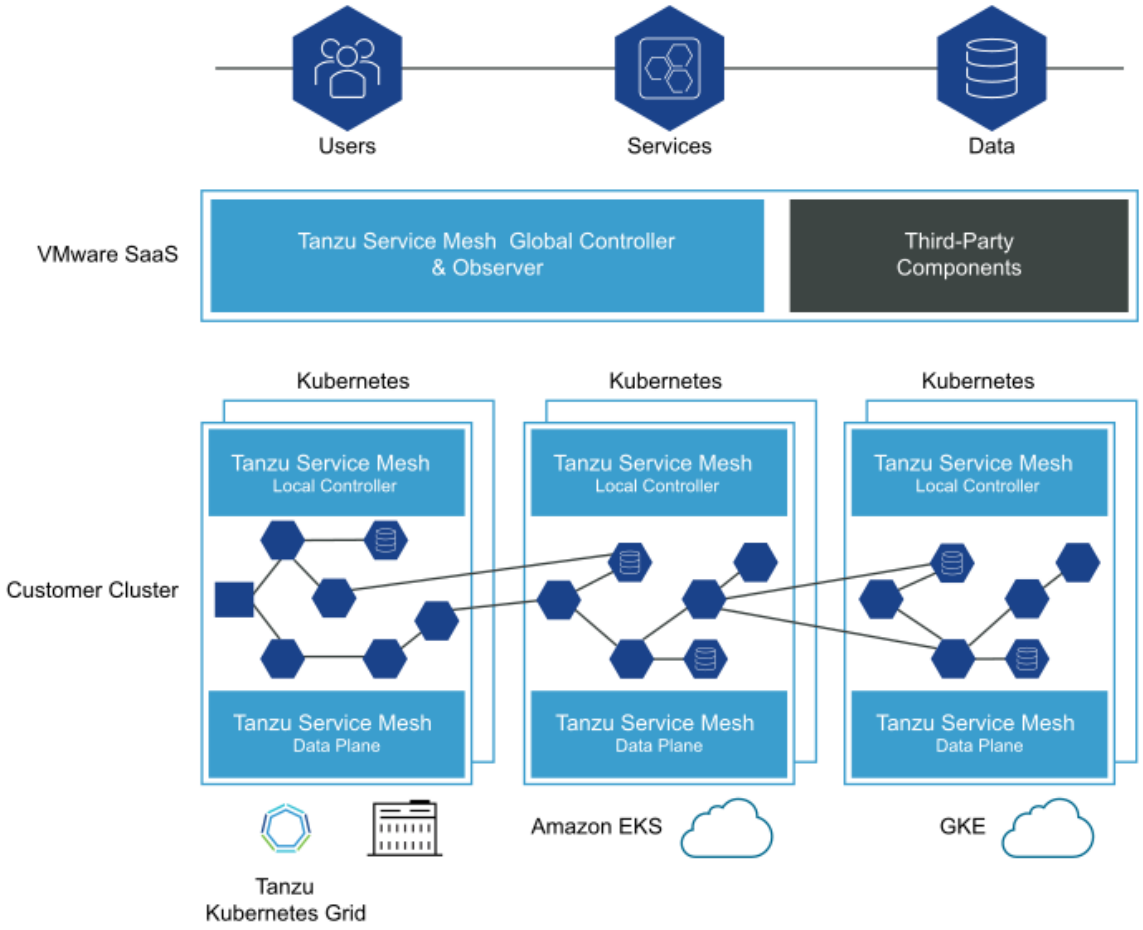

Figure 9: Tanzu Service Mesh Architecture

TSM addresses problems related to:

- users-to-service-to-data communication – processing full end-to-end request, from end users, through the services and on to the data
- applications running on highly distributed heterogenous platforms – managing traffic, security and observability of apps deployed in a multi cloud, multiplatform environment
- service autoscaling – scaling up or down of microservices that have different levels of demand based on metrics (CPU, memory, storage usage)
- service level objectives – formalizing the way to describe, measure and monitor the performance, quality and reliability of microservices apps across the organization by using real-time graphs and out-of-the box metrics, without needing of additional plugins or code changes

From a high architectural point of view, Tanzu Service Mesh has the following components:

- Global Controller and Observer – a collection of microservices that resides on VMware SaaS which deliver control, security, observability and autoscaling capabilities
- Local Controller – local control plane components that run on each cluster both on-premises or in a public cloud which provides fault tolerance in case a cluster cannot communicate with the Global Controller
- Sidecars – data-plane components running on each cluster on-premises or in a public cloud, that handles east-west traffic inside the service mesh

- Ingress/egress Gateways – data-plane components running on each cluster on both on-premises or public cloud providing north-south traffic from and to the service mesh

A unique concept in Tanzu Service Mesh that defines an application boundary is Global Namespace which provides consistent traffic routing, connectivity, resiliency and security for applications across multiple clusters and clouds by connecting the resources and workloads into a single virtual unit regardless of where the resources are located. A Global Namespace manages the following functions:

- Identity – A Global Namespace has its own certificate authority that provisions identities for each resource
- Discovery (DNS) – provides a registry that controls how the resources can locate each other
- Connectivity – defines how the communications between resources are established and how the traffic is routed within itself or outside the Global Namespace
- Security – Enforces the encryption of the traffic between the resources by using mTLS (mutual Transport Layer Security) authentication
- Observability – the telemetry data (metrics for services, clusters, nodes inside a GN) is aggregated and can be distributed to external endpoint using plugins

Resource Groups are used to enforce policies and monitor the performance of resources in a single GN or across the entire organization. A service level objective can also be added to a resource group and each SLI (service limit indicator) violation can be tracked in the TSM console.

A Service is an abstract concept that encloses business logic running in a distributed application and can be mapped to multiple service versions and each service version can be mapped to multiple service instances. This allows for easy management of life cycle changes within a service.

5.7 VMware Private AI

VMware private AI is a combined solution from partners and open source ecosystem across AI infrastructure, machine learning models and operations, frameworks and deployment. VMware Private AI with Lenovo ThinkAgile VX is an ideal solution deploying enterprise-wide AI/ML use cases and generative AI workloads in on-premises infrastructure. On-premises deployment in the private cloud provides compete control over data and compliance requirements.

VMware Private AI comprises of Lenovo ThinkAgile VX systems, VMware Cloud Foundation, Tanzu, NVIDIA Enterprise AI and MLOps frameworks and the solution can be extended by adding more software and tools from AI ecosystem.

VMware and NVIDIA software are well integrated and provide AI Ready enterprise platform to deliver end-end AI/ML workloads. With ThinkAgile VX and latest generation processors Intel and AMD, VMware Private AI solution accelerates data analytics and AI/ML development.

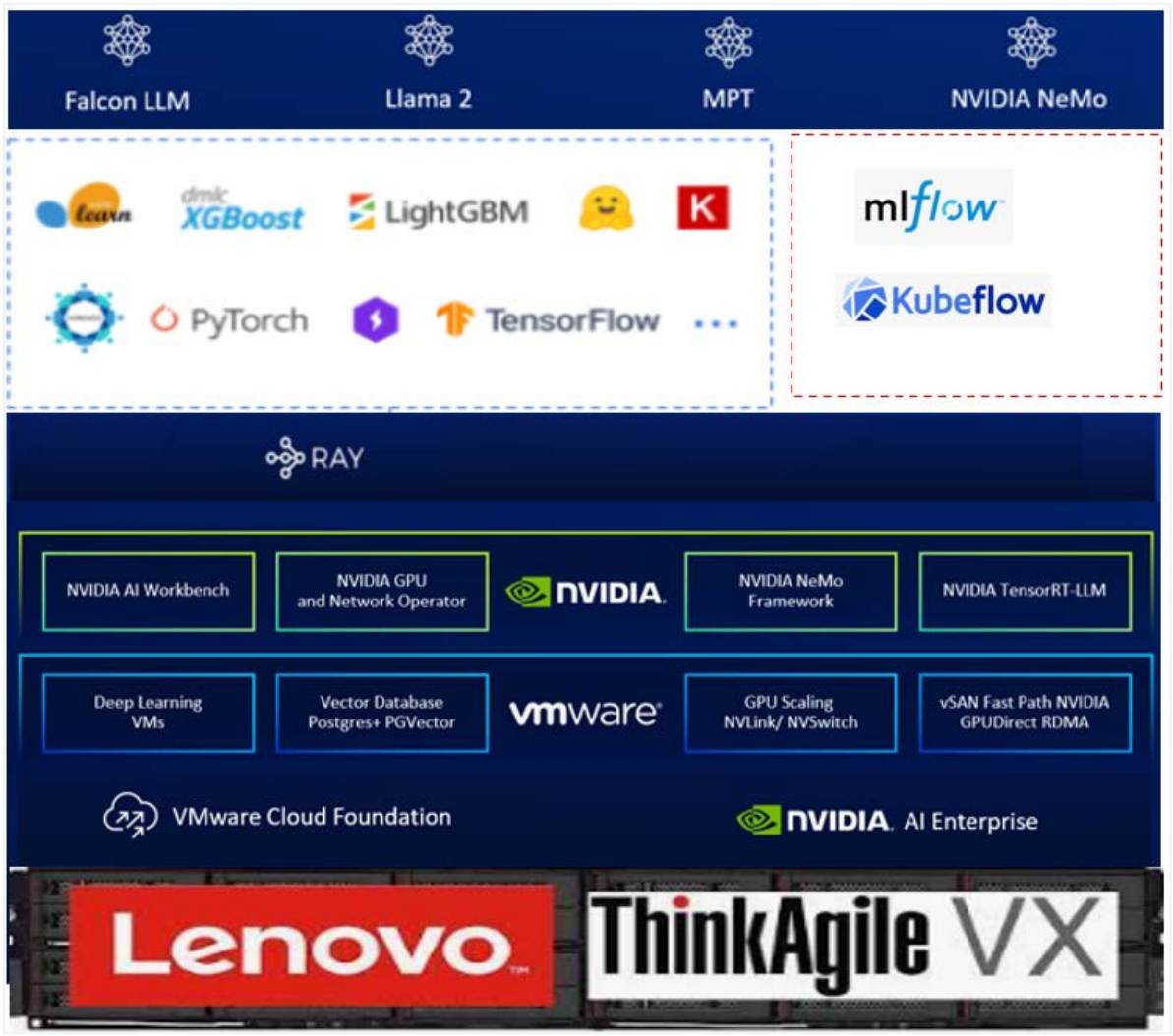

Figure 10: VMware Private AI Foundation with ThinkAgile VX

5.7.1 NVIDIA AI Enterprise

NVIDIA AI Enterprise is an end-to-end, cloud native software platform that accelerates the data science pipeline and streamlines development and deployment of production-grade AI applications, including generative AI, computer vision, speech AI, and more. Enterprises that run their businesses on AI rely on the security, support, and stability provided by NVIDIA AI Enterprise to improve productivity of AI teams, reduce total cost of AI infrastructure, and ensure a smooth transition from pilot to production.

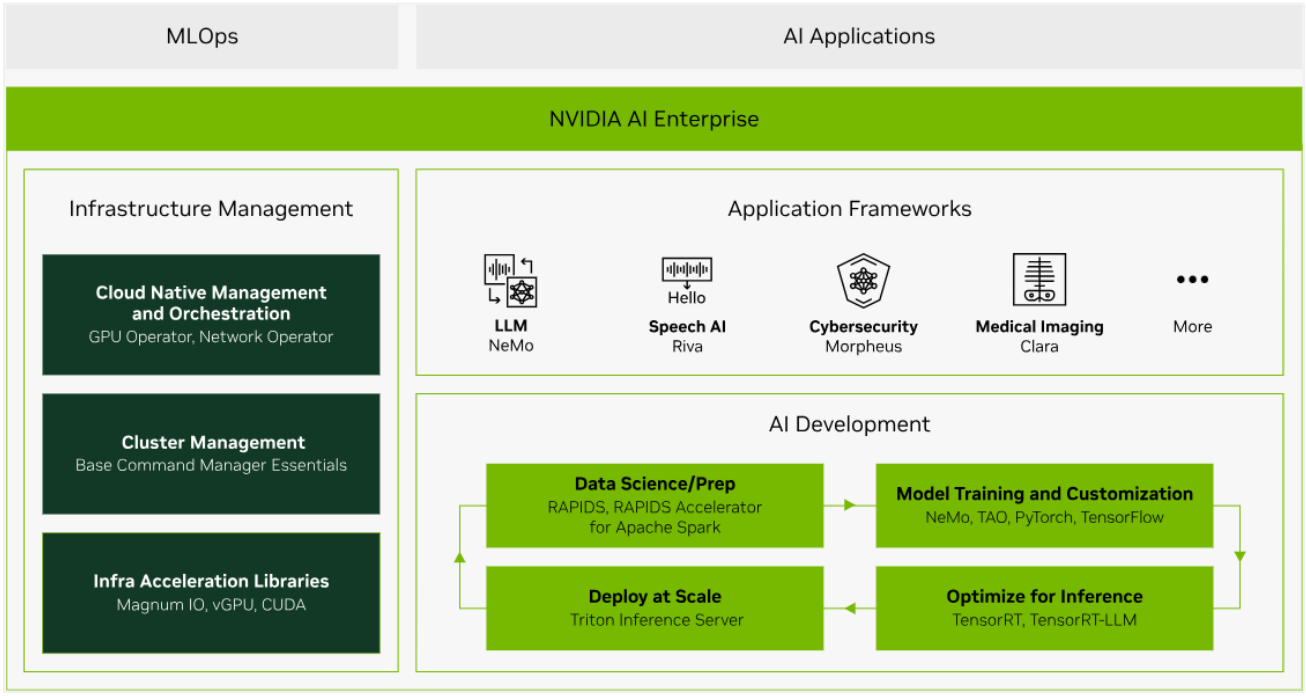

Figure 11: NVIDIA AI Enterprise Components

NVIDIA AI Enterprise includes the following components.

NVIDIA NeMo an end-to-end framework for building, customizing, and deploying enterprise-grade generative AI models; NeMo lets organizations easily customize pretrained foundation models from NVIDIA and select community models for domain-specific use cases.

NVIDIA RAPIDS is an open-source suite of GPU-accelerated data science and AI libraries with APIs that match the most popular open-source data tools. It accelerates performance by orders of magnitude at scale across data pipelines.

NVIDIA TAO Toolkit simplifies model creation, training, and optimization with TensorFlow and PyTorch and it enables creating custom, production-ready AI models by fine-tuning NVIDIA pretrained models and large training datasets.

NVIDIA TensorRT, an SDK for high-performance deep learning inference, includes a deep learning inference optimizer and runtime that delivers low latency and high throughput for inference applications. TensorRT is built on the NVIDIA CUDA parallel programming model, enables you to optimize inference using techniques such as quantization, layer and tensor fusion, kernel tuning, and others on NVIDIA GPUs.

NVIDIA TensorRT-LLM is an open-source library that accelerates and optimizes inference performance of the latest large language models (LLMs). TensorRT-LLM wraps TensorRT's deep learning compiler and includes optimized kernels from FasterTransformer, pre- and post-processing, and multi-GPU and multi-node communication.

NVIDIA Triton Inference Server optimizes the deployment of AI models at scale and in production for both neural networks and tree-based models on GPUs.

5.8 VMware Licensing

The licensing for vSphere is based on a CPU metric and licensing for other products is based on the number of OS instances. Other components, such as NSX, have their own separate licenses and are optional add-ons. Table 9 lists the standard and optional components that are provided with VCF editions. However, add on licenses can be added as long as they meet compatibility.

Table 9: VMware VCF Editions

Component	Starter	Standard with Tanzu	Advanced with Tanzu	Enterprise with Tanzu
SDDC Manager				
vSphere	Enterprise Plus	Enterprise Plus	Enterprise Plus	Enterprise Plus
vSAN	Advanced	Advanced	Advanced	Enterprise
NSX-T	Advanced	Advanced	Advanced	Enterprise Plus
vRealize Network Insight	Advanced		Advanced	Enterprise
vRealize Suite	Standard		Enterprise	Enterprise
Tanzu		Standard	Standard	Standard
Tanzu standard can be upgraded to Tanzu Advanced separately				

5.9 HyTrust Security

HyTrust provides a suite of security-oriented products for vSphere environment. These products are HyTrust KeyControl, DataControl, and CloudControl. Note that the HyTrust products are currently supported on ESXi 7.0U2 and NSX-T Data Center 3.1..2.

5.9.1 HyTrust KeyControl

HyTrust KeyControl (HTKC) enables enterprises to easily manage all their encryption keys at scale, how often they rotate them, and how they are shared securely. HyTrust KeyControl capabilities include:

- VMWare Certified Key Manager Server (KMS) for:
  - vSphere 7.0u2 and vSAN 7.0
- Universal key management for KMIP-compatible encryption agents
- Enterprise scalability and performance
- KeyControl can run in an active-active, high availability cluster
- FIPS 140-2 Level 1 validation and FIPS 140-2 Level 3 hardware security module (HSM)

5.9.2 HyTrust DataControl

HyTrust DataControl (HTDC) secures multi-cloud workloads throughout their lifecycle. DataControl helps manage workloads and encryption keys from a central location to reduce complexity, comply with regulations such as the GDPR.

DataControl provides granular encryption for better multi-cloud security. The protection boundary does not stop at the hypervisor or at the data store; VMs are individually encrypted. Inside the VM, unique keys can be assigned to encrypt individual partitions, including the boot (OS) disk. Encryption and rekeying can be done on the fly and there is no need to take workloads off-line.

Table 10 compares the data encryption features of vSphere, vSAN, and HyTrust DataControl/KeyControl.

Table 10: Comparison of Encryption Features

Encryption	vSphere VM Encryption	vSAN Encryption	HyTrust DataControl
Protection level	Data at rest and in motion	Data at rest	Data at rest
Encryption Approach	Hypervisor does the encryption	Disk based encryption	In Guest encryption
Components	KMS, vCenter, ESXi Host	KMS, vCenter, ESXi Hosts in vSAN Cluster, Disks	KMS, HyTrust DataControl Agent
Encryption Cipher	AES-XTS-256	AES-XTS-256	AES-XTS-512,AES-XTS-256, AES 128
Encrypted objects	Virtual machine files, virtual disk files, and ESXi core dump files	All files in the vSAN datastore	All data in the drives
Interface	vSphere Web Client, vSphere Web Services SDK	vSphere Web Client	HyTrust DataControl UI in the Guest OS. HyTrust KeyControl UI to manage VM Set, VMs and users
Enabling Option	Per VM level through vSphere Encryption Storage Policy	Enabled at cluster or VSAN datastore level	Enabled within Guest OS
Access Control	Users with vSphere Cryptographic Operations Privileges	Users with vSphere Cryptographic Operations Privileges	Guest OS User uses KeyControl admin user. Authorization can also be done by HyTrust CloudControl
Interoperability Limitations	vSphere Fault Tolerance, vSphere Replication, Content Library	N/A	N/A
Platform Support	All Guest OS running on the Hypervisor	All Guest OS running on the Hypervisor	Most Windows and Linux flavors and version running on vSphere, KVM, Hyper-V, or XenSever

5.9.3 HyTrust CloudControl

HyTrust CloudControl (HTCC) provides a variety of security and policy enhancements without impacting the existing GUI of vSphere, NSX and ESXi. CloudControl is deployed as a transparent proxy and mediates the actions taken by administrators using familiar interfaces. CloudControl provides the following security features:

- **Role Based Access Control** (RBAC) to control which functions have access to what resources and allows a much closer alignment of access rights to governance and compliance requirements.
- **Policy Control including Two Man Rule** to define and more importantly enforce policy including requiring secondary approval for potentially disruptive actions, reducing potential impact of human error or intentional malevolent behaviour.
- **Access Control including Two Factor Authentication** to significantly enhance the overall security posture of an organization without the traditional weaknesses of using even strong passwords.
- **Forensic grade logs** to provide an in-depth perspective on what has happened as well as what has not happened in your virtual environment.

Table 11 compares the access control features of vCenter and HyTrust CloudControl.

Table 11: Comparison of access control features

Access Control Feature	vCenter	HyTrust CloudControl
vSphere Web Client Access	vCenter URL	Published IP (PIP) associated with vCenter
Authentication	vCenter SSO, IWA	vCenter SSO, IWA, HTCC Service Account, Two factor authentication with RSA Secure ID, RADIUS, or TACACS+
Authorization	Predefined permissions to access various vCenter components	Uses permissions defined in vCenter
vCenter Users	SSO users from multiple AD Domain and vSphere local domain. Predefined solution users for vSphere services.	Users from Single AD Domain which includes configured HTCC Service Account
vCenter User Access Setup	Directory users/group need to be added in vCenter SSO users/group	Directory users need to be added to respective HTCC directory group which is associated with HTCC role
User Groups	14 predefined SSO groups. Directory users/group is mapped to SSO groups.	16 predefined rules for vSphere. HTCC directory group is mapped to HTCC rule.
Role Based Access Control	14 predefined roles with respective privileges	16 predefined roles for vSphere with appropriate privileges
Custom Roles Creation	Supported	Supported
Secondary Approval	Not Available	Available for set of compute and network operations

Access Control Feature	vCenter	HyTrust CloudControl
Auditing	Integrated with vRealize Log Insight. Auditing dashboard is available based on the event type. User's session details can be monitored in vSphere web client.	Has its own Log Viewer and dashboard. Logs can be redirected to use vRealize Log Insight as syslog server.

5.9.4 Compliance Management

An important part of security is compliance management. VMware vRealize Configuration Manager has twenty built-in compliance templates and others can be added. HyTrust CloudControl (HTCC) supports customizing built-in compliance templates but does not provide any out of the box.

Table 12 compares the compliance management features of vRealize Configuration Manager and HTCC.

Table 12: Comparison of compliance management features

Compliance Management Feature	vRealize Configuration Manager	HTCC
ESXi Host Compliance	Yes	Yes
Guest Virtual Machine Compliance	Yes	Limited
NSX Manager Compliance	No	Yes
Patching assessment and Deployment	Yes	No
Active Directory Compliance	Yes	No
Software Asset Management	Yes	No
Integration with vRealize Operation Manager	Yes	No
Manage Virtual Machines	Yes	No

6 Operational model

This chapter describes the options for mapping the logical components of SDDC onto Lenovo ThinkAgile VX servers. The following section describes the hardware components in a SDDC deployment.

6.1.1 ThinkAgile VX Servers

You can use various rack-based Lenovo ThinkAgile VX3 server platforms to implement edge, management, or compute clusters with VMware vSAN and supports All Flash and Hybrid configurations.

VX System	vSAN Support	Base System	CPU	Max drives	Max Possible Capacity
VX630 V3 IS	Hybrid All Flash	ThinkSystem SR630 V3	Intel Xeon 4th Gen	12	2 disk group 10 drives
VX630 V3 CN	Hybrid All Flash	ThinkSystem SR630 V3	Intel Xeon 4th Gen	12	2 disk group 10 drives
VX650 V3 IS	Hybrid All Flash	ThinkSystem SR650 V3	Intel Xeon 4th Gen	32	4 disk group 28 drives
VX650 V3 CN	Hybrid All Flash	ThinkSystem SR650 V3	Intel Xeon 4th Gen	32	4 disk group 28 drives
VX650V3 DPU IS	Hybrid All Flash	ThinkSystem SR650 V3	Intel Xeon 4th Gen	32	4 disk group 28 drives
VX650V3 DPU CN	Hybrid All Flash	ThinkSystem SR650 V3	Intel Xeon 4th Gen	32	4 disk group 28 drives
VX655 V3 Integrated System	Hybrid All Flash	ThinkSystem SR655 V3	AMD EPYC 4th Gen	40	5 disk group 35 drives
VX655 V3 Certified Node	Hybrid All Flash	ThinkSystem SR655 V3	AMD EPYC 4th Gen	40	5 disk group 35 drives
VX665 V3 Integrated System	Hybrid All Flash	ThinkSystem SR665 V3	AMD EPYC 4th Gen	40	4 disk group 28 drives
VX665 V3 Certified Node	Hybrid All Flash	ThinkSystem SR665 V3	AMD EPYC 4th Gen	40	4 disk group 28 drives

6.2 Edge cluster servers

The edge cluster runs NSX services for all tenants in the SDDC infrastructure, provides internal and external routing, and also runs tenant workloads.

The shared edge and compute cluster uses its own dedicated vCenter server and NSX-T manager which are deployed in the management cluster. The NSX controllers and edge gateway services VMs are deployed on the shared cluster. The tenant VMs can be deployed in the shared edge and compute cluster or in a separate compute cluster leveraging the vCenter server and NSX services in the shared edge and compute cluster.

6.2.1 Edge and Infrastructure Services VMs

The VMs used for infrastructure services such as Active Directory, DNS/DHCP, firewalls, proxy and anti-virus are deployed in the shared edge and compute cluster. Table 13 lists each infrastructure service VM with the recommended sizes in terms of virtual CPUs, RAM, storage, and networking.

Table 13: Infrastructure services VMs

VM description	CPU (vCPUs)	Memory (GB)	Storage (GB)	Network bandwidth	High availability
AD, DHCP, DNS server	2	4	70	1 GbE	clustered
http proxy server	2	4	30	1 GbE	clustered
NSX Controller (odd # deployment; min 3)	4	4	20	1 GbE	Built-in/vSphere HA

Table 14 lists the NSX service VMs with the recommended sizes in terms of virtual CPUs, RAM, storage, and networking.

Table 14: Edge services VMs for NSX

VM description	CPU (vCPUs)	Memory (GB)	Storage (GB)	Network bandwidth	High availability
Compact (also used for logical router)	1	0.5	0.5	1 GbE	Yes, Optional
Large	2	1	0.5	1 GbE	Yes, Optional
Quad Large	4	1	0.5	1 GbE	Yes, Optional
X-Large	6	8	4.5	1 GbE	Yes, Optional

The actual VM size (compact, large, quad-large, and X-large) depends on the number of type of services that are deployed in the VM. A logical router is always deployed by using a compact VM. A quad large is required for a firewall and an X-large is used for more than one service (for example, firewall, load balancer, and router).

6.2.2 Hybrid cloud VMs

Table 15 lists the cloud connectivity VMs with the recommended sizes in terms of virtual CPUs, RAM, storage, networking, and location. Note that these VMs do not have options for high availability.

Table 15: Cloud connectivity VMs

VM description	CPU (vCPUs)	Memory (GB)	Storage (GB)	Network bandwidth	Location
VMware HCX	2	4	300	1 GbE	On-Premise

6.3 Management cluster servers

The number of VMware SDDC components in the management cluster increases as capabilities are added. This section addresses the SDDC management components that could be used. Third party add-ons must be sized separately.

6.3.1 Management cluster VMs

There are several considerations that contribute to an end-to-end sizing of an entire VMware vCloud environment including Lenovo software for systems management. This section is intended to provide some high-level guidance for management cluster configuration sizing. The recommended number of virtual CPUs, memory size, storage size, and network bandwidth are given for each VM and the VMs are grouped by each major component or appliance.

An essential part of the infrastructure is load balancing of the server VMs and recognizing when a server is down and failing over to another server. The following cases are available for VMs in the management cluster:

- vSphere HA: vCenter automatically restarts the VM on another server, but there is some downtime while the VM starts up.
- Microsoft SQL server clustering: The SQL server cluster automatically handles failover.
- Clustering within component to provide built-in high availability.

Load balancing: An external load balancer such as a Big-IP switch from F5 and/or VMware NSX load balancers can be used.

Table 16 lists each management cluster VM for vSphere with its recommended size in terms of virtual CPUs, RAM, storage, and networking.

Table 16: Management cluster VMs for vSphere

VM description	CPU (vCPUs)	Memory (GB)	Storage (GB)	Network bandwidth	High availability
SDDC Manager	4	16	1000	1 GbE	vSphere HA
vCenter Server Appliance(1) Management Cluster	8	24	50	1 GbE	load balancer
vCenter Server Appliance(2) Edge and Compute Cluster	8	24	50	1 GbE	load balancer
vCenter Server Database (MS SQL)	4	8	200	1 GbE	SQL AlwaysOn Availability Group

VM description	CPU (vCPUs)	Memory (GB)	Storage (GB)	Network bandwidth	High availability
SDDC Manager	4	16	1000	1 GbE	vSphere HA
vSphere Replication	2	4	20	1 GbE	not required
vSphere Data Protection	4	4	1600	1 GbE	not required
vRealize Orchestrator Appliance	2	3	12	1 GbE	Clustered

Table 17 lists each management cluster VM for vRealize Automation with its size in terms of virtual CPUs, RAM, storage, and networking.

Table 17: Management cluster VMs for vRealize Automation

VM description	CPU (vCPUs)	Memory (GB)	Storage (GB)	Network bandwidth	High availability
vRealize Suite Lifecycle Manager	4	16	135	1 GbE	N/A
vRealize Automation Appliance	4	16	30	1 GbE	load balancer
IaaS Database (MS SQL)	8	16	100	1 GbE	SQL AlwaysOn Availability Group
Infrastructure Web Server	2	4	40	1 GbE	load balancer
Infrastructure Manager Server	2	4	40	1 GbE	load balancer
Distributed Execution Manager (DEM)	2	6	40	1 GbE	load balancer
vSphere Proxy Agent	2	4	40	1 GbE	load balancer
vRealize Application Services	8	16	50	1 GbE	vSphere HA

Table 18 lists each management cluster VM for vRealize Operations Manager with its size in terms of virtual CPUs, RAM, storage, and networking.

Table 18: Management cluster VMs for vRealize Operations Manager

VM description	CPU (vCPUs)	Memory (GB)	Storage (GB)	Network bandwidth	High availability
vRealize Operations Manager – Master	4	16	500	1 GbE	clustered
vRealize Operations Manager – Data	4	16	500	1 GbE	not required
vRealize Configuration Manager – Collector	4	16	150	1 GbE	load balancer

VM description	CPU (vCPUs)	Memory (GB)	Storage (GB)	Network bandwidth	High availability
vRealize Configuration Manager Database (MS SQL)	4	16	1000	1 GbE	SQL AlwaysOn Availability Group
vRealize Hyperic Server	8	12	16	1 GbE	load balancer
vRealize Hyperic Server - Postgres DB	8	12	75	1 GbE	load balancer
vRealize Infrastructure Navigator	2	4	24	1 GbE	not required

Table 19 lists the management VMs that are needed for NSX.

Table 19: NSX-T Management cluster VMs

VM description	CPU (vCPUs)	Memory (GB)	Storage (GB)	Network bandwidth	High availability
NSX-T Manager Management Cluster	4	12	300	1 GbE	vSphere HA
NSX-T Controller Management Cluster (odd # deployment; min 3)	4	4	20	1 GbE	Built-in/vSphere HA
NSX-T Manager Edge and Compute Cluster	4	12	60	1 GbE	vSphere HA

Table 20 lists each management cluster VM for HyTrust with its size in terms of virtual CPUs, RAM, storage, and networking.

Table 20: Management cluster VMs for HyTrust

VM description	CPU (vCPUs)	Memory (GB)	Storage (GB)	Network bandwidth	High availability
HyTrust CloudControl	4	16	70	1 GbE	Clustered
HyTrust KeyControl	2	8	20	1 GbE	Clustered

Table 21 lists the VMs that are needed for Lenovo software for systems management.

Table 21: Lenovo System Management VMs

VM description	CPU (vCPUs)	Memory (GB)	Storage (GB)	Network bandwidth	High availability

Lenovo XClarity Administrator	2	4	64	1 GbE	not required
Lenovo XClarity Orchestrator	4	16	500	1 GbE	not required
Lenovo XClarity Integrator (Windows OS)	1	2	30	1 GbE	not required

6.4 Systems management for Lenovo servers

Lenovo XClarity™ family of system management software and tools provide centralized resource management, monitoring and analytics solution that reduces complexity, speeds up response, and enhances the availability of Lenovo® server systems and solutions. The XClarity integrator plugins are designed to work with VMware VCF components as an extension to simplify the operations. For more information, see this website: https://www.lenovo.com/us/en/data-center/software/management/.

6.4.1 Lenovo XClarity Administrator(LXCA)

The Lenovo XClarity Administrator provides agent-free hardware management for Lenovo's ThinkAgile, ThinkSystem® rack servers, System x® rack servers, and Flex System™ compute nodes and components, including the Chassis Management Module (CMM) and Flex System I/O modules. Figure 12 shows the Lenovo XClarity administrator interface, in which Flex System components and rack servers are managed and are seen on the dashboard. Lenovo XClarity Administrator is a virtual appliance that is quickly imported into a virtualized environment server configuration. Lenovo XClarity Administrator supports auto discovery of endpoints, inventory, monitoring, firmware compliance, firmware updates, Windows device driver updates, configuration management and compliance, user management, deployment of operating systems and hypervisors to bare metal servers.

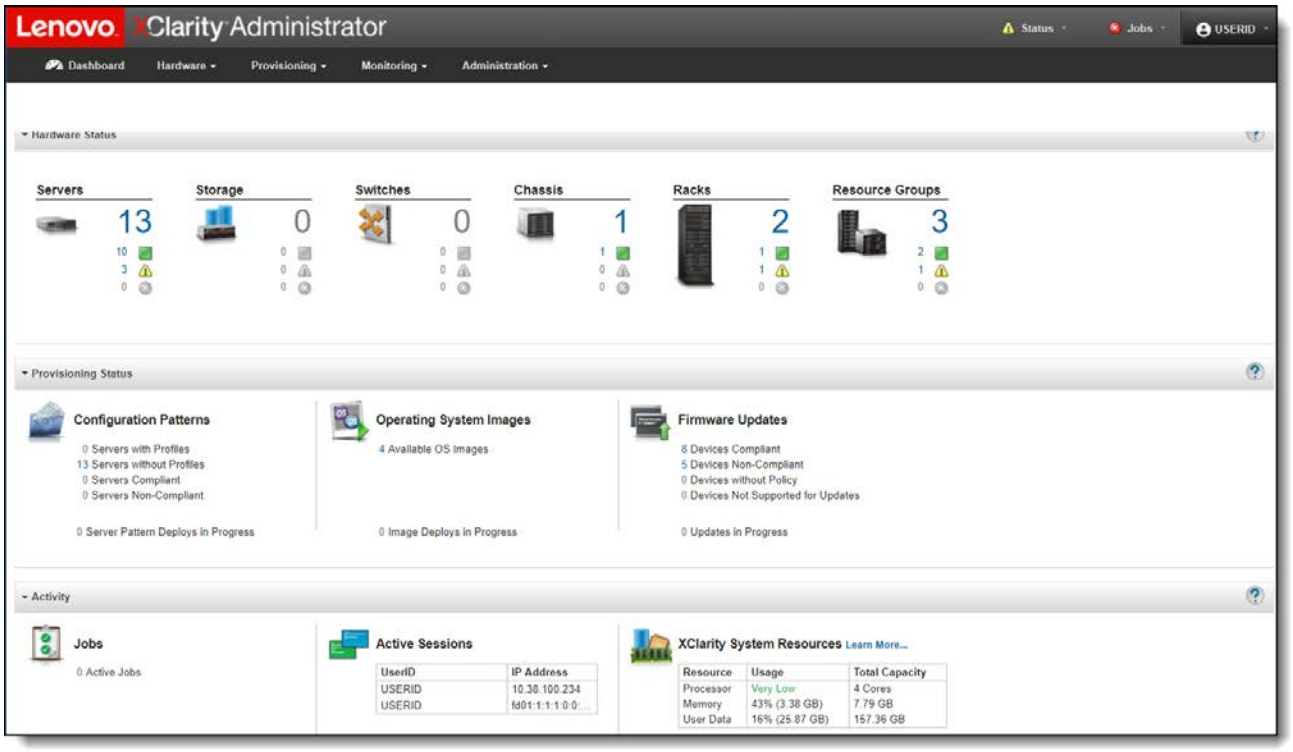

Figure 12: XClarity Administrator dashboard

6.4.2 Lenovo XClarity Orchestrator(LXCO)

XClarity Orchestrator provides a single interface to monitor and manage multiple Lenovo XClarity Administrators and the devices managed by them. LXCO supports deploying updates to Lenovo Lenovo XClarity Administrator and firmware updates to devices that are managed. LXCO can connect to third-party services (such as Splunk) for business intelligence machine learning and predictive analytics to collect resource utilization data and uses metric data to predict failures, create reports and custom alert rules that, when enabled, raise alerts when specific conditions exist in your environment.

6.4.3 Lenovo XClarity Integrators (LXCI) for VMware

Lenovo provides XClarity integration modules for VMware vCenter, VMware vRealize Automation, VMware vRealize Orchestrator and VMware vRealize Log Insight.

By using the Lenovo XClarity Integrator for VMware vCenter, administrators can consolidate physical resource management in VMware vCenter, which reduces the time that is required for routine system administration.By using the Lenovo XClarity Integrator for VMware vCenter, administrators can consolidate physical resource management in VMware vCenter, which reduces the time that is required for routine system administration.

The Lenovo XClarity Integrator for VMware vCenter provides the following features and benefits:

- Extends Lenovo XClarity Administrator features to the virtualization management console
- Enables management of legacy infrastructure from the virtualization management console
- Reduces workload downtime by dynamically triggering workload migration in clustered environments during rolling server reboots or firmware updates, and predicted hardware failures

Figure 13 shows Lenovo XClarity Integrator deployed in the vCenter and displays ThinkAgile VX nodes.

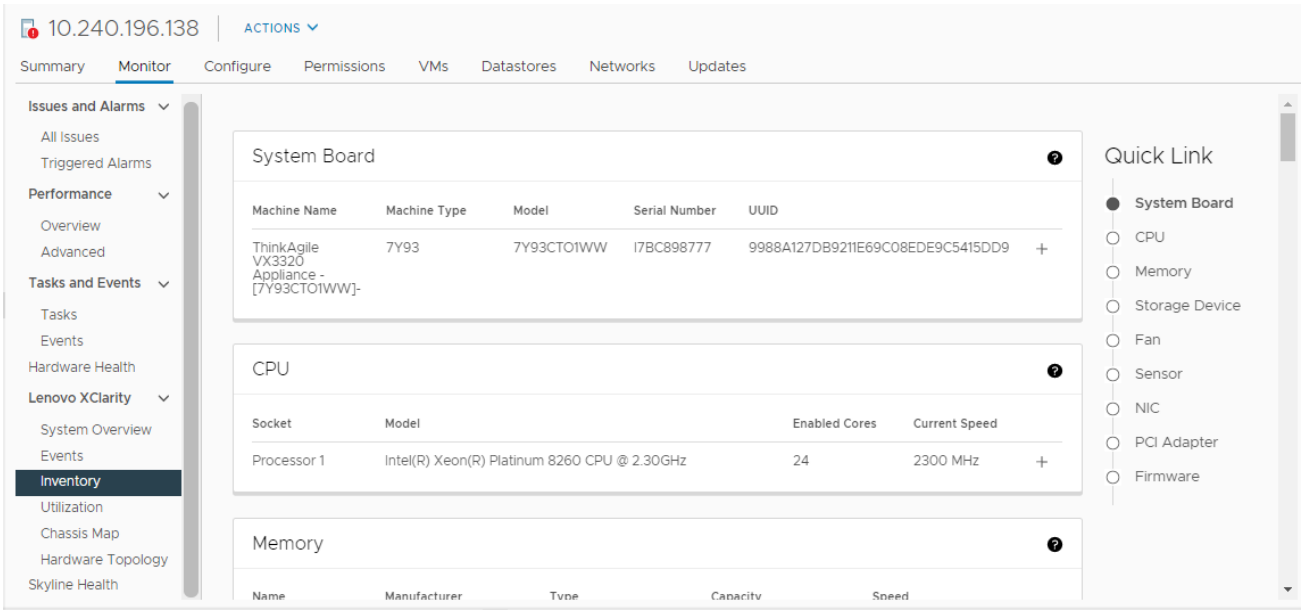

Figure 13: Lenovo XClarity Integrator for VMware vCenter

The Lenovo XClarity Integrator for VMware vRealize Orchestrator provides IT administrators with the ability to coordinate physical server provisioning features of Lenovo XClarity Pro with broader vRealize Orchestrator workflows. Lenovo XClarity Integrator for VMware vRealize Orchestrator provides a library of simple yet robust and customizable workflow routines and actions designed to automate complex, repetitive IT infrastructure tasks such as system discovery and configuration, hypervisor installation, and addition of new hosts to vCenter. Figure 14 shows the Lenovo XClarity Integrator for vRealize Orchestrator workflow interface.

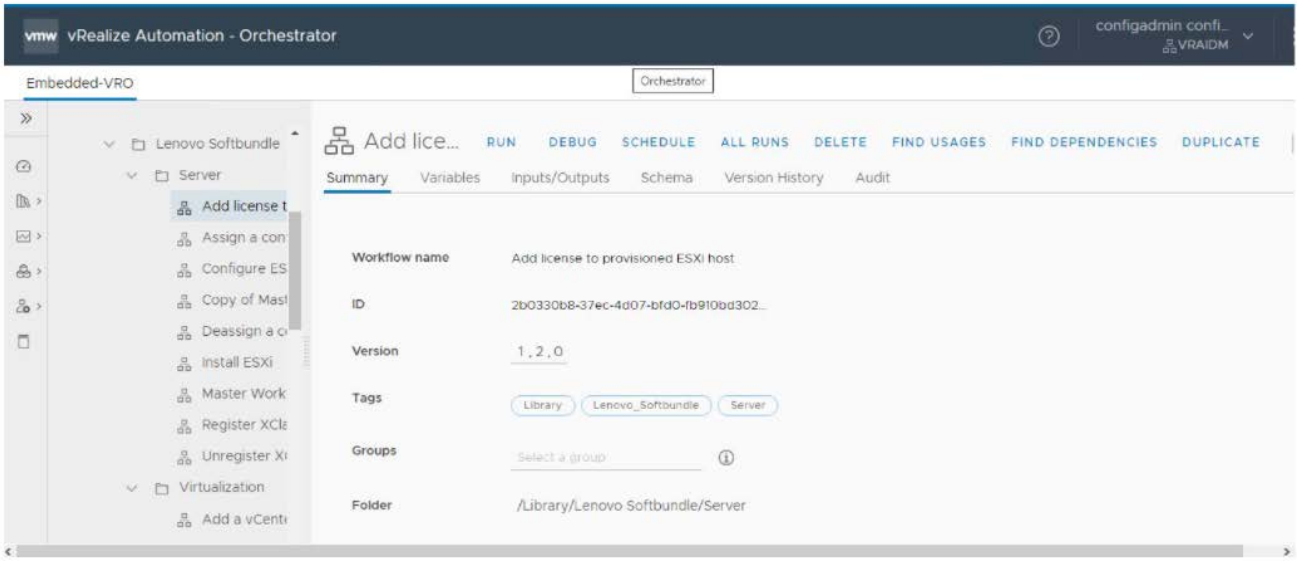

Figure 14: Lenovo XClarity Integrator for VMware vRealize Orchestrator interface

The Lenovo XClarity Administrator Content Pack for VMware vRealize Log Insight simplifies the collection and forwarding of Lenovo XClarity Administrator logs to VMware vRealize Log Insight for powerful processing and analytics, and displaying insightful information in an intuitive format.

The VMs for VMware vCenter, vRealize Orchestrator, Lenovo XClarity Administrator and Lenovo XClarity Administrator Integrator should have access to the management network used for managing servers, storage and networking.

Lenovo XClarity Integrator for vRealize Automation provides a set of blueprints to provision infrastructure services based on Lenovo servers, network switches and vSphere. This eases provisioning a new Lenovo server with vSphere installed, network isolation parameters configured on the Lenovo switches, apply vSphere distributed switch configurations and adding the server to the existing or new vSphere Cluster. The Lenovo vRealize content pack for vRealize Automation needs to imported into vRealize Orchestrator and then the Blueprints package is imported using the vRealize Cloud Client command line utility by Tenant Administrators and it creates catalog items automatically. The catalog items are created under Lenovo Servers, Lenovo Network, and Lenovo Virtualization services. Figure 15 shows Lenovo XClarity Integrator template items for vRealize Automation.

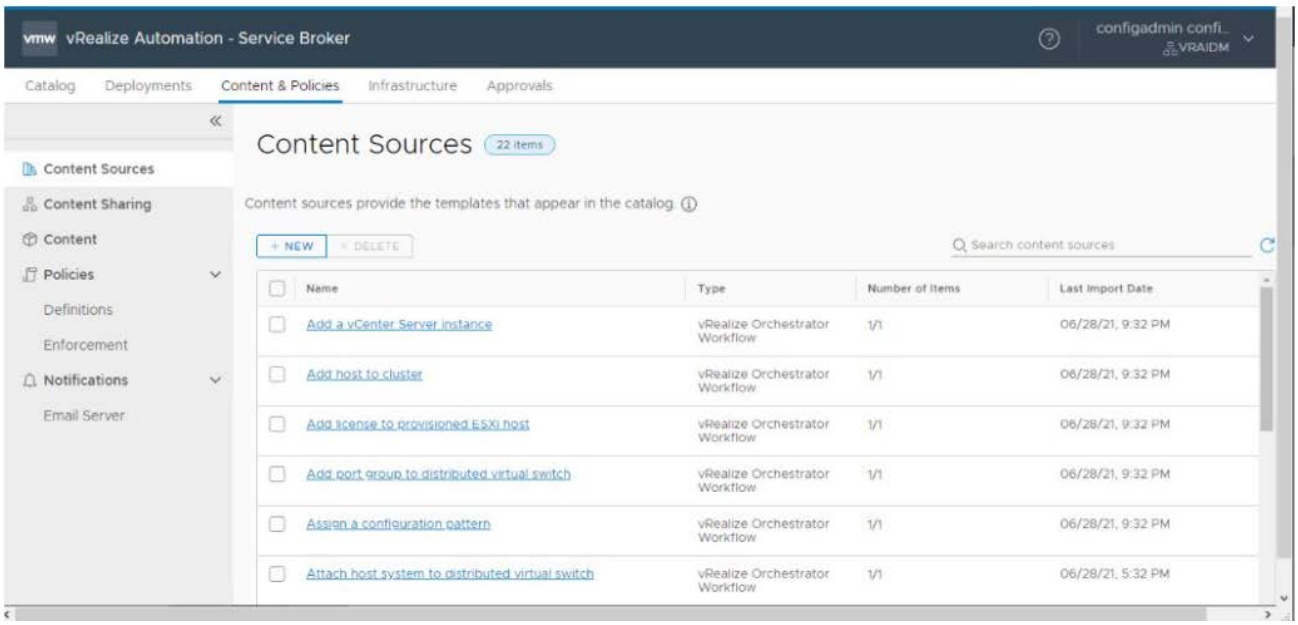

Figure 15: Lenovo XClarity Integrator for vRealize Automation template Items

Lenovo XClarity Adapter for vRealize Operations Manager provides a global view of the relationship between resources, such as connected chassis, servers, power supplies, and ESXi connectivity. The plugin helps to monitor the hardware events in a Lenovo XClarity Administrator-managed environment. Quickly identify trends based on hardware events received, including hardware failures, power/thermal thresholds that exceeded, and PFAs (predicted failure alerts). These events categorize by source, type of hardware surfacing the events, and whether service is required. This information can help identify issues in your data centers so that you can react before more serious issues occur. Figure 16 shows the XClarity Adapter for vRealize Operations Manager interface summary tab contains alerts and recommended actions.

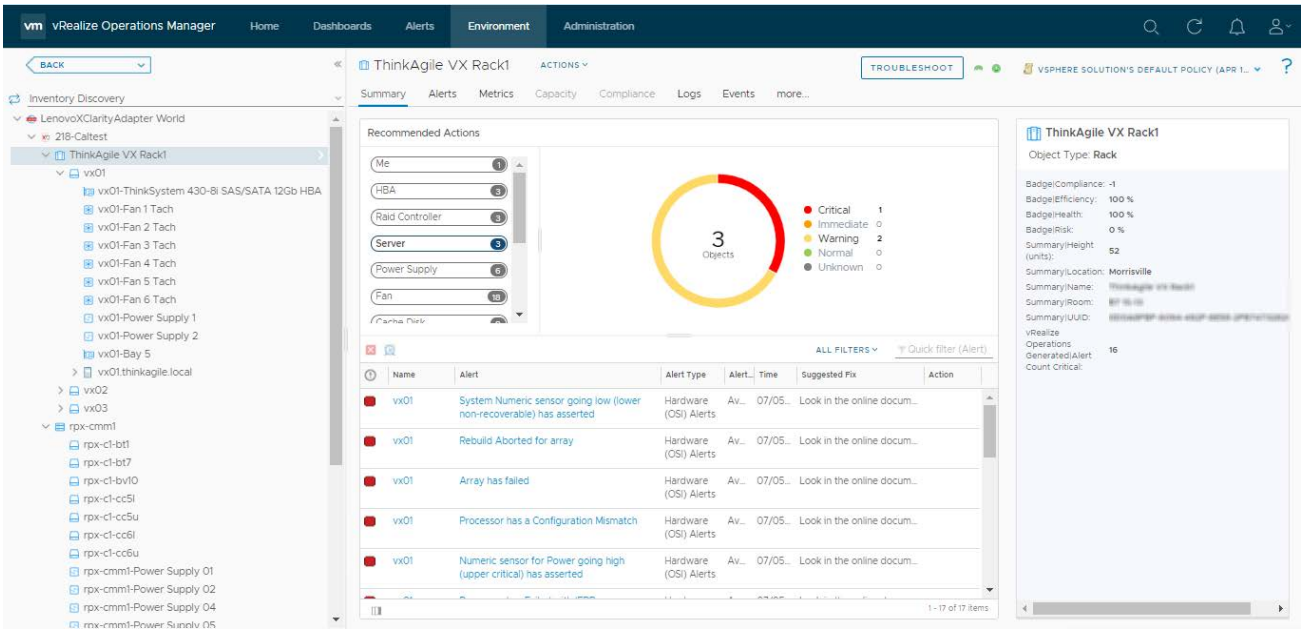

Figure 16: Lenovo XClarity Integrator for vRealize Operations Manager

The vRealize Log Insight content pack for provides analysis of events from the Lenovo XClarity Administrator, Lenovo XClarity Orchestrator, and the resources managed by XClarity. These insights helps to monitor hardware events, resource alerts, auditing security changes, firmware upgrades and configuration management. Figure 17 shows the events insight page for Lenovo XClarity content pack for vRealize Log Insight.

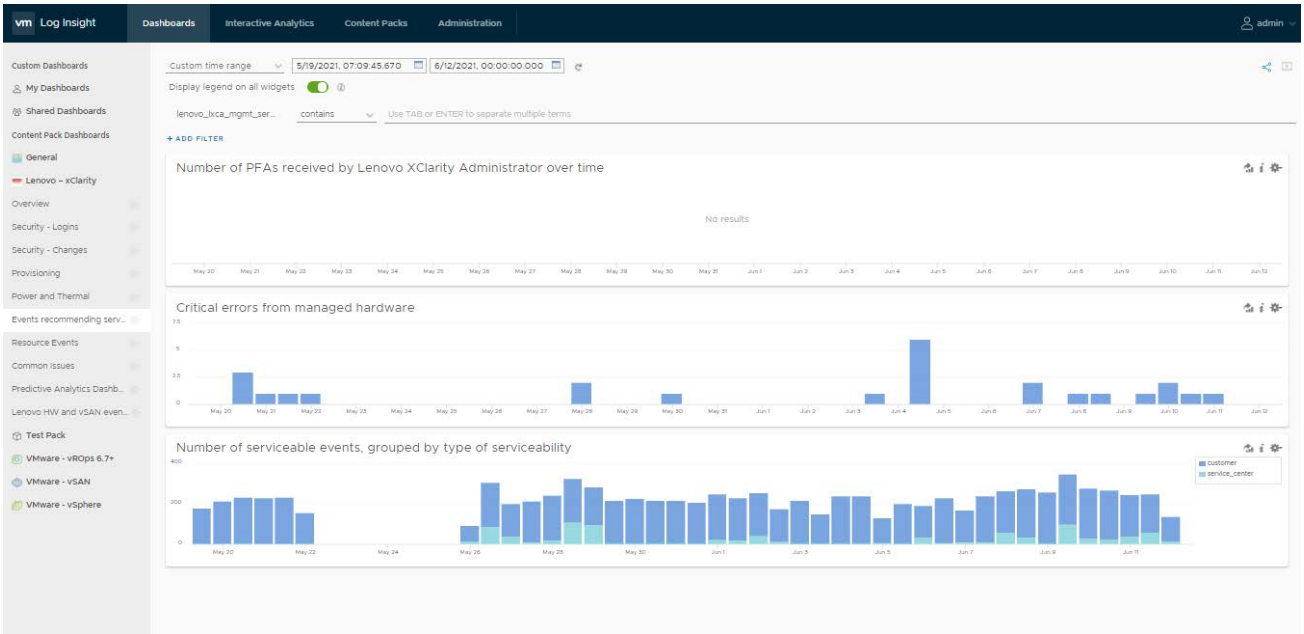

Figure 17: Lenovo XClarity Integrator for vRealize Log Insight

6.9.3 Lenovo XClarity plugins compatibility

Table 22 below lists current versions of Lenovo integration plugins and the required or supported VMware vCenter and vRealize Suite products.

Table 22: Plug-in compatibility

Component Name	Version	Supported Product Versions
Lenovo XClarity Administrator (LXCA)	3.4	VMware vCenter 6.0U2/6.5/6.7, ESXi 6.0U2/6.5 U1/6.7/7.0U2
Lenovo XClarity Integrator (LXCI) for vCenter	7.4	Lenovo XClarity Administrator 1.4.x, 2.x VMware vCenter 5.x U1/U2/U3, 6.0 U1/U2/U3, 6.5 U1/U2,6.7(U1,U2,U3), 7.0(U1, U2,U3)
Lenovo XClarity Administrator content pack for VMWare vRealize Log Insight	1.0	Lenovo XClarity Administrator 1.1 or higher VMware vRealize Log Insight 2.5 or higher
Lenovo XClarity Integrator for VMware vRealize Automation	1.2	VMware vRealize Automation 8.3 or higher
Lenovo XClarity Integrator for VMware vRealize Orchestrator	1.2	VMware vRealize Automation 7.0 VMware vRealize Orchestrator 6.0/7.0
Lenovo Network Plugin for VMware vRealize Orchestrator	1.4.0	VMware vRealize Orchestrator 7.4.x
Lenovo XClarity Content Pack for vRealize Operations Manager	1.2	vRealize Operations Manager 8.0, 8.1, 8.2, and 8.3

7 Deploying SDDC

This chapter provides an introduction to deploying SDDC in your data center.

7.1 VMware Validated Design

The VVD documentation (version 6.1) provides a family of solutions for data center designs that span compute, storage, networking, and management, serving as a blueprint for an SDDC implementation.

This reference design is based on VVD. For more details on VVD, please see this website: vmware.com/support/pubs/vmware-validated-design-pubs.html.

7.2 VMware Cloud Foundation

VMware Cloud Foundation (VCF) is a hybrid cloud platform to deploy VMware SDDC for private cloud based on the VMware Validated Design and to integrate with public clouds running VMware SDDC clouds. It provides software defined services for compute, storage, networking, and cloud management to run different workloads. It simplifies installation, upgrade and patch management of SDDC components through lifecycle management either through online or offline.

VCF supports deploying SDDC components on broad range of physical servers (vSAN Ready Nodes) to have flexible customer defined heterogeneous infrastructure to support variety of workloads.

7.2.1 SDDC Manager

The SDDC Manager provides the core management software for VCF. It automates the installation and lifecycle management of the vSphere, vSAN, and NSX from bring-up and configuration to patching and upgrading, making it simple for the cloud admin to build and maintain the SDDC. SDDC Manager also automates the installation and configuration of vRealize Log Insight, vRealize Operations, and vRealize Automation by using vRealize Suite Lifecycle Manager. SDDC Manger uses same vCenter sso login. The cloud administrator uses vCenter Server as the primary management interface for the virtualized environment.

7.2.2 Workload Domain

A workload domain is a dedicated environment with servers, storage and networking managed by dedicated vCenter and NSX Manager. The management workload domain is created automatically and virtual infrastructure workload domains are created by cloud administrators based on requirements. The resource maximums, limits and scalability for each workload domain is same as the limits applicable for vCenter. The SDDC Manager deploys and configures one vCenter Server and NSX manager per workload domain automatically when the workload domain is created.

Figure 18 shows an example of a management workload domain and two virtual infrastructure workload domains.

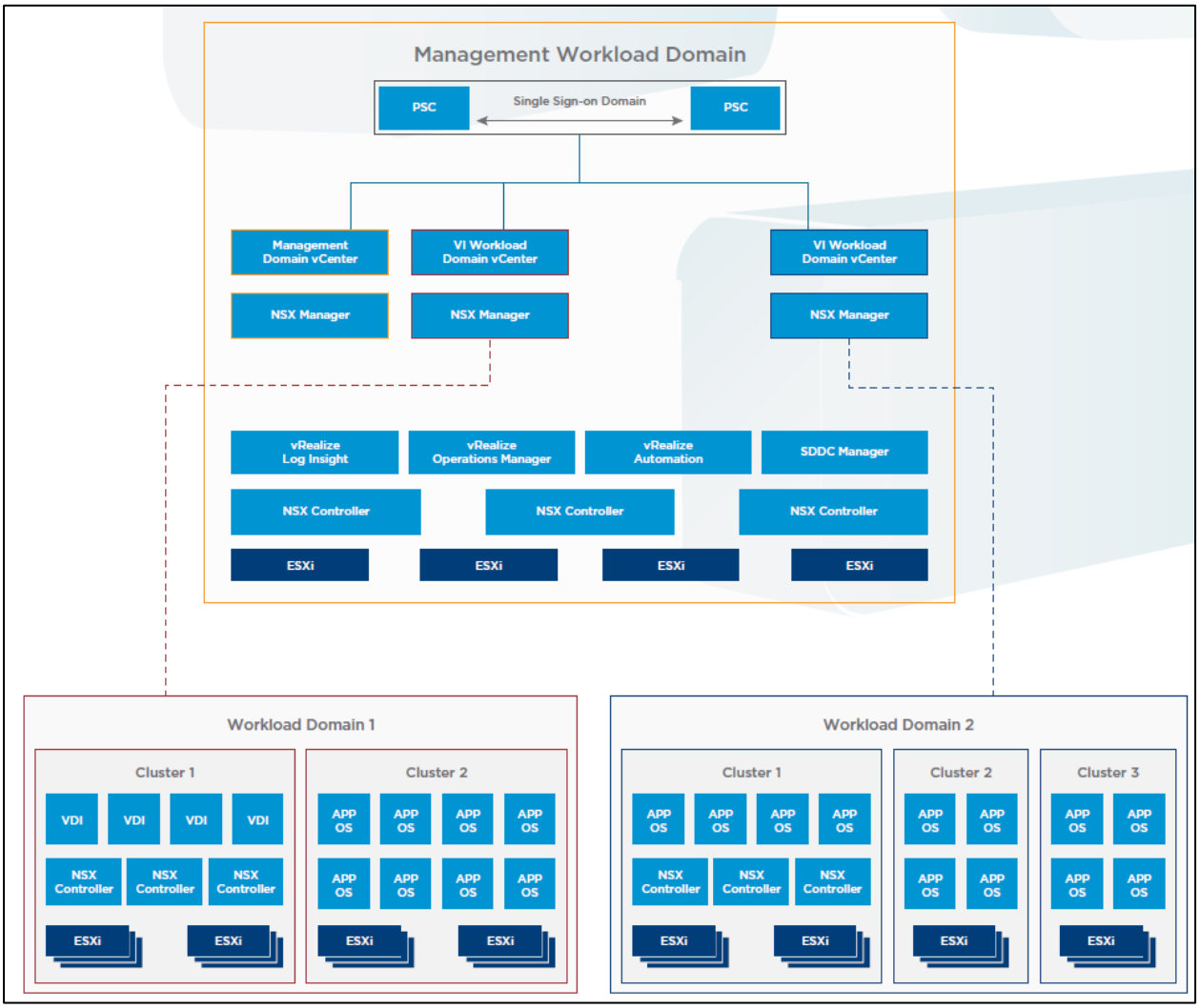

Figure 18: VCF Workload domains

7.3 Lenovo VX Appliance

Lenovo ThinkAgile VX appliances are preloaded with a wizard-based deployment tool to accelerates the greenfield vSAN deployment or new clusters with ESXi 7.0u2 or later. A 4-Node vSAN cluster can be deployed in less than an hour and it works with All Flash and Hybrid vSAN deployments. It discovers the Lenovo ThinkAgile VX nodes over the network, installs ESXi, deploys vSAN and vCenter and install vRealize plugins in the vCenter. Figure 19 shows logical network architecture for deploying ThinkAgile VX cluster to setup vSAN using VX Deployer tool. Please refer this page to use VX Deployer to setup vSAN cluster on VX Appliances and verify the deployed components https://thinkagile.lenovofiles.com/help/index.jsp?topic=%2Fcom.lenovo.thinkagile.vx%2Fcluster_deployment_with_vx_deployer.html

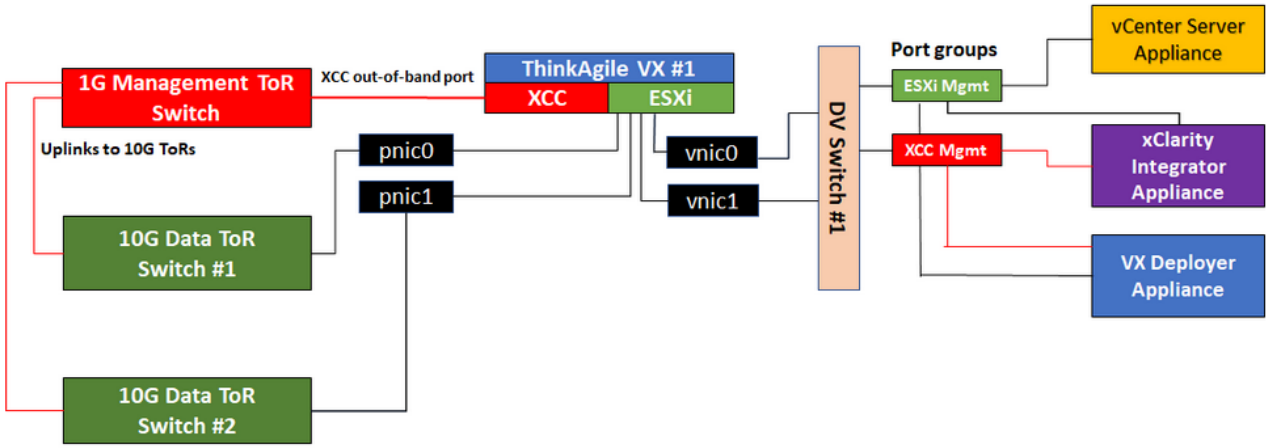

Figure 19: Logical network architecture for Lenovo ThinkAgile VX cluster

7.3.1 Deploying VCF with ThinkAgile VX Appliances

VCF can be installed on Lenovo ThinkAgile VX certified nodes or VX appliances as both have been vSAN certified. Lenovo has validated the install of VCF 3.5.

Table 23 describes the steps to install a complete SDDC environment using VCF 3.5 and ThinkAgile VX appliance.

Table 23: VCF 3.5 installation steps with ThinkAgile VX

#	SDDC Deployment Sequence	Deployed Component/Feature	Comment
1	Run VX Appliance on Lenovo ThinkAgile VX Servers	Deploy ESXi, vSAN, vCenter	Management cluster and compute cluster *(Lenovo XClarity Administrator can also be used to image ESXi servers manually)*
2	Install Cloud Builder Appliance		
3	Run Cloud Builder to deploy Management Workload Domain	SDDC Manager, NSX-T Manager	Input file is used with all configured parameter
4	Deploy VI Workload Domain(s) with SDDC Manager	vCenter, NSX-T	
6	Deploy vRealize Suite (Management Workload Domain)	vRealize Operations, vRealize Suite Lifecycle Manager, vRealize Automation, vRealize Load balancers (NSX Edges)	

7	Deploy Tanzu using VCF	Create Edge services, Tanzu Supervisor Cluster and Kubernetes Cluster	

8 Deployment example

This chapter describes an example deployment of vRealize Suite 8.5, VMware Tanzu 1.x, vSAN 7.0U2, and NSX-T 6.3.1 following the guidance in the VMware Validated Design (VVD) documentation. Four physical servers are used for each of the shared edge and compute, management, and additional compute clusters. Lenovo ThinkAgile VX servers are used for the shared edge and compute cluster and management cluster.
Hardware views

The various hardware views are described in this section.

8.1.1 Rack view

Figure 20 shows a view of the rack with the twelve servers and switches.

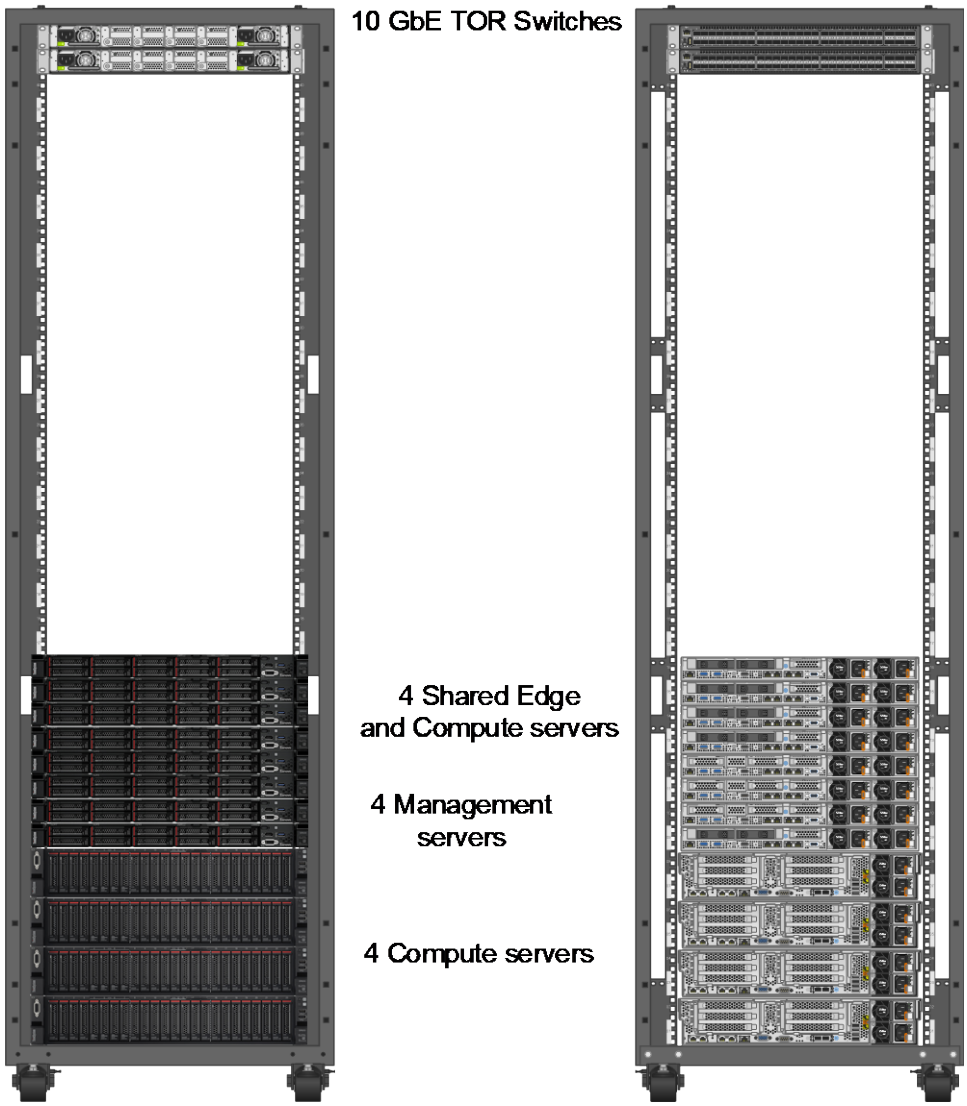

Figure 20: Rack Layout

8.1.2 Network view

Figure 21 shows a view of the physical 10 GbE network and connections to the external internet.

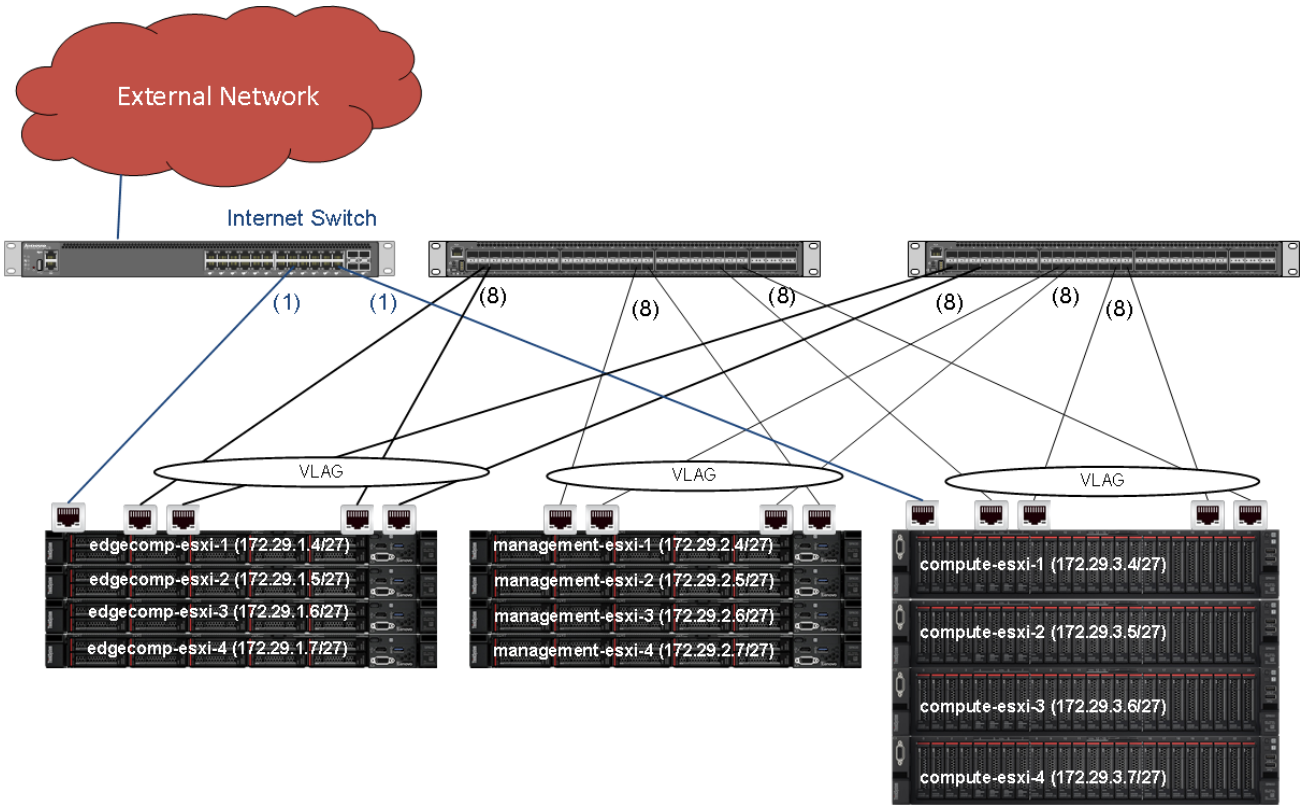

Figure 21: Networking Overview

For the shared edge and compute, management and additional compute clusters, the nodes use VLAG technology and as such are using a LAG configuration within the vSphere Distributed Switches. It is recommended to use VLAG for all the clusters connected to the same set of switches.

The servers in the shared edge and compute cluster and the additional compute cluster are connected to a 1G switch. This switch in turn is connected to the internet via a gateway and firewall (not shown).

8.2 IP/VLAN mapping

This example deployment uses the following nine VLANs:

- Management
- vMotion
- FT
- Storage
- VTEP
- vSAN
- vRA1
- vRA2 (for second region)
- Compute VMs
- vSphere Pod

Table 24 lists example IP address ranges for the VLANs in each cluster where RID means Rack ID.

Table 24: Network Segments

Traffic	**Shared Edge and Compute (RID 1)**		**Management (RID 2)**		**Compute (RID 3)**	
	Subnet	**VLAN**	**Subnet**	**VLAN**	**Subnet**	**VLAN**
Manage	172.29.1.0/27	101	172.29.2.0/27	201	172.29.3.0/27	301
vMotion	172.29.1.32/27	102	172.29.2.32/27	202	172.29.3.32/27	302
FT	172.29.1.64/27	103	172.29.2.64/27	203	172.29.3.64/27	303
Storage	172.29.1.96/27	104	172.29.2.96/27	204	172.29.3.96/27	304
TEP	172.29.1.128/27	105	172.29.2.128/27	205	172.29.3.128/27	305
vSAN	172.29.1.160/27	106	172.29.2.160/27	206	172.29.3.160/27	306
vRA1	N/A	107	172.29.2.192/27	207	N/A	307
vRA2	N/A	108	172.29.2.224/27	208	N/A	308
Comp VMs	172.29.2.192/27	109	N/A	209	172.29.2.192/27	309
vSphere Pod	172.29.2.224/27	110	N/A		172.29.3.224/27	310

In this example, each cluster needs a minimum of five network segments within the 172.29.RID.x address range. Each segment does not require more than 30 IP addresses; therefore, a 255.255.255.224 (/27) netmask provides enough addresses. The vSphere pods and virtual machines uses dedicated VLAN to address appropriate workloads running on them. The same VLAN IDs can be used across racks with different IP segments. In this example, that option is not available because the switches and routers are shared across the three clusters. For more information about how customers can plan their network segments, see the VMware NSX-T Design Guide.

8.3 Cluster Deployment

This section describes list of underlay and overlay virtualized networking used for the clusters. Multiple transport zones are used to segregate the clusters and logical switches that participate in each cluster. With NSX-T, there are flexible options chosen to use either underlay or overlay for different tenants and workloads. NSX-T and VMware Validated Design provides flexibility to deploy edge, management and compute VMs on the same cluster or shared cluster or dedicated cluster.

8.3.1 Deploying vSphere with Tanzu with dedicated cluster

vSphere with Tanzu deployment is done from VCF console. vSphere with Tanzu can be deployed in two clusters, one cluster for the Management and Edge functions, and another one dedicated to Workload Management. Figure 22 shows shared management and edge cluster and dedicated Kubernetes cluster.

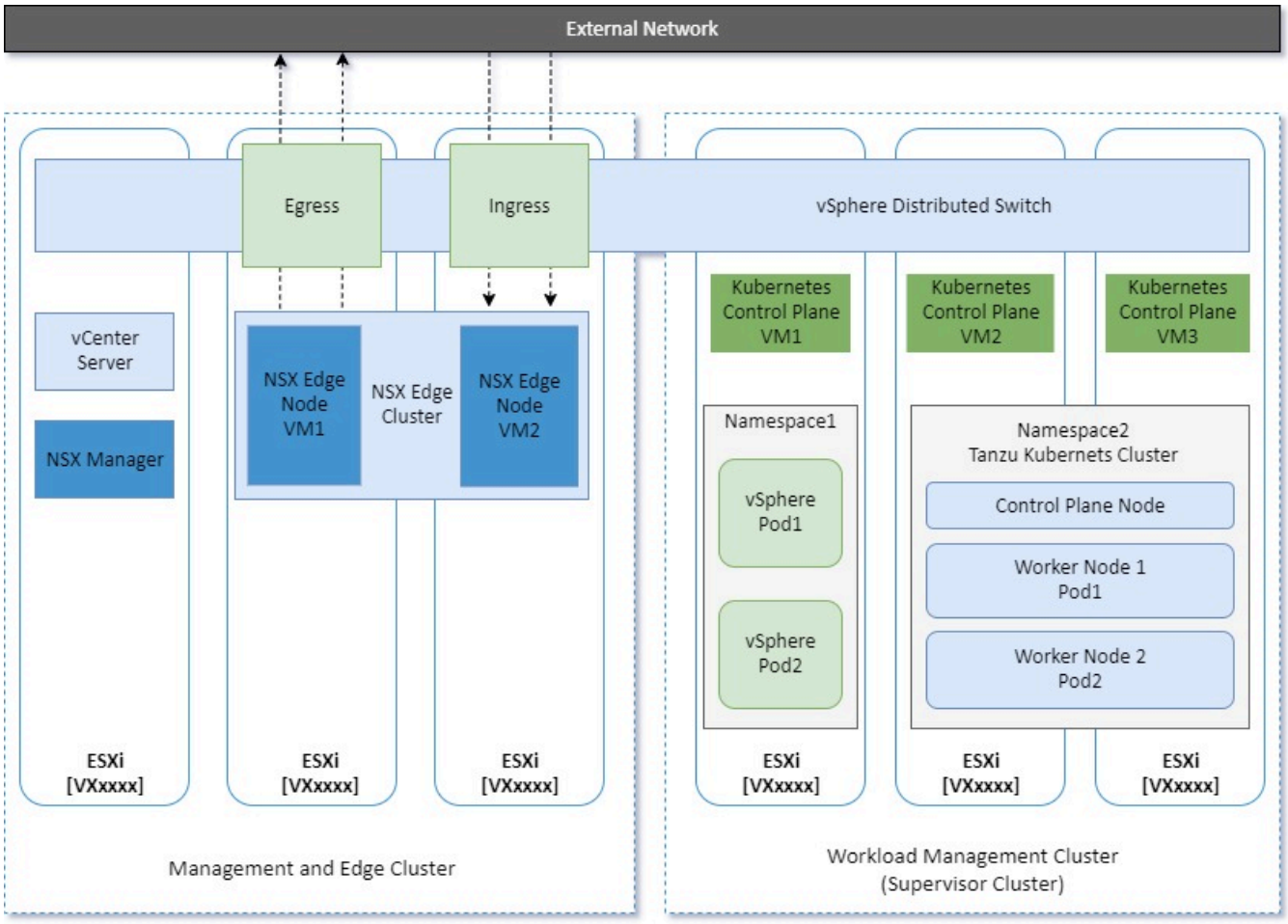

Figure 22: vSphere Tanzu with Dedicated Clusters

Deploying Tanzu Kubernetes Grid(TKG) is network sensitive operation and proper configurations need to be set in the SDDC Manager deployment input file. Since it is an automated deployment, any small configuration

issue will cause failure. The prerequisites are a running VI Workload Domain for Tanzu Kebernetes Cluster, NSX-T Edge on the VI Workload domain, a subnet for pod networking (non-routable, minimum /22), a subnet for Service IP (non-routable, minimum /24), a subnet for ingress(routable, minimum /27), a subnet for egress(routable, minimum/27) and MTU set to 9000 for all VLANs in physical switches. Also the content library needs to be available in vSphere to download and install components.

8.3.2 Deploying vSphere with Tanzu Consolidated Architecture

vSphere with Tanzu can be deployed in a single vSAN cluster where VCF workload management domain is running. This consolidated architecture hosts management, workload and edge components on a single cluster which ideally suit for development and SMB environment, but it can also be used for large environments where isolation needed based on teams or organization group. Figure 23 shows consolidated architecture to run all domains on the same cluster.

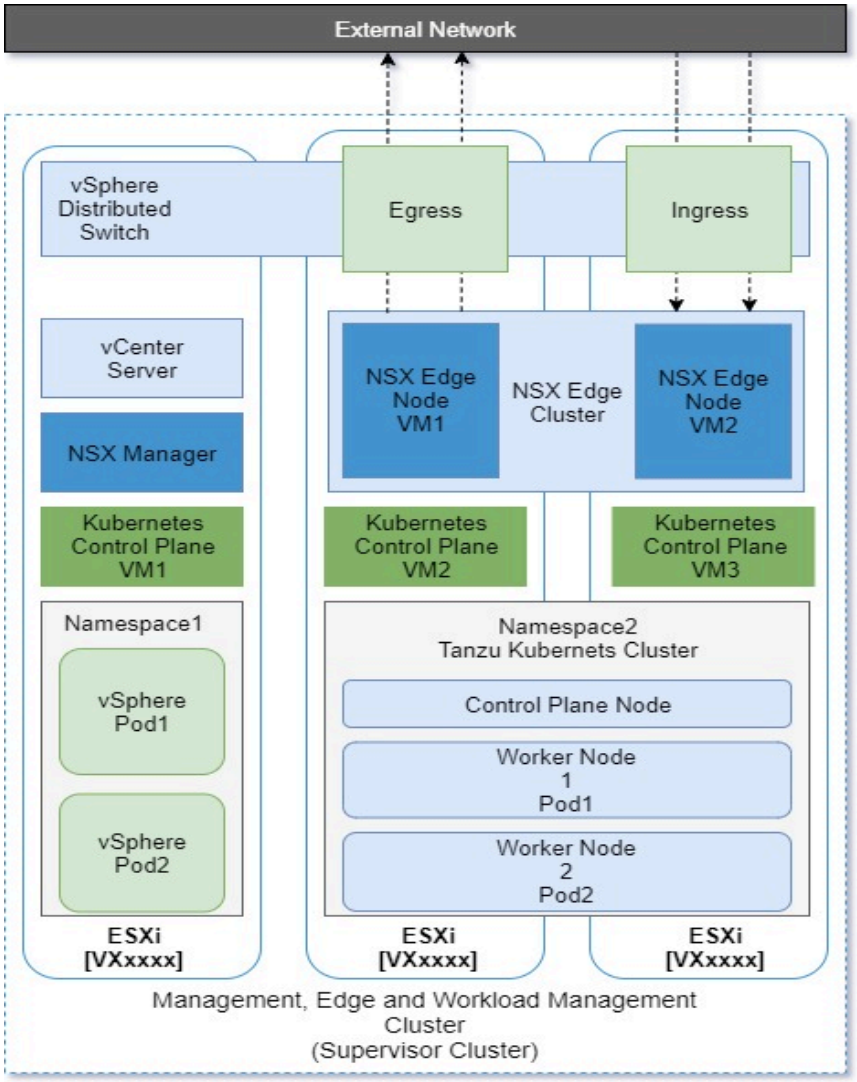

Figure 23: vSphere Tanzu with consolidated architecture

8.3.3 vSphere with Tanzu Deployment Best Practices

The SDDC Manager automates the deployment of Tanzu and the following networking considerations taken care before starting the deployment.

- The MTU should be set to 9000 for all VLANs except the management VLAN
- The NSX-T Edge cluster should use EBGP
- The routable VLANs should be configured on physical switches for ingress and egress traffic for Workload Domain
- The first IP of the subnet should not be used as default gateway IP on the physical switches for ingress and egress VLANs. Because this is being used by NSX-T for SNAT rule for subnets.
- The DNS server should be accessible from egress VLAN
- At least 5 IP addresses need to be reserved for the control plane VMs on the management VLAN
- The VLANs used for NSX-T Edge VTEP and Host TEP should be inter-routable.

9 Microsoft SQL Server

Microsoft SQL Server instances can be successfully deployed on the vCF Management Cluster, using the Compute, Storage and Network resources available across the Management Cluster. In parallel with the existing virtual machines (NSX-T, vCenter etc.) that were already deployed during the vCF build-up, we can deploy Microsoft SQL Server VMs that will use the afore-mentioned resources. In order to deploy a vCF Management Cluster, a minimum of four bare metal server must be used (see vCF specifications). The servers that host the vCF Management Cluster and the Microsoft SQL server have the following configuration:

- Server Model: ThinkAgile VX
- CPUs: 2x 3rd generation Intel Xeon Scalable Processors
- RAM: 768GB RAM
- Network: 2 x 10Gbps network interfaces (minimum requirement for vCF deployment)
- 2 x 128GB NVMe drives (ESXi OS)
- vSAN storage: 2 x Disk Groups (1 x 800GB SSD for vSAN Cache and 3 x 7.68TB SSD for vSAN Capacity per Disk Group)

Microsoft SQL Server 2019 was deployed on 8 VMs (2 VMs per each ESXi host) with the following settings, per VM:

1. OS: Windows Server 2019 Standard
2. MSSQL version: Microsoft SQL Server 2019
3. vCPUs: 26 vCPUs (2 NUMA nodes, 13 vCPUs per NUMA node)
4. RAM: 256GB

A professional Microsoft SQL database benchmarking tool called HammerDB was used to compare the performance of the Microsoft SQL Servers that run on each ESXi server. The same database was used across all MSSQL Servers.

Note: The HammerDB VMs are running on a separate cluster so they will not interfere with the MSSQL Servers. Enough Compute resources have also been allocated to the HDB VMs.

CASE I – HammerDB tests running on in sequence on each Microsoft SQL Server 2019:

ESXi Hostname	MSSQL VM Hostname	ESXi CPU Core utilization (%)	MSSQL VM CPU usage (%)	Transactions per minute	vSAN Read IOPS	vSAN Write IOPS
ESXi01	MSSQL2019-01	45	65	1,032,237	134	1,690
ESXi01	MSSQL2019-02	35	50	737,898	63	1,402
ESXi02	MSSQL2019-03	44	55	**1,053,828**	56	1,614
ESXi02	MSSQL2019-04	35	53	754,177	1	1,620
ESXi03	MSSQL2019-05	45	35	841,115	145	1,214
ESXi03	MSSQL2019-06	45	35	779,311	108	863

ESXi04	MSSQL2019-07	45	53	750,877	1	1,648
ESXi04	MSSQL2019-08	53	58	770,458	2	1,751
Maximum TPMs				**1,053,828**		

A total of 8 tests were ran in sequence with the **maximum TPMs** on the vCF Workload Management Cluster running on a single MSSQL VM achieved on MSSQL2019-03: **1,053,828**

CASE II.a – HammerDB tests running simultaneously on 4 MSSQL VMs (one on each host):

ESXi Hostname	MSSQL VM Hostname	ESXi CPU Core utilization (%)	MSSQL VM CPU usage (%)	Transactions per minute	vSAN Read IOPS	vSAN Write IOPS
ESXi01	MSSQL2019-01	22	13	**517,544**	67	936
ESXi01	MSSQL2019-02					
ESXi02	MSSQL2019-03	30	33	**536,838**	1	619
ESXi02	MSSQL2019-04					
ESXi03	MSSQL2019-05	30	33	**451,501**	127	976
ESXi03	MSSQL2019-06					
ESXi04	MSSQL2019-07	40	33	**458,694**	1	729
ESXi04	MSSQL2019-08					
Maximum TPMs				**1,964,577**		

Four HammerDB tests with 4 MSSQL Servers were run in parallel and the maximum TPMs on the vCF Management Cluster was obtained by adding all the results **1,964,577**.

CASE II.b – HammerDB tests running simultaneously on 4 MSSQL VMs (one on each host):

ESXi Hostname	MSSQL VM Hostname	ESXi CPU Core utilization (%)	MSSQL VM CPU usage (%)	Transactions per minute	vSAN Read IOPS	vSAN Write IOPS
ESXi01	MSSQL2019-01					
ESXi01	MSSQL2019-02	22	25	**398,570**	73	862
ESXi02	MSSQL2019-03					
ESXi02	MSSQL2019-04	30	33	**459,582**	1	1,494
ESXi03	MSSQL2019-05					
ESXi03	MSSQL2019-06	30	15	**437,060**	119	1,525
ESXi04	MSSQL2019-07					
ESXi04	MSSQL2019-08	38	31	**376,585**	1	877
Maximum TPMs				**1,671,797**		

Four HammerDB tests with 4 MSSQL Servers were run in parallel with the maximum TPMs on the vCF Management Cluster obtained by adding all the results: **1,671,797.**

CASE III – HammerDB running in sequence, on two MSSQL VMs on the same ESXi host:

ESXi Hostname	MSSQL VM Hostname	ESXi CPU Core utilization (%)	MSSQL VM CPU usage (%)	Transactions per minute	vSAN Read IOPS	vSAN Write IOPS
ESXi01	MSSQL2019-01	75	66	786,252	80	1,280
ESXi01	MSSQL2019-02	75	54	522,032	52	1,188
ESXi02	MSSQL2019-03	75	72	**875,197**	1	1,154
ESXi02	MSSQL2019-04	75	58	**657,593**	1	1,158
ESXi03	MSSQL2019-05	80	41	635,015	92	1,366
ESXi03	MSSQL2019-06	80	39	619,173	81	813
ESXi04	MSSQL2019-07	75	58	627,030	1	1,293
ESXi04	MSSQL2019-08	75	60	612,029	1	1,389
Maximum TPMs				**1,532,790**		

Four HammerDB tests were run in sequence, using two Microsoft SQL Server VMs residing on the same host per each run with the maximum TPMs achieved on ESXi02: **1,532,790**

CASE IV – HammerDB test running on all MSSQL Server VMs

ESXi Hostname	MSSQL VM Hostname	ESXi CPU Core utilization (%)	MSSQL VM CPU usage (%)	Transactions per minute	vSAN Read IOPS	vSAN Write IOPS
ESXi01	MSSQL2019-01	52	39	**541,296**	61	1,659
ESXi01	MSSQL2019-02	52	38	**371,921**	68	836
ESXi02	MSSQL2019-03	50	42	**528,757**	1	1,169
ESXi02	MSSQL2019-04	50	35	**436,537**	1	925
ESXi03	MSSQL2019-05	53	22	**445,418**	111	1,424
ESXi03	MSSQL2019-06	53	20	**424,757**	52	1,803
ESXi04	MSSQL2019-07	57	37	**453,999**	1	1,240
ESXi04	MSSQL2019-08	57	37	**382,274**	1	858
Maximum TPMs				**3,584,959**		

The maximum TPMs achieved on the Management Cluster was obtained by adding the results from the 8 tests that were run in parallel: **3,584,959.**

10 AI/ML Workloads

Enterprise wide data science and AI/ML technologies and applications can leverage vSAN, VMware Cloud Foundation and Tanzu services for simplified infrastructure management, performance and security. Lenovo ThinkAgile VX systems are resilient and perfect choice for consolidating enterprise applications and AI/ML workloads on the shared infrastructure. Lenovo ThinkAgile VX3 systems with Intel 4th Generation Scalable Processors and AMD EPYC 4th Generation processors support running high performance data science and AI/ML inference workloads.

For detailed enterprise AI use cases, system design and guidelines refer “Reference Architecture: Enterprise AI/ML Workloads on Lenovo ThinkSystem and ThinkAgile Platforms” at https://lenovopress.lenovo.com/
For Generative AI use cases refer Reference Architecture for Generative AI Based on Large Language Models (LLMs)

10.1 Intel Accelerators for AI/ML

Intel provides many optimization libraries which can accelerate AI/ML workloads and Intel Xeon 4th Generation Scalable Processors provide significant performance improvement than previous generations.
Refer more information here https://www.intel.com/content/www/us/en/developer/topic-technology/artificial-intelligence/overview.html

Intel® Advanced Vector Extensions 512 (Intel® AVX-512) instruction set can accelerate performance for workloads such as scientific simulations, financial analytics, artificial intelligence (AI)/deep learning, 3D modelling and analysis, image and audio/video processing, cryptography and data compression. AVX-512 provides significant performance improvement over AVX2, AVX and SSE.
Intel® Advanced Matrix Extensions (Intel® AMX) is a set of matrix multiply instructions that can operate on a tile using a variety of datatypes including bfloat16 (BF16) and int8 to accelerate deep learning inference and training tasks
Intel® Data Streaming Accelerator (Intel® DSA) is a high-performance data copy and transformation accelerator that is integrated in several Sapphire Rapids processors, targeted for optimizing streaming data movement and transformation operations common with applications for high-performance storage, networking, persistent memory, and various data processing applications.

10.1.1 Image Classification Performance on Intel Xeon 4th Generation Scalable Processors

The ResNet50 (Residual Network) benchmark is for testing deep learning model for image recognition and the test is performed with Intel 4th Generation Xeon Scalable Processors. The table below shows throughput (images per second) achieved for FP32, BF6 and INT8.

Test Configuration	
Server	4x ThinkAgile VX650 V3
CPU	2x Intel® Xeon® Platinum 8468V, 48 cores

Memory	1 TB, 16x64GB DDR5 4800 MT/s
NVMe	10x 3.2 TB
NIC	2 x 100 GbE
vSAN	vSAN ESA, RAID-5
vSphere	VMware 8.0U1c, 22088125
Guest OS	Ubuntu Server 22.04, Kernel 5.15
VM Configuration	96vCPU+64GBRAM, 4 Cores per instance
Benchmark	ResNet50 v1.5
Configuration	intel-optimized-tensorflow:2.11.0 Batch size=128 Dataset (synthetic, autogenerated (no accuracy measurement))
Throughput Results	
Intel AVX-512 for FP32	1300 images per second
Intel AMX for BF16	5700 images per second
Intel AMX for INT8	11100 images per second

10.1.2 Natural Language Processing (NLP) Performance on Intel Xeon 4th Generation Scalable Processors

The BERT-large benchmark is for testing NLP inference performance with Stanford Question Answering Dataset(SQuAD) and the test is performed with Intel 4th Generation Xeon Scalable Processors. The table below shows throughput (samples per second) achieved for FP32, BF6 and INT8.

Test Configuration	
Server	4x ThinkAgile VX650 V3
CPU	2x Intel® Xeon® Platinum 8468V, 48 cores
Memory	1 TB, 16x64GB DDR5 4800 MT/s
NVMe	10x 3.2 TB
NIC	2 x 100 GbE
vSAN	vSAN ESA, RAID-5
vSphere	VMware 8.0U1c, 22088125
Guest OS	Ubuntu Server 22.04, Kernel 5.15
VM Configuration	96vCPU+64GBRAM, 4 Cores per instance
Benchmark	Bert-Large
Configuration	intel-optimized-tensorflow:2.11.0 Batch size=128, 28x2, 32x2 and 48x2 instances Dataset (SQuAD 1.1)
Throughput Results	
Intel AVX-512 for FP32	32 samples per second
Intel AMX for BF16	135 samples per second
Intel AMX for INT8	197 samples per second

11 Conclusion

The combination of Lenovo ThinkAgile VX nodes, VMware Cloud Foundation and Tanzu provides an ideal hybrid cloud platform for a customer to start their application modernization journey. The integration with NVIDIA AI Enterprise provides ideal foundation to deploy machine learning and deep learning workloads at scale.

Resources

For more information about the topics that are described in this document, see the following resources:

- Software Defined Data Center:
  vmware.com/software-defined-datacenter
- VMware Validated Designs Documentation (VVD):
  vmware.com/support/pubs/vmware-validated-design-pubs.html
- vSphere Hypervisor (ESXi):
  vmware.com/products/vsphere-hypervisor
- vCenter Server:
  vmware.com/products/vcenter-server
- vSAN:
  vmware.com/products/virtual-san
- VMware Compatibility Guide (VCG)
  vmware.com/resources/compatibility
- NSX:
  vmware.com/products/nsx
- VMware NSX-T Reference Design Guide
  nsx.techzone.vmware.com/resource/nsx-t-reference-design-guide-3-0
- VMware Cloud Foundation:
  vmware.com/products/cloud-foundation
- VMware Tanzu:
  tanzu.vmware.com/products
- vRealize Automation:
  vmware.com/products/vrealize-automation
- vRealize Automation Reference Architecture:
  vmware.com/files/pdf/products/vCloud/vRealize-Automation-6x-Reference-Architecture.pdf
- vRealize Operations:
  vmware.com/products/vrealize-operations
- vRealize Business:
  vmware.com/products/vrealize-business
- vRealize Log Insight:
  vmware.com/products/vrealize-log-insight

Document history

Version 1.0	29 September 2021	• First version for Lenovo ThinkAgile VX with VCF
Version 1.1	02 February 2022	• Added more ThinkAgile VX appliances • Added VMware Tanzu Advance edition features • Revised SDDC deployment components • Added VMware Tanzu deployment best practices
Version 1.2	02 April 2022	• Added SQL Server performance data
Version 1.3	28 June 2022	• Added Tanzu Mission Control and Tanzu Service Mesh
Version 1.4	27 September 2023	• Added AMD EPYC 9004 (Genoa) ThinkAgile V3 IS CN
Version 1.5	26 December 2023	• Added Section 5.3 vSAN Data Persistence Platform • Added Section 5.7 VMware Private AI Foundation with NVIDIA Enterprise AI • Updated Section 6.1.1 ThinkAgile VX Server with V3 servers • Added Section 10 AI/ML Workloads with Intel benchmark results

Trademarks and special notices

References in this document to Lenovo products or services do not imply that Lenovo intends to make them available in every country.

The following terms are trademarks of Lenovo in the United States, other countries, or both:

Lenovo®
Flex System
System x®
ThinkAgile
ThinkSystem
XClarity®

The following terms are trademarks of other companies:

Intel® and Xeon® are trademarks of Intel Corporation or its subsidiaries.

Linux® is the trademark of Linus Torvalds in the U.S. and other countries.

Active Directory®, Azure®, Hyper-V®, Microsoft®, SQL Server®, and Windows® are trademarks of Microsoft Corporation in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.

Information is provided "AS IS" without warranty of any kind.

All customer examples described are presented as illustrations of how those customers have used Lenovo products and the results they may have achieved. Actual environmental costs and performance characteristics may vary by customer.

Information concerning non-Lenovo products was obtained from a supplier of these products, published announcement material, or other publicly available sources and does not constitute an endorsement of such products by Lenovo. Sources for non-Lenovo list prices and performance numbers are taken from publicly available information, including vendor announcements and vendor worldwide homepages. Lenovo has not tested these products and cannot confirm the accuracy of performance, capability, or any other claims related to non-Lenovo products. Questions on the capability of non-Lenovo products should be addressed to the supplier of those products.

All statements regarding Lenovo future direction and intent are subject to change or withdrawal without notice, and represent goals and objectives only. Contact your local Lenovo office or Lenovo authorized reseller for the full text of the specific Statement of Direction.

Some information addresses anticipated future capabilities. Such information is not intended as a definitive statement of a commitment to specific levels of performance, function or delivery schedules with respect to any future products. Such commitments are only made in Lenovo product announcements. The information is presented here to communicate Lenovo’s current investment and development activities as a good faith effort to help with our customers' future planning.

Performance is based on measurements and projections using standard Lenovo benchmarks in a controlled environment. The actual throughput or performance that any user will experience will vary depending upon considerations such as the amount of multiprogramming in the user's job stream, the I/O configuration, the storage configuration, and the workload processed. Therefore, no assurance can be given that an individual user will achieve throughput or performance improvements equivalent to the ratios stated here.

Photographs shown are of engineering prototypes. Changes may be incorporated in production models.

Any references in this information to non-Lenovo websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this Lenovo product and use of those websites is at your own risk.