zadara

zCompute Administrator User Guides

Release 25.06

Zadara
Jun 14, 2026

CONTENTS

CHAPTER

ONE

INTRODUCTION TO ZADARA CLOUD SERVICES

Zadara Cloud Services is a full stack of cloud services: Compute, Networking and Storage. It is protected by Zadara security controls and provisioned as managed service.

Zadara cloud services gives any MSP the ability to have an AWS-compatible on-premise private or public cloud within its own data center. It gives enterprise customers the ability to consume above services in the exact same way they consume the public cloud using AWS-compatible API.

Zadara cloud services are developed, deployed and maintained by Zadara on behalf of its partners and customers.

Zadara cloud services always run on Zadara owned equipment (either in public or on premise deployments), and delivered as a service with a “per consumption” business model.

At a high level:

- Zadara compute cluster is added to Zadara storage cloud to provide the full range of Zadara Cloud Services.
- Zadara Software that runs on the above cloud configures the servers to be either compute or storage nodes, thus providing the experience of a private cloud region.

The Zadara Cloud Service includes the following:

- A **fully-featured compute layer** that allows for creation, management, and manipulation of x86-based virtual machines. Zadara provides intelligent load balancing. If it finds an overloaded node, it automatically moves workloads away from that node to a node with more capacity - typically moving a VM from source to destination in about half a second.
- The ability to use **any type of storage** - Block, File and Object Storage, all provided by Zadara VPSA.
- A **complete virtual networking layer** allowing for the creation of objects ranging from internal private networks, external NAT communication networks, fixed IPs for specific VMs, and security groups to control what may flow across those networking objects and paths.
- A **fully distributed automation and control architecture** that exists across the cloud. Management processes and state are replicated to every node in the region, This make the usage of management resources much more efficient, and much more reliable in case of a planned or unplanned outages.
- An **API** and **CLI** for scripting, automation, and control.
- A fully-featured, browser-based **GUI** to handle all day-to-day operations without having to drop to the command line.

CHAPTER

TWO

ZADARA CLOUD SERVICES CATALOG

2.1 Cloud Services Administration

Each customer that orders Cloud Services from the Zadara Provisioning Portal, gets an account on a specific cloud in the requested location. The account contains one Project by default, but additional projects can be created. A Project is similar to an AWS project or VMware folder, and is basically a container of resources (VPCs, VMs, ELB, etc.). Definition of different projects allows separation of resources to different application groups or environments. The Tenant Admin can assign roles and policies to users per project.

2.2 Users and Roles

On Zadara IaaS platform, users are defined with one of the following roles:

2.2.1 Account Member

This role allows the user to use the console, policies and APIs for creating, viewing, modifying and deleting virtual resources (e.g. VMs, volumes, etc.) belonging to projects to which the user has been assigned. This is the standard role for most users.

2.2.2 Tenant Admin (Account Admin)

In addition to allowing the use all functions which are granted to a Member, the Tenant Admin role also allows the user to use policies and APIs for creating and managing new projects and users within a specific account. The user that is registered on the Zadara Provisioning Portal for Compute services, has the role of Tenant Admin and is responsible to manage the account on behalf of his organization. The Tenant Admin can create additional users with this role.

The following services are available for regular users (account members):

2.3 Compute

2.3.1 Compute Instances (VMs)

Virtual Machines of any type (built-in or user defined) can be created based on OS image with the use of a cloud-optimized hypervisor (Enhanced KVM). This service has AWS EC2 compatible API that offers standard AWS API for application integration.

2.3.2 Snapshots

A snapshot is a copy of the volumes of a virtual machine at a given point in time. Snapshots provide a change log for the volume and are used to restore a VM to a particular point in time when a failure or system error occurs. You can create snapshots of the VMs that are in the system, and use these snapshots to restore any VM if needed.

2.3.3 Key Pairs (security keys settings)

Public–key cryptography is used to encrypt and decrypt login information for Linux instances. (Instances created from a Windows image are accessed by a password instead.) A public key is used to encrypt a piece of data, and then the recipient uses the private key to decrypt the data. These two keys are known as a *key pair*. Key pairs enable you to securely access your instances using a private key instead of a password.

2.3.4 Auto Scaling

The system monitors applications and automatically adjusts allocated resources to meet evolving requirements. It simplifies scaling operations to a minimal effort based on a simple and powerful interface for planning and defining launch configurations and scaling policies. You define the policy, and the system ensures performance and availability while keeping resource utilization optimized to reduce costs.

2.4 Networking

2.4.1 VPC

The Virtual Public Cloud (VPC) is your own private data center within the cloud infrastructure. You get to select the network addresses that you will use throughout your infrastructure. Since this is your network, you can decide to slice it up any way you prefer. The VPC is a networking resource with a logical router at its core. When you create a VPC, you specify a CIDR block. All subnets that you will create in the VPC will be carved out from this CIDR block, and the router will ensure IP connectivity between them.

2.4.2 Subnets

Subnets are used for configuration of networking within a VPC. As compared to using one big network, use of subnets in your VPC allows segregation between private and public facing networks, definition of availability zones, and other similar advantages.

2.4.3 Network Interfaces

A **virtual network interface** (VIF) is an abstract virtualized representation of a computer network interface. Each virtual machine (VM) may have one or more virtual network interfaces that act as virtual NICs.

2.4.4 Route Tables

The basic tenet of networking is that everything inside your subnet stays inside your subnet – and if you want to go outside of your subnet, you need to go through the default gateway through which traffic is routed to an external destination network. The route tables are associated with each subnet to allow the flow of traffic according to the VPC policies and configured options.

2.4.5 Internet Gateways (IGW)

Your connection to the outside world is the Internet Gateway. Having instances running in a cloud is great, but if you cannot get to them from the outside world they might be useless. Without an IGW, manageability would be very challenging – if not completely impossible.

2.4.6 Elastic IP (EIP)

Elastic IPs are used to expose instances outside of the Zadara compute cloud. An EIP will be used in the network address operation (NAT) of all traffic to/from the virtual network interface with which it is associated.

2.4.7 VPC Peering (in the same cluster)

VPC peering lets you create direct IP connectivity between any two VPCs. Direct connectivity between VPCs means that servers in a VPC can be reached from the other VPC without the need for elastic IPs or traffic flowing through external routing. VPC peering is simply L3 connectivity realized using routing tables and IP connectivity.

2.4.8 NAT Gateways (NGW)

There will be cases where you don't want instances to be exposed to the outside world and you don't want them to have public IP addresses. However, these instances may still need access to the outside world to get updates or to send information. By using a NGW, you can allow outbound access to the internet while limiting inbound access, thus providing an additional layer of abstraction and protection for your workloads.

2.4.9 DNS

The DNS Service provides an easy-to-manage DNS solution for your private network. It offers high availability and a cost-effective way to connect applications and services and make them available to users. The DNS service is seamlessly integrated with the resources and services of the cloud. You can map domain names to your compute instances and other services.

2.4.10 Security Groups

Security groups are essentially whitelists applied to the virtual network interfaces to control the inbound and outbound traffic. Traffic that does not match any rule in the security group will be discarded.

For each security group, you include one set of rules that controls the inbound traffic to the instances, and a separate set of rules that controls the outbound traffic from the instances.

2.5 Storage

2.5.1 Block Storage (EBS)

EBS (Elastic Block Store) is native block storage, designed to be extremely scalable, agile and flexible, while being fully compatible with the requirements of cloud computing. The Block Storage service offers comprehensive block storage capabilities, enabling you to reduce latency and overhead and to increase throughput. The service is based on Zadara VPSA block services and fully compatible with AWS EBS API for application integration.

Different capabilities are available, based on *Volume types*.

2.5.2 Volume types

From version 23.08, zCompute simplifies storage management with volume types that provide users with a range of options to meet their storage requirements, balancing factors such as performance, cost, and specific workload demands.

✔ **Note**

The cloud administrator needs to explicitly activate the volume types capability for specific edge clouds.

To check if this capability is configured, go to **Storage > Block Storage** and see whether the **Volume Type** column is displayed instead of the **Storage Pool** column.

By abstracting the underlying storage infrastructure, volume types simplify storage management and allow users to focus on selecting the appropriate type for their zCompute needs.

Different volume types offer varying levels of input/output operations per second (IOPS) and throughput, which determine the storage performance. Higher performance volumes are typically associated with faster data transfer rates and lower latency, enabling applications to handle demanding workloads or perform intensive operations.

Volume types may offer additional features such as encryption, snapshot capabilities, or integration with other services, enhancing data security, backup, and data management workflows in zCompute.

2.5.3 File Shares

NAS service is available to provide scalable files shares via NFS or SMB protocols. The service is based upon the Zadara VPSA NAS offering that allows file access to all associated VM instances. (This service is currently provisioned via the Zadara VPSA interface, and is not compatible with AWS EFS API).

2.5.4 Object Storage

Both private and public object storage services are available on the Zadara IaaS platform. You can create your own private Object Storage, where you have full control, or just get an account on a Zadara regional public service. Both can be accessed via AWS S3 API, or Swift interface. A private Object Storage is directly connected to your VPC, while the public Object Storage is accessible via a public IP. (This service is currently provisioned via the Zadara Object Storage interface).

2.6 Load Balancing

2.6.1 Load Balancers

The Managed Load Balancer service offers the ability to spin up, customize and scale load balancers to support fault tolerance and ensure high availability and application scalability over time. To simplify operations even further, this service supports AWS ELB APIs. It routes traffic according to application or network considerations and provides the required amount of load balancing capacity needed and distributes it to meet high availability and network performance requirements.

2.6.2 Target Groups

A target group is a group of instances to which a load balancer directs application traffic. The instances in this group collectively do the processing work that the application requires.

2.7 Machine Images

2.7.1 Images

The system comes with a number of VM images ready to use. In addition you can create your own images as needed. Two image formats are supported - VMware OVA image and the KVM compatible image in RAW or QCOW2 formats. The latter means that most OpenStack KVM images can be uploaded easily. All of the above are converted into the RAW format and preserved in the image repository in the cluster.

2.7.2 Images Marketplace

Machine images of the different OSs can be downloaded from the Images Marketplace, and be used to create Instances. For OSs that require activation (and vendor fees) you will need to get the activation codes directly from the vendor.

2.8 Monitoring

2.8.1 Events

Zadara compute cloud provides information about various events in your system. Events are any actual changes or attempts at changes in the state of the services, whether initiated manually or automatically by the system. This service provides an events viewer with the ability to filter by date, severity, time, and other relevant parameters.

The system offers visualized metrics, logs and events of all resources, including CPU utilization, memory, storage and network usage. Insights are available on a broad system view, as well as on a per-instance and per-volume view, to quickly and effectively resolve issues by analyzing diagnostics to understand the root cause.

You can configure the system events to be sent to a remote syslogs/logstash server located at some external endpoint.

2.8.2 Alerts

The cloud's automated monitoring helps you focus on proactively solving problems, rather than wasting time to identify them. You can set alarms on all relevant metrics and reduce the response time for facilitating automation. You can easily customize a set of metrics-based alarms with multiple conditions to notify you on threshold crossing. You will receive your alerts in real-time.

2.9 Certificates

2.9.1 Certificates

To secure your cloud domain name you need to procure an SSL Certificate from a trusted SSL Certificate Authority and install it on your domain. Once you get the certificate file from the CA, this service allows you to apply the certificate to the system.

2.10 Protection

2.10.1 Groups (policies)

The Protection Group basically consists of a schedule for automatic local backup of a project. It is a backup policy that defines the backup window, recurring schedule, and its retention period.

In addition to the member services above, the following services are available for Tenant Admins:

2.11 Identity and Access

2.11.1 Account

Tenant Admins have the added ability to monitor and modify the account state, create and manage projects within the account, and member user's access rights.

2.11.2 Access Keys

Tenant Admins can manage account-specific credentials that authenticate users before they can access the system. For example, to access the GUI or Management console, users sign in with a username and password. Similarly, for programmatic access, users provide an access key and a secret key. As the API's are AWS compatible, the authentication mechanism is also identical.

2.11.3 AWS Roles

AWS Roles are policy-based tokens with temporary credentials allowing a user temporary access to AWS services and actions which the user is normally not permitted to access. These users may be from different projects or even different accounts. These roles can also be embedded into specific instances allowing these instances access to the necessary actions.

The AWS roles are independent of the Zadara user roles.

2.11.4 AWS API Policies

Usage of all AWS compatible services and actions are governed by their corresponding AWS-managed policies. These policies can be assigned per project to Users, Groups of users and STS Roles. System usage is governed by Zadara Cloud Services policies together with the Zadara roles.

2.11.5 Zadara Cloud Services API Policies

Note

This section is under development.

CHAPTER

THREE

MAPPING ZADARA CLOUD SERVICES TO AWS SERVICES

Zadara Cloud Services deliver a software-defined data center platform that enables true IaaS, PaaS and CaaS in data centers and edge locations. Zadara Cloud Services runs on any hardware and is combined with cloud management features such as centralized user access management, self-service portals, integrated metering for showback / chargeback, and more.

In addition, Zadara Cloud Services delivers a suite of managed open source platforms for developers to accelerate application development and delivery. By offering AWS compatible APIs, Zadara Cloud Services enables multi-cloud and hybrid applications, and supports advanced DevOps and Infrastructure-as-Code in enterprise environments.

The following table provides a high-level mapping of Zadara Cloud Services products and AWS services.

3.1 IaaS Services

SERVICE	AWS	ZADARA CLOUD SERVICES	ZADARA-CS-SUPPORTED AWS APIs
Compute	Amazon Elastic Compute Cloud (EC2)	Zadara Compute Service (zCompute)	EC2
	Amazon EC2 Auto Scaling	Zadara Auto Scaling	EC2 Auto Scaling
Virtual Networks	Amazon Virtual Private Cloud (VPC)	Zadara Networking	VPC
Load Balancer	Elastic Load Balancer (ELB)	Zadara Load Balancers	ELB
Object Storage	Amazon Simple Storage Service (S3)	Zadara VPSA Object Storage	S3
Block Storage	Amazon Elastic Block Storage (EBS)	Zadara VPSA Storage Array	EBS
File Storage	Amazon Elastic File System (EFS)	Zadara VPSA Storage Array	

3.2 Management Services

SERVICE	AWS	ZADARA CLOUD SERVICES	ZADARA-CS-SUPPORTED AWS APIs
Monitoring	Amazon CloudWatch	Zadara Monitoring Service	CloudWatch
Security, Identity & Compliance	Amazon Identity & Access Management (IAM)	Zadara Identity Management	IAM
	AWS Security Token Service (STS)		STS
	AWS Certificate Manager (ACM)	Zadara Certificate Manager	ACM
Application & Service Catalog	AWS Service Catalog	Zadara CS Application Catalog	

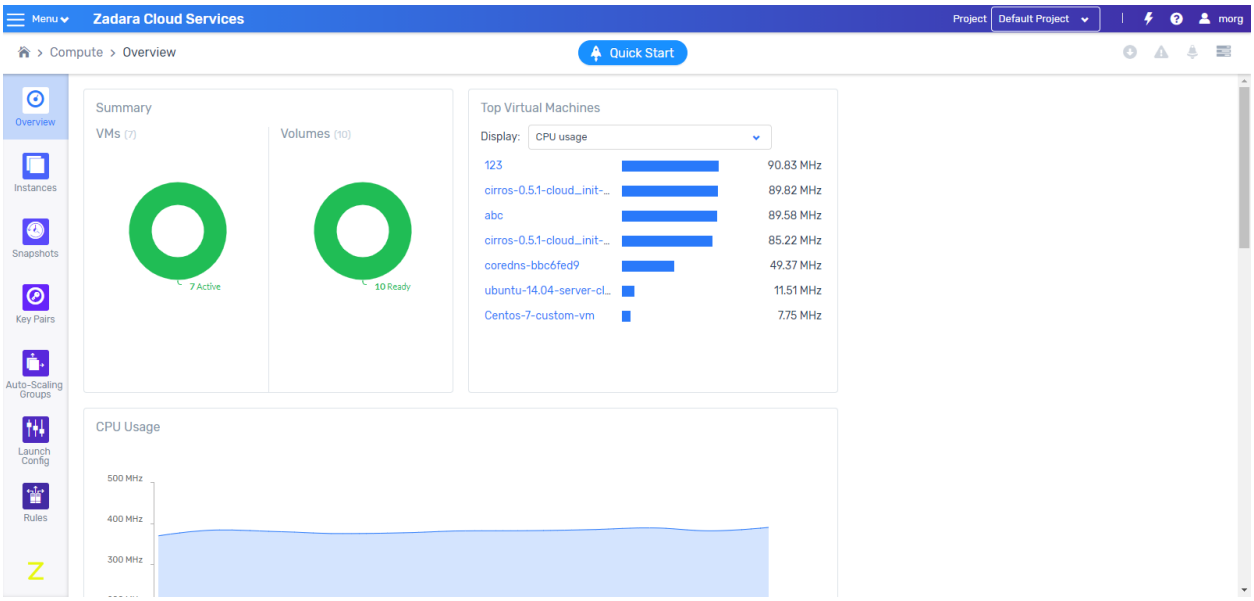
CHAPTER

FOUR

COMPUTE PROJECT OVERVIEW

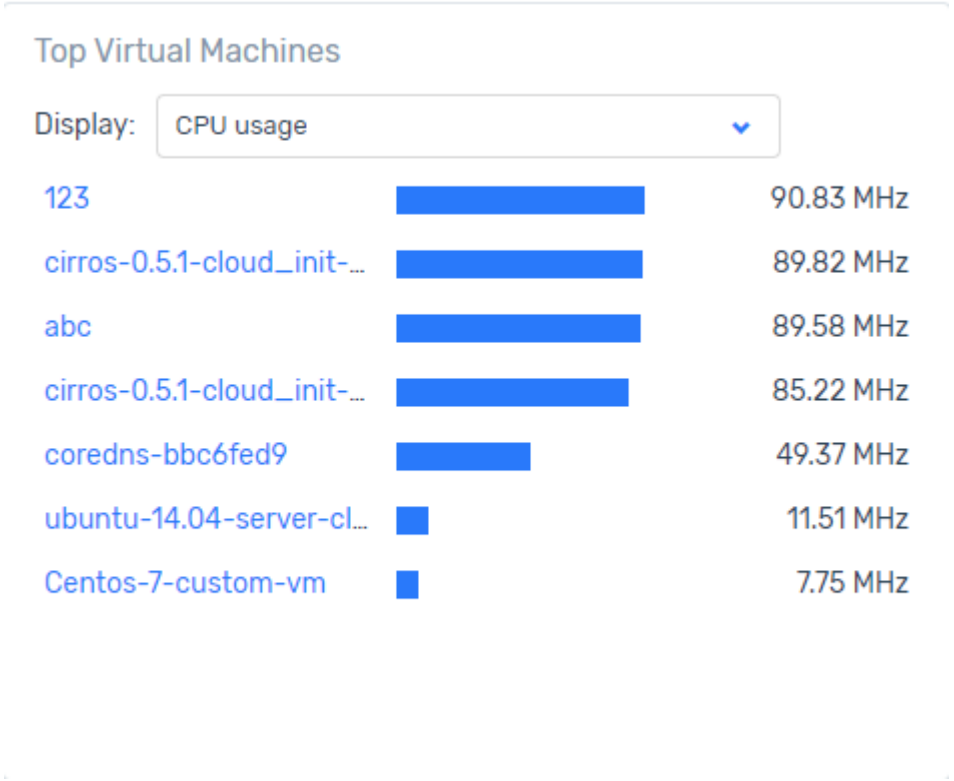
Zadara Cloud Services enables creating, configuring, running, accessing, and managing Virtual Machines (VMs).

The **Compute Project Overview** provides a number of widgets and graphs which summarize key resource usage statistics of VM instances used in your project.

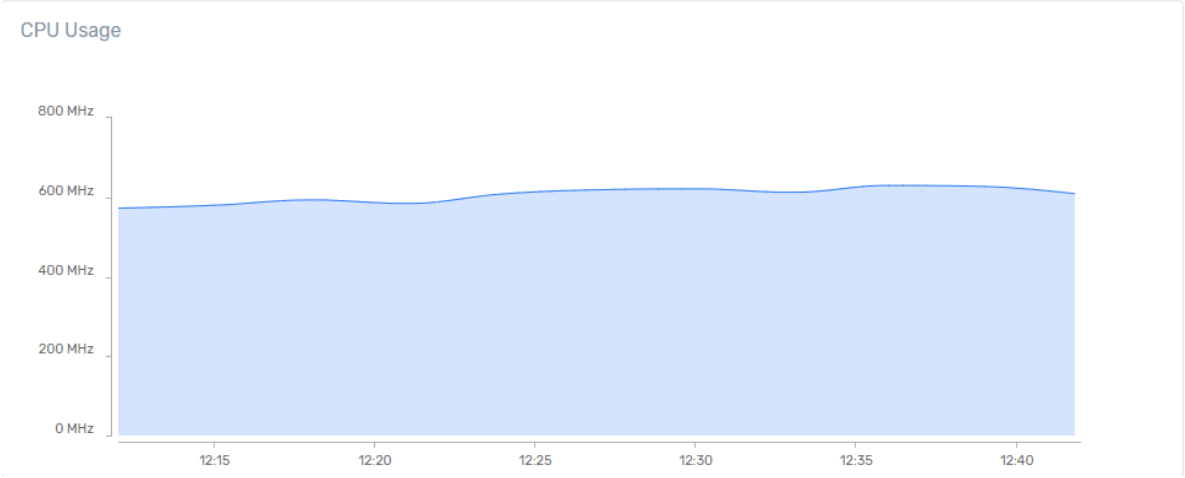


1. **Summary - VMs** - provides the number of VM instances used, including a count per VM instance state (running, active, error, shutoff).
2. **Summary - Volumes** - provides the number of volumes used and their status.

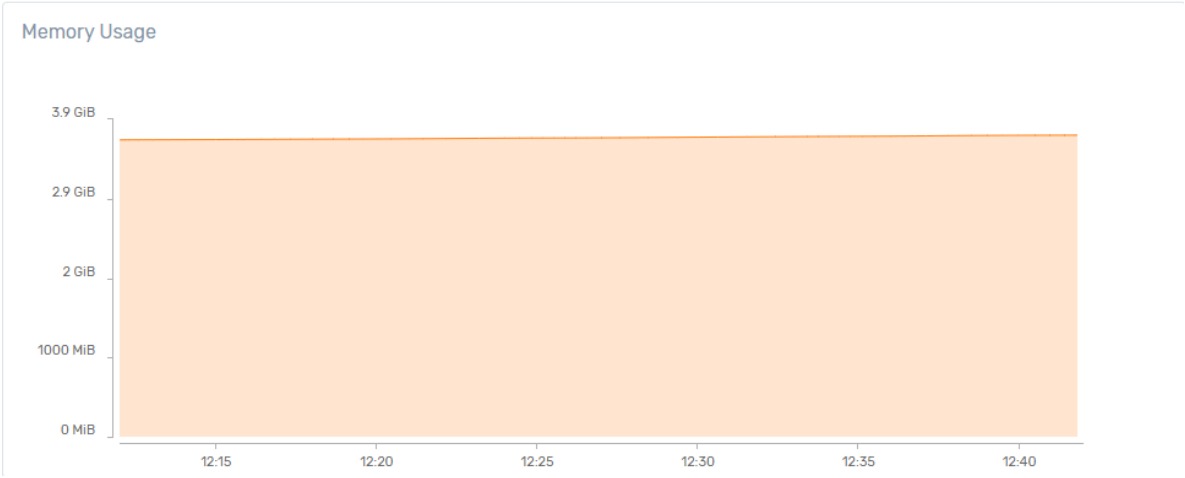
3. **Top Virtual Machines** - displays the top 10 VM instances based on resource usage: CPU Usage, Memory Usage. Network usage (receive), Network usage (send).



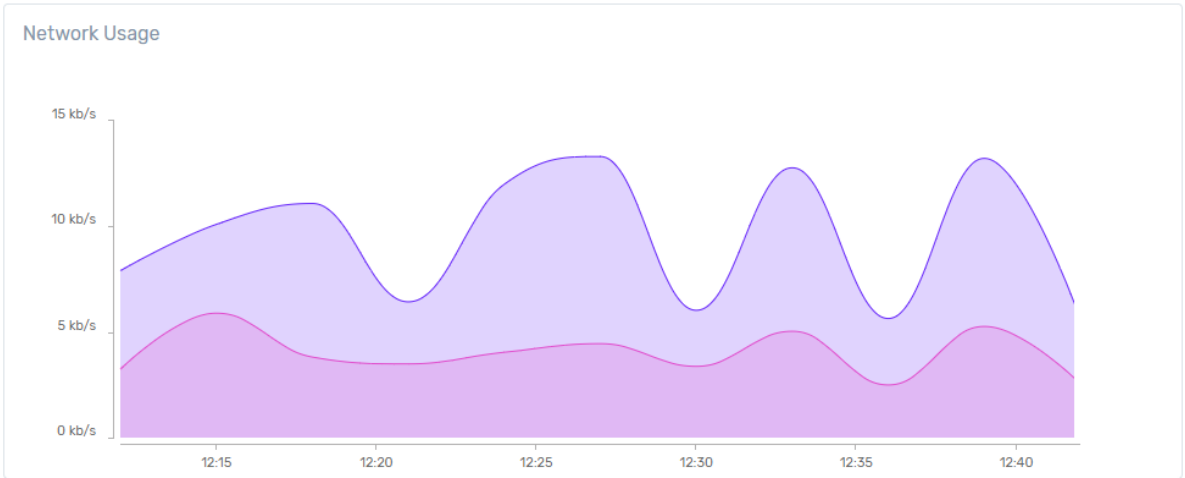
4. **CPU Usage** - displays the total CPU usage of all VM instances in your project.



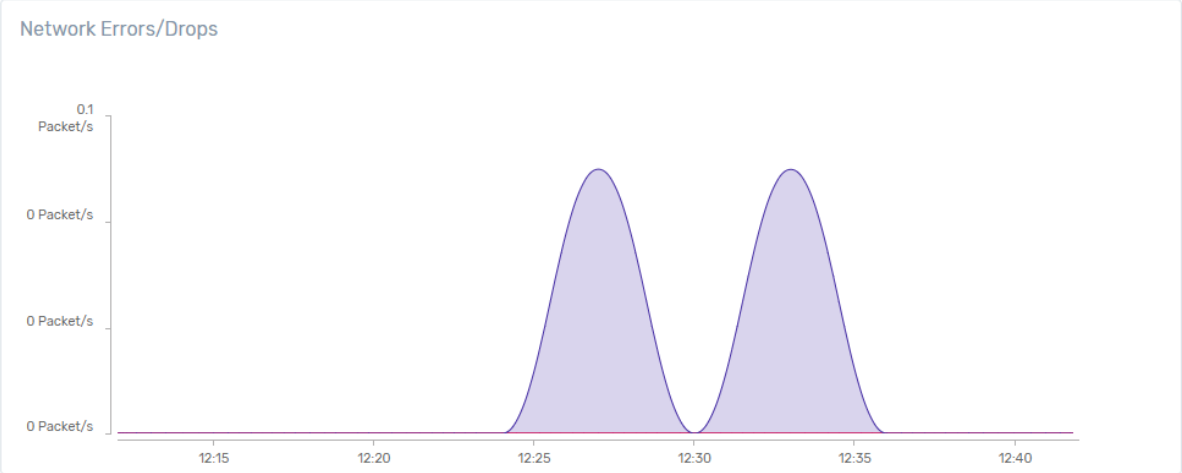
5. **Memory Usage** - displays the total memory usage of all VM instances in your project.



6. **Network Usage** - displays the total network transmit and receive rates.



7. **Network Drop** - displays the total network packet drops.


Network Errors/Drops
0.1 Packet/s
0 Packet/s
0 Packet/s
0 Packet/s
12:15
12:20
12:25
12:30
12:35
12:40

CHAPTER

FIVE

GLOSSARY

Affinity Rules

Affinity Rules allow you to control the placement of virtual machines on hosts within a cluster by using affinity rules.

API

Application program interface (API) is a set of routines, protocols, and tools for building software applications. An API specifies how software components should interact and APIs are used when programming graphical user interface (GUI) components. API makes it easier to develop a program by providing all the building blocks. A programmer then puts the blocks together.

AWS

Amazon Web Services (AWS) is a secure Cloud services platform, offering compute power, database storage, content delivery and other functionality to businesses.

Cloud

Cloud computing means storing and accessing data and programs over the Internet instead of your computer's hard drive. The cloud is just a metaphor for the Internet.

Cloud Computing

A computing capability that provides an abstraction between the computing resource and its underlying technical architecture (e.g., servers, storage, networks), enabling convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction." This definition states that clouds have five essential characteristics: on-demand self-service, broad network access, resource pooling, rapid elasticity, and measured service. Narrowly speaking, cloud computing is client-server computing that abstract the details of the server away;one requests a service (resource), not a specific server (machine). Cloud Computing enables Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS). Cloud computing means that infrastructure, applications, and business processes can be delivered to you as a service, over the Internet (or your own network).

Clusters

A cluster is a collection of nodes. The zCompute system runs over a cluster and uses it as a logical unit to provide:

- Node Management – the cluster is used for suspending and removing nodes, as well as monitoring the node status and activity.
- High Availability – The cluster is responsible for automatically detecting hardware failures, and for handling them by moving workloads from failed nodes to active ones. By performing High Availability procedures, zCompute assures that no data is lost and that the overall cluster operation remains intact.
- Cluster health and operational status – the cluster has pre-defined events and alarms that reflects the state of the system. The events provide information about monitored actions, such as the creation and stopping of a workload or a disk failure, and they include a severity level. Alarms are notifications that are activated in response to an event, or in response to a state of a certain object in the system. They indicate a potentially major issue in the system.

Data Center

A data center is a centralized repository, either physical or virtual, for the storage, management, and dissemination of data and information organized around a particular body of knowledge or pertaining to a particular business.

Infrastructure as a service (IaaS)

Cloud infrastructure services or "Infrastructure as a Service (IaaS)" delivers computer infrastructure, typically a platform virtualization environment, as a service. Rather than purchasing servers, software, data center space or network equipment, clients instead buy those resources as a fully outsourced service. The service is typically billed on a utility computing basis and amount of resources consumed (and therefore the cost) will typically reflect the level of activity. It is an evolution of web hosting and virtual private server offerings.

Interfaces

You can communicate with your environment by using one of the three existing zCompute interfaces: REST API - the main programmable interface into the system, the CLI, or the web-based graphical User Interface.

Hyperconverged

Hyperconverged is a type of infrastructure system with a software-centric architecture that tightly integrates compute, storage, networking and virtualization resources and other technologies from scratch in a commodity hardware box supported by a single vendor.

Hypervisor

A hypervisor or virtual machine monitor (VMM) is a piece of computer software, firmware or hardware that creates and runs virtual machines.

A computer on which a hypervisor is running one or more virtual machines is defined as a host machine.

Each virtual machine is called a guest machine.

Monitoring Agent

A Monitoring Agent exists on each node, and it communicates with the Monitoring Service. The Agent provides the Monitoring Service with information, such as how much CPU is being used, the amount of allocated and used disk space on a given node, and more. Based on this information, the Monitoring Service is able to keep track of the system status across all nodes, and initiate changes when resources are running low.

Multi Tenancy

Multi Tenancy refers to a software architecture in which a single instance of software runs on a server and serves multiple tenants. A tenant is a group of users who share a common access with specific privileges to the software instance.

Nodes

Nodes are the physical machines that provide the infrastructure for storage, compute, and memory, and on which zCompute is installed. The minimal zCompute configuration requires 3 nodes. Nodes can be added, removed, or put in a maintenance mode after the initial installation and configuration. Virtual Machines are running on the nodes of the cluster. The placement of the Virtual Machines on the nodes is determined by a zCompute management sub-system, which also migrates Virtual Machines from one node to another according to the changing state of the Virtual Machines, nodes, and the cluster as a whole.

Platform as a service (PaaS)

Cloud platform services, whereby the computing platform (operating system and associated services) is delivered as a service over the Internet by the provider.

The PaaS layer offers black-box services with which developers can build applications on top of the compute infrastructure. This might include developer tools that are offered as a service to build services, or data access and database services, or billing services.

Private Cloud

Private cloud is a type of cloud computing that delivers similar advantages to public cloud, including scalability and self-service, but through a proprietary architecture.

Unlike public clouds, which deliver services to multiple organizations, a private cloud is dedicated to a single organization.

Provisioning Services

zCompute Provisioning Services are responsible for:

- Placement – determines where to place resources.
- Monitoring – monitors node resources and workload demands.
- Live Migration – migrates workloads between nodes according to the node loads, the resources required by the workloads, and the priority and profile of the workloads.

Self Service

A feature that allows customers to provision, manage, and terminate services themselves, without involving the service provider, via a Web interface or programmatic calls to service APIs.

Server Sprawl

Server sprawl is a situation in which multiple, under-utilized servers take up more space and consume more resources than can be justified by their workload.

Software as a service (Saas)

Cloud application services, whereby applications are delivered over the Internet by the provider, so that the applications don't have to be purchased, installed, and run on the customer's computers. SaaS providers were previously referred to as ASP (application service providers). In the SaaS layer, the service provider hosts the software so you don't need to install it, manage it, or buy hardware for it. All you have to do is connect and use it. SaaS Examples include customer relationship management as a service.

Software Defined Everything

With Software Defined Everything, the computing infrastructure is virtualized and delivered as a service. In a Software-Defined Everything environment, management and control of the networking, storage and/or data center infrastructure is automated by intelligent software rather than by the hardware components of the infrastructure. its all done with API calls.

Storage

zCompute supports multiple storage systems. By default, the zCompute system comes with a proprietary storage solution.

In addition, you can also attach external storage systems that will serve as additional storage pools.

Tenant

A Tenant is a group of users that has access to certain logical resources in a cluster. The logical resources that a user can access, are the resources that were created by all of the users who belong to the same tenant. Users in the same tenant can have different roles.

Each role defines the type of privileges the user will have. Every user in the system must belong to at least one tenant, and each tenant can include one or more users.

Vendor lock

Dependency on the particular cloud vendor and difficulty moving from one cloud vendor to another due to lack of standardized protocols, APIs, data structures (schema), and service models.

Workloads

Workloads refer to Virtual Machines (VMs).

CHAPTER

SIX

ZCOMPUTE HOME PAGE

The zCompute UI home page is the main entry point for platform administration. It brings the main management areas together in one place for MSP administrators.

Within the home page, **Region Management** is the main workspace for day-to-day administration. It contains the region-level areas used to monitor the platform and work with core infrastructure resources.

The **Region Management** section organizes these functions into tiles. Each tile groups related options so admins can move directly to the correct management area.

6.1 Monitoring

The **Monitoring** tile groups pages used to review system state, activity, and service health.

- **Overview**

  The Overview dashboard shows high-level monitoring data for the cluster.

  Review it to check current status, alerts, events, resource consumption, and health summaries.

- **Events**

  An event is a record of a system activity or state change.

  Review it to inspect event history, apply filters, and export event data.

- **Diagnostics**

  Diagnostics are system validator checks that run on the platform.

  Review past runs or start an immediate run to inspect validator results.

- **Data Collection**

  Data collection creates a downloadable set of logs or a cluster dump.

  Use it to collect support data, download the result, or remove old collections.

- **System Alerts**

  A system alert is an automated notification or action triggered by a threshold or event in the cluster infrastructure.

  Review it to see alert summaries and the current state of each alert.

- **Alarms**

  An alarm is a configurable monitor that changes state when a rule threshold is met.

  Use it to view alarms and to create, modify, delete, or manually set alarm state.

- **SNS Topics**

  An SNS topic is a notification target with one or more email subscribers.

  Use it to create topics, update topic details, and manage subscribers.

- **System**

  The System screen shows the real-time health and status of platform services.

  Review it to inspect service tiles, filter service views, and check service events.

6.2 Configuration

The **Configuration** tile groups core region configuration tasks and hardware-related administration.

- **Nodes**

  A node is a physical server that provides compute, storage, and networking resources in the region.

  Use this page to view nodes and run supported lifecycle actions such as add, join, maintain, activate, remove, and inspect.

- **Upgrades**

  An upgrade item tracks software upgrade work for the region.

  Review it to inspect upgrade records and current upgrade state.

- **External Endpoints**

  An external endpoint is a configured connection to an external target, such as object storage.

  Use it to view endpoint records and create or modify endpoint settings.

- **Settings**

  Settings control global system behavior, security, branding, networking, and external integrations.

  Update them to manage platform-wide configuration values from a central area.

- **Cluster Certificates**

  A cluster certificate is used for server-side authentication by the cluster, allowing secure, HTTPS-encrypted operation with the region's UI and APIs.

  Use this page to upload a certificate and private key.

- **PCI Devices**

  A PCI device is hardware that can be exposed to virtual machines through PCI passthrough.

  Use this page to review device state and enable or disable devices for passthrough use.

6.3 Region Networking

The **Region Networking** tile groups network resources that are managed at the region level.

- **Switch Domains**

  A switch domain defines a Layer 2 network domain in Region Networking.

  Use it to review switch domains, inspect node links, and work with cluster networks and VLANs in that domain.

- **Cluster Networks**

  A cluster network defines Layer 2 and Layer 3 connectivity within a switch domain.

Use it to view or update network details and work with routes, applications, VIPs, and IP addresses.

- **Networking Applications**

  Networking applications define logical networking functions within a cluster network.

  Use them to review application details or create and delete application entries.

- **VIPs**

  A VIP is a virtual IP resource attached to a cluster network.

  Use this page to review VIP details and update supported VIP settings.

- **GPU Network Switches**

  A GPU network switch is a switch record used for the GPU networking layer.

  Use this page to add, modify, or delete switch records and review their connection settings.

- **GPU Network Ports**

  A GPU network port is a port record in the GPU networking layer.

  Use this page to review port identity, addressing, and linked resources.

- **Edge Networks**

  An edge network brings edge-network configuration, subnet values, IP pools, and router details together.

  Use this page to view, create, modify, configure, or delete edge network records.

6.4 Account Networking

The **Account Networking** tile groups network resources that are assigned to accounts.

- **VLANs Management**

  A VLAN is a Layer 2 identifier used to segment network traffic.

  Use this page to review VLAN records and manage VLAN allocation and ownership.

- **Direct Subnets**

  A Direct Subnet connects a VPC directly to the hosting-datacenter's physical network over a designated VLAN ID. Direct Subnets allow "stretching" (directly connecting) VPCs into the datacenter.

  Use this page to create subnets, run subnet operations, and test subnet connectivity.

- **GPU Networks Management**

  Modern AI workloads demand a dedicated, high-performance network fabric that connects GPUs within and across compute nodes, separate from general-purpose infrastructure networks.

  The **NVIDIA Spectrum-X Compute East-West (E-W) network** provides a purpose-built Ethernet fabric optimized for GPU-to-GPU communication at cloud scale.

  From version 25.06, **zCompute GPU Network (GPU-Net)** exposes this fabric as a first-class, self-service resource. Tenants can provision and manage GPU networks through zCompute's standard interfaces. Tenants can have multiple GPU networks, and can allocate these networks across multiple projects that provide project isolation and tenant isolation.

  Use this page to add, assign, release, delete, or force delete GPU network entries.

6.5 Storage Management

The **Storage Management** tile groups region storage resources and storage configuration tasks.

- **Service Controller**

  The service controller dashboard shows storage summaries and VSC events for the cloud.

  Review it to inspect storage counts, set event filters, and open related events.

- **Storage Classes**

  A storage class defines the underlying hardware and performance expectations used to store data.

  Use this page to review storage classes, inspect linked resources, and attach a VPSA when needed.

- **Volume Types**

  A volume type defines the performance profile and storage behavior of block storage volumes.

  Use this page to view, create, modify, delete, enable, disable, or set the default volume type.

- **Backup / Restore Tasks**

  A backup or restore task is a task record associated with backup protection group activity.

  Use this page to monitor currently running backup and restore work and open related resource records.

6.6 Service Engines

The **Service Engines** tile groups engine categories used by region services.

- **Networking**

  A networking service engine is a deployed engine used for regional networking services.

  Use this page to review engine entries and manage supported engine state actions.

- **Load Balancer**

  A load balancer service engine is a deployed engine used for load balancing services in the region.

  Use this page to review engine entries and manage supported engine state actions.

6.7 Identity & Access

The **Identity & Access** tile groups account, policy, and role management tasks.

- **Accounts**

  An account is an administrative container for projects, users, and related resources.

  Use this page to review account information and perform basic account configuration.

- **Symp API Policies**

  A Zadara Cloud Services policy is a rule set that grants access to Zadara Cloud Services functionality.

  Use this page to review and work with platform-specific policy definitions.

- **AWS API Policies**

  An AWS API policy governs the use of supported AWS services and actions.

  Use this page to view managed policy definitions and review their assignments.

- **AWS Roles**

  An AWS IAM role is a policy-based token with temporary credentials for AWS services and actions.

  Use this page to review role definitions and their trust and permission settings.

6.8 Consoles

The **Consoles** tile groups command-line console access options.

- **SYMP**

  The Symp Console provides access to the zCompute Symp CLI from within the UI.

  Open it to run Symp commands and manage zCompute services from the shell.

- **AWS**

  The AWS Console provides access to an AWS CLI shell from within the UI.

  Open it to run supported AWS commands, such as EC2 and S3 operations.

CHAPTER

SEVEN

FIRST STEPS IN ZCOMPUTE

7.1 zCompute UI

The zCompute web application interface is tested and supported in Google Chrome.

7.1.1 zCompute UI Language support

By default, the zCompute UI display language is English.

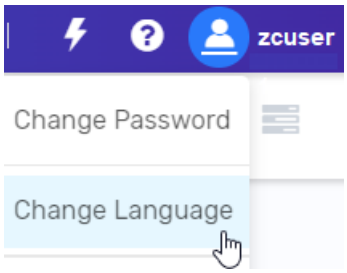
From version 23.08.1, the zCompute UI supports the following languages:

- English (default)
- German
- Japanese
- Korean
- Portuguese
- Spanish

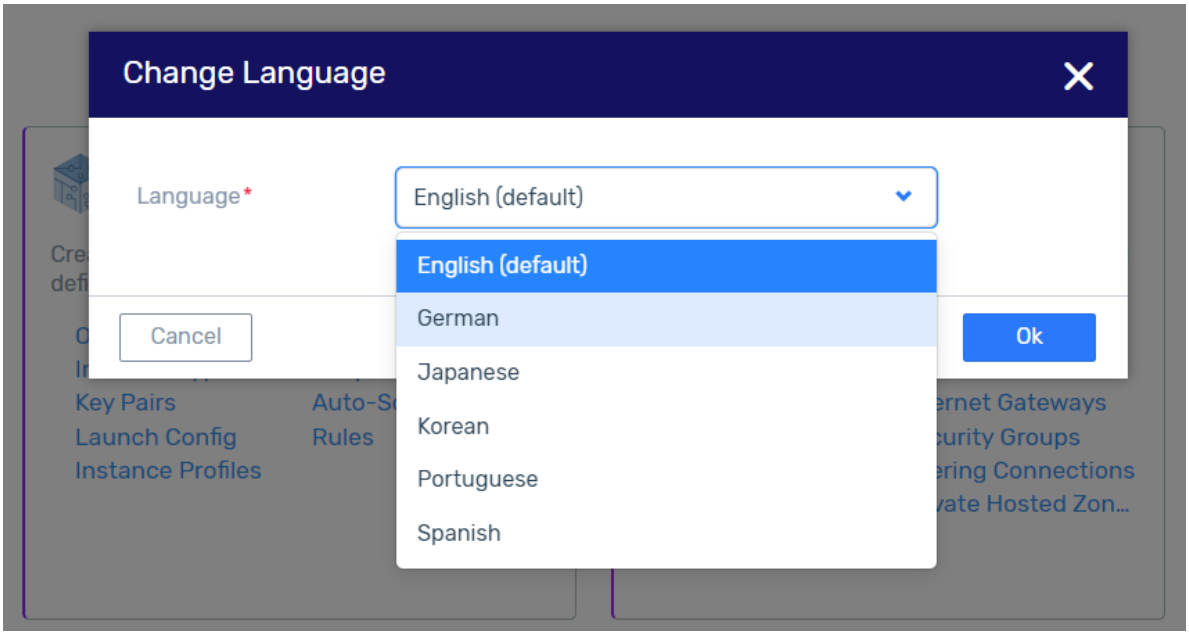
To change the UI display language:

1. Click the user icon at the top right.

   In the dropdown, select **Change Language**.

   The **Change Language** dialog opens.



2. In the **Change Language** dialog, expand the **Language** dropdown, scroll and select your preferred display language.
3. Click **OK**.

   The zCompute UI display will appear in the selected language.

Note

This configuration is session-based and will revert to the system default upon a browser page reload.

Administrators can update the system default at Configuration Settings > General tab > Language.

7.1.2 zCompute UI Navigation

The zCompute UI provides a rich user experience for cloud compute management.

See zCompute's main user resources and interaction features in the zCompute UI Navigation video:

7.2 Quick Start - Basic

7.2.1 Upload Your Own Image

Images provide the information required to launch an instance. A single image can be used to launch multiple instances with the same configuration. You can create multiple images to allow various instance configurations. There are multiple ways to creates an image:

- File Upload
- URL
- Using a volume or a snapshot of existing instance

Learn more about Machine Images in the zCompute User Guide.

See the video on the basics of zCompute Images and ISOs:

7.2.2 Create a Compute Instance

An instance is a virtual machine hosted on Zadara Cloud Services. An instance is launched using a copy of image or volume of your choice, which provides an initial configuration for that instance. There are multiple instance types of which you can choose when you launch an instance. An instance type determines the used resources required by your instance. Instance's type can also be changed after launch.

Learn more about Creating VM Instances in the zCompute User Guide.

See the video demonstrating the basics of creating and management of zCompute Instances:

7.2.3 Add a Data Disk to Your VM

Volumes are block devices which you can mount as devices on your instances, and persist independently from the life of the instance. You can create a file system on top of these volumes, or use them in any way you would use a block device. Volumes can also be used to launch new instances.

Learn more about Adding Storage to VM Instances in the zCompute User Guide.

See the zCompute Volumes, Snapshots and Protection Groups overview video:

7.2.4 Make Your Instance Accessible

By default, your instances do not have a public IP address, which makes them unreachable from the internet. Elastic IP is a public IP which is accessible to the internet. Elastic IPs can be attached and detached to your instance as you choose, to allow internet access.

Learn more about Associating Security Groups with VM Instances in the zCompute User Guide.

See the video demonstrating the basics of creating and configuring zCompute Security Groups and Source/Destination checks:

7.3 Quick Start - Advanced

7.3.1 Add Load Balancer

Load Balancer allows you to automatically distributes incoming traffic across multiple targets, such as instances, or IP addresses. There are two types of load balancers:

- **Application Load Balancer (ALB)** - Distributes HTTP or HTTPS traffic using application layer protocol based routing.
- **Network Load Balancer (NLB)** - Distributes TCP traffic regardless of the application layer protocol.

Learn more about zCompute's Load Balancing capabilities.

See the video demonstrating the basics of creating and configuring zCompute Load Balancers, Target Groups and Listeners:

7.3.2 Create Your Own Virtual Private Cloud

Virtual Private Cloud allows you to create an isolated section of the Zadara Cloud Services where you can launch resources in a virtual network that you define. VPC offers you a complete control of you network environment, including selection of an IP address range of your choice, creation of subnets, and configuration of route tables, network gateways and security groups.

Learn more about VPCs and zCompute networking capabilities.

See the video demonstrating the basics of creating and configuring zCompute VPCs:

CHAPTER

EIGHT

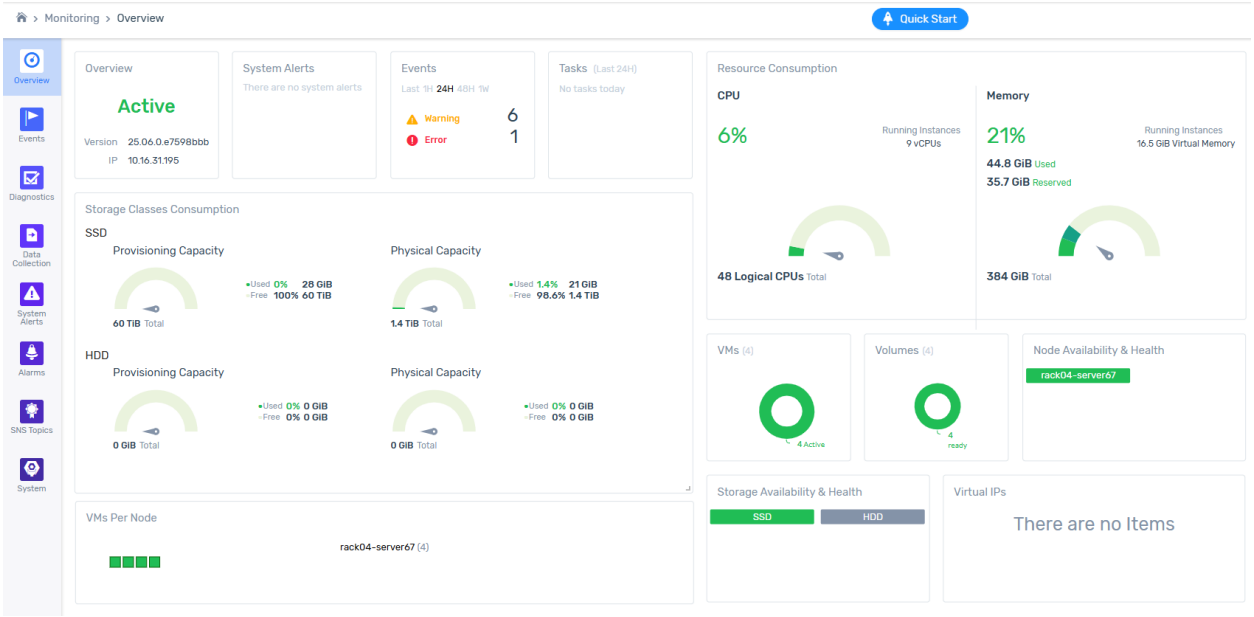
ZCOMPUTE MONITORING

8.1 Introduction

zCompute provides monitoring tools that help administrators track and manage system and workload activities.

8.2 Monitoring Overview dashboard

Each zCompute cluster has a main **Monitoring > Overview** dashboard, that displays several tiles.



- **Overview**

  Displays the cluster's basic information:

  - The status of the cluster.
  - The zCompute software version operating the cluster.
  - The cluster's IP address.

- **System Alerts**

  Indicates if there are any current system alerts.

  In the event of a current alert, hovering the mouse over the reported alert provides a clickable link to navigate directly to the **System Alerts** screen, for further drilldown to details.

- **Events**

  Displays the number of events that occurred in the most recent time period, grouped by severity. The time period in the dashboard's **Events** tile can be adjusted to reflect the past:

  - Hour
  - 24 hours
  - 48 hours
  - Week

  Hovering the mouse over a severity group of events provides a clickable link to navigate directly to the **Events** screen listing the events, filtered according to the selected severity level and time period.

- **Tasks**

  Displays tasks of the past 24 hours.

- **Resource Consumption**

  Display system resource consumption for the cluster:

  - **CPU**

    Shows average CPU usage of this host and the total vCPUs of running VMs that are scheduled on the underlying Logical CPUs (hardware threads).

    * **CPU %**: The percentage of current CPU utilization of all logical CPUs on the cluster.
    * **Running Instances**: The sum of virtual CPUs (vCPUs) allocated to all VM instances currently active on the cluster.
    * **Logical CPUs**: The sum of hardware execution threads that exist on the physical cluster.

  - **Memory**

    Shows average memory usage of this cluster and how the total virtual RAM of running VMs consumes the underlying physical memory capacity.

    * **Memory %**:

      Overall memory utilization of the cluster, showing how much of the physical RAM is currently in use or reserved.

    * **Used**:

      Amount of physical RAM actively consumed by the cluster and currently running VMs.

    * **Reserved**:

      Physical RAM that has been committed to VMs and system services but might not be fully used yet. It is set aside and unavailable for other allocations.

    * **Running Instances (Virtual Memory)**:

      Sum of RAM configured for all VMs currently running on this cluster (total of their virtual memory sizes).

    * **Total**:

      The total physical RAM installed on the cluster.

    Clicking anywhere in the Memory tile opens a Memory Breakdown popup chart displaying total used and available memory capacity by:

    * Used by System
    * Reserved for System
    * Used by VMs

        * Rerserved for VMs
        * Free for VMs

- **Storage Classes Consumption**

  Shows the cluster's disk allocation capacities, in use and free:

  - **SSD**
    * **Provisioned Capacity**:

      The total, used and free amounts of SSD storage space assigned to the cluster, as the host OS sees as its drive size, regardless of whether the capacity physically exists.

    * **Physical Capacity**:

      The total, used and free amounts of actual SSD storage space available to the cluster. In many configurations, physical capacity is likely to be significantly lower than the provisioned capacity.

  - **HDD**
    * **Provisioned Capacity**:

      The total, used and free amounts of HDD storage space assigned to the cluster, as the host OS sees as its drive size, regardless of whether the capacity physically exists.

    * **Physical Capacity**:

      The total, used and free amounts of actual HDD storage space available to the cluster. In many configurations, physical capacity is likely to be significantly lower than the provisioned capacity.

- **VMs**

  Shows the total of the cluster's VMs, and a chart grouping totals of VMs by status: active and shutoff.

  Clicking on the one of the chart's VM status groups (active or shutoff) navigates directly to the **Compute > Instances** screen, displaying the table of VMs filtered according to the selected VM status.

- **Volumes**

  Shows the total of the cluster's volumes, and a chart grouping totals of volumes by their readiness status.

  Clicking on the one of the chart's volume readiness status groups navigates directly to the **Storage > Block Storage** screen, displaying the table of volumes filtered according to the selected volume status.

- **Node Availibilty & Health**

  Shows the total of the cluster's nodes, and lists the nodes individually.

  Clicking on one of the nodes opens a popup displaying:

  - Link that navigates directly to the node's details screen: **Configuration > Nodes > <node name>**.
  - **Status**: The node's health status.
  - **Uptime**: Time since the node's last reboot.
  - **Memory**: The node's total, used and reserved physical RAM.
  - **CPU**: The node's total of logical CPUs, and current percentage usage.

- **Storage Availibilty & Health**

  Shows the combined total of SSD and HDD storage volumes. Color-coded entries indicate the health status of each of the SSD and HDD storage class groups.

  Clicking on a storage class navigates to the selected **Storage Management > Storage Classes > SSD** or **HDD** details screen.

- **Virtual IPs**

  A virtual IP (VIP) exists only in software and can move between different physical machines, ensuring connectability to a service using a consistent address irrespective of the underlying infrastructure.

  Zadara Operations team provisions clusters with preconfigured VIPs.

- **VMs Per Node**

  Each node in the cluster is listed with icons representing each of its VMs, color-coded to indicate the VM's status.

  Hovering over a VM icon displays the VM's name and status.

  Clicking a VM navigates directly to the selected VM's **Compute > Instances > <VM>** details screen.

8.3 Monitoring Diagnostics

Diagnostics run on system validators at hourly intervals.

For the latest diagnostics run, and each previous run, the **Monitoring > Diagnostics** screen displays a Diagnostics Results summary row with the following columns:

- The number of system validators checked.
- A color-coded bar indicating the status.
- A **Details** link that navigates to the selected diagnostic run's **Periodic Validators** details screen, which lists the name and status of each system validator.
- Timestamp of the diagnostics run.
- Completion status of the diagnostics run.
- In the event of a user-invoked diagnostics run, the username appears in the User column, otherwise the column entry remains blank for a regular scheduled system diagnostics run.

8.3.1 Invoking a diagnostics run

To invoke an immediate non-scheduled diagnostics run:

1. Go to **Monitoring > Diagnostics**.
2. In the top toolbar click **Run**.

   A **Run Diagnostics** dialog opens.

   Enter the parameters:

   1. **Type**:

      Select one of:

      - **Default**: Diagnostics run on all system validators.
      - **Custom**: The dialog expands, listing each system validator separately, with a checkbox to mark to include it in the custom diagnostics run.

   2. **Decription**: Optional description that is displayed in the diagnostic run's **Periodic Validators** details screen.
   3. Click **Ok** to launch the immediate diagnostics run.

8.4 Data Collection

Zadara Operations and Support may request data for analysis of performance, issues or other symptoms.

To download the requested data:

1. Go to **Monitoring > Data Collection**.

   Existing data collection downloads for the cluster appear in a table.

2. In the top toolbar click **+ Create**.

   The **Collect Data for Download** dialog opens.

   Enter the parameters:

   1. **Data Type**:

      Select one of the **Data Type** options:

      - **Logs** (default):

        Enter the following parameters:

        1. **Log Files**: Enter the file pattern as a string that comprises the file names, or matches files that contain the string.

           ✔ **Note**

           Wildcards are not supported.

           The entered file pattern string is displayed in the data collection run's entry on the **Monitoring > Data Collection** list.

        2. **Decription**: Optional description that is displayed in the data collection run's entry on the **Monitoring > Data Collection** list.

        3. **Nodes**:

           Select one of:

           - **All**: From all nodes in the cluster, collect files with names matching the entered file pattern.
           - **Only Selected**: From the dropdown, select the nodes in the cluster from which to collect files with names matching the entered file pattern.
           - **Excluded Selected**: From the dropdown, select the nodes to ignore, and from all other nodes in the cluster collect files with names matching the entered file pattern.

      - **Cluster Dump**:

        Select the **Cluster Dump** option to download a comprehensive extract of information about a cluster's state, including configurations, networking, storage, and VM metadata.

   2. Click **Collect Data** to create the selected data collection set for download.

      A new entry appears in the Data Collection table, displaying the filename pattern, progress status of the data collection process, creator username, and timestamp.

   3. After the data collection's status changes to **Done**, click its row in the table to display additional operations for that data collection.

      - To download the selected data collection to the local machine, on the upper tool bar click **Download**.

        The download file is a zipped tar with the namimg convention: `logs-dump-<YYYYMMDDHHmmss>.tar.gz`.

- To delete the selected data collection from the cluster, on the upper tool bar click **Delete**.

  To confirm deletion of the selected data collection, in the **Delete Data Collection** confirmation window, click **Delete**.

> **Tip**
>
> **Best practice**
>
> After downloading, delete the data collections to free up occupied space on the cluster.

8.5 System Alerts

A system alert is an automated notification or action triggered when specific thresholds are reached or events occur in the cluster's infrastructure.

To view a summary of system alerts, go to **Monitoring > System Alerts**.

The following information is available for each system alert:

- **Name**: A descriptive name for the alert.
- **Entity**: The specific resource name that triggered the alert, such as a cluster.
- **Entity Type**: The category of the resource that triggered the alert.
- **Status**: The current state of the alert. If the issue is resolved, the status is **Closed**.
- **Updated At**: The date and time when the alert status last changed.

8.6 Alarms

An alarm is a configuration that monitors data points relative to a set of rules. On reaching the threshold of a specified condition, an alarm's status changes, triggering reporting of an alarm event.

SNS topics can be used to notify relevant target audiences according to different alarm events.

Use the **Alarms** screen to view, monitor and manage alarms for your managed resources.

8.6.1 Viewing Alarms

To view alarms:

1. Go to **Monitoring > Alarms**.

   A list of alarms appears, displaying the following columns:

   - **Name**: The alarm name.
   - **Condition**: The metric calculation condition that triggers the alarm. The condition is configurable according to the entity type.
   - **Account**: The associated MSP account.
   - **Project**: The project that contains the monitored entity.
   - **Entity**: The monitored resource, that is a specific instance of the entity type.
   - **Entity Type**: The resource type, for example:
     - Autoscale Group

  - Cluster
  - Compute Instance
  - Node
  - Project
  - Volume
- **State**: The current alarm status, for example:
  - **OK** indicates the condition is not currently met.
  - **ALARM** indicates that a condition threshold is met.
  - **INSUFFICIENT DATA** typically indicates a manual status change, using the **Set State** option.
- **Created At**: The date and time the alarm was created.
- **Updated At**: The most recent alarm reporting update time.

View Alarm Details

1. Go to **Monitoring > Alarms**.

   A list of alarms appears.

2. Click an alarm name to view its details.

   The selected alarm's detail screen appears with two tabs:

   - **Overview** tab:

     In addition to the columns displayed in the **Monitoring > Alarms** list (above):

     - **Trigger**
       - **Period**: Sampling window time frame in seconds.
       - **Statistic**: Calculation on the measured metric:
         - Average
         - Sum
         - Minimum
         - Maximum
       - **Details**: Additional information.
       - **ID**: The alarm's internal UUID.
     - **Topics**

       SNS topics are used to communicate alarm status changes to configured audiences:

       - On alarm active
       - On alarm off
       - On insufficient data
   - **Events** tab:

     Displays a list of Events, filtered to display alarm events triggered when the alarm status changed.

The toolbar provides options to:

- **Create** a new alarm.

- **Modify** the selected alarm.
- **Set State** manually.
- **Delete** the selected alarm.

8.6.2 Creating an Alarm

To create an alarm:

1. Go to **Monitoring > Alarms**.

   A list of alarms appears.

2. On the top toolbar, click **+ Create**.

   The **Create Alarm** dialog opens.

   1. Enter the alarm's details:

      - **Name**: The alarm name.
      - **Description**: A brief description for the alarm.

      Trigger Conditions

      The parameters of the conditions that trigger the alarm:

      - **Type**: The resource type, for example:
        - Autoscale Group
        - Cluster
        - Compute Instance
        - Node
        - Project
        - Volume

        Based on the selected resource type, the next prompt provides a dropdown list of the selected resource type entities.

        Select the entity instance to monitor for the alarm.
      - **Metric**: Based on the selected resource type, from the dropdown select the metric that will be monitored.
      - **Statistic**: From the dropdown, select the calculation method on the selected metric:
        - Average
        - Sum
        - Minimum
        - Maximum
      - **Threshold**:
        - Select the operator (**>=, >, <, <=**) for the comparison calculation.
        - Enter the threshold value for the selected metric.

        The alarm is triggered when the metric statistic, compared to the threshold value using the selected operator, evaluates to true.
      - **Sampling Window**: Enter the duration of the statistic sampling window in seconds.

2. Click **Next** to go to the **Notifications** tab.
3. In the **Notifications** tab, from the dropdown select the *SNS Topics* or click **+** for *Creating an SNS Topic*.

   This configuration determines the notification distribution for each change in the alarm's **Status**.

   Repeat the dropdown selection step to add multiple SNS topics to any of the alarm's status change notification distributions:

   - **On alarm active**:

     To send out a notification to the relevant distribution list when the alarm is triggered, and its **State** transitions to **ALARM**.

     This indicates indicates that the alarm's condition threshold is met.

   - **On alarm off**:

     To send out a notification to the relevant distribution list when the alarm is deactivated, and its **State** transitions to **OK**.

     This indicates indicates that the alarm's condition threshold is no longer met.

   - **On insufficient data**:

     To send out a notification to the relevant distribution list when the alarm is triggered, and its **State** transitions to **INSUFFICIENT DATA**.

     This typically indicates a manual status change for the alarm, where the **Set State** option is used.

4. To confirm completion of the alarm's configuration and to save it, click **Finish**.

   The **Create Alarm** dialog closes, and the alarm row appears in the **Monitoring > Alarms** list.

8.6.3 Modifying an Alarm

To change the configuration of an existing alarm:

1. Go to **Monitoring > Alarms**.

   A list of alarms appears.

2. Click the row or the alarm name to select it.
3. In the top toolbar, click **Modify**.

   The **Modify Alarm** dialog opens.

   1. Optionally, update any of the alarm's details:

      - **Name**: The alarm name.
      - **Description**: A brief description for the alarm.

      Trigger Conditions

      The parameters of the conditions that trigger the alarm:

      - **Type**: The resource type, for example:
        - Autoscale Group
        - Cluster
        - Compute Instance
        - Node
        - Project

  - Volume

  Based on the selected resource type, the next prompt provides a dropdown list of the selected resource type entities.

  Select the entity instance to monitor for the alarm.

- **Metric**: Based on the selected resource type, from the dropdown select the metric that will be monitored.
- **Statistic**: From the dropdown, select the calculation method on the selected metric:
  - Average
  - Sum
  - Minimum
  - Maximum
- **Threshold**:
  - Select the operator (**>=, >, <, <=**) for the comparison calculation.
  - Enter the threshold value for the selected metric.

  The alarm is triggered when the metric statistic, compared to the threshold value using the selected operator, evaluates to true.
- **Sampling Window**: Enter the duration of the statistic sampling window in seconds.

2. Click **Next** to go to the **Notifications** tab.
3. In the **Notifications** tab, from the dropdown select the *SNS Topics* or click **+** for *Creating an SNS Topic*.

   This configuration determines the notification distribution for each change in the alarm's **Status**.

   Repeat the dropdown selection step to add multiple SNS topics to any of the alarm's status change notification distributions:

   - **On alarm active**:

     To send out a notification to the relevant distribution list when the alarm is triggered, and its **State** transitions to **ALARM**.

     This indicates indicates that the alarm's condition threshold is met.
   - **On alarm off**:

     To send out a notification to the relevant distribution list when the alarm is deactivated, and its **State** transitions to **OK**.

     This indicates indicates that the alarm's condition threshold is no longer met.
   - **On insufficient data**:

     To send out a notification to the relevant distribution list when the alarm is triggered, and its **State** transitions to **INSUFFICIENT DATA**.

     This typically indicates a manual status change for the alarm, where the **Set State** option is used.
4. To confirm completion of the alarm's configuration and to save it, click **Finish**.

8.6.4 Set Alarm State

It is possible to manually change the alarm's state.

Reasons could include:

- An alarm threshold condition was reached but the threshold measurement is back to normal levels some time before the automatic “OK” notification is distributed. In such a case there might be a decision to manually reset the state and trigger a notification that can lower the support team’s alert level.
- Although an alarm condition is not met, other symptoms might be a cause for concern. Manually changing the alarm state will trigger a notification that will raise the support team awareness level.
- Immediately after an alarm notification is triggered due to meeting a threshold condition, the team would likely analyze the status. Analysis might determine that although the threshold has been reached and might not return to OK for a while, the scenario is under control, and manually changing the status will trigger another notification that can reduce the urgency level.

To manually set an alarm’s state:

1. Go to **Monitoring > Alarms**.

   A list of alarms appears.

2. Click the row or the alarm name to select it.
3. In the top toolbar, click **Set State**.

   The **Set Alarm State** dialog opens.

4. In the **Set Alarm State** dialog:

   1. **State**:

      From the dropdown, select the state:

      - **OK** to indicate the threshold condition is not currently met.
      - **ALARM** to indicate that the threshold condition threshold is met.
      - **INSUFFICIENT DATA** to indicate something other than **OK** or **ALARM**.

   2. **Reason**:

      Enter an explanation for manually setting the alarm’s state.

5. To confirm the alarm state change, click **OK**.

8.6.5 Deleting an Alarm

Deleting an alarm deactivates it and removes its configuration.

However, its events are still available for viewing in the Events Log for the duration of event retention. For further information, see *Monitoring zCompute Events*.

To delete an alarm:

1. Go to **Monitoring > Alarms**.

   A list of alarms appears.

2. Click the row or the alarm name to select it.
3. In the top toolbar, click **Delete**.

   The **Delete Alarm** dialog opens, displaying the alarm name and ID.

4. To confirm deletion of the selected alarm, in the **Delete Alarm** dialog click **Delete**.

   The **Delete Alarm** dialog closes, and the alarm row disappears from the **Monitoring > Alarms** list.

8.7 SNS Topics

SNS topics are used to communicate notifications to relevant audiences.

8.7.1 Viewing SNS Topics

To view SNS topics:

1. Go to **Monitoring > SNS Topics**.

   A list of SNS topics appears, displaying the following columns:

   - **Name**: The SNS topic's name.
   - **Subscriptions**: The number of email addresses subscribing to this SNS topic.
   - **User** The admin username who created the SNS topic.
   - **Account**: The account in which the SNS topic was created.

View SNS Topic Details

1. Go to **Monitoring > SNS Topics**.

   A list of SNS topics appears.

2. Click an SNS topic name to view its details.

   The selected SNS topic's detail screen appears:

   - **Overview** section
     - **Name**: The SNS topic's name.
     - **User**: The admin username who created the SNS topic.
     - **Project**: The project for which the SNS topic was created.
     - **Account**: The account in which the SNS topic was created.
     - **Subscriptions size**: The number of email addresses subscribing to this SNS topic.
     - **ID**: The SNS topic's internal ID.
   - **Subscribers** section

     Lists the email addresses of subscribers to this SNS topic.

8.7.2 Creating an SNS Topic

To create an SNS topic:

1. Go to **Monitoring > SNS Topics**.

   A list of SNS topics appears.

2. On the top toolbar, click **+ Create**.

   The **Create Topic** dialog opens.

   1. Enter the topic's details:
      - **Name**: The topic's name.

- **Subscribers**: The email addresses that will receive this topic's notification messages.

  To add a second or more email addresses, after entering the first email address press enter. The email address appears on a greyed background with an x control to delete it. Add a new email address next to it and press enter.

3. To save the topic, click **OK**.

8.7.3 Modifying an SNS Topic

There are two separate aspects for modifying an SNS topic:

- *Modifying an SNS Topic Name*
- *Adding or removing SNS Topic Subscribers*

Modifying an SNS Topic Name

1. Go to **Monitoring > SNS Topic**.

   A list of SNS topics appears.

2. Click the row or the SNS topic name to select it.
3. In the top toolbar, click **Modify**.

   The **Modify Alarm** dialog opens.

4. Enter the new name for the topic.
5. To save the SNS topic name change, click **OK**.

Adding or removing SNS Topic Subscribers

A subscriber is an email address that is subscribed to an SNS topic to receive the topic's email notification messages.

1. Go to **Monitoring > SNS Topic**.

   A list of SNS topics appears.

2. Click the SNS topic name to select it.

   The SNS topic's detail screen opens.

   The **Subscribers** list of email addresses appears in the lower half of the screen.

   - **Adding a new subscriber**:

     1. On the **Subscribers** toolbar, click **+ Create**.

        The **Subscriptions** dialog opens.

     2. Enter the subscriber's email address.

        To add a second or more email addresses, after entering the first email address press enter. The email address appears on a greyed background with an x control to delete it. Add a new email address next to it and press enter.

     3. To save the new subscriber email addresses, click **OK**.

        - The new email addresses appear in the SNS topic's **Subscribers** list.
        - The **Subscriptions size** in the **Overview** section displays the updated total of subscriber email addresses.

   - **Removing an existing subscriber (unsubscribing)**:

1. To select one or more subscriber email addresses to remove, mark their email address check box.

   To remove all subscriber email address for an SNS topic, click the top **Address** checkbox, and all checkboxes will be marked.

   To deselect marked email addresses, click the relevant marked checkbox, or the the top **Address** checkbox to select/deselect all email addresses.

   If there are any marked subscriber email addresses, the **Delete** options appears in the **Subscribers** toolbar.

2. On the **Subscribers** toolbar, click **Delete**.

   The **Delete Subscription** dialog opens, listing the subscriber email addresses that are marked for deletion.

   Caution

   The **Delete** action will remove **ALL** of the **marked** subscriber email addresses, and does not prompt for each one separately.

3. To confirm deletion of the subscriber email addresses listed in the **Delete Subscription** dialog, click **Delete**.
   - The marked email addresses no longer appear in the SNS topic's **Subscribers** list.
   - The **Subscriptions size** in the **Overview** section displays the updated total of subscriber email addresses.

8.7.4 Deleting an SNS Topic

Deleting an SNS topic removes the topic and its subscribers.

To delete an SNS topic:

1. Go to **Monitoring > SNS Topic**.

   A list of SNS topics appears.

2. Click the row or the SNS topic name to select it.

3. In the top toolbar, click **Delete**.

   The **Delete Topic** dialog opens, displaying the topic name and ID.

   Caution

   If the topic is connected to any alarms, those alarms are listed in the the **Delete Topic** dialog.

   Before proceeding with the SNS topic deletion, it is recommended to check the alarm and verify that the deletion of the topic should proceed.

   Connected alarms' **Notification** tabs should be updated. See *Modifying an Alarm*.

4. To confirm deletion of the selected topic, in the **Delete Topic** dialog click **Delete**.

   The **Delete Topic** dialog closes, and the alarm row disappears from the **Monitoring > SNS Topics** list.

8.8 System

The **Monitoring > System** screen to enables monitoring the real-time health and status of platform services. It provides a visual summary of service state and a detailed service events breakdown view.

8.8.1 Monitoring System screen layout

The **Monitoring > System** screen includes:

- A left control pane for filtering and display options
- A service overview panel (top)
- A Service events panel (bottom)

Monitoring System screen default view

By default:

- Services are grouped by **Group**.
- Tiles are colored by **Status**.
- Active services appear in green.

The overview displays service categories, such as:

- General
- API Services
- Region Management
- Compute Services
- Region Monitoring
- Cloud Network
- Cloud Compute
- Policy Engine

The number shown next to each category indicates the total number of services in that category.

Each small square tile represents a single service instance.

If all services are operating normally, tiles appear in the color associated with the **Active** or **Normal** state.

8.8.2 Monitoring System screen left control pane

The left pane includes controls to change how services are displayed:

- **Filter**

  Enter text to restrict the display to matching services.

- **Group by**

  Use **Group by** to organize services by none, one or more attributes:

  - **Group** (default)
  - Name
  - Status
  - Importance
  - Node

  When selecting multiple attributes, the system nests groupings in the order selected. For example:

  Selecting **Group, Status, Importance, Node** creates a hierarchical layout:

  - Service category
  - Status (for example, Active)
  - Importance (for example, Normal)
  - Node (<node name>)

  Selecting **Node** displays services grouped by host, which helps identify node-level issues.

  Selecting **Importance** highlights service priority within each group.

  Removing all group selections displays all services in a flat layout.

- **Color by**

  From the dropdown, select one of the **Color by** options to control how service tiles are colored:

  - **Status** (default)

    Tiles reflect the operational state of each service.

  - **Importance**

    Tiles reflect the configured priority level.

  - **CPU**

    Tiles reflect current CPU utilization.

    The legend displays utilization ranges.

    Hover over a tile to view the exact percentage.

  - **Memory**

    Tiles reflect current memory utilization.

    A legend displays memory usage ranges.

  Coloring by CPU or Memory helps you quickly identify resource pressure across services.

- **Reset All**

  Select **Reset All** to clear filters and restore the default view.

8.8.3 Service overview panel (top)

The top service overview main panel displays services according to the layout configured in the left control pane.

Hover over a service tile to display:

- Service name
- Current metric value (for example, CPU percentage)

This allows quick inspection without changing screens.

8.8.4 Service events panel

The lower **Service events** panel displays events for the selected time range.

Use the date and time selectors to define the reporting period.

Use the **Filter** box to search for events.

Select **More filters** to apply additional criteria.

If no events exist for the selected time range, the screen displays **There are no events**.

8.8.5 Best practices

Use this screen as part of regular operational monitoring.

- Use **Color by CPU** or **Memory** to detect resource saturation.
- Group by **Node** to isolate host-level problems.
- Group by **Importance** to focus on critical services.
- Review events after upgrades or configuration changes.

CHAPTER

NINE

MONITORING ZCOMPUTE EVENTS

9.1 Introduction

zCompute provides event monitoring tools that help both administrators and members keep track of system and workload activities. Monitoring events can be done through the Event Logs, APIs, and the Symp command line interface.

9.1.1 Event Retention

Event retention is not configurable in zCompute. The cloud-wide Event Log has fixed limits. Events are retained until one of the following limits is reached:

- **Duration**: 60 days

  When an event exceeds 60 days, it is automatically purged, irrespective of whether the maximum number of events has accumulated.

- **Accumulated events limit**: 250,000 events

  When the 250,000 events limit is reached, the oldest events are purged.

  If there are fewer than 250,000 events, old events are purged when they exceed 60 days.

9.2 Viewing and Filtering Event Logs

The Event Log keeps a searchable and exportable history of compute events. It is useful for troubleshooting and compliance audits.

9.2.1 zCompute UI

To view events:

1. Go to **Monitoring > Events**.

   An initial table of the most recent events displays.

2. Click an event row to display all of its details in the lower details pane.
3. To export events to a `.csv` file, click the speadsheet export icon at the top right.
4. To filter the Events table:
   - **Date and time range**
   - **Severity level** dropdown (Critical, Error, Warning, Info)
   - Click **More filters** to display and select additional filters from the following dropdowns:
     - **Severity**

- **Entity Type**
- **Event Name**
- **Account**
- **Project**
- **User**
- **Time Of Day** (range)

9.2.2 Events API

The **events** API can be used to programatically retrieve selected events according to parameters, such as the filters available in the zCompute UI's **Monitoring > Events** page.

zCompute's built-in API Explorer provides usage documentation on zCompute APIs, such as the **events** API.

9.2.3 Symp Command Line

The Symp CLI runs inside the zCompute Toolbox VM.

Use the Symp `event` command to view event details.

Example: Query recent events

```
symp event query \
  --severity ERROR WARNING \
  --limit 20
```

Example: Count events by type

```
symp event count \
  --severity ERROR WARNING \
  --group-by event-type
```

For more information, see the Symp CLI Client and the event command options in the Symp CLI Reference Guide in the zCompute User Guide.

9.3 Best Practices

- **Admins** should review dashboards and export logs, and set up notifications.
- **Members** should monitor dashboards, filter logs for troubleshooting, and escalate recurring issues to admins.

CHAPTER

TEN

NODE MANAGEMENT

10.1 Introduction

Nodes are the physical servers that provide compute, storage, and networking resources in your cloud region. As an MSP administrator, you manage the full lifecycle of these nodes through the **Region Management > Nodes** page.

A node passes through several states during its lifecycle:

10.1.1 Status lifecycle

- **Candidate**

  A newly discovered server that has not yet joined the cluster.

- **Approved**

  A candidate that was accepted but has not finished activation.

- **Starting**

  The node is booting and initializing services.

- **Activating**

  The node is completing its activation process.

- **Active**

  The node is online and running workloads.

- **Failed**

  The node encountered an error and is not functioning correctly.

- **Fenced**

  The node was isolated from the cluster to protect data integrity.

- **Entering maintenance**

  Workloads are migrating off the node in preparation for maintenance.

- **In maintenance**

  The node is offline for servicing. No workloads run on it.

- **Unfencing**

  The node is rejoining the cluster after a fencing event.

- **Inadequate**

  The node has unsupported hardware and cannot join the cluster.

- **Rejected**

  An administrator declined this candidate node.

- **Out**

  The node has been removed from the cluster.

10.2 The Nodes page

The Nodes page shows all nodes and node candidates in a single table.

You can switch between two views:

10.2.1 List view

A table with sortable columns:

- **Name**

  The display name of the node.

- **Candidate**

  Indicates whether the node is still a candidate and has not yet joined the cluster.

- **IP**

  The management (access) IP address of the node.

- **CPU**

  The current CPU usage of the node.

- **vCPUs**

  The total number of virtual CPU cores available on the node.

- **Memory**

  The total physical memory installed on the node.

- **VM Memory (Total)**

  The total amount of memory allocated to virtual machines on this node.

- **VM Memory (Available)**

  The remaining memory available for additional virtual machines.

- **Network Rx**

  The amount of inbound network traffic received by the node.

- **Network Tx**

  The amount of outbound network traffic transmitted by the node.

- **Storage**

  The total storage capacity available on the node.

- **Uptime**

  The amount of time the node has been running since its last restart.

- **Status**

  The current lifecycle state of the node (for example, Active, In maintenance, or Failed).

- **Tags**

  Custom labels assigned to the node for identification and filtering.

CPU, Memory, and Network columns include sparkline charts that show recent usage trends.

10.2.2 Heatmap view

A visual map of nodes grouped and color-coded by resource or status.

You can group nodes by:

- **Status**

  Groups nodes according to their lifecycle state, such as **Active** or **In maintenance**.

- **CPU**

  Groups nodes based on CPU capacity or usage.

- **Memory**

  Groups nodes based on memory capacity or usage.

- **Storage**

  Groups nodes based on storage capacity.

You can color the heatmap by:

- **Status**

  Applies colors based on lifecycle state.

- **Used Memory**

  Colors nodes according to the amount of memory currently used.

- **Memory**

  Colors nodes based on total memory capacity.

- **Used CPU**

  Colors nodes according to current CPU utilization.

- **CPU**

  Colors nodes based on total CPU capacity.

- **Storage**

  Colors nodes based on storage capacity.

- **Throughput**

  Colors nodes according to network activity levels.

10.2.3 Built-in filters

Use filters to narrow the table to specific node states.

- **Active**

  Shows only nodes that are currently online and running workloads.

- **Candidate**

  Shows only nodes that were discovered but have not yet joined the cluster.

- **Inadequate**

  Shows only nodes that do not meet hardware requirements.

- **In Maintenance**

  Shows only nodes that are currently in maintenance mode.

10.3 Node candidates management

A node candidate is a physical server that has booted, completed hardware discovery, and appeared in the system but has not yet joined the cluster.

Before a candidate can run workloads, you must add it, review its hardware, and join it to the cluster.

10.3.1 Add a candidate node

Add a candidate node when you have a new physical server that you want to register with the cluster. This operation tells the system the server's network address and credentials so it can connect and inspect the hardware.

Prerequisites:

- The physical server is powered on and network-reachable.
- The server's access IP address, SSH username, and SSH password.
- The MAC addresses of the server's converged network interfaces.

Steps:

1. Go to **Region Management > Nodes**.
2. Select **+ Add Candidate** in the top toolbar.
3. In the **Add Candidate Node** dialog, complete the **Info** step:
   - **Candidate Access IP**: Enter the server's IPv4 address.
   - **SSH User**: Enter the SSH username.
   - **SSH Password**: Enter the SSH password.
   - **Node Type**: From the dropdown, select the node type:
     - **Worker**: Runs tenant workloads.
     - **Control**: Runs cluster management services.
4. Select **Next** to continue to the **MAC Addresses** step.
5. Enter one or more **Converged MAC Addresses**.

   These identify the network interfaces the node will use for converged traffic.

   Select **Add** to enter another **Converged MAC Address**.
6. To confirm adding the candidate node, select **Finish**.

   The system displays "Candidate node creation is in progress."

   When complete, the system displays "Candidate node added successfully."

10.3.2 Modify a candidate node

Modify a candidate node when you need to update SSH credentials or change the converged MAC addresses. You cannot change the access IP address after the candidate is created.

Steps:

1. Go to **Region Management > Nodes**.
2. In the nodes table, locate the candidate node.
3. In the top toolbar, select **Modify Candidate**.
4. In the **Modify Candidate Node** dialog, update:
   - SSH User and SSH Password.
   - Converged MAC Addresses (add or remove).

   Candidate Access IP and Access MAC Addresses are read-only.
5. Select **Save**.

   The system displays “Candidate node modification is in progress.”

   When complete, the system displays “Candidate node modified successfully.”

10.3.3 Join a candidate node to the cluster

Join a candidate node when you are satisfied with its hardware configuration and want it to begin serving workloads. During this step, you choose the node’s role in the cluster.

Steps:

1. Go to **Region Management > Nodes**.
2. In the nodes table, locate the candidate node.
3. In the top toolbar, select **Join**.
4. In the **Join Candidate Node** dialog, review:
   - **IP**: Displays the node’s access IP (read-only).
   - **Node Link Details**:

     Expand sections to review interface details including Name, Address (MAC), Vendor, Duplex, Speed, Driver, Device Name, Carrier status (up/down), PCI Slot, Type, and Device ID.
5. In the Node Type dropdown, select a role:
   - **Control**: Runs cluster management services.
   - **Worker**: Runs tenant workloads.
6. Select **Join** and confirm.

   The system displays “Candidate node join is in progress.”

   When complete, the system displays “Candidate node joined successfully.”

10.3.4 Delete a candidate node

Delete a candidate node when the server is no longer needed or was added in error. This removes the candidate record from the system. It does not affect the physical server.

Steps:

1. Go to **Region Management > Nodes**.
2. Locate the candidate node.
3. In the top toolbar, select **Delete Candidate**.
4. In the **Delete Node** dialog, review the node's IP address.
5. To confirm deletion, select **Delete**.

   The system displays "Candidate node deleted successfully."

10.4 Node management activities

After a node joins the cluster and becomes active, you manage it through a different set of operations.

10.4.1 Accept (approve) a candidate node

When a node appears with **Candidate** status, you can accept it to allow it to proceed toward activation.

Steps:

1. Go to **Region Management > Nodes**.
2. Locate the node with **Candidate** status.
3. In the top toolbar, select **Accept**.

The node status changes to **Approved**, then proceeds through **Starting** and **Activating** automatically.

10.4.2 Reject a candidate node

Reject a candidate node when you do not want it to join the cluster. This marks the node as rejected without deleting its record.

Steps:

1. Go to **Region Management > Nodes.**
2. Locate the node with **Candidate** status.
3. Open the actions menu and select **Reject**.

The node status changes to **Rejected**.

10.4.3 Rename a node

Rename a node to give it a descriptive name or update its description. This does not affect the node's hostname or functionality.

Steps:

1. Go to **Region Management > Nodes**.
2. Select the node name to open the detail page, or open the actions menu and select **Rename**.
3. In the **Rename Node** dialog, enter:
   - **Name**: Required.
   - **Description**: Optional.
4. Select **OK**.

10.4.4 Put a node into maintenance mode

Put a node into maintenance mode when you need to perform hardware repairs, firmware updates, or other servicing. The system migrates running workloads off the node before it enters maintenance.

Prerequisites:

- The cluster has more than one node.
- No active VMs with GPU passthrough devices are running.
- Remaining nodes have enough memory for migrating VMs.

Steps:

1. Go to **Region Management > Nodes**.
2. Locate a node with **Active**, **Failed**, or **Fenced** status.
3. Open the actions menu and select **Maintenance**.
4. In the **Node Maintenance** dialog, review the node's **ID** and **Name**.
5. Select **OK** to confirm.

The status changes to **Entering maintenance** while VMs migrate, then to **In maintenance** when complete.

10.4.5 Force maintenance

Use force maintenance only as a last resort. It bypasses the normal workload migration process.

Severe warning: Forcing a node into maintenance mode prevents the system from completing some housekeeping actions and may result in full system failure and data loss.

Steps:

1. In the **Node Maintenance** dialog, select **Force Maintenance**.
2. Read the warning carefully.
3. Wait for the safety timer to complete.
4. Confirm the action.

The system displays: "Node [name] is being forcefully deactivated."

10.4.6 Activate a node (exit maintenance mode)

Activate a node to return it to service after maintenance is complete.

Steps:

1. Go to **Region Management > Nodes**.
2. Locate the node with **In maintenance** status.
3. Open the actions menu and select **Activate**.

The node returns to **Active** status.

10.4.7 Remove a node from the cluster

Remove a node when you want to permanently take it out of the cluster. The node must be in maintenance mode before removal.

Prerequisites:

- The node is in **In maintenance** status.

- The cluster has more than one node.

Steps:

1. Go to Region Management > Nodes.
2. Locate the node.
3. Open the actions menu and select Remove.
4. Review the node's IP address.
5. Select Remove.

Caution

If the node has unhealthy (degraded) data pools, removal may result in data loss. Contact support before proceeding.

10.4.8 Configure a node (open web installer)

Use Configure to open the node's built-in installer interface in a new browser tab. This is useful for initial hardware setup or advanced configuration.

Steps:

1. Go to **Region Management > Nodes**.
2. Locate the node.
3. Open the actions menu and select **Configure**.

A new tab opens to: `http://<node-ip>/installer`.

10.5 Manage node tags

Tags help you organize and identify nodes by purpose, location, hardware type, or other categories.

Add a tag:

1. Go to **Region Management > Nodes**.
2. Select the node name to open the detail page.
3. Select **Add tag**.
4. Enter the tag name and confirm.

Remove a tag:

1. On the node detail page, locate the tag.
2. Select the remove action on the tag.

10.6 Create an alarm for a node

You can set up CloudWatch-style alarms to monitor node metrics and receive notifications when thresholds are exceeded.

Steps:

1. Go to **Region Management > Nodes**.
2. Locate the node.

3. Open the actions menu and select **Create Alarm**.
4. Complete the alarm creation form. See *Creating an Alarm*.

10.7 View node details

Select any node name in the table to open its detail page.

Tabs:

- **Overview** tab

  Displays the node name, description, status, IP address, uptime, memory, disk capacity, and CPU details.

- **Events** tab

  Shows a chronological log of events related to this node.

- **Node Links** tab

  Lists network interfaces with MAC address, admin state, operational state, MTU, vendor, model, and bond type. Expand rows to view cluster network interfaces with IP address, VLAN ID, and state.

- **Monitoring** tab

  Provides performance charts for CPU, memory, and network usage over time.

- **Disks** tab

  Displays physical disks attached to the node.

- **VMs** tab

  Lists virtual machines currently running on the node.

- **Services** tab

  Shows cluster services running on the node, including CPU and memory usage per service.

- **PCI Devices** tab

  Lists PCI passthrough devices including device name, PCI slot, vendor, device ID, enabled or blocked status, kernel driver, and associated VM if assigned.

10.8 Recommended best practices

10.8.1 Cluster sizing

- Maintain at least two nodes in every cluster.
- Keep enough spare memory so any single node can enter maintenance without capacity issues.

10.8.2 Before maintenance

- Check the node's VMs tab for running workloads.
- Shut down or migrate GPU-attached VMs before maintenance.
- Verify sufficient free memory on remaining nodes.

10.8.3 Monitoring

- Use the **Heatmap** view to spot imbalances.
- Review sparkline charts for sustained trends.
- Set up alarms for critical metrics.
- Monitor partition free space indicators such as `/var/log`, `/mnt/data`, and `/mnt/containers`.

10.8.4 Tags

- Create a consistent tagging strategy.
- Use tags to help filter and identify nodes quickly.

10.8.5 Node removal

- Always put a node into maintenance mode before removal.
- If degraded data pools are detected, contact support.

10.8.6 Network verification

- Verify converged MAC addresses when adding candidates.
- Review the **Node Links** tab after a node joins.

10.9 Troubleshooting

10.9.1 A candidate node does not appear in the table

Possible causes:

- Hardware discovery not complete.
- Server is not network-reachable.
- Table filter excludes Candidate view.

10.9.2 “Failed to add candidate node” error

Possible causes:

- Candidate Access IP is incorrect or unreachable.
- SSH User or SSH Password credentials are wrong.
- SSH service is not running.

Resolution: Verify IP and credentials. Confirm SSH connectivity.

10.9.3 A node shows “Inadequate” status

Hardware does not meet cluster requirements. Review specifications on the Overview tab.

10.9.4 Maintenance mode is blocked by GPU VMs

Shut down affected GPU VMs before retrying.

10.9.5 Maintenance mode is blocked by insufficient memory

Shut down or migrate VMs to free cluster memory.

10.9.6 A node shows “Failed” status

Hover over Status to see the failure reason.

If recoverable: Put into maintenance, fix the issue, then activate.

If not recoverable: Put into maintenance and remove from cluster.

10.9.7 A node shows “Fenced” status

The node was isolated to protect data integrity.

Possible next steps:

- Wait for recovery.
- If not recovering, use maintenance mode and investigate.

10.9.8 “Failed to join candidate node” error

Possible causes:

- Network links not in expected state.
- Selected Node Type is invalid for cluster configuration.

10.9.9 Node removal shows a “data loss” warning

Do not proceed. Contact support to assess data pool health.

10.9.10 The “Remove” and “Maintenance” actions are not available

- **Remove** requires **In maintenance** status and more than one node.
- **Maintenance** requires **Active**, **Failed**, or **Fenced** status and more than one node.
- Single-node clusters do not support these actions.

CHAPTER

ELEVEN

EXTERNAL ENDPOINTS AND B2OS

In zCompute, external endpoints are primarily used to connect your virtual environment to resources outside the local cloud, particularly for backing up local snapshots to remote object storage. They enable data protection by defining destination storage, such as remote Object Storage containers, to securely transfer data away from the zCompute origin.

11.1 Backup to Object Storage (B2OS)

zCompute Backup to Object Storage (B2OS) extends backup and restore capabilities beyond local block storage. It enables backing up and restoring VMs and volumes that are protected by protection-group to and from Zadara Object Storage systems.

These Zadara Object Storage systems can also reside in different physical locations than the source zCompute cloud, allowing recovery to any zCompute cloud in the event of a site-level failure.

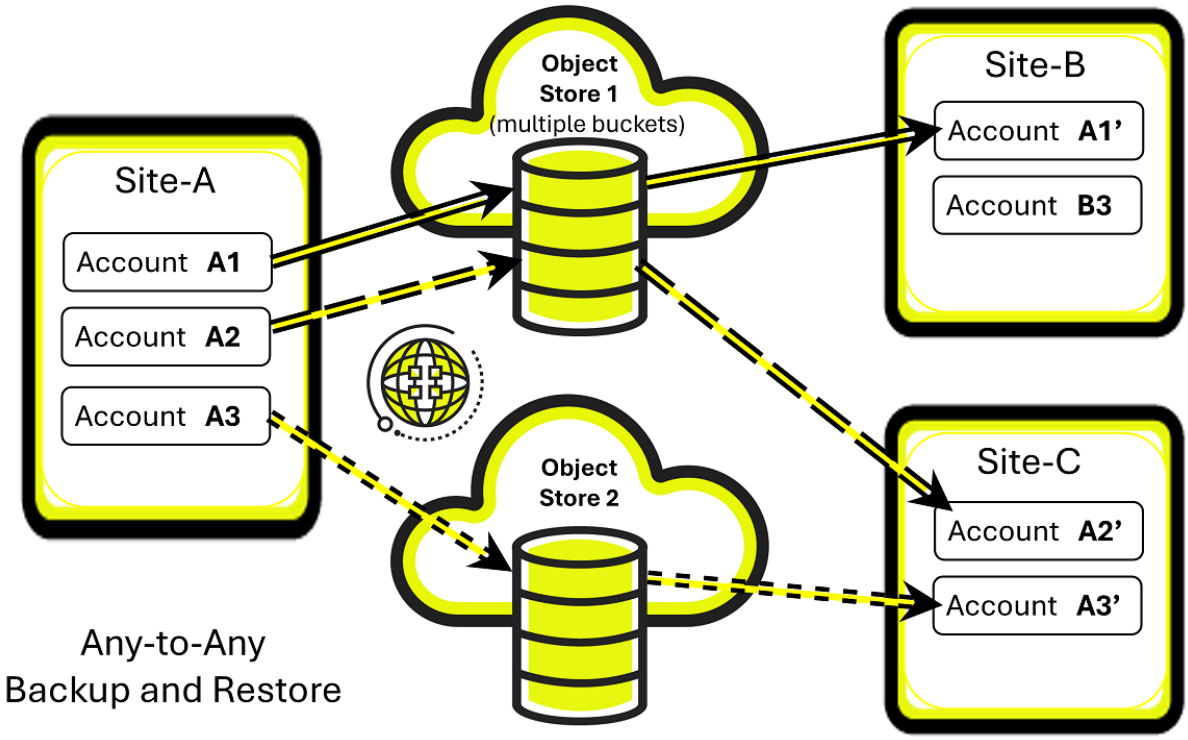
On top of this capability, from zCompute v24.03, Protection Groups also provide VM-level crash-consistent backups. Backup snapshots are taken as an atomic operation on all volumes of a protected VM, treating the VM's volumes as a consistency group.

zCompute B2OS is an integral feature of zCompute. It provides full backup and restore functionality without requiring third-party software or installing software agents on protected VMs.

Protection groups are backup policies that define the backup schedule for protected VMs and volumes, including the backup interval and retention period.

With the introduction of B2OS, you can optionally configure protection groups to back up protected VMs and volumes to Zadara Object Storage.

The following diagram depicts an example scenario of zCompute's B2OS any-to-any backup and restore capability:



In this example, Site-A is a site that has several zCompute accounts. Each account has one or multiple *Data Protection using Protection Groups*. Snapshots of a Protection Group's VM instances and volumes are taken according to schedules and stored locally.

Individual, multiple or all of an account's Protection Groups local snapshots can also be scheduled for backup to an Object Storage.

Each Protection Group has an exclusive Object Storage bucket or container, used solely for that Protection Group's snapshots.

An account's backups can be restored from the Object Storage back to the original account or to other sites, maintaining full data integrity and crash-consistency.

In the example in the diagram above, Site-A's selected accounts' Protection Groups are restored to Site-B and Site-C. Each Protection Group's VM instances and volumes are restored as an integral unit from a snapshot in the Protection Group's dedicated Object Storage container.

11.2 Backup to Object Storage (B2OS) Configuration Flow

The zCompute Backup to Object Storage (B2OS) configuration high-level flow:

1. Create an External Endpoint to the Object Storage.

   See *Creating an External Endpoint*.

2. Create a Backup Protection Group:

   See *Creating a Backup Protection Group*.

3. Verify that Protection Group's Backup to Object Storage is enabled.

   See *Enabling or Disabling a Backup Protection Group*.

4. Add VMs and volumes to a Backup Protection Group.

See *Adding or Removing Protected Resources*.

Snapshots of the Protected Resources of the Backup Protection Group occur according to the configured schedule.

5. Optionally, take an initial immediate snapshot with *Backup Protection Group Trigger Now*.

Caution

Manual changes must never be made directly to the Zadara Object Storage container!
The Zadara Object Storage container is managed exclusively by zCompute.

Any manual change to the Object Storage container or to its contents, including deletions, can cause severe malfunctions.

Deletion of a Protection Group's protected data should be accomplished by either changing the retention period or deleting the Protection Group. It can take a while until deletion of the Protection Group's containers and their contents completes. After that, the external endpoint pointing to the container can also be deleted.

If in doubt, please contact Support for guidance.

11.3 Remote Snapshots

Tenant administrators can configure External Endpoints for the purpose of saving snapshots to remote Object Storage destinations.

Tenant administrators can use *Data Protection using Protection Groups* to configure sets of protected resources comprising volumes and VM instances for scheduled backups at the same specified periodic intervals for all members of a group. They can also trigger additional immediate backups of a group.

11.4 External Endpoints

11.4.1 Viewing External Endpoints

1. Navigate to **Configuration > External Endpoints**.

   A list of configured External Endpoints displays.

2. Click an External Endpoint to display its details.

   The External Endpoint's details display in the lower pane.

11.4.2 Creating an External Endpoint

Note

- At the cloud level, accessing an Object Storage container via an External Endpoint requires the combination of the container name and the Object Storage user's Access Key to be unique within the cloud.

  To create more than one External Endpoint for the same Object Storage container within a cloud, you must configure the additional External Endpoint with the Access Key and Secret of a different user in the Object Storage.

  A separate external B2OS endpoint must be defined for each backup protection group, forming a one-to-one mapping.

  Backup protection groups are project-scoped. Each protection group belongs to a single project and protects only the VMs and volumes within that project.

- Before creating an External Endpoint, consult with your MSP regarding configuration values, in particular, **Network Topology** and **Endpoint URL**.

1. Navigate to **Configuration > External Endpoints**.

   A list of configured External Endpoints displays.

2. In the top menu bar, click **+ Create**.
3. In the **Create External Endpoint** dialog, enter the External Endpoint's parameters:
   - **Name**: A unique meaningful name for the External Endpoint.
   - **Description**: Optional description.
   - **Endpoint Type**: From the dropdown, select **B2OS**.

     Currently, B2OS only supports Zadara Object Storage (NGOS).

     Caution

     To create a new External Endpoint for the purpose of Backup to Object Store, the target Zadara Object Store container must already exist and must be empty.

     Any object in the container, including empty folders, will cause the creation of a B2OS endpoint to fail.

     Provide the following details from the Zadara Object Storage's **User Information** and **Console** screens.

     - **Network Topology**: Based on input from your MSP, select the topology:
       - **Frontend Network**
       - **Outbound Network**
     - **Region**: Copy the **User Information > Authentication > Region**.
     - **Endpoint URL**:

       Typically, copy the **User Information > Connectivity - Public Network > Public API Endpoint**.

       Consult and verify this with your MSP, as this endpoint value can depend on your **Network Topology**.

     Important

     The configuration requires a URL beginning with `https://`.

     Prefix the B2OS **Endpoint URL** string with `https://`, if it is missing in the URL copied from the Zadara Object Storage's **Public API Endpoint** configuration.

     For example:

     If Zadara Object Storage's **Public API Endpoint** is `abc00000123-public-zadara.zadarazios.com`
     then the B2OS **Endpoint URL** is `https://abc00000123-public-zadara.zadarazios.com`

     - **Bucket**: Enter the **Container** name as it appears in the **Console** screen.
     - **Access Key**: Copy the **User Information > Authentication > S3 Access Key**.
     - **Secret**: Copy the **User Information > Authentication > S3 Secret Key**.

- **Verify SSL**: Toggle switch to enable or disable checking whether the SSL certificate is valid.

Caution

Manual changes must never be made directly to the Zadara Object Storage container!
The Zadara Object Storage container is managed exclusively by zCompute.

Any manual change to the Object Storage container or to its contents, including deletions, can cause severe malfunctions.

Deletion of a Protection Group's protected data should be accomplished by either changing the retention period or deleting the Protection Group. It can take a while until deletion of the Protection Group's containers and their contents completes. After that, the external endpoint pointing to the container can also be deleted.

If in doubt, please contact Support for guidance.

11.4.3 Modifying an External Endpoint

1. Navigate to **Configuration > External Endpoints**.

   A list of configured External Endpoints displays.

2. Click an External Endpoint to display its details.
3. In the top menu bar, click **Modify**.
4. In the **Modify External Endpoint** dialog, the following fields can be updated:
   - **Name**: A unique meaningful name for the External Endpoint.
   - **Description**: Optional description.
   - **Access Key**: Object Storage user's S3 Access Key
   - **Secret**: Object Storage user's S3 Secret Key

Caution

Manual changes must never be made directly to the Zadara Object Storage container!
The Zadara Object Storage container is managed exclusively by zCompute.

Any manual change to the Object Storage container or to its contents, including deletions, can cause severe malfunctions.

Deletion of a Protection Group's protected data should be accomplished by either changing the retention period or deleting the Protection Group. It can take a while until deletion of the Protection Group's containers and their contents completes. After that, the external endpoint pointing to the container can also be deleted.

If in doubt, please contact Support for guidance.

CHAPTER

TWELVE

DATA PROTECTION USING PROTECTION GROUPS

12.1 Introduction

Backup Protection Groups are backup policies that define backup schedule intervals and retention periods.

VMs and volumes that are a protection group's protected resources are backed up according to the group's backup schedule and retention settings.

A Protection Group can optionally be configured to back up local snapshots to a Remote Object Storage.

There are no prerequisites for the creation of a Protection Group.

Protected resources such as volumes and VM instances can be added to a Protection Group at any time.

Note

To configure a Protection Group to back up local snapshots to a remote Object Storage on completing local snapshot creation, an External Endpoint for the Object Storage container must first be configured, if it does not already exist.

See *Creating an External Endpoint*.

Restore Protection Groups are required only when restoring to another site (restore only, without backup) or when restoring to another account on the same cloud.

Caution

A Protection Group's snapshot is skipped if one of its protected volumes isn't in a **Ready** state. This could also happen when the volume is migrating.

12.2 Backup Protection Group Operations

Note

Backup Protection Groups back up and recover within the same project.

Restore Protection Groups are intended only for recovery from a remote Object Store to a different project, account or cloud.

12.2.1 Viewing Backup Protection Groups

1. Navigate to **Protection > Protection Groups**.
2. Click the **Backup Protection Groups** tab.

   A list of configured Backup Protection Groups displays, with the following columns:

Column	Description
Name	The Protection Group's name
Remote Retention Days	Number of days the backup is retained in the remote Object Storage
Local Retention Days	Number of days the local backup is retained
Last Triggered	The last date and time a snapshot was taken
User	The user that triggered the last snapshot
Resources	The number of protected resources backed up in the last snapshot
Enabled	Indicator whether the Protection Group is currently enabled
External Endpoint	Name of the External Endpoint to Object Storage
Admin Only	Indicator whether the Protection Group is managed only by the MSP
Health	The Protection Group's health status
State	The Protection Group's readiness status for triggering snapshots

3. To view a Backup Protection Group's details, click its **Name**, to display the following:

   - **Top Menu Bar**

     The Backup Protection Group's top menu bar displays the following option buttons:

     - **Disable**: See *Enabling or Disabling a Backup Protection Group*.
     - **Modify**: See *Modifying a Backup Protection Group*.
     - **Schedule**: See *Rescheduling a Backup Protection Group*.
     - **Trigger Now**: See *Backup Protection Group Trigger Now*.
     - **Delete**: See *Deleting a Backup Protection Group*.

     If the Backup Protection Group is not configured for backup to remote Object Store, the top menu bar also display's the following option button:

     - **Associate Object Storage**: See *Associating a Backup Protection Group with an Object Storage*.

   - **Top Pane**

     The Backup Protection Group's top pane displays the following sections:

     - **Backup Protection Group** basic information: Name, Description, Status and Health
     - **Schedule** details for local Snapshots and Remote Object Storage Snapshots
     - **Protected Resources** summary count of protected resources, grouped by type (VMs, volumes)
     - **External Endpoint** parameter values

   - **Lower Pane** tabs:

     The Backup Protection Group's lower pane has the following tabs:

     - **Overview** tab:

       Basic information about the Backup Protection Group:

Column	Description
Name	The Protection Group's name
Type	The type of Protection Group (Backup or Restore)
Creation Date	The date and time the Protection Group was created
Last Update	The date and time of the last change to the configuration
External Endpoint	Name of the External Endpoint to Object Storage
Enabled	Indicator whether the Protection Group is currently enabled
Admin Only	Indicator whether the Protection Group is managed only by the MSP
ID	The Protection Group's UUID

- **Events** tab

  Allows applying filters to view selections of the Protection Group's events log.

- **Protected Resources** tab

  See *Adding or Removing Protected Resources*.

- **Local Protection Group Snapshots** tab

  List of the Protection Group's local snapshots, with the option for authorized users to select and delete snapshots.

- **Remote Protection Group Snapshots** tab

  List of the Protection Group's snapshots on remote Object Storage, with the option for authorized users to select and delete snapshots.

12.2.2 Creating a Backup Protection Group

To create a Backup Protection Group and configure its backup snapshot schedule:

1. Navigate to **Protection > Protection Groups**.

   A list of configured Backup Protection Groups displays in the **Backup Protection Groups** tab.

2. In the top menu bar, click **+ Create**.

3. In the **Create Backup Protection Group** dialog:

   1. In the **Group** tab, enter the parameters:

      - **Name**: A unique meaningful name for the Backup Protection Group.
      - **Description**: Optional description.
      - **Backup to Object Store**: Toggle switch determining whether local snapshots are backed up to a remote Object Store.

        Select:

        - **Off** (default): The snapshots are created locally.
        - **On**:

          If selected, from the **External Endpoint** dropdown select the B2OS endpoint for the remote Object Storage backup container.

          The snapshots are created locally.

          Local snapshots are backed up to the selected remote Object Storage destination container, as specified in the selected remote Object Storage's External Endpoint.

Click **Next**.

2. In the **Schedule** tab, enter the parameters:

   Local Snapshots:

   - **Recurrence**: The frequency interval units, as one of:
     - **Minute**
     - **Hour**
     - **Day**
     - **Week**
     - **Month**
   - **Every**: The frequency interval as a number of the selected unit, between each snapshot.
     - For **Week** intervals, click the days of the week that weekly snapshots are scheduled.
     - For **Month** intervals, select one of the **Repeat by** options, and the **Start date**:

       Day of the month:

       Snapshots are scheduled for the selected day of the month, starting from the **Start date**, and repeating according to the interval defined as the number of months (*n*) in **Every** *n* **Month(s)**.

       For example, if **Every** = 2 and **Start date** = 17 Jan 2025, snapshots are scheduled for the 17th of every second month, starting 17 Jan 2025.

       Day of the week:

       Snapshots are scheduled for the selected day of the week, and week of the month, starting from the **Start date**, and repeating according to the interval defined as the number of months (*n*) in **Every** *n* **Month(s)**.

       For example, if **Every** = 2 and **Start date** = Fri 17 Jan 2025, since the start date occurs on the 3rd Friday of the month, snapshots are scheduled for the 3rd Friday of every second month, starting 17 Jan 2025.
     - **Start Time**: The start time of the schedule, in Hours and Minutes in 12-hour format, and either AM or PM.
     - **Retention for Local snapshots**: The duration in **days** to retain local snapshots.

       Local snapshots are deleted automatically after this duration.

       Note

       A maximum of 50 accumulated local snapshots can be stored.

   Additional parameters for Backup Protection Groups that are configured for **Backup to Object Store**:

   - **Remote Object Storage Snapshots**:
     - **Remote snapshots every**: The number oflocal snapshots that accumulate, after which they are backed up to the remote Object Store.
     - **Retention for Remote snapshots**: The duration in **days** to retain Remote snapshots.

       Remote snapshots are deleted automatically after this duration.

       Note

       A maximum of 100 accumulated remote snapshots can be stored.

Click **Finish**.

Note

Continue at any time with *Adding or Removing Protected Resources*.

12.2.3 Adding or Removing Protected Resources

1. Navigate to **Protection > Protection Groups**.

   A list of configured Backup Protection Groups displays in the **Backup Protection Groups** tab.

   Note

   Volumes that are attached to a VM instance are automatically included in that VM's Backup Protection Group.

   Similarly, volumes attached to a VM that is in a Backup Protection Group are automatically removed from the Backup Protection Group upon removal of that VM.

   If a volume is independently protected in a Protection Group and also attached to a VM that is later added to the same Protection Group, the volume's Protection Group snapshots will be deleted. However, you can still recover an individual volume from a VM snapshot or from the Protection Group.

   Caution

   Attempting to add a VM to a Protection Group will fail, if a volume on that VM is already previously defined as a (standalone) protected resource in the **same** Protection Group.

   In this case, to successfully add the VM and all of its volumes to the Protection Group, it is necessary to first **Remove Protection** of the conflicting volume from the Protection Group. Then, adding the VM to the Protection Group automatically includes all its volumes.

2. Click the Backup Protection Group to add or remove protected resources.

   The Backup Protection Group's details display.

3. In the lower pane, click the **Protected Resources** tab.

   The Backup Protection Group's protected resources are listed.

   - **Adding a VM instance to the Backup Protection Group**

     1. In the lower pane menu bar click **+ Add VM**.
     2. In the **Add instance to protection group** dialog, from the **VMs** dropdown, select the VM instance to add to the Backup Protection Group, and click **OK**.

        The VM instance appears in the Backup Protection Group's protected resources list.

        Future snapshots of the Backup Protection Group will include the VM instance and its attached volumes.

   - **Adding a Volume to the Backup Protection Group**

     1. In the lower pane menu bar click **+ Add Volume**.
     2. In the **Add volume to protection group** dialog, from the **Volumes** dropdown, select the volume to add to the Backup Protection Group, and click **OK**.

        The volume appears in the Backup Protection Group's protected resources list.

Future snapshots of the Backup Protection Group will include the volume.

- **Removing a VM or Volume from the Backup Protection Group**

  To remove a protected resource:

  1. On the row of the VM or volume to remove from the Backup Protection Group, click anywhere **except** on the Resource Name.

     The **Remove Protection** option appears on the lower pane menu bar.

  2. Click **Remove Protection**.

     In the **Remove Protection** confirmation dialog, click **OK** to proceed with removing the selected VM or volume from the Backup Protection Group.

     The VM or volume disappears from the Backup Protection Group's protected resources list.

     Future snapshots of the Backup Protection Group will no longer include the removed VM or volume.

12.2.4 Enabling or Disabling a Backup Protection Group

1. Navigate to **Protection > Protection Groups**.

   A list of configured Backup Protection Groups displays in the **Backup Protection Groups** tab.

   The **Enabled** column displays the Backup Protection Group's Enabled or Disabled status.

2. Click a Backup Protection Group to select it for Enabling or Disabling.

   The Backup Protection Group details display.

3. In the top menu bar, click the **Enable/Disable** toggle.

   The **Enabled** status field displays the Backup Protection Group's updated Enabled or Disabled status.

12.2.5 Modifying a Backup Protection Group

To modify a Backup Protection Group:

1. Navigate to **Protection > Protection Groups**.

   A list of configured Backup Protection Groups displays.

2. Click a Backup Protection Group to modify it.

   The Backup Protection Group details display.

3. In the top menu bar, click **Modify**.

4. In the **Update Backup Protection Group** dialog, optionally modify one or more of the modifiable parameters:

   - **Name**: A unique meaningful name for the Backup Protection Group.
   - **Description**: Optional description.
   - **Backup to Object Store**: Toggle switch determining whether local snapshots are backed up to a remote Object Store.

     Note

     If the Protection Group is not already associated with a remote Object Storage, it is also possible to configure backup to remote Object Storage.

A Protection Group that is associated with a remote Object Storage cannot be disassociated with its remote Object Storage, except for deletion, which also removes all accumulated local snapshots.

Select:

- **Off** (default): The snapshots are created locally.
- **On**:

  The snapshots are created locally.

  Local snapshots are backed up to the selected remote Object Storage destination container, as specified in the selected remote Object Storage's External Endpoint.

  If selected, the **Associate Object Storage** dialog opens.

  Group tab:

  1. From the **External Endpoint** dropdown select the B2OS endpoint for the remote Object Storage backup container.
  2. Click **Next** to continue to the **Schedule** tab.

  Schedule tab:

  1. Optionally modify the **Local Snapshots** schedule parameters as described in *Rescheduling a Backup Protection Group*.
  2. Accept or modify the **Remote Object Storage Snapshots** parameters, as described in *Rescheduling a Backup Protection Group*.
  3. Click **Finish**.

12.2.6 Rescheduling a Backup Protection Group

To change the frequency or time of a Backup Protection Group's local or remote Object Storage snapshots:

1. Navigate to **Protection > Protection Groups**.

   A list of configured Backup Protection Groups displays.

2. Click a Backup Protection Group to select it for snapshot rescheduling.

   The Backup Protection Group details display.

3. In the top menu bar, click **Schedule**.
4. In the **Schedule Backup Protection Group** dialog, modify the relevant parameters:

   - **Local Snapshots**:
     - **Recurrence**: The snapshot frequency interval units, as one of:
       - **Minute**
       - **Hour**
       - **Day**
       - **Week**
       - **Month**
     - **Every**: The snapshot frequency interval as a number of the selected unit, between each snapshot.
       - For **Week** intervals, click the days of the week that weekly snapshots are scheduled.

        * For **Month** intervals, select one of the **Repeat by** options, and the **Start date**:

          day of the month:

          Based on the **Start date**, snapshots are scheduled to repeat on the selected day of the month at the interval frequency of the number entered for **Every** number of months.

          For example, if **Every** = 2 and **Start date** = 24 Jan 2025, snapshots are scheduled for the 24th of every second month, starting 24 Jan 2025.

          day of the week:

          Based on the **Start date**, snapshots are scheduled to repeat on the selected day of the week, and week of the month, at the interval frequency of the number entered for **Every** number of months.

          For example, if **Every** = 2 and **Start date** = Fri 24 Jan 2025, since the start date occurs on the 4th Friday of the month, snapshots are scheduled for the 4th Friday of every second month, starting 24 Jan 2025.

      - **Start Time**: The start time of the schedule, in Hours and Minutes in 12-hour format, and either AM or PM.
      - **Retention for Local snapshots**: The duration in **days** to retainlocal snapshots.

        Local snapshots are deleted automatically after this duration.

  Additional parameters for Backup Protection Groups that are configured for **Backup to Object Store**:

  - **Remote Object Storage Snapshots**:
    - **Remote snapshots every**: The number of local snapshots that accumulate, after which they are backed up to the remote Object Store.
    - **Retention for Remote snapshots**: The duration in **days** to retain remote snapshots.

      Remote snapshots are deleted automatically after this duration.

  Click **Finish**.

12.2.7 Backup Protection Group Trigger Now

To trigger an immediate snapshot of a Backup Protection Group's protected resources, in addition to its scheduled snapshot:

1. Navigate to **Protection > Protection Groups**.

   A list of configured Backup Protection Groups displays.

2. Click a Backup Protection Group to select it for triggering an immediate snapshot.

   The Backup Protection Group details display.

3. In the top menu bar, click **Trigger Now**.

4. In the **Trigger Backup Protection Group** dialog:

   1. Optionally, enable **Remote Object Storage Snapshots** to send the snapshot to remote Object Storage.
   2. Click **OK** to progress with creating the snaphot of the protected resources immediately.

      The snapshot and its progress are listed in the Backup Protection Group's **Local Protection Group Snapshots** tab in the lower pane.

      If **Remote Object Storage Snapshots** is enabled, the snapshot and its progress are also listed in the Remote Protection Group's **Local Protection Group Snapshots** tab in the lower pane.

12.2.8 Deleting a Backup Protection Group

To delete a Backup Protection Group:

1. Navigate to **Protection > Protection Groups**.

   A list of configured Backup Protection Groups displays.

2. Click a Backup Protection Group to select it for deletion.

   The Backup Protection Group details display.

3. In the top menu bar, click **Delete**.

4. In the **Delete Protection Group** dialog, click **Delete** to confirm deletion of the Backup Protection Group.

12.2.9 Associating a Backup Protection Group with an Object Storage

If a Backup Protection Group is configured for **Local Snapshots** only, it is also possible to configure backup to remote Object Storage by associating the Backup Protection Group with an Object Storage:

Note

A Protection Group that is associated with a remote Object Storage cannot be disassociated with its remote Object Storage, except for deletion, which also removes all accumulated local snapshots.

To associate a Backup Protection Group with an Object Storage:

1. Navigate to **Protection > Protection Groups**.

   A list of configured Backup Protection Groups displays.

2. Click a Backup Protection Group to select it for association with an Object Storage.

   The Backup Protection Group details display.

3. In the top menu bar, click **Associate Object Storage**.

   The **Associate Object Storage** dialog opens:

   In the **Group** tab:

   1. From the **External Endpoint** dropdown select the B2OS endpoint for the remote Object Storage backup container.

      Note

      On completion of snapshot creation, local snapshots are backed up to the remote Object Storage destination container, as configured in the remote Object Storage's External Endpoint.

   2. Click **Next** to continue to the **Schedule** tab.

   In the **Schedule** tab:

   1. Optionally modify the **Local Snapshots** schedule parameters as described in *Rescheduling a Backup Protection Group*.
   2. Accept or modify the Remote Object Storage Snapshots parameters, as described in *Rescheduling a Backup Protection Group*.

4. Click **Finish**.

12.2.10 Restoring from a Backup Protection Group

Restoring a VM instance

To restore a **VM instance**, see *Recover VM Instances from Snapshots* in the *Snapshots* page.

The VM's attached volumes are restored as an integral part of the VM instance's recovery process.

Restoring a Volume

To restore a **volume**:

1. Navigate to **Storage > Snapshots**.
2. Select the **Local Snapshots** tab.
3. Click **Create Volume** in the upper menu bar.
    1. In the **Create Volume** dialog, accept or update the values for:
        - **Name**: Volume name
        - **Volume Type**
    2. Click **OK**.

The restored volume displays in the **Storage > Block Storage** list.

See *Volume Snapshot Operations* in the *Snapshots* page for other snapshot operations.

12.3 Restore Protection Group Operations

Note

Restore Protection Groups are intended only for recovery from a remote Object Store to a different project, account or cloud. In contrast, Backup Protection Groups back up and recover within the same project.

12.3.1 Viewing Restore Protection Groups

1. Navigate to **Protection > Protection Groups**.
2. Click the **Restore Protection Groups** tab.

   A list of configured Restore Protection Groups displays, with the following columns:

   Name, **User**, **External Endpoint**, **Health** and **State**.
3. To view a Restore Protection Group's External Endpoint details, click a Restore Protection Group.

   The Restore Protection Group's External Endpoint details display in the lower pane.

12.3.2 Creating a Restore Protection Group

To create a Restore Protection Group:

1. Navigate to **Protection > Protection Groups**.
2. Click the **Restore Protection Groups** tab.

   A list of configured Restore Protection Groups displays.
3. In the top menu bar, click **+ Create**.

4. In the **Create Restore Protection Group** dialog:
    1. Enter the parameters:
        - **Name**: A unique meaningful name for the Restore Protection Group.
        - **Description**: Optional description.
        - **External Endpoint**: From the dropdown, select the remote Object Storage from the list of B2OS endpoints.
    2. Click **Finish**.

12.3.3 Modifying a Restore Protection Group

To modify a Restore Protection Group:

1. Navigate to **Protection > Protection Groups**.
2. Click the **Restore Protection Groups** tab.

   A list of configured Restore Protection Groups displays.
3. Click a Restore Protection Group to modify it.

   The Restore Protection Group details display.
4. In the top menu bar, click **Modify**.
5. In the **Update Restore Protection Group** dialog:
    1. Optionally modify one or both modifiable parameters:
        - **Name**: A unique meaningful name for the Restore Protection Group.
        - **Description**: Optional description.
    2. Click **Finish**.

12.3.4 Deleting a Restore Protection Group

To delete a Restore Protection Group:

1. Navigate to **Protection > Protection Groups**.
2. Click the **Restore Protection Groups** tab.

   A list of configured Restore Protection Groups displays.
3. Click a Restore Protection Group to select it for deletion.

   The Restore Protection Group details display.
4. In the top menu bar, click **Delete**.
5. In the **Delete Protection Group** dialog, click **Delete** to confirm deletion of the Restore Protection Group.

12.3.5 Restoring from a Restore Protection Group

Note

A VM's attached volumes are restored as an integral part of the VM instance's recovery process.

If a Protection Group comprises more than one VM instance (e.g. 2 VMs and two unattached volume), in the UI, each VM instance and each unattached volume must be recovered individually, separately.

The UI does not provide a wizard or dialog for recovering all of a Protection Group's Protected Resources in a single-phase interaction or transaction.

12.3.6 Restoring a VM instance from Object Storage

To restore a **VM instance** from a remote Object Storage, see *Recover VM Instances from Snapshots* in the *Snapshots* page.

The VM's attached volumes are restored as an integral part of the VM instance's recovery process.

12.3.7 Restoring a Volume from Object Storage

To restore a **volume** from a remote Object Storage:

1. Navigate to **Storage > Snapshots**.
2. Select the **Remote Snapshots** tab.
3. Click **Create Volume** in the upper menu bar.
    1. In the **Create Volume** dialog, accept or update the values for:
        - **Name**: Volume name
        - **Volume Type**
    2. Click **OK**.

The restored volume displays in the **Storage > Block Storage** list.

See *Volume Snapshot Operations* in the *Snapshots* page for other snapshot operations.

CHAPTER

THIRTEEN

SNAPSHOTS

Snapshots can be taken of a specific volume or a VM instance which can include a boot volume and multiple data volumes. The snapshot takes a copy of the volume or a VM instance at a specific time, and then subsequent snapshots create a change log thus allowing for restoration of a single volume or an entire VM instance. Volume snapshots can also be used to create images.

Caution

Snapshots configured in the VM and Storage UI are individual, and not crash-consistent.

For crash-consistent backups, we highly recommend configuring snapshots using *Data Protection using Protection Groups*.

For remote snapshots, see *Remote Snapshots* under *Backup to Object Storage (B2OS)*.

13.1 Creating Snapshots of VM Instances

1. Navigate to the **Compute > Instances** view.
2. From the displayed list, select the VM instance for which the snapshot is to be created and click **More**.
3. In the More menu, select **Snapshot**.
4. In the displayed **Create Snapshot** window, enter a name for the new snapshot or accept the default name consisting of the original VM instance name and the date-time stamp of the snapshot creation.
5. Enter the description of the snapshot.
6. Click **OK**. A new snapshot is created. It is displayed in the **Compute** > **Snapshots** view.

13.2 Recover VM Instances from Snapshots

1. Navigate to the **Compute > Snapshots** view.

   To recover from Object Storage, select the **Remote Snapshots** tab.
2. From the displayed list, select a VM instance snapshot.
3. From the top toolbar, click **Restore**.
4. In the **Restore Instance > Compute** tab, enter the following:
   - **Name** - the display name for the VM instance.
   - **Instance Type** - defines the amount of compute resource of the VM instance (CPU and RAM).
   - **Key Pair** - set of security credentials for ensuring the identity of the user connecting to the VM instance.
   - **Options:**

    - **Power Up** - launches the VM instance post-creation.
  - **High Availability** - check this option to ensure the instance is restarted in the event of a failure.
  - **Protect from Deletion** - check this option to protect the instance from accidental deletion.
5. In the **Restore Instance > Storage** tab, enter the following:
   - **Boot Volume** - size of boot volume in GB.
   - **Volume Type** - from the dropdown list, select the volume type that will be used for the boot volume of the VM instance.
6. In the **Restore Instance > Networking** tab, select from the following options:
   - **Same Network + IP** - the recovered VM instance will have the same IP as that in the snapshot.
   - **Same Network** - the recovered VM instance will be on same network as that in snapshot, but with different IP.
   - **Manual** - select any network.
7. In the **Restore Instance > Config** tab, enter the following:
   - **Tags** - attach one or more tags to this VM instance.
   - **Add** - to enter metadata key and value pairs.
8. Click **Finish**.

13.3 Other VM Instance Snapshot Operations

The following other operations are available from top toolbar in **Compute > Snapshots** view.

- **Modify** - create updated snapshot.
- **Delete** - delete a snapshot.
- **Copy** - copy a snapshot.

13.4 Volume Snapshot Operations

1. Navigate to the **Storage > Snapshots** view.

   To recover from Object Storage, select the **Remote Snapshots** tab.
2. Select a snapshot from the displayed list.

   The following operations are available from the top toolbar.
   - **Create** - create updated snapshot.
   - **Rename** - rename existing snapshot.
   - **Create Volume** - restore volume from a snapshot.
   - **Create Image** - create image from a snapshot.
   - **Delete** - delete a snapshot.

CHAPTER

FOURTEEN

CONFIGURATION SETTINGS

Use **Configuration > Settings** to manage global system behavior, security policies, branding, networking, and external integrations.

Settings are organized into tabs:

- *General*

  System titles, links, language, and session timeout settings.

- *Security*

  Password policies, expiration, MFA enforcement, and reset email settings.

- *Branding*

  Customized logos and the password reset email template.

- *Services & Support*

  Enable or disable cluster services and related features.

- *Proxy*

  Semi-offline cluster management proxy settings and exclusions, generally set up by Zadara Operations.

- *DNS & SMTP*

  DNS name servers and SMTP settings for system email delivery for alarms, generally set up by Zadara Operations.

- *External Monitoring*

  External endpoints for resources tracking, logs, events, and inventory data, mainly set up by Zadara Operations.

Changes apply only after you select **Save** on the active tab.

14.1 General

Use the **General** tab to view or configure titles, links, language, and session settings.

14.1.1 Titles and links

Display names and external references.

- **Title**

  Application title shown in the UI.

- **Product Name**

  Product display name.

- **Support Link**

  Zadara support site's URL.

- **API Explorer**

  Relative path to the zCompute API explorer.

- **Knowledge Base**

  Zadara documentation guides portal's URL.

- **Support Email**

  Zadara support contact email address.

- **Custom Link Button Text**

  Label on the optional **Custom Link Button**, when enabled.

- **Custom Link Button URL**

  Target URL on clicking the optional **Custom Link Button**, when enabled.

Options:

- **Login Services Info**:

  Reserved

- **Powered By**

  Toggle to display or hide the **Powered by Zadara** caption during sign-on.

  Default: Enabled.

- **Custom Link Button**: Toggle to enable or disable displaying the Custom Link Button. When enabled, the Custom Link Button appears at the top of the screen next to the **Quick Start** button.

14.1.2 Language

The UI display language.

Default Language:

From the dropdown, select the default UI language:

- English (default)
- German
- Japanese
- Korean
- Portuguese
- Spanish

14.1.3 UI settings

Session Idle Timeout:

The sessions inactivity timeout in minutes.

Select **Save** to apply changes on this tab.

14.2 Security

Use the **Security** tab to view or configure password policies and authentication requirements.

14.2.1 Password policy

The password requirements for user sign-on to the cloud.

- **Min. Password Length**

  The minimum number of characters required in a user password.

  Default: 8 characters.

- **Must include a letter**

  Toggle to enable or disable whether the password must include a letter.

  Default: Enabled - required.

- **Must include a special character**

  Toggle to enable or disable whether the password must include a special character.
  (! ' @ # $ % ^ & * ( ) _ + | = { } [ ] -)

  Default: Enabled - required.

- **Must include a digit**

  Toggle to enable or disable whether the password must include a digit.
  (0 - 9)

  Default: Enabled - required.

14.2.2 Settings

- **Passwords expire after**

  The number of days that a password is valid from the time it is created until it expires.

  Default: 90 days.

- **Enforce password history**

  The number of most recent passwords that a user cannot reuse.

  Default: 6 immediate previous passwords cannot be reused.

- **Enforce Multi-Factor Authentication**

  Toggle to enable or disable whether MFA is required for user sign-on.

  Default: Disabled - not required.

- **Password Reset SMTP Endpoint**

  The SMTP endpoint that is used for password reset emails.

  Default: Not configured.

Select an existing SMTP endpoint from the dropdown, or click "**+**" for *Creating an External Endpoint*. In addition, see the *Branding* tab, for the *Password reset email template*.

Select **Save** to apply changes on this tab.

14.3 Branding

Use the **Branding** tab to view and customize logos and the password reset email template.

14.3.1 Logos

Each logo has required dimensions and a maximum file size. Ensure that images meet the specified requirements before uploading.

You can upload:

- **Login Logo**

  Displayed on the login screen.

- **Main Navigation Pane Logo**

  Shown in the navigation pane and in the **About** dialog.

- **Favicon Logo**

  Displayed in the browser tab.

- **External Object Storage Logo**

  The logo appears in the Object Storage Screen.

14.3.2 Password reset email template

The template for the email subject line and body, sent to users notifying them to reset their password.

- **Subject**

  Email subject line.

- **Text Template**

  Plain-text version of the email body.

- **HTML Template**

  HTML version of the email body.

  The HTML template must include required variables as indicated in the UI.

Select **Save** to apply changes on this tab.

14.4 Services & Support

Use this tab to enable or disable cluster services.

When you disable a service, related UI functionality is hidden from non-admin users.

14.4.1 Manage cluster services

Toggle availability of the following services and functionality for non-admin users:

- **Marketplace**

  Enable or disable access to the collection of downloadable machine images in the Marketplace.

- **Data Protection**

  Enable or disable access to configuration of backup and recovery for protection groups.

- **Load Balancing**

  Enable or disable access to configuration of distribution of traffic and workloads across resources, to improve availability and performance.

- **Certificates**

  Enable or disable access to certificate management for secure communication (HTTPS protocol) between the outside consumer and certain services.

Changes take effect immediately when toggled.

14.5 Proxy

The **Proxy** tab displays semi-offline cluster managment proxy settings, generally configured by the Zadara Operations team on provisioning the cloud.

14.5.1 General proxy settings

- **Enabled**

  Toggle to enable or disable proxy usage.

  Default: Disabled.

- **Exclusion list**

  Direct addresses that bypass the proxy.

14.5.2 Region proxy settings

- **HTTP proxy**

  Specifies the address of the HTTP proxy server.

- **Port**

  Defines the port used to connect to the proxy server.

- **Username** and **Password**

  Authentication credentials for the proxy.

- **Use this proxy server for all protocols**

  Toggle to apply the HTTP proxy settings to all outbound traffic.

  Default: Disabled.

14.5.3 SSL proxy

- **HTTPS proxy**

  Specifies the address of the secure HTTPS proxy server.

- **Port**

  Defines the port used for the HTTPS proxy connection.

- **Username** and **Password**

  Authentication credentials for the HTTPS proxy.

Select **Save** to apply changes on this tab.

14.6 DNS & SMTP

The **DNS & SMTP** tab contains the server configurations for alarms. These settings are generally configured by the Zadara Operations team.

14.6.1 SMTP configuration

- **Server Name**

  Specifies the address of the SMTP server.

- **Port**

  Defines the port used for SMTP communication.

- **Username** and **Password**

  Authentication credentials for the SMTP server.

- **From Address**

  The sender address for system-generated emails.

- **TLS Enable**

  Toggle to enable or disable TLS encryption for secure email transmission.

Select **Test Settings** to validate the configuration.

14.6.2 System alert subscribers

Add email recipients who receive system alerts.

Select **Save** to apply changes on this tab.

14.7 External Monitoring

Use the **External Monitoring** tab to view and configure external endpoints for logs, events, and inventory data.

14.7.1 External endpoints

From the dropdown for each category, select an endpoint or click the plus icon (+) for *Creating an External Endpoint*.

- **Resource Tracker**

  S3 endpoint for exporting resource usage data.

- **Events**

  S3, Rsyslog, Zendesk, or SMTP endpoint for forwarding system events.

- **Logs**

  S3 or Rsyslog endpoint for log export.

- **Metrics**

  Not in use. Replaced with the **Observability** endpoint.

- **Hardware Inventory**

  S3 endpoint for hardware inventory data.

- **Logical Inventory**

  S3 endpoint for logical inventory data.

- **API Trail**

  S3 or Rsyslog endpoint for API audit logs.

- **Observability**

  Selects an S3 endpoint for observability data export.

Select **Save** to apply changes on this tab.

14.7.2 Creating an External Endpoint

Use this procedure to create a new external endpoint.

1. Go to **Configuration > Settings**.
2. Select the **External Monitoring** tab.
3. Select the plus icon (+) next to the required category.

   The **Create External Endpoint** dialog opens.
4. In the **Create External Endpoint** dialog, enter the external endpoint's initial details:
   1. **Name**

      A unique name for the external endpoint.
   2. **Description**

      An optional description of the endpoint.
   3. **Endpoint Type**

      The type of external integration.

      Note

      zCompute provides configurations for several types of external endpoints.

      Depending on the selected endpoint type, the dialog prompts for parameters specific to that endpoint type:

      From the dropdown, select the endpoint type and enter values for its parameters:

      - *S3 endpoint type*
      - *Rsyslog endpoint type*

- *Zendesk endpoint type*
- *SMTP endpoint type*

S3 endpoint type

For an **S3** endpoint type, configure the following:

- **Region**

  Select one of AWS Region or Custom Region.

  - **AWS Region**

    From the dropdown, select the AWS region for the bucket.

  - **Custom Region**

    * **Custom Region**

      A name to identify or describe the custom region.

    * **Endpoint URL**

      The URL custom region's external endpoint.

  Enter values for the parameters that are common to both **AWS Region** and **Custom Region** external endpoints:

- **Bucket**

  The target S3 bucket name.

- **Access Key**

  The S3 access key ID.

- **Secret**

  The S3 secret access key.

- **Verify SSL**

  Toggle to enable or disable certificate verification, that checks whether the SSL certificate is valid.

For all endpoint types, continue with *Endpoint Permissions*.

Rsyslog endpoint type

For an **Rsyslog** endpoint type, configure the following:

- **Hostname/IP**

  The Rsyslog server address.

- **Port**

  The port used by the server.

- **Protocol**

  - **TCP**

    Used for traffic requiring high reliability, such as HTTP/HTTPS, database connections, and file transfers.

  - **UDP**

    Used for speed-sensitive applications, such as streaming, voice over IP (VoIP), or DNS, where minimal latency is preferred over perfect reliability.

For all endpoint types, continue with *Endpoint Permissions*.

Zendesk endpoint type

For an **Zendesk** endpoint type, configure the following:

- **Zendesk full address**

  The full Zendesk URL.

- **Zendesk username**

  The Zendesk account username.

- **Zendesk password**

  The Zendesk account password.

For all endpoint types, continue with *Endpoint Permissions*.

SMTP endpoint type

For an **SMTP** endpoint type, configure the following:

- **SMTP Hostname/IP**

  The SMTP server address.

- **SMTP Port**

  The SMTP server port.

- **SMTP From**

  The sender email address.

- **SMTP To**

  The recipient email address.

- **SMTP Username**

  The SMTP authentication username.

- **SMTP Password**

  The SMTP authentication password.

- **Verify SSL**

  Toggle to enable or disable certificate verification, that checks whether the SSL certificate is valid.

For all endpoint types, continue with *Endpoint Permissions*.

Endpoint Permissions

1. **Permissions**

    1. **Scope**

       One of:

       - **Public**

         The endpoint is public.

       - **Account**

         From the **Account** dropdown, select the account to which the endpoint must be accessible to all projects in that account.

         From all other accounts, the endpoint is inaccessible.

- **Project**

  From the **Project** dropdown, select the project to which the endpoint is accessible.

  From all other projects, the endpoint is inaccessible.

2. **Access**

   Determine access via the endpoint:

   - **Read/Write**
   - **Read Only**

2. Select **Finish** to create the endpoint.

14.7.3 Modifying an External Endpoint

1. Navigate to **Configuration > External Endpoints**.

   A list of configured External Endpoints displays.

2. Click an External Endpoint to display its details.
3. In the top menu bar, click **Modify**.
4. In the **Modify External Endpoint** dialog, the following fields can be updated:

   - **Name**: A unique meaningful name for the External Endpoint.
   - **Description**: Optional description.

   Depending on the endpoint type, some of the endpoint type-specific fields can be updated.

CHAPTER

FIFTEEN

CLUSTER CERTIFICATES

Use **Configuration > Cluster Certificates** to manage certificates for the cluster.

You can enable automatic certificate renewal or upload a custom certificate and private key.

15.1 Cluster certificate auto renewal

Use this option to enable automatic renewal of the cluster certificate.

To configure auto renewal:

1. Go to **Configuration > Cluster Certificates**.
2. From the top toolbar, select **Auto Renewal**.

   The **Cluster Certificate Auto Renewal** dialog opens.
3. In the **Cluster Certificate Auto Renewal** dialog:
   1. To enable automatic renewal, mark the **Enabled** checkbox.
   2. To confirm automatic renewal, select **Ok**.

Note

Auto renewal is performed only if:

- No custom certificate is installed.
- The remote certificate is newer than the installed one.
- The remote certificate is valid.

15.2 Creating a custom cluster certificate

Use this option to upload a custom certificate.

To create or replace a cluster certificate:

1. Go to **Configuration > Cluster Certificates**.
2. From the top toolbar, select **+ Create**.

   The **Create Cluster Certificate** dialog opens.
3. In the **Create Cluster Certificate** dialog:

1. In **Certificate**, upload the certificate file.

   Drag and drop the file, or select **Browse** to open a file browser window.

2. In **Private Key**, upload the private key file.

   Drag and drop the file, or select **Browse** to open a file browser window.

3. Optional: Enter the **Private Key Passphrase** if the private key is encrypted.
4. To confirm applying the certificate, select **Ok**.

Note

When a custom certificate is installed, automatic renewal does not occur.

CHAPTER

SIXTEEN

PCI DEVICES

In virtual computing platforms, Peripheral Component Interconnect (PCI)) devices are the hardware or software components that provide critical functionality, such as networking, storage, and graphics to virtual machines (VMs). Their operating protocol allows diverse hardware to communicate with a system's processor.

For Managed Service Provider (MSP) admins, PCI is the industry-standard interface that allows VMs to communicate directly with high-performance physical hardware.

In the zCompute platform, PCI is primarily leveraged through PCI Passthrough, which maps physical devices and high-speed NVMe storage directly to a tenant's instance. This bypasses the hypervisor's emulation layer to deliver near-native hardware performance, which is crucial for supporting modern client demands such as AI/ML training, VDI, and video analytics.

By managing these assignments through the zCompute interface, admins can offer specialized high-throughput storage tiers, enabling them to operate high-end workloads with the same multi-tenancy efficiency as their standard compute offerings.

16.1 PCI devices management

The PCI Devices screen lets system administrators view and manage PCI passthrough devices in a region.

Use this screen to view devices and their statuses, see where a device is installed, check whether a device is already tied to a VM, and change whether a device is enabled or disabled for passthrough use.

The following tasks are supported:

- Viewing the PCI devices that are available in the region.
- Checking device state and related details before making a change.
- Enabling a single or multiple devices in one action.
- Disabling a single or multiple devices in one action.

16.1.1 Viewing PCI devices

To view the available PCI devices:

1. Select **Configuration > PCI Devices**.

   The list of PCI devices appears, with the following columns:

   - **Device Name**

     The display name of the PCI device.

     Use it to identify the device in the table.

   - **PCI Slot**

     The PCI slot for the device.

Use it to identify the hardware slot on the node.

- **Node**

  The node that owns the device.

  When a value is displayed, it is a link to the node screen.

- **Vendor**

  The vendor ID for the device.

  Use it to identify the device vendor.

- **Device ID**

  The device ID for the hardware.

  Use it to identify the device type.

- **Enabled**

  Indicates whether the device is enabled.

- **Blocked**

  Indicates whether the device is blocked.

- **Kernel Driver**

  The current kernel driver for the device.

- **Original Kernel Driver**

  The original kernel driver for the device before the current driver was applied.

- **Associated VM**

  The VM associated with the device.

  When a value is displayed, it is a link to the VM screen.

- **Family**

  The device family value.

16.1.2 Enabling PCI devices

Enabling a PCI device makes that device available for PCI passthrough management on its node, so it can be used in later passthrough workflows.

To enable PCI devices:

1. Select **Configuration > PCI Devices**.

   The list of PCI devices appears.

2. In the list, locate the PCI devices to enable.

   Verify that the **Enabled** column of the devices to enable shows that they are not enabled.

3. Select the device rows.

   The **Enable Device** option appears in the upper toolbar.

4. From the upper toolbar, select **Enable Device**.

5. Confirm that the selected devices now appear as enabled.

16.1.3 Disabling PCI devices

Disabling a PCI device removes that device from the platform's PCI passthrough management path on its node, so it is no longer kept enabled for passthrough use.

Note

Disabling a PCI device is not possible for a device associated with a VM.

To disable PCI devices:

1. Select **Configuration > PCI Devices**.

   The list of PCI devices appears.

2. In the list, locate the PCI devices to disable.

   1. Verify that the **Enabled** column of the devices to disable shows that they are enabled.
   2. Verify that **Associated VM** is empty.

3. Select the device rows.

   The **Disable Device** option appears in the upper toolbar.

4. From the upper toolbar, select **Disable Device**.
5. Confirm that the selected devices now appear as disabled.

16.2 Recommended best practices

- Review **Associated VM** before you try to disable a device.
- Review **Enabled** before you run an enable or disable action.
- Review **Blocked** before you make a change so you can confirm the current device state.
- Review **Node** and **PCI Slot** when you need to identify a device on a specific node.
- Use the single-device actions for isolated changes.
- Use the batch actions only after you review all selected rows.
- Use sorting and the column picker to simplify large device lists.

16.3 Troubleshooting

16.3.1 No devices appear in the PCI devices list

Possible cause:

- No PCI devices are currently available.

What to check:

1. Open **Configuration** > **PCI Devices**.
2. Check whether the screen shows **There are no PCI Devices defined**.
3. If that message appears, confirm that PCI passthrough devices are available in the region.

16.3.2 The Enable Device action is unavailable

Possible cause:

- The selected device is already enabled.

What to check:

1. Review the **Enabled** value for the selected device.
2. If the device is already enabled, no enable action is needed.

16.3.3 The Disable Device action is unavailable

Possible causes:

- The selected device is not enabled.
- The selected device is associated with a VM.

What to check:

1. Review the **Enabled** value for the selected device.
2. Review **Associated VM** for the selected device.
3. If a VM is associated with the device, clear that association before you try to disable the device in this screen.

16.3.4 A batch action does not change every selected device

Possible cause:

- One or more selected devices are not eligible for the requested action.

What to check:

1. For batch enable, confirm that the selection includes disabled devices.
2. For batch disable, confirm that the selection includes enabled devices.
3. For batch disable, confirm that **Associated VM** is empty for the devices that you expect to disable.

CHAPTER

SEVENTEEN

SWITCH DOMAINS

A switch domain is a physical Layer 2 network representation in zCompute. Every switch domain may have a set of VLAN IDs, cluster networks and nodes' NICs attached to it.

Each switch domain has a Maximum Transmission Unit (MTU). The MTU defines the largest packet size allowed for networks in the domain.

Note

As of zCompute 25.06, it's not possible to add or remove a switch domain from the system.

Administrators use switch domains to:

- Separate access, storage, or data networks.
- Apply different MTU sizes for different traffic types.
- Control VLAN allocation and ownership.
- Monitor node network links.

Open the switch domains page from **Region Networking > Switch Domains**.

The **Switch Domains** page lists all switch domains in the region.

A switch domain has these core properties:

- **Name**

  A descriptive name for the switch domain.

- **MTU**

  The Maximum Transmission Unit (MTU) for the switch domain.

- **ID**

  The unique identifier for the switch domain.

17.1 Switch Domains management

This section describes the supported tasks and their steps.

17.1.1 View the Switch Domains list

Use the switch domains list to find a switch domain and confirm its MTU.

1. Go to **Region Networking**.
2. Select **Switch Domains**.

Note

This view is read-only.

It does not include actions.

Fields

- **Name**

  The switch domain name. Select it to open the domain.

- **MTU**

  The MTU value for the switch domain.

- **ID**

  The unique identifier for the switch domain. This column is hidden by default. Use the column picker to show it.

Note

- **Filtering**

  You can filter the table by using the filter input.

- **Export Spreadsheet**

  You can export the table by using **Export Spreadsheet**. The export creates a `.csv` file.

17.1.2 View a switch domain details page

Use the details page to manage VLANs and cluster networks.

1. Go to **Region Networking**.
2. Select **Switch Domains**.
3. In **Switch Domains**, select a domain name.

   The selected switch domain's details page opens.

   Each switch domain details page includes these tabs:

   - **Node Links**

     Read-only link information for nodes in the switch domain.

   - **Cluster Networks**

     Networks that use a VLAN and a subnet in this domain.

   - **VLANs**

     VLAN records for this domain, including allocation status.

   Use switch domain's details page to answer common questions:

- Which switch domains are configured for this region?
- What MTU does each switch domain use?
- Which nodes are linked to the switch domain?
- Which VLANs are free, reserved, or allocated?
- Which cluster networks exist in the switch domain?

View Node Links

View the **Node Links** tab to validate that nodes are connected to the domain.

1. Go to **Region Networking**.
2. Select **Switch Domains**.
3. Select a switch domain to view.

   The switch domain detail page displays.
4. Select the **Node Links** tab.

   Note

   This tab is read-only. It does not include actions.

   A list of node links is displayed in a table with the following columns:

   - **Node Name**

     The node name. Select it to open the node details page.
   - **MAC Addresses**

     MAC addresses used by this link.
   - **Admin State**

     The configured state of the link.
   - **Operational State**

     The current runtime state of the link.
   - **Vendor**

     The NIC vendor, if detected.
   - **Model**

     The NIC model, if detected.
   - **Bond Type**

     The bond mode for the link, if used.

Managing Cluster Networks

A switch domain's cluster networks can be created, viewed and deleted.

Viewing Cluster Networks

View the **Cluster Networks** tab's list to see the cluster networks in the switch domain.

1. Go to **Region Networking**.
2. Select **Switch Domains**.
3. Select a switch domain to view.

   The switch domain detail page displays.
4. Select the **Cluster Networks** tab.

   A list of cluster networks is displayed in a table with the following columns:

   - **Name**

     The cluster network name. Select it to open the network.
   - **VLAN ID**

     The VLAN tag used by the cluster network.
   - **MTU**

     The MTU configured for the cluster network.
   - **Subnet CIDR**

     The network range for the cluster network.
   - **Connected Nodes**

     The number of nodes connected to the cluster network.
   - **Switch Domain**

     The switch domain name for the cluster network.

Note

Use the column picker to show more fields, such as Nameservers.

Creating a Cluster Network

Create a cluster network to define subnet and VLAN settings for the switch domain's traffic.

Note

- A cluster network must belong to a specific Layer 2 switch domain.
- A cluster network consumes a VLAN in that switch domain.

To create a cluster network in a switch domain:

1. Go to **Region Networking**.
2. Select **Switch Domains**.
3. Select a switch domain.

   The switch doman's detail page opens.
4. Select the **Cluster Networks** tab.

5. Select **+ Create** from the tab's toolbar.

   The **Create Switch Domain** dialog opens in the **Info** step.

   1. Enter the **Info** details:

      1. Enter a **Name**.
      2. Select or create a **VLAN**.
      3. Enter a **Subnet CIDR**.
      4. Confirm or set the **MTU**.

   2. Click **Next** to configure IPs.

      The **IPs** dialog step opens.

   3. Enter the **IPs** details:

      For each IP Range enter:

      - **Start** IP address of the range.
      - **End** IP address of the range.
      - **Select Add IP Range to enter the Start and End IP** addresses of an additional IP range.

   4. Select **Next** to configure routes.

      The **Routes** dialog step opens.

   5. Enter the **Routes** details:

      - **Destination**: The destination network in CIDR format (for example, 0.0.0.0/0).
      - **Next Hop**: The IPv4 address of the next-hop gateway.
      - Select **Add Route** to enter the **Destination** and **Next Hop** an additional route.

   6. Select **Next** to configure **Nameservers**.

      The **Nameservers** dialog step opens.

   7. Enter the **Nameservers** details:

      Define DNS servers for the cluster network:

      - **Nameserver**: Enter an IPv4 address or DNS name.
      - Select **Add Nameserver** to configure an additional nameserver for the cluster network.

6. To save the cluster network configuration, select **Finish**.

After creation, verify:

- The network appears in the list.
- The VLAN is allocated.
- The connected nodes count is correct.

Deleting a Cluster Network

Delete a cluster network to remove an unused network.

Caution

Deleting a cluster network can disrupt traffic.

1. Go to **Region Networking**.
2. Select **Switch Domains**.
3. Select a switch domain.

   The switch doman's detail page opens.
4. Select the **Cluster Networks** tab.
5. Select the cluster network row to delete.
6. Select **Delete**.

   The **Delete Cluster Network** dialog opens, displaying the name of the cluster network to delete.
7. To confirm the deletion, select **Ok**.

   The cluster network is deleted and no longer displayed in the cluster networks table of the switch domain.

Managing VLANs

A switch domain's VLANs can be created, viewed and deleted.

Viewing VLANs

Use the VLAN list to understand VLAN availability in the domain.

1. Go to **Region Networking**.
2. Select **Switch Domains**.
3. Select a switch domain to view.

   The switch domain detail page displays.
4. Select the **VLANs** tab.

   A list of VLANs is displayed in a table with the following columns:

   - **Status**

     The VLAN status. Use it to find free VLANs.
   - **VLAN ID**

     The VLAN tag value (VID).
   - **Owner Name**

     The object that owns the VLAN. Select it to open details.
   - **Owner Type**

     The owner type.
5. Optional: Use a quick filter:
   - Free
   - Reserved
   - Allocated

Creating a VLAN

Create VLANs so you have free VLANs available for new networks.

1. Go to **Region Networking**.
2. Select **Switch Domains**.
3. Select a switch domain to view.

   The switch domain detail page displays.
4. Select the **VLANs** tab.

   The list of VLANs is displayed.
5. Select **Create**.
6. In **Create VLAN**, enter the VLAN ID.
7. Select **+ Create** from the tab's toolbar.

   The **Create VLAN** dialog opens.
8. Enter the **VLAN ID**:

   The VLAN tag value (VID).

   Valid range is 0 to 4094.
9. Select **Create**.

   The new VLAN appears in the VLAN table with the status **free**.

Deleting a VLAN

Delete a VLAN to remove an unused VLAN record.

Only unallocated VLANs (status is **free**) can be deleted.

Caution

Deleting a VLAN can affect automation that depends on it.

1. Go to **Region Networking**.
2. Select **Switch Domains**.
3. Select a switch domain to view.

   The switch domain detail page displays.
4. Select the **VLANs** tab.

   The list of VLANs is displayed.
5. Select the VLAN to delete.

   Ensure that the VLAN is not allocated, and that its status is **free**.
6. Select **Delete** from the **VLANs** tab's toolbar.

   The **Delete VLAN** dialog opens, displaying the selected VLAN, its ID and status.
7. Review the confirmation details.

8. To confirm deleting the VLAN, select **Ok**.

   The VLAN no longer displays in the switch domain's **VLANs** tab's list.

17.2 Recommended Best Practices

For any desired change, consult first with Zadara support. These switch domains are predefined, and tightly-coupled with specific functionality.

17.3 Troubleshooting

17.3.1 Cluster Network Not Reachable

Check:

- VLAN ID matches physical switch configuration.
- Subnet CIDR is correct and not overlapping.
- Node link operational state is **up**.
- MTU matches across the path.

17.3.2 MTU Related Issues

Symptoms:

- Intermittent connectivity.
- Large packet loss.
- Storage latency.

Actions:

- Confirm switch domain MTU.
- Validate physical switch MTU.
- Test with standard MTU 1500 if unsure.

17.3.3 VLAN Allocation Errors

If a VLAN cannot be allocated:

- Confirm the VLAN ID is not already in use.
- Check the VLAN list in the domain.
- Verify you are creating the network in the correct switch domain.

17.3.4 Node Link Down

If operational state is down:

- Verify cabling.
- Check LACP configuration on the physical switch.
- Confirm the correct bond type.
- Review node NIC status.

17.3.5 There are no switch domains

- **Symptom**

  The list shows an empty state message.

- **Possible actions**

  Confirm your account has access to Region Networking.

  If the issue persists, check the back-end service state.

17.3.6 No VLANs are available in the VLAN field

- **Symptom**

  The VLAN field shows no available VLANs.

- **Why this happens**

  The VLAN selector lists only VLANs with Status **Free**.

- **Possible actions**

  - Create a VLAN in the **VLANs** tab.
  - Free an existing VLAN by removing its owner.

17.3.7 Cannot delete a VLAN

- **Symptom**

  The Delete action is disabled.

- **Why this happens**

  You can delete a VLAN only when its Status is **Free**.

- **Actions**

  1. Identify the owner.
  2. Remove the owner first.
  3. Retry the delete.

17.3.8 General Validation Checklist

- Switch domain exists and is correct.
- Cluster network is assigned to the correct domain.
- VLAN ID is correct.
- Subnet does not overlap.
- MTU is consistent end to end.
- Node links are operational.

If issues persist, collect:

- Switch domain ID.
- VLAN ID.
- Node name and MAC addresses.

- MTU configuration.

Provide this data to support for further analysis.

CHAPTER

EIGHTEEN

CLUSTER NETWORKS

A cluster network defines Layer 2 and Layer 3 connectivity within a switch domain.

You use a cluster network to control VLAN tagging, IP addressing, MTU, and routing for cluster traffic.

Cluster networks are scoped to a single switch domain. This ensures traffic isolation and clear operational boundaries.

18.1 Cluster Networks Management

18.1.1 View Cluster Networks

The Cluster Networks page displays all defined cluster networks.

To view cluster networks:

1. Go to **Region Networking**.
2. Select **Cluster Networks**.

The list view includes the following columns:

- **Name**

  The unique name of the cluster network.

- **VLAN ID**

  The VLAN tag associated with the network.

- **MTU**

  The maximum transmission unit in bytes.

- **Subnet CIDR**

  The IPv4 subnet in CIDR notation.

- **Connected Nodes**

  The number of nodes attached to the network.

- **Switch Domain**

  The switch domain that owns the network.

Select a cluster network name to open its details page.

View a Cluster Network's details

1. Go to **Region Networking**.
2. Select **Cluster Networks**.
3. Select a cluster network name to open its details page.

Each cluster network is associated with:

- **Name**

  A descriptive identifier of the cluster network. This value is defined during creation and is displayed in list and detail views.

- **VLAN ID**

  The VLAN tag assigned to the cluster network. This value identifies the Layer 2 broadcast domain used for traffic isolation.

- **MTU**

  The maximum transmission unit, in bytes. This value defines the largest packet size allowed on the network.

- **Connected Nodes**

  The number of nodes currently attached to the cluster network. This value indicates active participation in the network.

- **Nameservers**

  DNS server entries assigned to the cluster network. These values define how hosts in the network resolve domain names to IP addresses. Each entry can be an IPv4 address or a DNS hostname.

- **Switch Domain** The switch domain that contains the cluster network. This value determines the logical networking boundary.

- **ID**

  The unique identifier of the cluster network. This value is defined during creation and is displayed in list and detail views.

- **Subnet CIDR**

  The IPv4 subnet in CIDR notation. This value defines the address space available for IP allocations.

A cluster network can also include associated components, displayed in the following tabs:

- **Nodes** tab

  Displays nodes connected to the cluster network, including their IP address, Node ID, admin state, and operational state.

- **Routes** tab

  Displays configured static routes for the cluster network, including destination networks and next-hop addresses.

- **Applications** tab

  Displays applications associated with the cluster network, including application type, switch domain, and linked cluster network.

- **VIPs** tab

  Displays virtual IP resources attached to the cluster network, including direction and sharing status.

- **IP Addresses** tab

  Displays individual IPv4 addresses and address ranges defined for the cluster network, including allocation status and ownership details.

18.1.2 Create a Cluster Network

Creating a cluster network defines connectivity parameters for a specific switch domain.

To create a cluster network:

1. Go to **Region Networking**.
2. Select **Cluster Networks**.
3. Select **Create**.
4. Complete the wizard steps.

Info Step

The Info step defines core network properties.

Fields:

- **Name** Enter a unique name for the cluster network.
- **Switch Domain** Select the switch domain that will contain the network. Only one switch domain can be selected.
- **VLAN** Select an existing VLAN. You can also create a new VLAN from this field.
- **Subnet CIDR** Enter the IPv4 subnet in CIDR format (for example, 10.16.24.0/21).
- **MTU** Enter a value between 1500 and 9216. The value is typically aligned with the switch domain.

IP Ranges Step

Define IP allocation ranges for the subnet.

You can:

- Create individual IP addresses
- Create IP address ranges

Create IP Address Range dialog fields:

- **Start IP** The first IPv4 address in the range.
- **End IP** The last IPv4 address in the range.

Create IP Address dialog fields:

- **IP Address** A single IPv4 address inside the subnet.

Routes Step

Define static routes for the cluster network.

Create Route dialog fields:

- **Destination** The destination network in CIDR format (for example, 0.0.0.0/0).
- **Next Hop** The IPv4 address of the next-hop gateway.
- **Policy Routing Table** The routing table identifier. The default value is 0.

Nameservers Step

Define DNS servers for the cluster network.

Fields:

- **Nameserver** Enter an IPv4 address or DNS name.

After completing all steps, submit the wizard to create the cluster network.

18.1.3 Update a Cluster Network

Updating a cluster network modifies its configuration.

To update a cluster network:

1. Go to **Cluster Networks**.
2. Select the required cluster network.

   The details screen opens.

   Some of the cluster network tabs permit updating associated components, as follows:

Tab	Update existing	Add new	Delete existing
Nodes	✗		
Routes	✗	✓	✓
Applications	✗	✓	✓
VIPs	✓		
IP Addresses	✗	✓	✓

View Nodes

The **Nodes** tab displays all nodes connected to the cluster network.

Note

There are no controls for MSP admins to create, modify or delete a cluster network's nodes.

To view the Nodes tab:

1. Go to **Region Networking**.
2. Select **Cluster Networks**.
3. Select the required cluster network name.
4. Select the **Nodes** tab.

The table lists the following details for each associated node:

- **Node Name** The display name of the connected node. This value identifies the node in the environment.
- **IP Address** The IPv4 address assigned to the node within the cluster network.
- **Node ID** The unique identifier of the connected node.
- **Admin State** The administrative state of the node interface, such as enabled or disabled.
- **Operational State** The runtime state of the node interface, such as up or down.

View or Manage Routes

The **Routes** tab displays all routes configured for the selected cluster network.

Each entry represents a routing rule that directs traffic to a specific destination network.

The Routes tab allows you to review existing routing rules and verify traffic forwarding behavior for the cluster network.

To view the Routes tab:

1. Go to **Region Networking**.
2. Select **Cluster Networks**.
3. Select the required cluster network name.
4. Select the **Routes** tab.

The table lists the following details for each associated route:

- **Destination** The target network in CIDR format. This defines the address range that the route applies to.
- **Next Hop** The IPv4 address of the gateway used to reach the destination network.

Creating a new route

To create a new route for the selected cluster network:

1. From the **Routes** tab's toolbar, select **+ Create**.

   The **Create Route** dialog opens.

2. Enter the required fields.

   - **Destination**

     Enter the destination network in CIDR format (for example, 0.0.0.0/0). This defines the network that the route will match.

   - **Next Hop**

     Enter the IPv4 address of the gateway. This is the address used to forward traffic to the destination.

   - **Policy Routing Table**

     Enter the routing table identifier. This value determines which routing table processes the route.

     The default value is 0.

3. Select **Create** to save the route.

Deleting a route

Caution

Removing a route can immediately stop traffic to the destination network. Services that depend on that route can become unreachable.

Before deleting a route:

- Verify the destination network is no longer required.
- Confirm that another valid route exists, if needed.
- Assess the impact on connected applications and VIPs.
- Schedule the change during a maintenance window if production workloads coulbd be affected.

To delete an existing route configured for the selected cluster network:

1. In the selected cluster network's **Routes** tab, select the route to delete.
2. From the **Routes** tab's toolbar, select **Delete**.

   The **Delete Route** dialog opens, displaying the selected route's ID.
3. Select **Delete** to confirm deleting the route.

   The route no longer appears in the **Routes** tab list.

View or Manage Applications

The **Applications** tab displays all applications that are attached to the selected cluster network.

Each entry represents a network-based service that uses the cluster network for connectivity.

Use this tab to review which services depend on the cluster network before making configuration changes, such as route modifications or deletion.

To view the Applications tab:

1. Go to **Region Networking**.
2. Select **Cluster Networks**.
3. Select the required cluster network name.
4. Select the **Applications** tab.

The table lists the following details for each associated application:

- **Name**

  The application name. This identifies the service in the environment.
- **Type**

  The application type, such as `network` or `virtual_ip`. This indicates the function of the application.
- **Switch Domain**

  The switch domain associated with the application. This confirms the logical networking boundary.
- **Cluster Network**

  The cluster network to which the application is attached. This shows the connectivity context for the service.

Creating a new application

To create a new application for the selected cluster network:

1. From the **Applications** tab's toolbar, select **+ Create**.

   The **Create Application** dialog opens.
2. Enter the required fields.
   - **Name**

     Enter a unique name for the application.

- **Application Type**

  The application type determines the behavior of the application.

  Select the application type.

  Note

  Currently, for cluster networks, MSP admins can select only **Network** as the application type.

3. Select **Create** to save the application.

Deleting an application

Deleting an application removes the service from the cluster network.

Caution

Considerations Before Deleting

Review the following before proceeding:

- Service Availability

  Deleting the application immediately stops theassociated service. Clients could lose connectivity.

- VIP Dependencies

  If the application is associated with a virtual IP, traffic directed to that VIP could fail.

- IP Address Allocation

  The application might own allocated IP addresses. Confirm whether those IPs should be released or reassigned.

- Route Dependencies

  The application could rely on specific static routes.

  Removing the application does not remove routes.

- Change Management

  Perform deletion during a maintenance window if production workloads could be affected.

After deletion, verify that:

- Required services are still reachable.
- No unintended IP allocations remain.
- Monitoring systems reflect the change.

To delete an existing application configured for the selected cluster network:

1. In the selected cluster network's **Applications** tab, select the application to delete.
2. From the **Applications** tab's toolbar, select **Delete**.

   The **Delete Application** dialog opens, displaying the selected application's name.
3. Select **Delete** to confirm deleting the application.

   The application no longer appears in the **Applications** tab list.

View or Manage VIPs

From the cluster network details page, select **VIPs** to view virtual IP resources.

VIPs table columns include:

- **Name** The VIP name.
- **Switch Domain** The associated switch domain.
- **Cluster Network** The linked cluster network.
- **Direction** Traffic direction, such as inbound or outbound.
- **Shared** Indicates whether the VIP is shared.

The **VIPs** tab displays all virtual IP resources that are attached to the selected cluster network.

Each entry represents a logical IP address that can be used to expose or receive traffic for a service.

The table includes:

- **Name**

  The name of the virtual IP resource. This identifies the VIP in the environment.

- **Switch Domain**

  The switch domain that contains the VIP. This confirms the networking boundary.

- **Cluster Network**

  The cluster network to which the VIP is attached. This defines the subnet and VLAN context used by the VIP.

- **Direction**

  The traffic direction associated with the VIP, such as inbound or outbound. This defines how the VIP handles traffic flow.

- **Shared**

  Indicates whether the VIP is shared across multiple services or dedicated to a single service.

Use the VIPs tab to review external exposure and service entry points before modifying routes, IP ranges, or deleting the cluster network.

Modifying a Virtual IP (VIP)

You can modify a virtual IP from the cluster network details screen.

Caution

Considerations Before Deleting

Modifying a VIP can affect traffic immediately.

- Disabling the VIP stops traffic to the IP address.
- Changing the direction can alter traffic behavior.
- Changing the associated service can redirect traffic.

After saving changes, verify that:

- The VIP responds as expected.
- Dependent services remain reachable.
- Monitoring reflects the updated state.

To modify a VIP:

1. Go to **Region Networking**.
2. Select **Cluster Networks**.
3. Select the required cluster network name.
4. Select the **VIPs** tab.
5. Select the required VIP.
6. Select **Modify**.

   The **Modify VIP** dialog opens.

   Some of the VIP fields can be modified:

   - **Name**

     The display name of the virtual IP. Update this value to change how the VIP appears in the system.

   - **Cluster Network** (Display only, cannot be modified)

     Displays the cluster network to which the VIP is attached. This value is read-only.

   - **IP Address**

     The IPv4 address assigned to the VIP. This value identifies the logical service endpoint. Update this value to assign a different IP address to the VIP.

   - **Enable**
     - **Select** to enable the VIP.
     - **Clear** to disable the VIP.

     Disabling a VIP stops traffic without deleting it.

   - **Direction**

     Defines how traffic is handled.

     Select the traffic direction:

     - **Inbound**
     - **Outbound**

   - **Service**

     This value links the VIP to a specific service.

     Enter or select the associated service.

   - **Shared**
     - **Select** to allow the VIP to be shared.
     - **Clear** to dedicate the VIP to a single service.

7. Select **Finish** to apply the changes.

View or Manage IP Addresses

The **IP Addresses** tab displays all IPv4 addresses defined for the cluster network.

This includes both allocated addresses and addresses that are available for allocation.

Each entry represents an IP resource within the subnet defined for the cluster network.

To view the cluster network's IP Addresses:

1. Go to **Region Networking**.
2. Select **Cluster Networks**.
3. Select the required cluster network name.
4. Select the **IP Addresses** tab.

   The table of IP Addresses displays all IPv4 addresses defined for the cluster network.

   This includes both allocated addresses and addresses that are available for allocation.

   Use this tab to review address usage, verify allocations, and confirm that sufficient IP capacity remains available within the subnet.

   Each entry in the table represents an IP resource within the subnet defined for the cluster network:

   - **IPv4 Address**

     The IP address assigned within the cluster network subnet.

   - **Node ID**

     The identifier of the node associated with the IP address, if the address is allocated.

   - **Status**

     The allocation state of the IP address, such as **allocated** or **free**.

   - **Owner ID**

     The identifier of the resource that owns the IP address.

   - **Owner Type**

     The type of resource that owns the IP address, such as an application or network interface.

Creating new IP Addresses

You can add individual IP addresses or IP address ranges to a cluster network from the **IP Addresses** tab.

Note

Considerations

- The IP address must belong to the cluster network subnet.
- The IP address must not already be allocated.
- Avoid adding addresses that overlap with existing ranges.

After adding addresses, verify that the new entries appear in the **IP Addresses** table.

To add IP addresses:

1. Go to **Region Networking**.

2. Select **Cluster Networks**.
3. Select the required cluster network name.
4. Select the **IP Addresses** tab.
5. Select one of:
    - **+ Create** to add a single IP address.

      The **Create IP Address** dialog opens.

        - **IP Address**

      Enter a single IPv4 address within the subnet defined for the cluster network. The address must be unique and must belong to the subnet CIDR of the network.
    - **+ Create Range** to add a range of addresses.

      The **Create IP Address Range** dialog opens.

        - **Start IP**

          Enter the first IPv4 address in the range. This value defines the beginning of the address block.
        - **End IP**

          Enter the last IPv4 address in the range. This value defines the end of the address block.
6. Select **Create** to add the address or address range.

Deleting an IP Address

You can delete individual IP addresses from a cluster network.

Caution

Considerations for deleting IP Addresses

Deleting an IP address removes it from the cluster network address pool.

Review the following before deleting:

- **Allocated Addresses**

  Do not delete an IP address that is currently allocated to a node, application, or VIP.
- **Service Impact**

  If a resource depends on the IP address, deleting it can interrupt connectivity.
- **Address Pool Capacity**

  Deleting address ranges reduces the number of IP addresses available for allocation.

After deletion, verify that:

- The address no longer appears in the **IP Addresses** table.
- Dependent services continue to function correctly.

To delete an IP address from a cluster network:

1. Go to **Region Networking**.
2. Select **Cluster Networks**.

3. Select the required cluster network name.
4. Select the **IP Addresses** tab.
5. Select the IP address to remove.
6. Select **Delete**.

   The **Delete IP Address** dialog opens, displaying the IP address and its ID.
7. Select **Ok** to confirm deleting the IP address.

18.1.4 Delete a Cluster Network

Deleting a cluster network removes it from the switch domain.

Before deletion, ensure no dependent resources remain attached to it.

Dependencies can include:

- Nodes
- Routes
- Applications
- VIPs
- IP allocations

To delete a cluster network:

1. Go to **Cluster Networks**.
2. Select the required network.
3. Select **Delete**.
4. Confirm the action.

18.2 Recommended Best Practices

- Use clear, descriptive names for cluster networks.
- Separate environments by using different switch domains.
- Avoid overlapping subnet CIDRs across cluster networks.
- Align MTU values with the physical network design.
- Define IP ranges before deploying applications.
- Review static routes carefully before applying them.
- Remove unused cluster networks to reduce complexity.

18.3 Troubleshooting Cluster Networks

18.3.1 Cluster Network Not Visible

- Verify you are in the correct region.
- Confirm the correct switch domain is selected.
- Check user permissions.

18.3.2 IP Allocation Fails

- Ensure the IP is inside the defined subnet CIDR.
- Confirm the IP is not already allocated.
- Verify the defined IP ranges include the address.

18.3.3 Route Does Not Work

- Validate the destination CIDR format.
- Confirm the next-hop IP is reachable.
- Check for conflicting routes.

18.3.4 Connectivity Issues

- Verify VLAN ID configuration.
- Confirm MTU consistency across nodes.
- Review nameserver entries.
- Check switch domain boundaries.

Use the cluster network overview page to verify:

- VLAN ID
- Subnet CIDR
- MTU

Correct configuration mismatches and retest connectivity.

CHAPTER

NINETEEN

NETWORKING APPLICATIONS

Networking applications define logical networking functions within a cluster network. They represent the networking purpose that a workload or service uses inside the regional network.

Each networking application maps networking behavior to a specific cluster network. This mapping allows the platform to apply the correct network domain, routing, and connectivity rules.

Networking applications appear in the **Region Networking > Networking Applications** page. The page lists all applications that exist in the region.

Each entry contains the following information:

- **Name**

  The unique descriptive name of the networking application.

- **Type**:

  The application networking type, that represents the networking behavior of the service.

  Examples include standard network connectivity and virtual IP services.

- **Switch Domain**

  The switch domain that routes the networking application traffic.

- **Cluster Network**

  The cluster network where the networking application operates.

Networking applications are created and managed by administrators to organize cluster networking and to ensure consistent connectivity policies across workloads.

19.1 Networking Applications Management

Administrators manage networking applications from the **Networking Applications** page in the Region Networking section.

The page displays the list of existing applications and provides the main lifecycle operations, such as creating, viewing and deleting a networking application.

19.1.1 Viewing Networking Applications

The networking applications table provides a view of all configured applications.

To view networking applications:

1. Open **Region Networking**.
2. Select **Networking Applications**.

The table displays the following properties for each entry:

- **Name**

  The unique descriptive name of the networking application.

- **Type**:

  The application networking type, that represents the networking behavior of the service.

  Examples include standard network connectivity and virtual IP services.

- **Switch Domain**

  The switch domain associated with the networking application.

- **Cluster Network**

  The cluster network where the networking application operates.

This view helps administrators verify the networking layout of the cluster.

Viewing a Network Application's details

The **Application Details** page provides detailed information about a specific networking application. Administrators use this page to review application configuration and to view the firewall rules associated with the application.

Applications represent networking functions that operate on a cluster network. Each application is linked to a cluster network and inherits the network behavior defined for that network.

To view details of a networking application:

1. Open **Region Networking**.
2. Select **Networking Applications**.

   The page displays the list of existing applications.

3. Select a networking application to view its details.

   The networking application details screen opens with the following sections:

   - **Info** section:

     Displays basic configuration properties for the application:

     - **Enabled/Disabled** status, indicating the networking application's availability.
     - **ID**

       The unique identifier assigned to the application.

       This value is generated by the system when the application is created. Administrators typically use this identifier when referencing the application in system operations or when reviewing configuration data.

     - **Name**

       The descriptive name of the application.

       The name identifies the networking function associated with the application. Administrators should use clear and descriptive names to simplify network management.

     - **Type**

       The networking type assigned to the application.

       The type indicates the role that the application performs within the cluster networking model.

       Example types :

       * **network**

    * **virtual_ip**

  - **Cluster Network ID**

    The identifier of the cluster network associated with the application.

    Each application operates on a specific cluster network. The cluster network defines the traffic role used by the application.

  - **Enabled**

    Indicates whether the application is currently enabled.

    An enabled application allows network behavior and firewall rules associated with the application to be applied.

Network applications of type **virtual_ip** expose a service through a virtual IP endpoint. In addition to the standard application properties, the details page includes several fields specific to virtual IP behavior.

These fields describe how the virtual IP is used, the service it represents, and whether the address can be shared across multiple services:

  - **Direction**

    Indicates the traffic direction associated with the virtual IP application.

    The direction defines whether the virtual IP accepts traffic from external systems or handles internal traffic within the cluster network.

    This property helps determine how the virtual IP participates in the networking flow and which services or clients are expected to connect to it.

  - **Service**

    Identifies the service associated with the virtual IP.

    The service value represents the networking function or platform service that is exposed through the virtual IP. Clients use the virtual IP address to reach the service instead of connecting directly to individual nodes.

    Associating a service with a virtual IP provides a stable endpoint that remains consistent even if the underlying service location changes.

  - **Shared**

    Indicates whether the virtual IP address can be shared by multiple services or applications.

    When enabled, the same virtual IP can be used by more than one service. This allows the platform to expose multiple services through a single logical address when required by the network design.

    When **Shared** is disabled, the virtual IP is dedicated to a single service.

  - **IP Address**

    Displays the virtual IP address assigned to the application.

    This address represents the network endpoint used by clients to access the associated service. The virtual IP remains stable even if the service moves between nodes or instances within the cluster.

    Administrators use this address when configuring clients, external systems, or network integrations that must reach the service.

- **Firewall Rules** section:

  Lists the network protocols and ports allowed for the application.

  Firewall rules control which types of traffic can access services that use the application.

  Each rule defines the protocol and port used for communication.

Displayed fields:

- **Protocol**

  The network protocol allowed by the firewall rule.

  Examples:

  * **tcp**
  * **icmp**

- **Port**

  The network port allowed for the protocol.

  These ports represent services that are reachable through the application.

  Example interpretation:

  * **22 (TCP)** commonly represents SSH access.
  * **80 (TCP)** commonly represents HTTP traffic.
  * **443 (TCP)** commonly represents HTTPS traffic.

  Firewall rules allow administrators to restrict network access so that only required protocols and ports are permitted.

19.1.2 Creating a Networking Application

Creating an application defines a new networking function for the cluster. This allows services or workloads to use the appropriate network connectivity.

To create a networking application:

1. Open **Region Networking**.
2. Select **Networking Applications**.

   The page displays the list of existing applications.
3. Select **+Create**.

   The **Create Application** dialog opens.
4. In the **Create Application** dialog, enter the required values.

   - **Name** – Enter a uniquie descriptive name for the application.
   - **Cluster Network** – Select the cluster network that will host the application.
   - **Application Type** – Select the networking type.

     Currently, MSP admins can only select **Network** for networking applications.
5. Select **Finish** to save the configuration.

The application appears in the networking applications list after creation.

19.1.3 Deleting a Networking Application

Deleting an application removes it from the regional networking configuration. This operation should be performed only when the application is no longer required.

To delete a networking application:

1. Open **Region Networking**.

2. Select **Networking Applications**.

   The page displays the list of existing applications.

3. Select the networking application to delete.

4. Select **Delete** from the top toolbar.

   The **Delete Application** dialog opens, displaying the name of the networking application selected to be deleted.

5. To confirm deletion, select **Ok**.

   The application is removed from the configuration and is no longer displayed in the networking applications list.

19.2 Recommended Best Practices

19.2.1 Use consistent naming conventions

Use clear and descriptive names for applications. This makes it easier to identify the networking purpose and reduces the risk of configuration errors.

19.2.2 Align applications with cluster network design

Create applications that reflect the network roles within the cluster. Examples include access, control, or data networks.

A structured approach simplifies troubleshooting and improves operational clarity.

19.2.3 Avoid unnecessary applications

Create applications only when they represent a real networking function. Excess applications increase administrative complexity and can make troubleshooting more difficult.

19.2.4 Validate cluster network selection

Confirm that the correct cluster network is selected when creating an application. Selecting the wrong network can prevent services from communicating correctly.

19.2.5 Review switch domain mappings

Verify that the application is associated with the correct switch domain. Incorrect domain mapping can affect connectivity and routing.

19.3 Troubleshooting

19.3.1 Application does not appear in the list

Possible causes include:

- The application was not successfully created.
- The page has not refreshed after the operation.

Possible actions:

- Refresh the page.
- Reopen the **Networking Applications** view.

19.3.2 Incorrect network connectivity

Services may fail to communicate if the application is associated with the wrong cluster network.

Possible actions:

- Verify the **Cluster Network** value in the application entry.
- Confirm the network role assigned to the application.

19.3.3 Unexpected switch domain assignment

If an application appears under the wrong switch domain, traffic may not follow the intended network path.

Possible actions:

- Review the application configuration.
- Confirm that the correct cluster network was selected during creation.

19.3.4 Application cannot be deleted

Deletion may fail if the application is still in use by active services.

Possible actions:

- Identify services that depend on the application.
- Remove or reconfigure those services before deleting the application.

CHAPTER

TWENTY

VIPS (VIRTUAL IPS)

A Region (Cluster) Network VIP (or Virtual IP) is an IP address servicing inbound traffic of a specific network application or function. It is virtual because it can reside on any physical node that also functions as a Control Node.

In Region Networking, VIPs have a dedicated page at **Region Networking > VIPs**.

This page gives MSP admins a focused view of VIP objects. It helps admins find service-facing IP entries without opening the full Networking Applications page.

20.1 VIPs management

20.1.1 Viewing VIPs

To view VIPs:

1. Go to **Region Networking**.
2. Select **VIPs**.

The VIPs list shows these columns:

- **Name**

  The VIP object's descriptive name.

- **Switch Domain**

  The switch domain that contains the VIP context.

- **Cluster Network**

  The cluster network linked to the VIP.

- **Service**

  The service value shown for the VIP.

- **Direction**

  The traffic direction shown for the VIP.

- **Shared**

  The shared state shown for the VIP.

Viewing a VIP's details

To view a VIP's details:

1. Open **Region Networking > VIPs**.
2. Select the VIP name.

   The VIP details screen opens.

   The details page shows the selected VIP's configuration, with the following sections:

   - **Info** section:

     Displays basic configuration properties for the application:

     - **Enabled/Disabled** status, indicating the VIP's availability.
     - **ID**

       The unique identifier assigned to the VIP.

       This value is generated by the system when the application is created. Administrators typically use this identifier when referencing the application in system operations or when reviewing configuration data.

     - **Name**

       The descriptive name of the VIP.

       The name identifies the function associated with the VIP.

     - **Type**

       The type is `virtual_ip`, that indicates the role that a VIP performs within the cluster networking model.

     - **Cluster Network ID**

       The identifier of the cluster network associated with the VIP.

     - **Enabled**

       Indicates whether the VIP is currently enabled.

       An enabled VIP allows network behavior and firewall rules associated with the VIP to be applied.

     - **Direction**

       Indicates the traffic direction associated with the VIP.

       The direction defines whether the VIP accepts traffic from external systems or handles internal traffic within the cluster network.

       This property helps determine how the VIP participates in the networking flow and which services or clients are expected to connect to it.

     - **Service**

       Identifies the service associated with the VIP.

       The service value represents the networking function or platform service that is exposed through the VIP. Clients use the VIP address to reach the service instead of connecting directly to individual nodes.

       Associating a service with a VIP provides a stable endpoint that remains consistent even if the underlying service location changes.

     - **Shared**

       Indicates whether the VIP address can be shared by multiple services or applications.

When enabled, the same VIP can be used by more than one service. This allows the platform to expose multiple services through a single logical address when required by the network design.

When **Shared** is disabled, the VIP is dedicated to a single service.

- **IP Address**

  Displays the VIP address assigned to the application.

  This address represents the network endpoint used by clients to access the associated service. The VIP remains stable even if the service moves between nodes or instances within the cluster.

  Administrators use this address when configuring clients, external systems, or network integrations that must reach the service.

- **Firewall Rules** section:

  Lists the network protocols and ports allowed for the VIP.

  Firewall rules control which types of traffic can access services that use the VIP.

  Each rule defines the protocol and port used for communication.

  Displayed fields:

  - **Protocol**

    The network protocol allowed by the firewall rule.

    Examples:

    * **tcp**
    * **icmp**

  - **Port**

    The network port allowed for the protocol.

    These ports represent services that are reachable through the VIP.

    Example interpretation:

    * **22 (TCP)** commonly represents SSH access.
    * **80 (TCP)** commonly represents HTTP traffic.
    * **443 (TCP)** commonly represents HTTPS traffic.

    Firewall rules allow administrators to restrict network access so that only required protocols and ports are permitted.

Use the details page to confirm the current state and network settings of the VIP.

20.1.2 Modifying a VIP

1. Open **Region Networking > VIPs**.
2. Select the VIP name.

   The VIP details screen opens.
3. On the top toolbar, select **Modify**.
4. Review the fields in the **Modify VIP** dialog.

   The **Modify VIP** dialog opens.

   Update the required values:

- **Name**

  The VIP name shown in the dialog.

- **Cluster Network**

  The related cluster network.

- **IP Address**

  The IP address value for the VIP.

- **Enable**

  The enable/disable control toggle for the VIP enabled state.

- **Direction**

  The direction value:

  - **Inbound**
  - **Outbound**

- **Service**

  Identifies the service associated with the VIP.

  The service value represents the networking function or platform service that is exposed through the VIP. Clients use the VIP address to reach the service instead of connecting directly to individual nodes.

- **Shared**

  The shared state control toggle indicating that the VIP address can be shared by multiple services or applications.

5. To save the changes, select **Finish**.

20.2 Recommended best practices

For any desired change, consult first with Zadara support. These VIPs are predefined, and tightly-coupled with specific functionality.

CHAPTER

TWENTYONE

GPU NETWORK SWITCHES

GPU network switches are high-performance networking platforms designed to accelerate AI, machine learning, and GPU computing workloads. They provide ultra-low latency, high-bandwidth per port, and in-network computing features to connect thousands of GPUs, preventing bottlenecks in data-intensive environments.

The zCompute GPU Network comprises NVIDIA Ethernet switches configured in a Spectrum-X Architecture. The GPU Network switches defined in zCompute are the leaf switches directly attached to the GPU's dedicated NICs (SuperNICs).

21.1 GPU network switches management

The GPU Network Switches page stores GPU network switch records in a central place.

You can use this page to review existing records, add a new record, update a record, or remove one.

21.1.1 Viewing GPU network switches

To view GPU network switches:

1. Go to **Region Networking > GPU Network Switches**.

   The GPU network switches list view displays.

   This view shows the current GPU network switch records with the the default columns:

   - **Name**

     The switch record name.

     Use it to identify the record in the list, in success or failure messages, and in the delete dialog.

   - **Status**

     The current state that the UI reports for the switch record.

     Use it for a quick state check from the list page.

   - **Hostname**

     The host address saved for the switch record.

     Use it to confirm that the record points to the intended switch host.

   - **API Port**

     The API TCP port saved for the switch record.

     Use it to confirm the port value used by the record.

- **Username**

  The saved user name for the switch record.

  Use it to confirm which account name the record is set to use.

- **Rail Group**

  The rail group value saved for the switch record.

  Use it to review the value that was set during creation.

- **SU ID**

  The scalable unit value saved for the switch record.

  Use it to review the value that was set during creation.

- **Router Distinguisher**

  The router distinguisher value saved for the switch record.

  Use it to verify the saved distinguisher from the list page.

- **Verify SSL**

  Indicates whether SSL verification is enabled for the switch record.

  Use it to confirm the setting that was chosen during creation.

You can also use the filter selection to display these optional columns:

- **ID**

  The record ID.

  Use it when you need the unique record value.

- **Created At**

  The record creation date and time.

  Use it for audit and tracking work.

- **Updated At**

  The record's last update date and time.

  Use it for audit and tracking work.

When no records exist, the page shows a message that no GPU network switches are defined.

Viewing GPU network switch details

To view a GPU network switch details:

1. Go to **Region Networking > GPU Network Switches**.

   The GPU network switches list view displays.

2. Select a GPU network switch to view its details.

   The GPU network switch details page opens.

21.1.2 Adding a GPU network switch

To add a GPU network switch:

1. Open **Region Networking > GPU Network Switches**.

   The GPU network switches list view displays.

2. Select **+ Add**.

   The **Add GPU Network Switch** dialog opens.

3. Enter new GPU network switch values:

   - **Name**

     The GPU network switch display name.

     Required.

   - **Hostname**

     The host address.

     Required.

   - **API TCP Port**

     The API port for the GPU network switch.

     Valid values: **1** to **65536**.

     Default: **8765**

   - **Verify SSL**

     Toggle to enable/disable SSL verification for the GPU network switch.

     Default: **Enabled**.

   - **Rail Group**

     Represents a collection of network switches that connect the same-index GPU from every server across multiple racks. For example, Rail Group 0 connects all GPU 0s, Rail Group 1 connects all GPU 1s, and so on.

     Valid values: 0 or higher.

     Required.

   - **Router Distinguisher**

     Used to create unique addresses from overlapping IP address spaces, ensuring they can be routed correctly across a virtualized network.

     Required.

   - **Scalable Unit**

     ID of a set of resources for rapid scaling purposes.

     Instead of building a node-by-node, organizations can use SUs to deploy a standardized set of resources simultaneously. This modularity allows for rapid deployment and ensures the resulting cluster is balanced for maximum performance.

     Valid values: **0** to **31**.

     Required.

- **Username**

  The username for GPU network switch.

  Required.

- **Password**

  The password for GPU network switch username.

  Required.

4. To save the configuration, select **Finish**.

21.1.3 Modify a GPU network switch

Update a switch record when connection details or credentials change:

1. Open **Region Networking > GPU Network Switches**.

   The GPU network switches list view displays.

2. Select a GPU network switch to modify.

   The GPU network switch details page opens.

3. Select **Modify**.

   The **Modify GPU Network Switch** dialog opens.

4. The following fields can be updated:

   - **Name**

     The GPU network switch display name.

   - **Hostname**

     The host address.

   - **API TCP Port**

     The API port for the GPU network switch.

   - **Router Distinguisher**

     IP address used to create unique addresses from overlapping IP address spaces.

   - **Username**

   Required. Use it to change the saved user name.

   - **Change Password**

     1. Select the checkbox to enable changing the password for GPU network switch username.

        The **Password** prompt appears.

     2. Enter a new password for the GPU network switch username.

5. To save the GPU network switch updates, select **Finish**.

21.1.4 Delete a GPU network switch

To remove a GPU network switch record when it is no longer needed.

1. Open **Region Networking > GPU Network Switches**.

   The GPU network switches list view displays.

2. Select a GPU network switch to delete.

   The GPU network switch details page opens.

3. Select **Delete**.

   The **Delete GPU Network Switch** dialog opens, displaying the name of the GPU network switch selected for deletion.

4. To confirm deletion of the selected GPU network switch, select **Delete**.

21.2 Recommended best practices

- Use a clear naming standard.

  The name appears in the list, in action messages, and in the delete flow.

- Check the hostname before you save the record.

  It is required and stays visible in the main table.

- Leave the API TCP port at the default only when that matches your switch configuration.

  The default value is **8765**.

- Set the **Verify SSL**, **Rail Group** and **Scalable Unit** values carefully during creation.

  The current UI displays these fields only in the **+ Add** dialog, and not when modifying or deleting.

- Review the router distinguisher value before saving a new entry or update.

  It is required in both **+ Add** and **Modify** dialogs.

- Change the password only when needed.

  In the **Modify** dialog, a new password key and value pair is setting with the update payload only when **Change Password** is selected.

- Use the optional columns during audits.

  ID, **Created At**, and **Updated At** can help with tracking.

21.3 Troubleshooting

21.3.1 The GPU Network Switches page does not appear in the menu

GPU Network Switches is part of the administrative Region Networking menu.

- Confirm that you are using an account with system admin access.
- Open **Region Networking** and check the left navigation again.

21.3.2 The GPU Network Switches list is empty

An empty list can be normal when no switch records exist.

- Check for the message that no GPU network switches are defined.
- Add a switch if this is a new setup.
- Refresh the page if you expect existing data.

21.3.3 The form does not let you finish

The dialog prevents completion when required fields are missing or invalid.

- Check **Name**.
- Check **Hostname**.
- Check **Router Distinguisher**.
- Check **Username**.
- Check **Password** in the **+ Add** dialog.
- Check **Rail Group** and **Scalable Unit** in the **+ Add** dialog
- Check the numeric range for **API TCP Port**.
- Check the numeric range for **Scalable Unit**.

21.3.4 You cannot change verify SSL, Rail Group, or Scalable Unit

This behavior is expected in the current UI.

The **Verify SSL**, **Rail Group**, and **Scalable Unit** fields are available only in the **+ Add** dialog.

Plan these values before you create the record.

21.3.5 Password is not updated

The modify dialog sends a password only when **Change Password** is selected.

- Reopen **Modify**.
- Select **Change Password**.
- Enter the new password.
- Click **Finish**.

21.3.6 A create, modify, or delete action fails

The UI defines success and failure messages for **+ Add**, **Modify** and **Delete** actions.

- Read the error message.
- Recheck the values you entered.
- Verify that you selected the intended switch.
- Retry the action after you correct the input.

CHAPTER

TWENTYTWO

GPU NETWORK PORTS

GPU Network Ports is a Region Networking page for system admins. Open it from **Region Networking > GPU Network Ports**.

The GPU Network Ports page shows GPU network port records in a central place. It helps you review port identity, network addressing, hardware mapping, and linked resources.

Use the GPU Network Ports page when you need to:

- Review GPU network port records across the region.
- Identify which VM uses a port.
- See which node and GPU network switch are linked to a port.
- Check addressing values such as port IP and peer IP.
- Review hardware values such as MAC address, PCI device ID, ENI device index, and rail ID.

22.1 GPU network ports management

22.1.1 Open the page

Use this page when you need to review GPU network port records.

1. In the left menu, open **Region Networking**.
2. Select **GPU Network Ports**.
3. Review the list of existing GPU network port records.

22.1.2 Review the list

The current UI provides a list view with a column picker.

1. Open **Region Networking > GPU Network Ports**.
2. Review the default columns.
3. Use the column picker when you need optional audit fields.
4. Select a linked VM, node, or GPU network switch when you need more details about a related resource.

22.1.3 Empty page behavior

When no records exist, the page shows an empty-state message.

1. Open **Region Networking > GPU Network Ports**.
2. Check whether the page shows that no GPU Network Ports are defined.

22.1.4 Linked views

This page includes links to related resources.

1. Select the **Associated VM** link to open the related VM.
2. Select the **Node** link to open the related node.
3. Select the **GPU Network Switch** link to open the related GPU network switch.

The same GPU network port collection can also appear in a VM page under its **Ports** tab.

22.2 List view fields

The list shows these default columns:

- **Name**

  Shows the GPU network port record name. Use it to identify the port in the list.

- **Switch Port Name**

  Shows the port name on the GPU network switch side. Use it to match the record to the switch-facing port name.

- **Associated VM**

  Shows the name of the linked VM when one is resolved in the UI. Use it to identify which VM is associated with the port. This value is a link to the VM details page.

- **Node**

  Shows the linked node name when one is resolved in the UI. Use it to identify the node related to the port. This value is a link to the node details page.

- **Project**

  Shows the linked project name when one is resolved in the UI. Use it to identify the project that owns or contains the related resource.

- **GPU Network Switch**

  Shows the linked GPU network switch name when one is resolved in the UI. Use it to identify the switch related to the port. This value is a link to the GPU network switch page.

- **Port IP**

  Shows the IP address assigned to the GPU network port. Use it to review the port-side addressing value.

- **Peer IP**

  Shows the peer IP address for the port. Use it to review the paired addressing value.

- **MAC Address**

  Shows the MAC address saved for the GPU network port. Use it to verify the Layer 2 hardware address.

- **PCI Device ID**

  Shows the PCI device identifier saved for the port. Use it to relate the port record to the hardware device identifier.

- **ENI Device Index**

  Shows the ENI device index saved for the port. Use it to review the device index value associated with the port.

- **Rail ID**

  Shows the rail identifier saved for the port. Use it to review the rail value assigned to the port record.

You can also add these optional columns:

- **Created At**

  Shows when the record was created. Use it for audit and tracking work.

- **Updated At**

  Shows when the record was last updated. Use it for audit and tracking work.

- **ID**

  Shows the unique record ID. Use it when you need the exact identifier for a specific GPU network port record.

22.3 What is not available on this page

The current collection view exposes no row or page actions.

- There is no verified add action on this page.
- There is no verified modify action on this page.
- There is no verified delete action on this page.
- There is no verified single-entity details component wired into the Region Networking service for GPU Network Ports.

22.4 Recommended best practices

Use this page as a review and cross-reference tool. The UI exposes linked names for the VM, node, and GPU network switch, so it works well for tracing a port to related resources.

Check both addressing fields together. **Port IP** and **Peer IP** are most useful when reviewed as a pair.

Use hardware identifiers during troubleshooting. **MAC Address**, **PCI Device ID**, **ENI Device Index**, and **Rail ID** help you match an entry in the UI to lower-level inventory or diagnostics.

Use optional audit columns when needed. **Created At**, **Updated At**, and **ID** can help with record tracking.

Open related records from the links in the table. This is the fastest way to move from a port record to its associated VM, node, or GPU network switch.

22.5 Troubleshooting

22.5.1 The page does not appear in the menu

GPU Network Ports is part of the Region Networking menu for system administrators.

- Confirm that you are using an account with system admin access.
- Open **Region Networking** and check the left navigation again.

22.5.2 The list stays empty

An empty list can be normal when no records are defined.

- Check whether the page shows that no GPU Network Ports are defined.
- Refresh the page if you expect existing data.
- Verify the related backend data source for GPU network ports.

22.5.3 A linked VM, node, or switch name is missing

The UI resolves related names from other models. A missing name can mean that the related resource was not resolved in the current data set.

- Refresh the page and allow the data to load.
- Check whether the related VM, node, project, or GPU network switch exists and is accessible.
- Review the raw port fields that are still visible in the list.

22.5.4 You need to create or edit a GPU network port

The current Region Networking page is read-only in the verified UI.

- Use this page to review records.
- Use the related linked resources for more investigation.
- Check whether port creation or update is handled in another workflow outside this page.

CHAPTER

TWENTYTHREE

EDGE NETWORKS

In zCompute, Edge Networks provide the connectivity between VPC entities (e.g VMs, load-balancers) and the Internet or any other local datacenter-routable network. Under an Edge Network one or more IP-address pools can be defined. These IP pools are then used to create Elastic-IPs, that when attached to a VM, allow traffic from outside the cloud (Internet or the hosting datacenter to reach the VM. A single zCompute region may have multiple Edge Networks, each with multiple IP pools. IP pools can be public, available for all tenant accounts, or private to specific accounts. Edge Networks use VLANs to connect the cloud region to the hosting datacenter's upstream switches and L3 gateways.

23.1 Viewing edge networks

Use the edge networks list view to find a network, review its state, and open the details page.

1. Go to **Region Networking** > **Edge Networks**.

   The edge networks are displayed.

   The list view shows these columns:

   - **Name**

     The descriptive display name of the edge network.

     This is the main identifier used in list views and action dialogs. It is also the link that opens the details page.

   - **Subnets**

     The subnet CIDR values that are attached to the edge network.

     This field shows the network address range or ranges already assigned to the edge network.

   - **Gateway**

     The gateway IP address for the edge subnet.

     This address is the routed exit point for systems on the subnet.

   - **VLAN**

     The VLAN value for the edge network.

     This is the layer 2 tag used by the network. If the network is untagged, the UI can show an untagged value.

   - **MTU**

     The maximum transmission unit for the edge network.

     This value controls the largest packet size allowed on the network path.

   - **Account**

     The account associated with the edge network.

This field helps an admin identify the owning account when the column is visible.

- **Project**

  The project associated with the edge network.

  This field helps an admin identify the linked project when the column is visible.

- **Switch Domain**

  The switch domain associated with the edge network.

  This is the switching domain that contains the edge network.

- **State**

  The current lifecycle state of the edge network.

  This field helps the admin decide whether change actions are available.

  - The UI only enables **Edit** when the state is **Available**.
  - The UI disables **Delete** when the state is **Pending**.

- **ID**

  The unique identifier of the edge network.

  Use this value when names are similar or when you must confirm the exact object in a prompt or a ticket.

2. Select the edge network name to open its details page.

23.2 Viewing an edge network's details

Use the details page to review the full configuration of one edge network.

23.2.1 Edge network main details

1. Go to **Region Networking** > **Edge Networks**.

   The edge networks are displayed.

2. In the edge networks list, select the edge network name.

   The top toolbar can show these controls:

   - **Edit**

   Opens the edge network edit flow.

   Use this control to update the editable edge network and subnet fields.

   - **Services Proxy**

   Opens the managed services proxy dialog.

   Use this control to store HTTP and HTTPS proxy settings for managed services.

   - **Delete**

   Opens the edge network delete confirmation prompt.

   Use this control to remove the edge network after you review dependencies.

   The details page displays the selected edge network's details.

   - **Info** section:

  - **Name**

    The descriptive display name of the edge network.

  - **ID**

    The unique identifier of the edge network.

  - **Description**

    Brief description of the edge network.

- **IP Pools** section:

  Each row in this block represents one non-internal IP pool.

  This means the block is meant to summarize allocatable pools that the admin can review at a glance, rather than internal-only pools.

  - **Pool name**

    The first column of each row is the IP pool name.

    This is the main label that identifies the pool in the summary block. It lets the admin match the summary row to the same pool in the **IP Pools** tab.

  - **Allocation bar**

    The value shown for each row is an allocation bar.

    It is a visual capacity summary for the pool indicating allocated and unallocated proportions. It can help admins estimate pool consumption without opening the full **IP Pools** tab.

- **Subnets**

  Each row in this block represents one subnet from the edge network.

  This means that this section is a summary of the subnets already attached to the edge network. It helps admins confirm the current subnet layout before opening the **Overview** tab to see the subnets and their details.

  - **CIDR**

    The first column of each row is the subnet CIDR block.

    This is the subnet address range in CIDR notation. It defines the IP space of that subnet and gives the admin a quick view of the network range in use.

  - **Gateway**

    The value shown for each row is the gateway IP of that subnet.

    This is the routed exit point for systems on the subnet. It helps admins confirm the main layer 3 value for the subnet without opening the detailed subnet record.

- **Internal Resources** section:

  The block summarizes the router-side internal resources that support the edge network.

  - **Edge Router Public IP**

    The public IP of the edge router.

    This is the external IP address assigned to the router. It helps admins identify the router's public-facing address quickly, without opening the **Overview** tab.

  - **Routers Subnet**

    The CIDR block of the routers subnet, when that subnet exists.

This is the subnet used for router-facing connectivity. It helps admins confirm the internal routed subnet that supports the edge router path.

- **Routers IP Pool**

  Shows an allocation bar for the routers IP pool, when that pool exists.

  This is a capacity summary for the internal routers IP pool. It helps admins review router-related address consumption at a glance.

23.2.2 Overview tab

The **Overview** tab shows the main configuration record for the selected edge network.

It provides the best single-page view of the network, subnets, proxy, router, and routers subnet records.

Network section

- **Network ID**

  The unique identifier of the edge network.

  Use it to confirm that you are working on the correct network object.

- **Name**

  The edge network name and description.

  This is the main human-readable label for the network. It is used across lists, tabs, and prompts.

- **Shared**

  The sharing state of the edge network.

  This shows whether the network is shared. It helps an MSP admin understand the intended scope of use.

- **VLAN**

  The VLAN value of the edge network.

  This is the layer 2 tag used by the network. It affects the network's underlay connectivity model.

- **Switch Domain**

  The switch domain that contains the edge network.

  This identifies the switching domain that the network belongs to.

- **MTU**

  The maximum transmission unit of the edge network.

  This controls the largest packet size allowed on the network path.

- **Account**

  The account associated with the edge network.

  This field is useful for MSP review when account-level ownership matters.

- **Project**

  The project associated with the edge network.

  This field helps identify the linked project context.

Subnets section

- **Subnet ID**

  The unique identifier of the subnet record.

  Use it to confirm the exact subnet object attached to the edge network.

- **Name**

  The subnet name.

  This is the human-readable label of the subnet. The UI can show no name when no value is set.

- **CIDR**

  The subnet address range in CIDR notation.

  This defines the IP space available on the subnet.

- **Gateway**

  The gateway IP address of the subnet.

  Systems on the subnet use this address to reach other networks.

- **DNS Servers**

  The DNS server IP addresses assigned to the subnet.

  Systems use these addresses for name resolution.

Managed Services Proxy section

This section appears when proxy data exists for the edge network.

- **HTTP proxy**

  The saved HTTP proxy value for the edge network.

  It identifies the proxy endpoint used for HTTP traffic.

- **HTTPS proxy**

  The saved HTTPS proxy value for the edge network.

  It identifies the proxy endpoint used for HTTPS traffic.

- **Exclusion list**

  The saved bypass list for proxy handling.

  It shows the destinations that should not use the proxy.

Edge Router section

- **Router ID**

  The unique identifier of the edge router.

  Use it to confirm the exact router object.

- **Name**

  The edge router name.

  This is the router label shown in the UI.

- **Internal Interfaces**

  The internal router interface address or addresses.

  These values show the router-side addresses used on internal connectivity paths.

- **Public IP**

  The public IP assigned to the edge router.

  Use it to identify the external address used by the router.

Routers Subnet section

- **Subnet ID**

  The unique identifier of the routers subnet.

  Use it to confirm the exact routers subnet record.

- **Name**

  The routers subnet name.

  This is the human-readable label of the routers subnet. The UI can show no name when no value is set.

- **CIDR**

  The routers subnet address range in CIDR notation.

  This defines the address space used by router-facing interfaces.

- **Gateway**

  The gateway IP address of the routers subnet.

  Use it to confirm the routed path on that subnet.

23.2.3 Events tab

The **Events** tab shows activity related to the selected edge network.

The screenshots show filter controls in this tab.

- **Date range**

  The time filter for the event list.

  Use it to narrow the visible events to a specific time window.

- **Filter**

  The text filter for the event list.

  Use it to find a specific event more quickly.

- **More filters**

  Additional event filters.

  Use them when the basic filters are not enough to narrow the event list.

23.2.4 IP Pools tab

The **IP Pools** tab shows the IP pools that belong to the edge network.

The list shows these fields:

- **Name**

  The display name of the IP pool.

  Use it to identify the pool in list views and confirmation prompts.

- **Usage**

  The current used and total address count in the pool.

  Use it to estimate remaining capacity.

- **IP Ranges**

  The start and end IP ranges assigned to the pool.

  These are the addresses that the pool can allocate.

- **Assigned Accounts**

  The assignment or sharing state of the pool.

  This helps the admin see whether the pool is shared or linked to specific accounts.

- **Type**

  The pool type.

  The screenshots show types such as VPC, Legacy, and Routers.

IP Pools tab actions

The IP Pools tab has its own toolbar menu of widgets, for creating a new IP pool, and modifying or deleting an existing IP pool.

Creating an IP pool

Use this action when the edge network needs a new address pool.

1. Go to **Region Networking** > **Edge Networks**.

   The edge networks are displayed.

2. In the edge networks list, select the edge network name.

   The selected edge network's details page displays.

3. Select the **IP Pools** tab.

4. From the tab's toolbar, select **+ Create**.

   The **Create IP Pool** dialog opens.

5. In the **Create IP Pool** dialog, enter the fields:

   - **Name**

     The descriptive display name for the IP pool.

     Use a name that helps admins identify the pool in list views and prompts.

- **Description**

  A short text description of the IP pool.

  Use it to explain the purpose of the pool.

- **Shared**

  The sharing state of the IP pool.

  Use it to mark the pool as shared when that matches the intended access model.

- **IP Ranges**

  One or more start and end IP values.

  These values define the addresses that belong to the pool.

- **Accounts**

  From the dropdown, select one or more accounts that will have use of the IP pool.

6. To save the new IP pool, select **Create**.

Modify an IP pool

Use this action when an editable IP pool setting must change.

1. Go to **Region Networking** > **Edge Networks**.

   The edge networks are displayed.

2. In the edge networks list, select the edge network name.

   The selected edge network's details page displays.

3. Open the **IP Pools** tab.
4. Select the IP pool.
5. Select **Modify**.

   The **Modify IP Pool** dialog opens.

6. In the **Modify IP Pool** dialog, the following fields can be updated:

   - **Name**

     The display name of the IP pool.

     It is recommended to change it only when the new name improves clarity.

   - **Description**

     The description of the IP pool.

     Change it when the pool purpose or operational notes need an update.

   - **Shared**

     The sharing state of the IP pool.

     Caution

     Review this carefully because it affects how the pool is intended to be used.

- **IP Ranges**

  The start and end IP values in the pool.

  Caution

  Change IP ranges only after you confirm the target address plan.

7. To save the changes to the IP pool, select **Modify**.

Delete an IP pool

Caution

Use this action only after you confirm that the pool is no longer needed.

1. Go to **Region Networking** > **Edge Networks**.

   The edge networks are displayed.

2. In the edge networks list, select the edge network name.

   The selected edge network's details page displays.

3. Open the **IP Pools** tab.
4. Select the IP pool to delete.
5. Select **Delete**.

   The **Delete IP Pool** dialog opens, displaying the name and ID of the selected IP pool.

   Review the prompt: **Are you sure you want to delete this IP pool?** and the IP pool's name and ID, before progressing.

6. To confirm deletion of the IP pool, select **Delete**.

23.2.5 VPC Projects tab

The **VPC Projects** tab shows which VPC projects use the edge network.

Note

- Review this tab before IP pool changes.

  It helps you estimate the effect of a pool change on dependent projects.

- Review this tab before deleting an edge network.

  It helps you see whether active projects still depend on the edge network.

The list displays these fields:

- **Name**

  The VPC project name.

  This is the project label shown in the list and used to identify the dependent project.

- **IP Pool**

  The linked IP pool name.

  This shows which IP pool the VPC project is using on the edge network.

- **Account**

  The owning account of the VPC project.

  Use it to identify the tenant or account that owns the project.

23.2.6 Elastic IPs tab

The **Elastic IPs** tab shows the Elastic IP records that are associated with VPC projects that use this edge network as their default edge network.

The edge-network view gets these records by collecting the VPC projects linked to the edge network and then filtering Elastic IPs by those project IDs.

The list displays these fields:

- **Elastic IP**

  The public Elastic IP address.

  This is the main public address assigned to the Elastic IP record. The table also shows an internal-resource tooltip next to this value.

- **Private IP**

  The private IP address currently associated with the Elastic IP.

  This value shows the internal address that the public address maps to. If a network interface is known, the table also shows a link to that interface and its name.

- **Instance**

  The attached VM instance.

  When an instance is associated, this field links to the VM and shows the instance name. This helps identify the workload that currently uses the Elastic IP.

- **VPC**

  The VPC that owns the Elastic IP record.

  This field links to the VPC and shows its name. It helps the admin see which VPC context the Elastic IP belongs

- **Public DNS**

  The public DNS name for the Elastic IP.

  This field shows the public DNS value when one exists.

- Non-defult display columns

  The following columns are not displayed by default, but can optionally be selected for display using the column filter selection.

  - **ID**

    The unique Elastic IP identifier.

    This field helps distinguish records when IP values are similar or when exact confirmation is needed.

  - **Subnet**

    The subnet linked to the Elastic IP record.

- **Tags**

  The tag list assigned to the Elastic IP.

23.3 Creating an edge network

Use this action when you need a new edge network.

Note

- The create flow has two steps: **Network** and **Subnets**.

  This keeps network properties separate from subnet addressing.
- Plan switch domain, VLAN, and MTU values before you create the network.

  Some of these fields are not editable later.

1. Go to **Region Networking** > **Edge Networks**.
2. Select **+ Create**.

   The **Create Edge Network** dialog opens.
3. In the **Network** step, enter the displayed fields:
   - **Name**

     The edge network name.

     This is the main display label used in list views, details pages, and prompts.
   - **Description**

     The short description of the edge network.

     Use it to explain the purpose of the network.
   - **Switch Domain**

     The switch domain for the edge network.

     This places the network in a specific switching domain.
   - **Shared**

     The sharing state of the edge network.
   - **VLAN ID**

     The VLAN value of the edge network.

     This is the layer 2 tag used by the network when the network is tagged.
   - **Untagged**

     The untagged option for the edge network.

     Use it when the network must not use a VLAN tag.
   - **MTU**

     The maximum transmission unit of the edge network.

     This controls the largest allowed packet size on the network path.

- **Public IP**

  The public IP for the edge router.

  Use it when you need to set a specific router public address. If you leave it empty, the form states that a random available IP address is allocated from the IP pool.

4. Select **Next** to progress to the **Subnets** step.
5. In the **Subnets** step, enter the displayed subnet fields:
   - **Subnet (CIDR)**

     The subnet address range in CIDR notation.

     This defines the IP space available on the edge subnet.

   - **Gateway**

     The gateway IP address of the subnet.

     Systems on the subnet use this address to reach other networks.

   - **DNS Servers**

     The DNS server IP addresses for the subnet.

     Systems use these addresses for name resolution.
6. To save the new edge network's configuration, select **Finish**.

23.4 Modifying an edge network

Use this action when you need to update the editable edge network settings.

Note

- The UI enables **Edit** only when the edge network state is **Available**.

  Review the current state before you plan the change.

- Expect some fields to be visible but not editable.

  The system disables switch domain, VLAN, and untagged values in edit mode.

1. Open the edge network details page.
2. Select **Edit**.

   The **Modify Edge Network** dialog opens.
3. In **Network** step, review and update the displayed fields:
   - **Name**

     The edge network name.

     Change it when the current name no longer describes the network clearly.

   - **Description**

     The short description of the edge network.

     Change it when the purpose or operational note needs an update.

- **Switch Domain**

  The switch domain of the edge network.

- **Shared**

  The sharing state of the edge network.

- **VLAN ID**

  The VLAN value of the edge network.

- **Untagged**

  The untagged option of the edge network.

- **MTU**

  The maximum transmission unit of the edge network.

- **Public IP**

  The public IP of the edge router.

4. Select **Next** to progress to the **Subnets** step.
5. In the **Subnets** step, review and update the displayed subnet fields:

   - **Subnet (CIDR)**

     The subnet address range in CIDR notation.

   - **Gateway**

     The gateway IP address of the subnet.

     The system checks whether the IP matches the subnet CIDR and excludes boundary addresses.

   - **DNS Servers**

     The DNS server IP addresses for the subnet.

6. To save the changes, select **Finish**.

23.5 Configuring a services proxy

There may be several reasons for configuring a services proxy, including:

- To force managed-services traffic through a controlled egress path.
- To use a proxy that requires authentication.
- To let some destinations bypass the proxy, by defining an exclusion list

Use this action when managed services for the edge network must use a proxy.

Note

- Keep the exclusion list focused.

  A smaller bypass list is easier to review and support.

- Review the Overview tab after saving proxy values.

  The tab shows the stored HTTP proxy, HTTPS proxy, and exclusion list when proxy data exists.

1. Open the edge network details page.

2. Select **Services Proxy**.

   The **Configure Managed Services Proxy** dialog opens.

3. Enter the fields:

   - **HTTP Domain**

     The HTTP proxy host or IP value.

     Use it to define the proxy endpoint for HTTP traffic.

   - **Port**

     The HTTP proxy port.

     Use it with the HTTP domain to define the full HTTP proxy endpoint.

   - **Username**

     The HTTP proxy user name.

     Use it when the proxy requires authentication.

   - **Password**

     The HTTP proxy password.

     Use it with the user name when the proxy requires authentication.

   - **HTTPS Domain**

     The HTTPS proxy host or IP value.

     Use it to define the proxy endpoint for HTTPS traffic.

   - **Port**

     The HTTPS proxy port.

     Use it with the HTTPS domain to define the full HTTPS proxy endpoint.

   - **Username**

     The HTTPS proxy user name.

     Use it when the HTTPS proxy requires authentication.

   - **Password**

     The HTTPS proxy password.

     Use it with the user name when the HTTPS proxy requires authentication.

   - **Exclusion list**

     The proxy bypass list.

     Use it to enter destinations that must bypass the proxy. The dialog hint says the values are domain suffixes that bypass the proxy and must start with a period.

4. To save the service proxy configuration, select **Ok**.

23.6 Deleting an edge network

Use this action only after you confirm that the edge network is no longer needed.

Note

- The UI disables **Delete** when the edge network state is **Pending**.

  Review the state before you plan the delete action.

- Review the **VPC Projects** and **IP Pools** tabs first.

  This helps you understand dependency and usage before removal.

---

1. Open the edge network details page.
2. Select **Delete**.

   The **Delete Network** dialog opens, displaying the edge network's name and ID.
3. Review the prompt:

   - **Are you sure you want to delete this edge network?**

     The confirmation prompt warns that the selected edge network will be removed.

     Use it as the final review step before deletion.

   - **ID**

     The unique edge network ID shown in the prompt.

     Use it to confirm the exact object to delete.

   - **Name**

     The edge network name shown in the prompt.

     Use it to confirm that the selected object is the correct one.
4. To delete the edge network, select **Delete**.

23.7 Recommended best practices for edge network management

- Plan **Switch Domain**, **VLAN ID**, and **Untagged** values before you create the edge network.

  The create/edit form allows these values at creation time, but the edit form disables **Switch Domain**, **VLAN ID**, and **Untagged**. That makes them design-time choices, not routine change-time values.

- Use the default **MTU** unless you have a clear reason to change it, and keep MTU changes controlled.

  The form always requires MTU and defaults it from the edge-network default. The input validation check also enforces minimum and maximum MTU values, which indicates that MTU is a core network setting that should be chosen carefully.

- Treat subnet **CIDR** as stable after creation.

  In the **Subnet** step, the CIDR field is disabled for editing once the subnet record already exists. That means you should choose the subnet range carefully before production use.

- Validate **Gateway** carefully against the subnet.

  The input validation checks the gateway IP against the subnet CIDR and excludes boundary addresses. A practical best practice is to review the subnet and gateway together before saving changes.

- Enter only valid IP addresses in **DNS Servers**, and keep the list intentional.

  The DNS Servers field accepts multiple values and validates them as IPs. That makes it a good place to keep only the resolvers that are actually required for that subnet.

- Use the edge network state as your first pre-change check.

  The UI enables **Edit** only when the edge network state is **Available**.

  It disables **Delete** when the state is **Pending**.

  Before any maintenance, verify that the network is in the right state.

- Keep **Public IP** assignment deliberate.

  If no public IP is provided in the **Network** step, a random available IP is allocated from the IP pool. Use an explicit value when you need predictable router addressing, and leave it empty only when automatic assignment is acceptable.

- Use the **Services Proxy** only when managed services actually need a proxy path, and keep the bypass list tight.

  The action stores proxy settings on the edge network, and the UI is specifically for managed-services proxy configuration. Because the **Services Proxy** dialog includes HTTP/HTTPS endpoints and an exclusion list, it is best to keep the configuration minimal and reviewable.

- Review free address space before creating an IP pool.

  The IP pool dialog fetches unused ranges for the edge network and only presents IP-range entry when there is an existing pool to modify or unused ranges to consume.

  Check available space before adding more pools.

- Use **Shared** IP pools only when that access model is intentional; otherwise assign accounts explicitly.

  In the IP pool dialog, when **Shared** is not selected, **Accounts** becomes required. That implies a clear operational split between shared pools and account-scoped pools.

- Do not plan routine changes to the routers IP pool.

  The IP pool actions disable **Modify** and **Delete** for internal pools, with the message **Routers IP Pool is managed internally**. Treat router pools as platform-managed resources.

- Use the built-in two-step workflow as your change checklist: **Network** first, then **Subnets**.

  The dialog explicitly separates these into two steps. In practice, this is a good admin habit as well: review network-wide properties first, then validate subnet addressing.

- Before deleting an edge network, confirm both identity and state.

  The delete prompt shows the network **ID** and **Name**, and the UI disables delete while the network state is **Pending**. That means the safest delete workflow is: verify state, then verify exact identity in the confirmation prompt.

23.8 Troubleshooting

23.8.1 The edge network is not editable

Check the edge network state first.

The UI enables **Edit** only when the state is **Available**.

23.8.2 A field is visible but cannot be changed

This can be expected behavior.

Switch Domain, **VLAN ID**, and **Untagged** are disabled in edit mode. The subnet **CIDR** field is also disabled when the subnet record already exists.

23.8.3 A gateway value is rejected

Review the subnet CIDR first.

The system validates the gateway value against the subnet CIDR and excludes boundary addresses.

23.8.4 A DNS server value is rejected

Review the entered value format.

The system validates the **DNS Servers** field as IP values.

23.8.5 You are not sure whether a delete action is safe

On the edge network's details page, review these tabs first:

- **Overview**

  Review the exact network, subnet, router, and routers subnet records.

- **IP Pools**

  Review current pool usage and ranges.

- **VPC Projects**

  Review dependent project usage.

- **Events**

  Review recent activity after the latest change.

CHAPTER

TWENTYFOUR

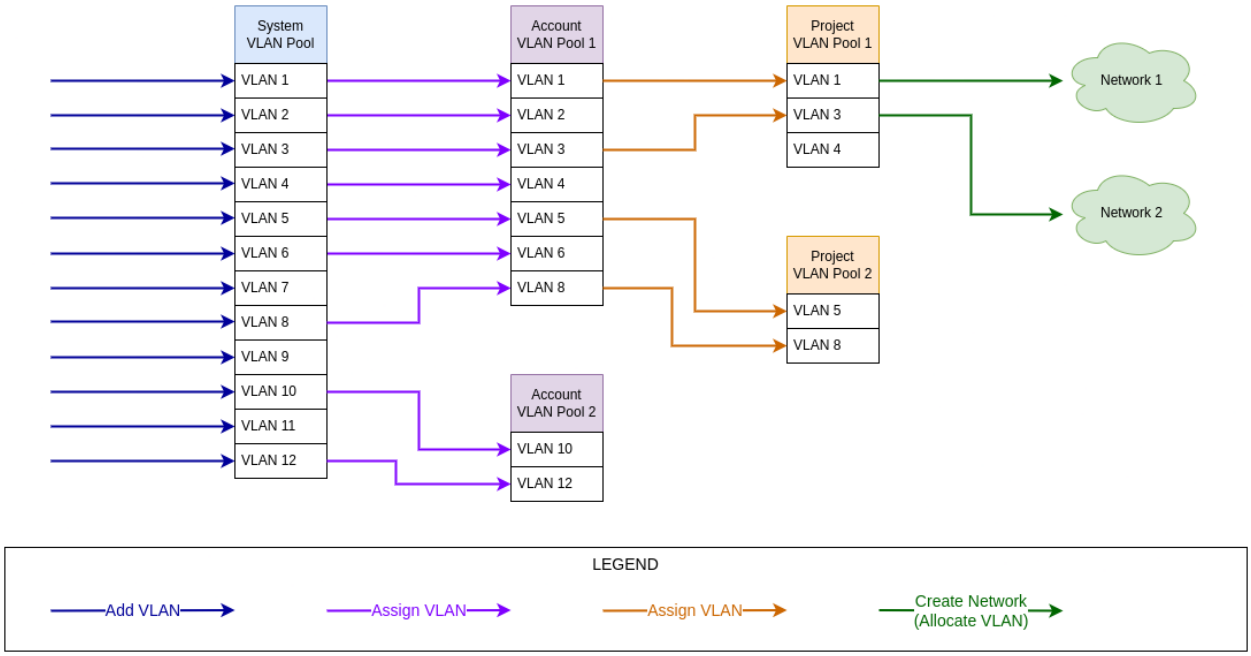
VLANS MANAGEMENT

24.1 VLANs

A VLAN is a numbered network segment used to separate traffic inside a switch domain. In zCompute, a VLAN helps organize network connectivity by linking a VLAN tag to a specific switch domain, and then associating that VLAN with an account, project, or network as needed.

Use the **Account Networking > VLANs Management** screen to manage VLANs for accounts, see where a VLAN is assigned, and to perform lifecycle tasks such as adding, assigning, releasing, or checking VLAN connectivity.

24.1.1 VLAN creation and allocation flow



1. Zadara Operations creates the system VLAN pool, and adds or removes VLANs in the pool.
2. MSPs or Zadara Operations allocate VLANs from the system VLAN pool to an account's VLAN pool.

Note

- An account's VLAN pool is automatically created when the account is created.
- An account can have only one VLAN pool.

- When an account is deleted, its VLAN pool is automatically deleted.
- When an account's VLAN pool is deleted:
  - All of the account's projects' VLAN pools are automatically deleted.
  - All of the account's VLANs are automatically deleted.
- A VLAN in a project's VLAN pool that is released from an account's VLAN pool is first automatically released from the project's VLAN pool.

3. The account administrator (MSP or tenant) assigns VLANs from the account's VLAN pool to a project's VLAN pool.

Note

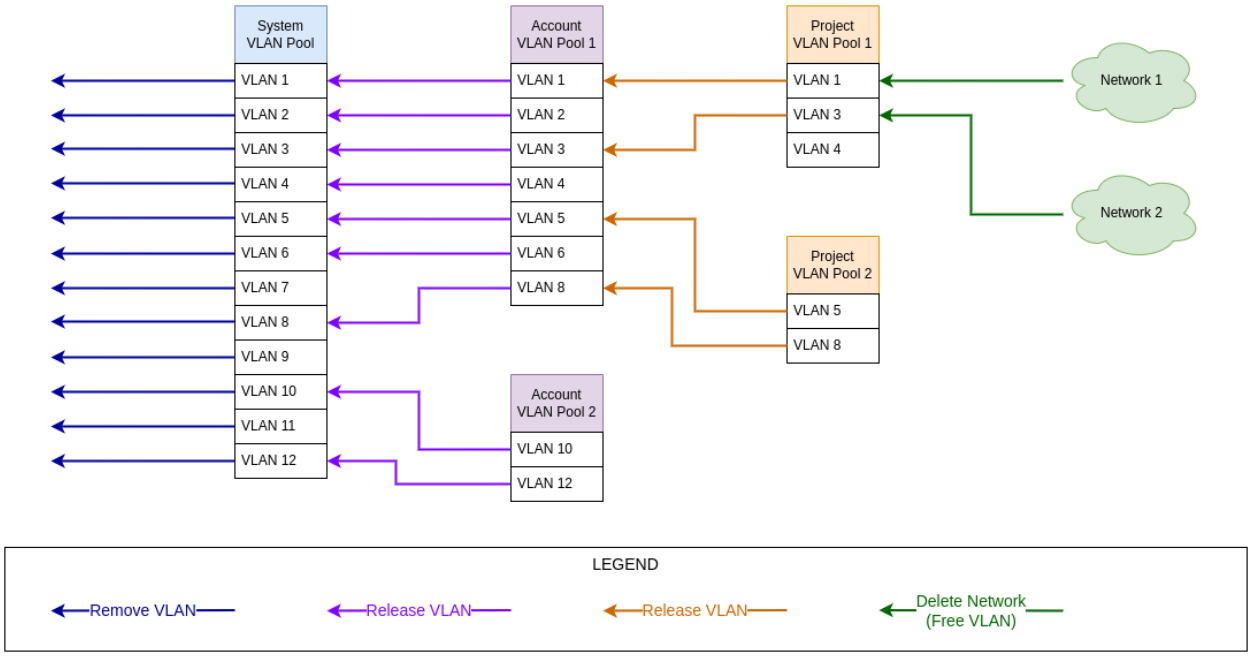
- A project's VLAN pool is automatically created when the project is created.
- A project can have only one VLAN pool.
- When a project's VLAN pool is deleted, all of its VLANs are automatically released.
- When a project is deleted, its VLAN pool is automatically deleted.

4. The account administrator can create a DVS network by allocating a VLAN to a DVS project, and associating a VM instance with the project's network.

Note

- A VLAN in a project's VLAN pool cannot be released if the VLAN is allocated to a network.

24.1.2 VLAN release and deletion flow



1. The account administrator (MSP or tenant) can release a VLAN that is allocated to a network, back to the project's VLAN pool, by deleting its network.

2. The account administrator can release VLANs that are not allocated to a network, from a project's VLAN pool back into the account's VLAN pool.

   Note

   - A VLAN in a project's VLAN pool cannot be released if the VLAN is allocated to a network.
   - When a project's VLAN pool is deleted, all of its VLANs are automatically released.
   - When a project is deleted, its VLAN pool is automatically deleted.

3. Zadara Operations can release VLANs that are not allocated to a network, from an account's VLAN pool back into the system VLAN pool.

   Note

   - When an account is deleted, its VLAN pool is automatically deleted.
   - When an account's VLAN pool is deleted:
     - All of the account's projects' VLAN pools are automatically deleted.
     - All of the account's VLANs are automatically deleted.
   - A VLAN in a project's VLAN pool that is released from an account's VLAN pool is first automatically released from the project's VLAN pool.

4. Zadara Operations can remove a VLAN from the system pool.

24.2 Managing VLANs

24.2.1 Viewing VLANs

To view the VLANs that are available:

1. Go to **Account Networking > VLANs Management**.

   The **VLANs Management** screen opens in **List** view.

2. Review the VLANs table.

   The table shows these fields:

   - **VLAN**

     The VLAN tag ID.

     Use it to identify the VLAN.

   - **Account**

     The account that owns the VLAN assignment.

   - **Project**

     The project that uses the VLAN.

   - **Network**

     The linked network.

     Use it to see whether the VLAN is already tied to a network.

- **VMs**

  The number of VMs that use the linked network.

- **Switch Domain**

  The switch domain linked to the VLAN.

24.2.2 Adding VLANs

A single VLAN, multiple individual VLANs, a range and multiple ranges of VLANs can be created in one request.

Assigning the VLANs to an account and project can be configured when adding new VLANs, or can be left to be assigned later.

To configure new VLAN entries:

1. Go to **Account Networking > VLANs Management**.

   The **VLANs Management** screen opens in **List** view.

2. From the top toolbar, select **+ Add**.

   The **Add VLANs** dialog opens.

   1. Enter the VLAN ranges:

   - **Switch Domain**

     From the dropdown, select the switch domain where the VLAN range will be created.

     Required.

   - **Account**

     From the dropdown, select the VLAN account pool that the VLANs can be assigned to.

     If this is not entered, it can be assigned later.

   - **Project**

     From the dropdown, select the project that can use the VLANs.

     The project list depends on the selected account.

     If this is not entered, it can be assigned later.

   - **VLAN Range**

     The section that defines one or more VLAN tag ranges to create.

     – **From**

       The first VLAN tag in the range.

       Required.

     – **to**

       The last VLAN tag in the range.

       Required.

       Note

       For a single VLAN, the **to** value is the same as **From**.

     – For an additional range, select **Add**, and enter the VLAN range.

- Selecting the **Delete row** icon removes the current VLAN range row from the list in the request.

To confirm creating the VLANs range, select **Ok**.

24.2.3 Assigning VLANs

A single VLAN or multiple VLANs can be assigned to an account or project in a single **Assign** request.

To select and assign VLANs to an account or project:

1. Go to **Account Networking > VLANs Management**.

   The **VLANs Management** screen opens in **List** view.

2. To select the VLANs to assign, mark their checkboxes.

   The **Assign** option is enabled on the top toolbar.

3. From the top toolbar, select **Assign**.

   The **Assign VLANs** dialog opens, prompting for the following:

   - **Account**

     From the dropdown, select the account pool that the VLAN will be assigned to.

     Required.

     Note

     If the VLAN is already assigned to an account, its value is displayed and only the **Project** can be selected.

   - **Project**

     From the dropdown, select the project that will use the selected VLAN.

     The project list depends on the selected account.

   To confirm assigning the VLAN, select **Ok**.

24.2.4 Removing VLANs

Use this action to remove one or more VLANs from the VLAN pool.

A VLAN cannot be removed when it is linked to a network.

To select VLANs to remove:

1. Go to **Account Networking > VLANs Management**.

   The **VLANs Management** screen opens in **List** view.

2. To select the VLANs to remove, mark their checkboxes.

   The **Remove** option is enabled on the top toolbar.

3. From the top toolbar, select **Remove**.

   The **Remove VLANs** dialog opens, displaying the VLANs selected for removal.

   To confirm removing the selected VLANs from the VLAN pool, select **Ok**.

24.2.5 Releasing VLANs from an account

Use this action to release a VLAN from an account's VLAN pool.

This action is enabled only when the VLAN is assigned to an account and is not assigned to a project.

To select a VLAN to release from an account:

1. Go to **Account Networking > VLANs Management**.

   The **VLANs Management** screen opens in **List** view.

2. To select the VLAN to release from an account, mark its checkbox.

   The **Release From Account** option is enabled on the top toolbar.

3. From the top toolbar, select **Release From Account**.

   The **Release VLAN From Account** dialog opens, displaying the **Name** and **ID** of the VLAN selected for release, and the **Account** from which to release it.

   To confirm releasing the selected VLAN from the account, select **Ok**.

24.2.6 Releasing VLANs from a project

Use this action to release a VLAN from a project's VLAN pool.

This action is available only when the VLAN is assigned to a project and is not linked to a network.

To select a VLAN to release from a project:

1. Go to **Account Networking > VLANs Management**.

   The **VLANs Management** screen opens in **List** view.

2. To select the VLAN to release from a project, mark its checkbox.

   The **Release From Project** option is enabled on the top toolbar.

3. From the top toolbar, select **Release From Project**.

   The **Release VLAN From Project** dialog opens, displaying the **ID** of the VLAN selected for release, and the **Project** from which to release it.

   To confirm releasing the selected VLAN from the project, select **Ok**.

24.2.7 Releasing VLANs from a network

To release a VLAN from a network, whether a direct subnet or a DVS network, deleting the network releases its assigned VLAN back to the project pool.

24.2.8 Checking VLANs

Use this action to run a VLAN connectivity check.

Check VLAN tests whether selected nodes can communicate with each other on the chosen VLAN. It shows the result as a pass/fail matrix, plus an overall **Connected**, **Not Connected**, or **Partially Connected** summary.

The action is available only when at least two nodes are present.

The check uses the VLAN's switch domain.

To select VLANs to check:

1. Go to **Account Networking > VLANs Management**.

   The **VLANs Management** screen opens in **List** view.

2. To select the VLANs to check, mark their checkboxes.

   The **Check VLAN** option is enabled on the top toolbar.

3. From the top toolbar, select **Check VLAN**.

   The **Check VLAN** dialog opens.

   Enter:

   - **MTU**

     The packet size that the check uses.

     Valid range: 1500 (default) to 16000.

   - **Nodes**

     The nodes that take part in the check.

     Select the nodes to test.

   Confirm proceeding with the connectivity test.

   The test results are displayed as a pass/fail matrix, plus an overall **Connected**, **Not Connected**, or **Partially Connected** summary.

24.3 Recommended best practices

- Review the **Network** column before you remove or release a VLAN.
- Review the **Project** and **Account** columns before you change VLAN ownership.
- Review the **Switch Domain** value to confirm that you are working on the correct VLAN context.
- Add VLAN ranges carefully to avoid overlap or incorrect tag entry.
- Run **Check VLAN** before and after a change when at least two nodes are available.
- Keep the **MTU** value within the supported range for the check.

24.4 Troubleshooting

24.4.1 You cannot assign a VLAN

Check whether the **Account** field is selected.

If the **Project** list shows no projects, verify that the selected account has available projects.

24.4.2 You cannot add VLANs

Check that **Switch Domain**, **From**, and **to** are completed.

If you add more than one range, review each row before you save.

24.4.3 You cannot remove a VLAN

A VLAN cannot be removed when it is linked to a network.

Check the **Network** column. If a network is shown, remove the network dependency first.

24.4.4 You cannot release a VLAN from a project

The project release action is disabled when the VLAN is not assigned to a project or when it is linked to a network.

Check the **Project** and **Network** columns before you try again.

24.4.5 You cannot release a VLAN from an account

The account release action is shown only when the VLAN is assigned to an account and is not assigned to a project.

Check the row values to confirm the current assignment state.

24.4.6 You cannot run Check VLAN

The check action requires at least two nodes.

If the action is unavailable, verify that the environment has at least two nodes.

CHAPTER

TWENTYFIVE

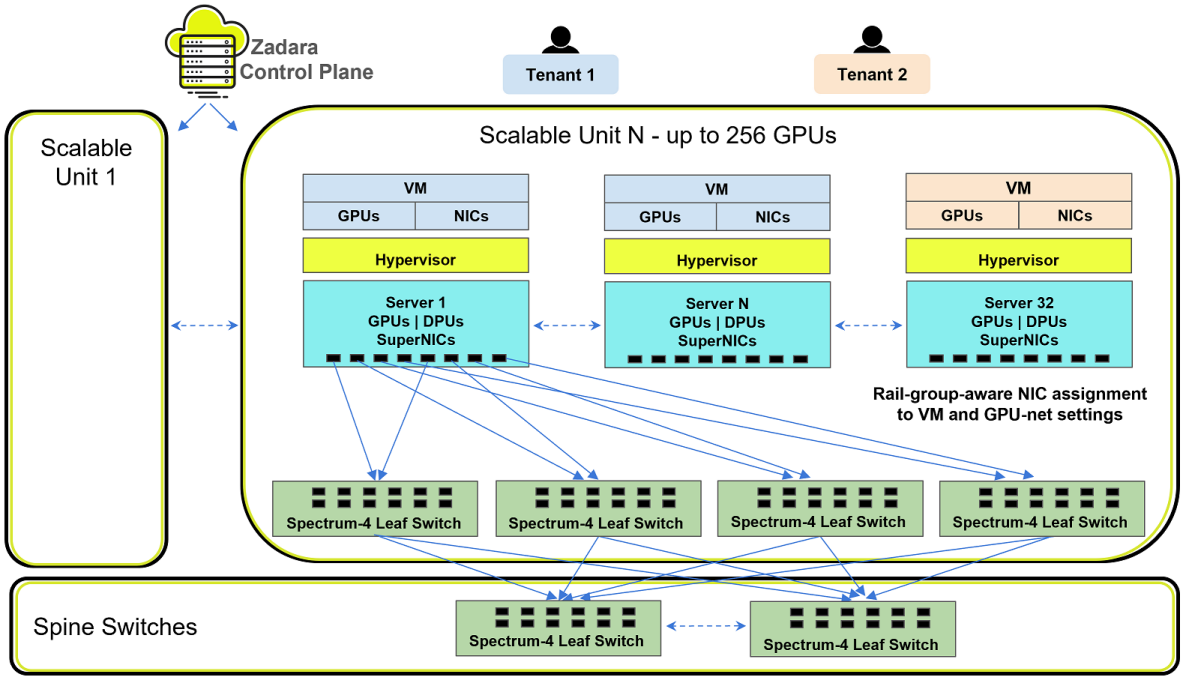
GPU NETWORKS OVERVIEW

25.1 Multi-Tenant GPU-to-GPU Network Provisioning and Management

Modern AI workloads demand a dedicated, high-performance network fabric that connects GPUs within and across compute nodes, separate from general-purpose infrastructure networks.

The **NVIDIA Spectrum-X Compute East-West (E-W) network** provides a purpose-built Ethernet fabric optimized for GPU-to-GPU communication at cloud scale.

From version 25.06, **zCompute GPU Network (GPU-Net)** exposes this fabric as a first-class, self-service resource. Tenants can provision and manage GPU networks through zCompute's standard interfaces. Tenants can have multiple GPU networks, and can allocate these networks across multiple projects that provide project isolation and tenant isolation.



25.2 Background: The NVIDIA Compute East-West Network (Spectrum-X)

25.2.1 What is the East-West Network?

In a GPU AI cloud, network traffic flows in two distinct directions:

- **North-South (N-S)**:

  Traffic between compute nodes and the outside world: user access, storage, management, and orchestration. This is a standard Ethernet fabric shared across the data center.

- **East-West (E-W)**:

  Traffic between GPUs, across nodes, within the AI compute fabric. This is the Spectrum-X network: a **dedicated, rail-optimized, high-bandwidth Ethernet fabric** built exclusively to serve GPU collective communication workloads (NCCL, RDMA, etc.).

The East-West network is entirely separate from the North-South fabric in both cabling and function. It carries no management or storage traffic: it exists solely for GPU-to-GPU communication.

25.2.2 Spectrum-X Architecture

NVIDIA Spectrum-X is the networking platform underlying the E-W fabric. It combines:

- **NVIDIA Spectrum-4 (SN5600) Ethernet switches**:

  64-port, 800 Gbps per port, delivering up to 128 x 400 Gbps of line-rate switching capacity per switch.

- **NVIDIA BlueField-3 B3140H SuperNICs**:

  Dual-port 400 GbE SmartNICs installed in each HGX compute node, providing the host-side GPU network interface and enabling hardware-accelerated network virtualization (VXLAN overlays, RoCE, adaptive routing, congestion control).

- **Cumulus Linux / NVUE**:

  The network operating system on the Spectrum-4 switches, providing an API-first, schema-driven configuration model used for automated fabric bring-up and multi-tenancy configuration.

- **NVIDIA NetQ**:

  Real-time network telemetry and validation tooling for the E-W fabric.

25.2.3 Rail-Optimized Topology

The E-W fabric uses a **rail-optimized topology**: the defining structural choice of NVIDIA's GPU network architecture. In an HGX server with 8 GPUs, each GPU is connected via a dedicated BlueField-3 SuperNIC to a specific rail of leaf switches. This means:

- GPU traffic between nodes within the same rail never traverses a spine switch, minimizing hop count and latency.
- Each HGX node connects to the leaf switches with two uplinks per rail, at speeds of 400 Gb/s or faster.
- The fabric scales from a single Scalable Unit (SU) of 32 HGX servers (256 GPUs) up to tens of thousands of GPUs using a two-tier (leaf + spine) or three-tier (leaf + spine + super-spine) topology, without requiring re-cabling.

The basic fabric building block is the **Scalable Unit (SU)**:

- 32 HGX nodes with 4 rail groups
- Each rail group served by its own set of leaf switches
- Multiple SUs connect to shared spine switches to form a cluster
- Multiple clusters form PODs
- PODs connect via super-spines for the largest deployments

Scale	GPUs	SUs	Topology
Single SU	256	1	2-tier (leaf only)
4 SUs	1,024	4	2-tier (leaf + spine)
16 SUs	4,096	16	2-tier (leaf + spine)

25.2.4 Key Network Technologies

The Spectrum-X E-W fabric is built around several technologies that work together to deliver near-lossless, high-throughput GPU communication:

- **RoCE (RDMA over Converged Ethernet)**:

  GPU collective communication libraries (NCCL) use RDMA to transfer data directly between GPU memory across the network, bypassing the CPU. This achieves maximum bandwidth and minimum latency.

- **Adaptive Routing**:

  The Spectrum-4 switches dynamically select the least-congested path on a per-packet basis, distributing GPU collective traffic evenly across all available links rather than relying on static ECMP hashing. This is critical for all-reduce and other collective patterns.

- **Spectrum-X Congestion Control (CC)**:

  A closed-loop congestion control mechanism that adjusts transmission rates proactively to prevent buffer buildup, maintaining low latency under heavy GPU workloads.

- **VXLAN Overlays (BGP-EVPN)**:

  The underlay is a pure layer-3 routed network (BGP). Multi-tenancy is implemented using VXLAN overlays with BGP-EVPN control plane, where each tenant gets its own VRF (Virtual Routing and Forwarding instance) with an associated VNI (VXLAN Network Identifier). Tenant traffic is fully isolated at the network layer.

- **W-ECMP (Weighted Equal-Cost Multi-Path)**:

  Load-balancing across multiple equal-cost paths in the fabric underlay, complementing adaptive routing for traffic distribution.

25.2.5 Multi-Tenancy in the E-W Fabric

Multi-tenancy on the Spectrum-X E-W network is achieved through a combination of:

1. **VRF-based isolation on the switches**:

   Each tenant is assigned a dedicated VRF on every leaf switch, with a unique VNI and Route Distinguisher (RD). Tenant routes are never leaked between VRFs, ensuring complete L3 isolation.

2. **VXLAN encapsulation**:

   Tenant GPU traffic is encapsulated in VXLAN as it traverses the fabric, logically separating it from other tenants' traffic even on shared physical links.

3. **BlueField-3 SuperNIC enforcement**:

   The SuperNIC on each host enforces tenant network membership at the endpoint, ensuring that a tenant's GPU workload only communicates within its assigned overlay network.

4. **Per-tenant interface assignment**:

   Leaf switch ports connected to HGX nodes are assigned to specific tenant VRFs based on the provisioned configuration. A tenant occupying specific GPUs on a node has precisely those GPU-facing switch ports mapped into its VRF.

The Reference Configuration Playbook (RCP), NVIDIA's Ansible-based automation framework for Spectrum-X, handles the translation of high-level tenant provisioning intent into the detailed switch and SuperNIC configuration across the entire fabric.

25.3 zCompute GPU-Net Feature

25.3.1 What Is GPU-Net?

GPU-Net is the zCompute capability that automates the provisioning, lifecycle management, and isolation of East-West GPU networks for cloud tenants. It bridges the zCompute control plane with the underlying NVIDIA Spectrum-X E-W fabric, giving tenants self-service access to high-performance GPU-to-GPU networking as a managed cloud resource.

When a tenant deploys GPU instances on zCompute, GPU-Net ensures that those GPUs are connected through a logically isolated, high-performance E-W network: independent of the tenant's regular VM networking and invisible to other tenants.

25.3.2 Why GPU-Net?

Without a managed GPU network layer, cloud operators face a difficult choice. Sharing the fabric across tenants isn't an option.

The choice is between one of the following options:

- Choosing a semi-manual allocation process characterized by rigidity, inefficiency, and a lack of self-service.
- A fully-automated solution such as GPU-Net provided by zCompute.

GPU-Net solves this by:

- **Automating per-tenant E-W network provisioning**:

  Eliminating manual fabric reconfiguration for every tenant GPU allocation.

- **Enforcing strict network isolation**:

  Tenants can only reach the GPUs within their own GPU-Net, regardless of physical co-location on the same nodes or switches.

- **Enabling multi-tenant GPU clusters**:

  Multiple independent tenants can share the same physical Spectrum-X infrastructure safely and simultaneously.

- **Abstracting Spectrum-X complexity**:

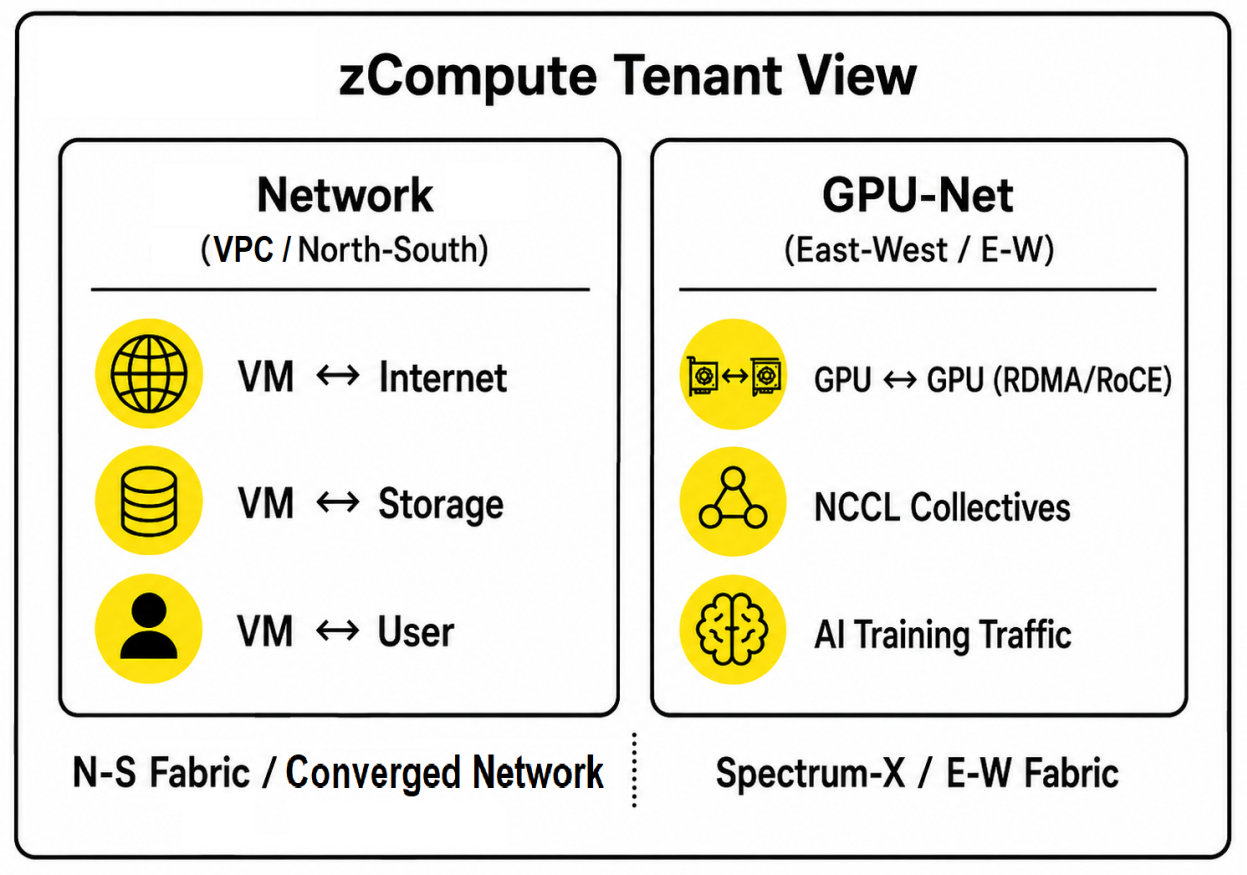
  Tenants interact with GPU-Net as a cloud networking primitive; the underlying VXLAN/VRF/BGP-EVPN configuration is invisible to them.

25.3.3 Key Capabilities

Capability	Description
Self-service provisioning	Tenants create and manage GPU networks through the zCompute API/UI, without operator intervention.
Multi-tenant isolation	Each GPU-Net is a dedicated overlay network (VRF + VNI); traffic is fully isolated from other tenants.
Scalable fabric integration	GPU-Net maps tenant networks onto the physical Spectrum-X E-W fabric across any supported topology (2-tier or 3-tier).
RoCE & RDMA support	The provisioned network is RDMA-capable, enabling GPU collectives (NCCL) to run at full bandwidth.
Lifecycle management	GPU-Net tracks the full lifecycle: create, attach, detach, and delete, of tenant GPU networks and their GPU instance memberships.
Network visibility	Tenants have visibility into their GPU-Net topology and health through the zCompute portal.

25.3.4 Relationship to Other zCompute Networks

GPU-Net is distinct from, and complements, regular zCompute tenant networking:



A GPU instance is attached to both networks simultaneously: its regular vNIC connects to VPC subnets, that are implemented on the North-South network. Its GPU NIC (BlueField-3 SuperNIC) connects to the GPU-Net for high-performance GPU-to-GPU communication.

25.4 Summary

The NVIDIA Spectrum-X East-West network is the industry-leading GPU fabric for AI cloud infrastructure: a rail-optimized, RoCE-capable Ethernet network with hardware-enforced multi-tenancy via VXLAN/BGP-EVPN overlays. It is physically and logically separate from all other data center networks, dedicated exclusively to the demands of GPU collective communication.

zCompute GPU-Net makes this fabric a managed, multi-tenant cloud resource. It automates the provisioning and isolation of per-tenant GPU networks on top of the Spectrum-X infrastructure, enabling cloud operators to efficiently and securely serve multiple independent AI tenants on shared GPU hardware, all through the standard zCompute control plane.

CHAPTER

TWENTYSIX

GPU NETWORKS MANAGEMENT AND OPERATIONS

A GPU network is a managed network object used to interconnect multiple GPUs within a single project on a high-bandwidth, low-latency network (Nvidia Spectrum-X East-West dedicated GPU network). GPU Networks provide tenancy isolation for such GPU-to-GPU dedicated inteconnectivity.

A GPU network has an ID, Name and VXLAN Network Identifier (VNI). In Nvidia Spectrum-X architecture, the CIDR is common to the whole fabric. A GPU network's lifecycle is managed as follows: System-level by admins, account-level by tenant admins, and project-level by members with permissions to the project.

The GPU Networks Management page shows the GPU networks that are available to the account and their current assignment state.

The list view includes these fields:

- **Name**

  The display name of the GPU network. Use it to identify the network in the list and in action dialogs.

- **VNI Number**

  The VNI value (VXLAN) for the GPU network. Use it to identify the network segment when you create a new GPU network.

- **State**

  The current state of the GPU network. Use it to confirm whether the network is active or in another status.

- **CIDR**

  The address range for the GPU network. Use it to see the network scope.

- **Project**

  The project currently associated with the GPU network, when one is assigned.

- **Account**

  The account currently associated with the GPU network.

- **ID**

  The system ID of the GPU network. Use it when you need a unique identifier.

- **Created At**

  The date when the GPU network entry was created.

- **Updated At**

  The date when the GPU network entry was last updated.

The page supports these lifecycle operations:

- Add a GPU network

- Assign a GPU network
- Release a GPU network from an account
- Release a GPU network from a project
- Delete a GPU network
- Force delete a GPU network
- Delete multiple GPU networks

26.1 GPU Networks Management

26.1.1 Viewing GPU networks

To see the GPU networks that are available to the account:

1. In the left navigation pane, go to **Account Networking**.
2. Select **GPU Networks Management**.

   The page opens in **List** view.
3. Review the GPU networks table.
4. Use the table columns to inspect each GPU network.

   The list view includes these fields:

   - **Name**

     The display name of the GPU network. Use it to identify the GPU network in the list.
   - **VNI Number**

     The numeric VNI value (VXLAN) for the GPU network. Use it to identify the GPU network record.
   - **State**

     The current state of the GPU network. Use it to check its current condition.
   - **CIDR**

     The network address range for the GPU network.
   - **Project**

     The project associated with the GPU network, if a project assignment exists.
   - **Account**

     The account associated with the GPU network, if an account assignment exists.
   - **ID**

     The system ID of the GPU network.
   - **Created At**

     The date and time when the GPU network was created.
   - **Updated At**

     The date and time when the GPU network was last updated.

After selecting a GPU network, the UI exposes actions for that item, based on the current state and role permissions:

- **Assign**

- **Release from Account**
- **Release from Project**
- **Delete**
- **Force Delete**

26.1.2 Adding a GPU network

Use this action to create a new GPU network entry.

1. Open **Account Networking** > **GPU Networks Management**.

   The list of GPU networks appears.

2. Click **Add**.

   The **Add GPU Network** dialog opens.

   Enter values for these fields:

   - **Name**

     A descriptive name for the new GPU network.

     Required.

   - **VNI Number**

     The VNI value (VXLAN) for the new GPU network.

     Required.

   - **Account**

     The account pool to assign after creation.

     Use this field to place the new GPU network in an account pool.

   - **Project**

     The project pool to assign after creation.

     Use this field to place the new GPU network in a project pool.

3. To save the configuration, select **Ok**.

26.1.3 Assign a GPU network

Use this action to place an existing GPU network into an account pool or project pool.

1. Open **Account Networking** > **GPU Networks Management**.

   The list of GPU networks appears.

2. Select the GPU network to assign to an account pool or project pool.

3. Select **Assign**.

   The **Assign GPU Network** dialog opens.

   In the dialog, review or set these fields:

   - **Account**

     The account pool for the GPU network. When the GPU network is already assigned to an account pool, this field is shown with the current value and is not editable.

- **Project**

  The project pool for the GPU network. The available project pools follow the selected account pool when an account is set.

4. To save the configuration, select **Ok**.

26.1.4 Release a GPU network from an account

Use this action to remove an account-level assignment from a GPU network.

Note

This action is available only when the GPU network is assigned to an account pool and is not assigned to a project pool.

1. Open **Account Networking** > **GPU Networks Management**.

   The list of GPU networks appears.

2. Select the GPU network to release from an account.
3. Select **Release from Account**.

   The **Release from Account** dialog opens.

   In the confirmation dialog, review these values:

   - **Name**

     The GPU network that will be released.

   - **ID**

     The system ID of the GPU network.

   - **Account**

     The account from which the GPU network will be released.

4. To confirm releasing the selected GPU network, select **Ok**.

26.1.5 Release a GPU network from a project

Use this action to remove a project-level assignment from a GPU network.

1. Open **Account Networking** > **GPU Networks Management**.

   The list of GPU networks appears.

2. Select the GPU network to release from a project.
3. Select **Release from Project**.

   The **Release from Project** dialog opens.

   In the confirmation dialog, review these values:

   - **ID**

     The system ID of the GPU network.

   - **Project**

     The project from which the GPU network will be released.

4. To confirm releasing the selected GPU network, select **Ok**.

26.1.6 Delete a GPU network

Use this action to remove one GPU network entry.

1. Open **Account Networking** > **GPU Networks Management**.

   The list of GPU networks appears.

2. Select the GPU network to delete.

3. Select **Delete**.

   The **Delete GPU Network** dialog opens.

   In the confirmation dialog, review these values:

   - **Name**

     The GPU network that will be deleted.

   - **ID**

     The system ID of the GPU network.

4. To confirm deleting the selected GPU network, select **Ok**.

26.1.7 Force delete a GPU network

Use this action when a standard delete does not complete.

Caution

This action cannot be undone and can cause data loss.

1. Open **Account Networking** > **GPU Networks Management**.

   The list of GPU networks appears.

2. Select the GPU network to delete.

3. Select **Force Delete**.

   The **Force Delete GPU Network** dialog opens.

   In the confirmation dialog, review these values:

   - **Name**

     The GPU network that will be deleted.

   - **ID**

     The system ID of the GPU network.

4. To confirm force deleting the selected GPU network, select **Ok**.

26.1.8 Delete multiple GPU networks

Use this action to remove more than one GPU network at the same time.

Note

The batch delete action is disabled when any selected GPU network is assigned to an account pool.

1. Open **Account Networking** > **GPU Networks Management**.

   The list of GPU networks appears.

2. Select the GPU networks to delete.

3. Select **Delete**.

   The **Delete GPU Network** dialog opens, displaying the list of selected GPU network names that will be deleted.

4. To confirm deleting the selected GPU networks, select **Ok**.

26.2 Recommended best practices

- Use clear, unique names so admins can identify GPU networks quickly.
- Record the **VNI** (VXLAN) number before you create the GPU network.
- Review the **Account** and **Project** columns before you release or delete a GPU network.
- Use standard delete first. Use force delete only when required.
- Review all selected entries before you run a batch delete.

26.3 Troubleshooting

- **Add** is not available.

  The add action is available to system admins.

- **Assign** does not let you change the account pool.

  When a GPU network already has an account pool assignment, the **Account** field is shown as disabled in the assign dialog.

- **Release from Account** is not shown.

  This action is hidden when the GPU network is not assigned to an account pool, or when it is already assigned to a project pool.

- **Release from Project** is disabled.

  This action is disabled when the GPU network is not assigned to a project pool.

- Batch delete is unavailable.

  Batch delete is disabled when any selected GPU network is assigned to an account pool.

- The delete action fails.

  Retry the delete. If the standard delete does not complete, review whether **Force Delete** is appropriate.

CHAPTER

TWENTYSEVEN

SERVICE CONTROLLER

The Service Controller page is the dashboard for the storage service controller in Storage Management.

The Volume Service Controller (VSC) is zCompute's storage control plane. It provides a volume-as-a-service interface to users while abstracting away the complexity of managing the underlying VPSAs. VSC is a control-plane component only. It runs as an independent microservice that sits between zCompute and the zStorage cloud. The VSC is the component responsible for managing and orchestrating storage subsystem components and objects such as VPSAs, Storage Classes, block storage volumes, volume cloning, snapshots, volume migrations, backup to Object Storage, and more.

Key Concepts:

- **Storage Class**

  An aggregation of one or more uniform VPSAs (same engine type, media type, similar drive counts). The Storage Class defines high-level capabilities, such as performance tier, compression, dedupe, encryption.

- **Volume Type**

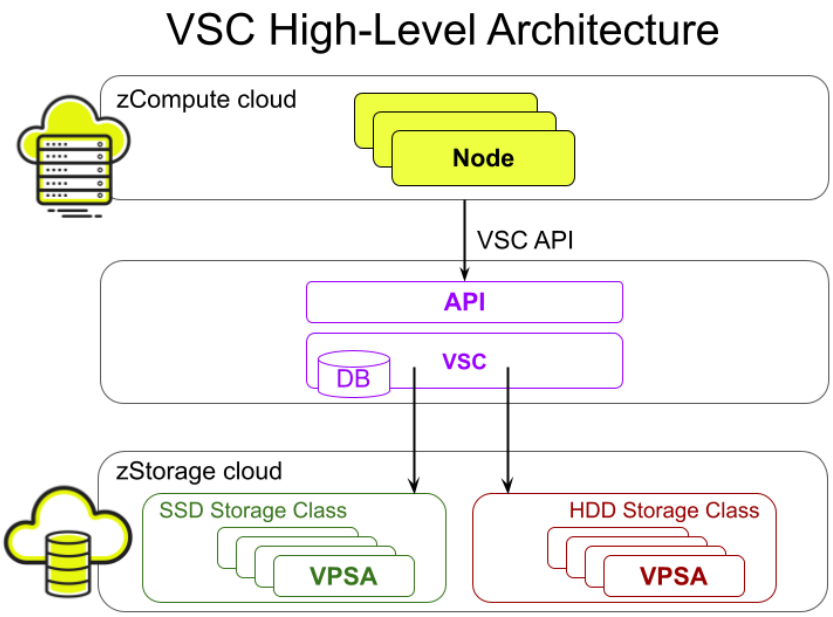
  The Volume Type is derived from a Storage Class with specific QoS limits, such as maximum read/write IOPS and maximum read/write MB/s.

- **Volume**

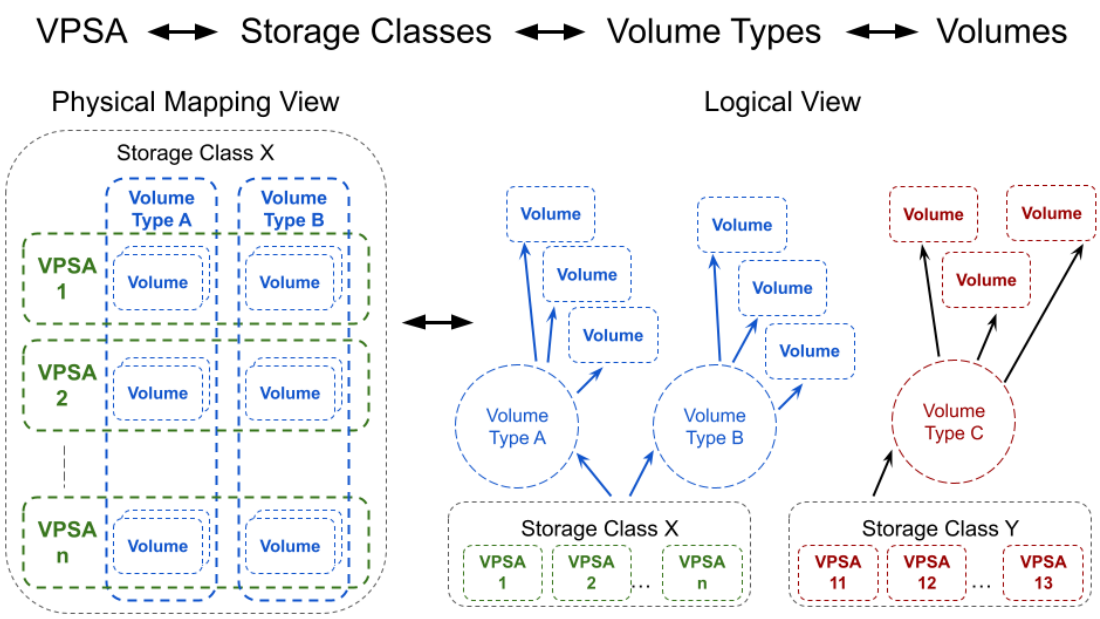
  A Volume is provisioned by the end user defined by a selected Volume Type. The user specifies capacity, and the VSC handles placement across member VPSAs.

- **VSC Hierarchy**

  - VSC high-level architecture:



- VSC abstraction: VPSA ↔ Storage Class ↔ Volume Type ↔ Volume:



Key Benefits

- Simplified provisioning: by specifying capacity and QoS, the VSC handles VPSA selection and lifecycle.
- Near infinitely scalable EBS via horizontal expansion - adding VPSAs to a Storage Class.
- Automatic monitoring of VPSA capacity, performance, and health, with volume redistribution to maintain QoS.
- AWS-style volume types and AWS API/CLI storage-pool to volume-type compatibility.

The Service Controller page brings command center details, storage summary values, and VSC events into one central dashboard.

Use this page to review the controller details that the platform reports, check the current storage counts, and inspect recent storage controller events.

The page also provides filtering, paging, and related-event access for the event list.

27.1 Viewing the service controller dashboard

To view the service controller dashboard:

1. Select **Storage Management > Service Controller**.

   The service controller dashboard opens, displaying the **Command Center Info** and **Summary** panes in the upper part of the screen.

   The lower section of the screen lists summary entries of recent **VSC Events**.

27.1.1 Viewing Command Center Info and Summary

The **Command Center Info** and **Summary** panes appear in the upper part of the service controller dashboard.

- **Command Center Info** details pane:
  - **IP**

    The command center IP address.
  - **Cloud Name**

    The command center name.
  - **ID**

    The command center's unique ID.
- **Summary** pane displays the totals of current storage objects and protection activity:
  - **Storage Classes**

    The number of storage classes in the cloud.
  - **VPSAs**

    The number of Virtual Private Storage Arrays (VPSAs) associated with the cloud.
  - **Volumes**

    The number of volumes in the cloud.
  - **Snapshots**

    The number of snapshots in the cloud, relecting the total sum of **Local snapshots** and **Obs snapshots**.
  - **Protection Groups**

    The number of protection groups.
  - **Backup jobs**

    The number of backup jobs.
  - **Restore jobs**

    The number of restore jobs.

  - **Local snapshot**

    The number of local snapshots.

  - **Obs snapshot**

    The number of Remote Object Storage Snapshots snapshots.

27.1.2 Viewing VSC events

The **VSC Events** pane filters events to display VSC events only.

The VSC Events list has the following columns:

- **Severity**

  The event severity value:

  - **Critical**
  - **Error**
  - **Warning**
  - **Info**

- **Time**

  Date and time that the event occured.

- **Entity Name**

  The entity name linked to the event.

- **Name**

  The event name.

- **Details**

  The event message text.

- **Account**

  The related account, when the event is account-specific.

- **Project**

  The related project, when the event is project-specific.

- **User**

  The username of the user that triggered the event.

- **Related Events**

  Provides access to events related to the event row.

  Selecting **View** redirects to the **Monitoring > Events** screen, filtered to events of the row's **[Event Name]**. See also *Monitoring zCompute Events*.

27.2 Recommended best practices

- Review **Command Center Info** first to confirm that the page is showing the expected command center record.
- Review **Summary** next. The counts provide a quick status view before deeper investigation.
- Set the date range before reading **VSC Events**. This keeps the event list focused.

- Use **Filter…** and **More filters** to reduce large result sets.
- Review **Severity**, **Name**, and **Details** together when you inspect an event row.
- Open **View** in **Related Events** when one row does not provide enough context.

27.3 Troubleshooting

27.3.1 The page opens, but key values are missing

If the panes load without expected values, confirm the visible page areas first.

1. Confirm that **Command Center Info** is shown.
2. Confirm that **Summary** is shown.
3. Reload the page and check the panes again.

27.3.2 The event list is too large to review

Use the visible list controls to reduce the scope.

1. Narrow the date range.
2. Enter text in **Filter…**.
3. Open **More filters** and apply additional limits.
4. Move to the correct page with the paging controls.

27.3.3 You cannot find the event that you need

The event might be outside the current list view.

1. Expand the date range.
2. Clear the value in **Filter…**.
3. Review more pages by using **1**, **2**, **3**, or the page arrows.

27.3.4 A single event row is not enough

Some events are easier to review with their linked records.

1. Find the event row.
2. Select **View** in **Related Events**.
3. Review the linked event records.

CHAPTER

TWENTYEIGHT

STORAGE CLASSES

Storage Classes provide seamless block-storage operations for end-users, by abstracting the underlying block storage implementation details, while providing cloud-operators with a performance-predictable, linear capacity growth.

- **SSD** (Solid State Drive) Class:

  This class uses one or more all-flash (SSD media) VPSAs as its building blocks. It is designed for high-speed performance, offering significantly lower latency and higher IOPS (Input/Output Operations Per Second) than HDDs.

  Suited to high, predictable performance, for example, transactional databases and real-time apps.

- **HDD** (Hard Disk Drive) Class:

  This class uses one or more HDD-based VPSAs as its building blocks. It is optimized for high-capacity and throughput rather than rapid random access, making it slower than SSDs due to the time needed for physical parts to move.

  A more cost-effective option, suited to scenarios for large datasets where occasional latency is acceptable, for example, backups or log processing.

The Storage Classes page shows each storage class and its current resource capacity, health, limits, and related resources.

From this page, you can:

- Review the available storage classes.
- Open a storage class to see its details.
- View the volumes, snapshots, and VPSAs linked to a storage class.
- Attach a VPSA to a storage class.

28.1 Storage Classes management

28.1.1 Viewing storage classes

To view the storage classes that are available in the system:

1. Select **Storage Management > Storage Classes**.

   The table of **Storage Classes** appears, showing the following fields:

   - **Name**

     The storage class name.

     This identifies the storage class in the list.

   - **Capacities**

     Note

  - **Physical Capacity** refers to the actual, usable space on the hard disk drives (HDDs) or solid-state drives (SSDs) in a storage pool.
  - **Virtual Capacity** refers to the capacity provisioned (assigned) to hosts or volumes, which can be larger than the physical capacity, due to thin provisioning.

  Zadara uses Virtual Capacity to present a specific size to a server, but that volume only consumes Physical Capacity when data is actually written to it.

  - **Physical Capacity - Usable**

    The amount of physical storage capacity that can be used.

  - **Physical Capacity - Used**

    The amount of physical storage capacity that is currently in use.

  - **Physical Capacity - Free**

    The amount of physical storage capacity that is still available.

  - **Virtual Capacity - Total**

    The total virtual capacity presented by the storage class.

  - **Virtual Capacity - Provisioned**

    The amount of virtual capacity that is already provisioned.

  - **Virtual Capacity - Free**

    The amount of virtual capacity that is still available for new use.

- **Max New Volume**

  The largest size allowed for a newly created volume in this storage class.

- **Available Volumes**

  The number of additional volumes that can still be created.

- **VPSAs**

  The number of VPSAs attached to the storage class.

- **Volume Types**

  The number of volume types associated with the storage class.

- **Default**

  Indicates whether this storage class is the default choice.

- **Built In**

  Indicates whether this storage class is built in, and not a custom storage class.

- **State**

  The current operational state of the storage class.

  Possible values:

  - **Ready**: Usable
  - **Deleting**: In the process of being removed
  - **Failed**: Not usable:Failed or unhealthy

- **Capacity Mode**

  The current capacity handling mode for the storage class, indicating whether the storage class is operating normally or approaching a capacity-related risk.

  Possible values:

  - **Normal**

    Indicates that the storage class is in its normal operating condition, regarded as **Healthy**.

  - **Degraded**

    Indicates that the storage class is in a reduced-capacity condition.

  - **Emergency**

    Indicates that the storage class is in a serious capacity condition, regarded as **Unhealthy**.

- **Health State**

  The current health condition of the storage class.

  Possible values:

  - **Healthy**
  - **PartiallyFailed**
  - **Failed**

- **Limit State**

  The current limit condition of the storage class, how close the storage class is to its configured limits, and whether the VPSAs in that storage class are still within normal operating bounds.

  Possible values, in increasing levels of severity:

  - **Normal**
  - **Alert**
  - **Degraded**
  - **Emergency**

2. Select a storage class to view further details.

28.1.2 Viewing a storage class details

To view a storage class details:

1. Select **Storage Management > Storage Classes**.

   The table of **Storage Classes** appears.

2. Select the storage class that you want to inspect.

   The storage class detail screen appears, with the following sections:

   - *Storage class details pane* in the upper part of the screen.
   - The **Tabs** section in the lower part of the screen:
     - *Storage class Volumes tab*
     - *Storage class Snaphots tab*
     - *Storage class VPSAs tab*

Storage class details pane

The main details section in the upper part of the screen displays the following panes:

- **Info** details displaying:
  - The storage class operational **State** and **Health State**.
  - **Limit State** - the current limit condition for the storage class.
  - **Name**: The storage class name.
  - **ID**: The storage class unique identifier.
  - **Encryption**: Indicates whether encryption is enabled for the storage class.
  - **Deduplication**: Indicates whether deduplication is enabled for the storage class.
  - **Compression**: Indicates whether compression is enabled for the storage class.
  - **Max. Volumes per VPSA**: The maximum number of volumes allowed on each VPSA of the storage class.
  - **Max New Volume**: The largest size allowed for a new volume in this storage class.
- **Capacity** guages:

  Guages dashboard displaying **Used** and **Free** capacity for:
  - **Provisioning Capacity**
  - **Physical Capacity**
- **Resources** summary, providing the totals of:
  - **User Volumes**
  - **Snapshots**
  - **VPSAs**

  A detail tab for each type of resource is available in the lower part of the screen.

Storage class Volumes tab

The **Volumes** tab displays the list of volumes of this storage class, with the following columns:

- **Name**: The volume name.
- **VPSA**: The VPSA that hosts the volume.
- **Owner**: The ownership type indicates whether the volume is **User**-defined (regular volume) or **System**-assigned (such as a volume marked for deletion).
- **Size**: The provisioned size of the volume.
- **Volume Type**: The volume type assigned to the volume.
- **Account**: The name of the account that owns the volume.
- **User**: The username of the user associated with the volume.
- **Status**: The current operational status of the volume.
- **Health**: The current health status of the volume.
- **Attached VM**: The name of the VM to which the volume is attached, if any.
- **Creation Date**: The date and time when the volume was created.

Storage class Snaphots tab

The **Snaphots** tab displays a table of the storage class snaphots, and quick filters for rapid location and listing of snapshots matching selected criteria.

The list of snapshots of this storage class, are displayed with the following columns:

- **Name**: The snapshot name.
- **Source name**: The name of the source object from which the snapshot was taken.
- **Source ID**: The source object unique identifier.
- **Created At**: The date and time when the snapshot was created.
- **Project**: The project in which the snapshot's source object resides.
- **Account**: The name of the account that owns the snapshot's project.
- **User**: The username of the user that triggered the snapshot.
- **Status**: The current operational status of the snapshot's storage class.
- **Health**: The current health status of the snapshot's storage class.
- **Protection Group**: The protection group in which the snapshot's source object is a member.
- **Storage Classes**: The storage classes of the snapshot's source objects.
- **VPSAs**: The VPSA hosting the snapshot.
- **Creation Date**: The date and time when the snaphot was created.

Storage class VPSAs tab

The **VPSAs** tab displays the list of VPSAs of this storage class, with the following columns:

- **Name**

  The VPSA name.

- **Nova ID**

  The VPSA's Nova identifier, which might differ from the display name.

- **State**: The current operational state of the VPSA.
- **Physical Capacity - Usable**: The amount of usable physical capacity on the VPSA.
- **Physical Capacity - Used**: The amount of physical capacity already in use on the VPSA.
- **Physical Capacity - Free**: The amount of unused physical capacity still available on the VPSA.
- **Virtual Capacity - Total**: The total virtual capacity available on the VPSA.
- **Virtual Capacity - Provisioned**: The virtual capacity already provisioned on the VPSA.
- **Virtual Capacity - Free**: The virtual capacity still available on the VPSA.
- **Volume Count**: The number of volumes currently on the VPSA.
- **Version**: The VPSA's zStorage version.
- **Pool Type**: The VPSA's storage pool type.
- **Engine**: The VPSA's IO engine type.
- **Snapshot Count**: The number of snapshots on the VPSA.
- **Capacity Mode**: The VPSA's current current capacity handling mode for its storage class.

28.1.3 Attaching a VPSA to a storage class

There may be a number of reasons for attaching a VPSA to a storage class, such as extending the storage capacity or for providing for more than the current maximum number of volumes.

To attach a VPSA to a storage class:

1. Select **Storage Management > Storage Classes**.

   The table of **Storage Classes** appears.

2. Select the storage class to which you want to attach a VPSA.

   The storage class detail screen appears.

3. In the top toolbar, select **Attach VPSA**.

   The **Attach VPSA** dialog opens.

4. In the **Attach VPSA** dialog, enter the required fields:

   - **IP Address**: The IP address of the VPSA that you want to attach.
   - **Token / Temporary password**: The token or temporary password used to authenticate the attach action.

     One of:

     – **Token**: The **Access Key**, available in the VPSA's admin **User Information**:

     > **Note**
     >
     > Relevant when the initial temporary password has been changed.

     1. In the VPSA UI, select the username at the top right.

        The **User Information** screen opens.

     2. In the **Authentication** section, copy the **Access Key** value.

     – **Temporary password** emailed to the admin on creation of the VPSA.

     > **Note**
     >
     > Relevant only if this temporary password wasn't yet changed.

   - **VPSA Nova ID**: The Nova ID of the VPSA that you want to attach.

5. To apply attaching the VPSA to the storage class, select **Finish**.

28.2 Recommended best practices

- Storage class limits are preset to provisioned capacity too, which are most likely reached first, before physical capacity is reached. Monitor your storage class total provisioned capacity and add capacity as additional VPSAs in advance, to avoid volume creation failures. The same applies to volume-count per VPSA limits. Make sure that you add capacity in the form of additional VPSAs to the storage class before you hit the limit.
- When VPSAs are added to a storage class, VSC doesn't re-distribute the volumes across VPSAs. In case there's a need to do that, for example, in case of a single VPSA performance/overload issue, administrators can live-redistribute volumes across VPSAs within the same storage class.
- Check the health, state, and limit indicators before you make changes.

- Use the available filters when lists become large.
- Before attaching a VPSA to a storage class, verify the IP address, token or temporary password, and VPSA Nova ID carefully.

28.3 Troubleshooting

28.3.1 You cannot find the storage class you need

- Use the **Filter** field on the main page to narrow the list.
- Check whether the page is shown in the view that you expect.

28.3.2 The storage class details you need are not visible

- Make sure the correct storage class row is selected.
- Open the correct tab: **Volumes**, **Snapshots**, or **VPSAs**.

28.3.3 A linked resource is difficult to locate

- Use the filter fields available on the tab.
- Review the status and health columns to identify the correct row.

28.3.4 A VPSA attach action does not complete as expected

- Recheck the values entered in **IP Address**.
- Recheck the values entered in **Token / Temporary password**.
- Recheck the values entered in **VPSA Nova ID**.
- Confirm that you selected the intended storage class before starting the attach action.

CHAPTER

TWENTYNINE

VOLUME TYPES

Volume types define the characteristics and performance profile of block storage volumes that you can attach to your virtual machines (VMs). Each volume type determines:

- **Underlying storage hardware**, for example, SSD and HDD
- **Performance limits**, such as IOPS and throughput
- **Features**, such as encryption, deduplication, and compression
- **Volume Retype**

  You can sometimes change a volume's type after creation, for example migrate from HDD to SSD, but there may be restrictions based on storage class, encryption, or capacity.

Certain applications such as databases and analytics require high IOPS and can benefit from SSDs, while others, such as backups and logs can performance adequately with HDDs.

Example Volume Types and Their Effects

Volume Type	Media	Typical Use Case	Performance	Features
zvtssd1 (gp2)	SSD	General purpose	Medium	Encryption, no dedup
zvtssd2 (gp3)	SSD	High-performance DB	High	Encryption, no dedup
zvtssdmx	SSD	Burst workloads	Very High	No IOPS limit
zvthdd1 (sc1)	HDD	Archival, backups	Low	Encryption, no dedup
zvthddmx	HDD	Burst archival/logging	Medium	No IOPS limit

29.1 Volume Types Management

The **Volume Types** screen provides the following actions:

- Viewing volume types
- Creating a new volume type
- Modifying an existing volume type
- Delete a volume type
- Enable or disable provisioning for a volume type
- Setting a volume type as the default for new volumes

29.1.1 Viewing volume types

To view the volume types:

1. Select **Storage Management > Volume Types**.

   The **Volume Types** screen opens, displaying the list of volume types with the following columns:

   - **Name**

     The volume type name. When a description exists, it is also displayed.

   - **API Alias**

     An alternate name for API-based access. This value can be used instead of the main volume type name.

   - **Provisioned Capacity**

     The amount of capacity already provisioned through this volume type.

   - **Built In**

     Indicates whether the volume type is a built-in system volume type.

     Note

     Built In volume types cannot be modified.

   - **Default**

     Indicates whether the volume type is the current default for new volumes.

   - **Storage Class**

     Shows which storage class the volume type uses.

   - **Health**

     Shows the health status of the volume type.

   - **Provisioning**

     Shows whether provisioning is enabled or disabled for the volume type.

   - **State**

     Shows the current state of the volume type:

     - **Normal**
     - **No Capacity**

2. When you mark a volume type's checkbox, the lower panel shows summary details:

   The **Info** area shows:

   - **Storage Class**

     Shows which storage class the volume type uses.

   - **Default**

     Indicates whether the volume type is the current default for new volumes.

   - **Provisioning Enabled**

     Shows whether new provisioning is allowed for the selected volume type.

- **Built In**

  Shows whether the selected volume type is built in.

- **Creation Date**

  The date and time when the volume type was created.

- **Modification Date**

  The date and time when the volume type was last updated.

- **Provisioned Capacity**

  The capacity already provisioned through the selected volume type.

- **ID**

  The unique identifier of the selected volume type.

  The **Capping** area shows whether the following features are enabled for the selected volume type:

- **Encryption**
- **Dedupe**
- **Compression**

3. Select a storage class name to view further details.

29.1.2 Viewing a volume type's details

1. Select **Storage Management > Volume Types**.

   The **Volume Types** screen opens, displaying the list of volume types.

2. In the volume types list, select the volume type **Name** to view that volume type's details.

   The volume type's detail screen appears, with the following sections:

   - *Volume type details pane* in the upper part of the screen.
   - The **Tabs** section in the lower part of the screen:
     - *Volume type More Info tab*
     - *Volume type Volumes tab*
     - *Volume type Events tab*

Volume type details pane

The **Info** section gives a quick summary of the selected volume type.

It shows the current health and state, and the main settings and identifiers for the volume type:

- **Health**: The current health status of the volume type.
- **State**: The current operating state of the volume type.
- **Storage Class**: The storage class assigned to the volume type.

  This value is a link to the related storage class screen.

- **Default**: Indicates whether this volume type is the default choice for new volumes.
- **Provisioning Enabled**: Indicates whether new volumes can be provisioned with this volume type.
- **Built in**: Indicates whether the volume type is a built-in system volume type.

- **Creation Date**: The date and time that the volume type was created.
- **Modification Date**: The date and time that the volume type was last updated.
- **Provisioned Capacity**: The total capacity already provisioned by volumes that use this volume type.
- **ID**: The unique identifier of the volume type.

Volume type More Info tab

The **More Info** tab shows additional settings for the selected volume type.

Use this tab to review descriptive details, sharing settings, data services, and performance limits for the volume type:

- **Description**: The descriptive text for the volume type.
- **Alias**: The alternate API name for the volume type.

  When an alias is set, it is the volume type name presented through the AWS compatibility APIs. If no alias is set, the original volume type name is used.
- **Shared**: Indicates whether the volume type is shared.
- **Dedupe**: Indicates whether deduplication is enabled for the volume type.
- **Encryption**: Indicates whether encryption is enabled for the volume type.
- **Compression**: Indicates whether compression is enabled for the volume type.
- **Read IOPS limit**: The maximum read IOPS limit for volumes that use this volume type.
- **Write IOPS limit**: The maximum write IOPS limit for volumes that use this volume type.
- **Read bandwidth limit**: The maximum read bandwidth limit, in MB/s, for volumes that use this volume type.
- **Write bandwidth limit**: The maximum write bandwidth limit, in MB/s, for volumes that use this volume type.

Volume type Volumes tab

The **Volumes** tab lists the volumes that use the selected volume type. It supports row selection, which in turn shows per-volume actions in the toolbar and the **More** menu.

The table includes the columns:

- **Name**: The volume name. When a volume description exists, it is displayed under the name.
- **Size**: The volume's configured size.
- **Volume Type**: The volume type assigned to the volume.

  This value is a link to the related volume type screen.
- **Account**: The account that owns the volume.
- **User**: The user associated with the volume.
- **Status**: The current volume status.

  Possible values include:

  - **attaching**
  - **deleted**
  - **destroying**
  - **detaching**
  - **downloading**

  - **error**
  - **extending**
  - **locked**
  - **ready**
  - **uninitialized**
  - **uploading**
- **Health**: The current health value for the volume.
- **Attached VM**: The VM to which the volume is attached.

  If a VM is attached, the value is a link to the screen of that VM.

  If no VM is attached, the entry shows **Not attached to any VM**.
- **Creation Date**: The date and time that the volume was created.

Available volume actions

The toolbar and row action menu in the **Volumes** tab expose the following actions for a selected volume:

- **Attach**: Connect an unattached volume to a VM.
- **Detach**: Disconnects an attached volume from its VM.
- **Clone**: Creates a clone from the selected volume.
- **Snapshot**: Creates a snapshot from the selected volume.
- **Delete**: Deletes the selected volume.
- **Create Image**: Creates an image from the selected volume.
- **Launch**: Opens the VM creation dialog, with the selected volume passed as the source volume.
- **More** submenu:
  - **Extend**: Opens the volume extension dialog for the selected volume.
  - **Protect**: Opens the resource protection flow for the selected volume.
  - **Change Volume Type**: Opens the volume type change flow for the selected volume.
  - **Create Alarm**: Opens alarm creation dialog for the selected volume.

Volume type Events tab

The **Events** tab shows event records for the selected volume type.

By default, the general event columns **Entity Name** and **Entity Type** are hidden, because the page is already scoped to the specified volume type.

Columns include:

- **Severity**: The event severity level.
- **Time**: Date and time that the event was recorded.
- **Name**: The event type name.
- **Details**: The event description.
- **Account**: The account associated with the event.
- **Project**: The project related to the event.

- **User**: The username of the user that triggered the event.
- **Related Events**: When the event has a related request ID, it shows a **View** to open related event records.

29.1.3 Creating a volume type

To create a new volume type:

1. Select **Storage Management > Volume Types**.

   The **Volume Types** screen opens, displaying the list of volume types.

2. From the top toolbar, select **+ Create**.

   The **Create Volume Type** dialog opens.

3. In the **Details** step, enter the following:

   - **Name**

     The volume type name.

     Required.

   - **Description**

     Optional text that helps identify the purpose of the volume type.

   - **Alias**

     Optional alternate name for AWS compatibility APIs. If no value is provided, the original name is used.

   - **Storage Class**

     From the dropdown, select the storage class that backs the volume type.

     Required.

   - **Set as Default for New Volumes**

     Select this volume type as the default choice for new volumes.

   - **Disable Provisioning**

     Creates the volume type with provisioning disabled.

4. Select **Next** to continue to the **Performance** step.

5. In the **Performance** step, enter the following:

   - **Max read IOPs**

     Sets the read IOPS limit for volumes that use this volume type.

   - **Max write IOPs**

     Sets the write IOPS limit for volumes that use this volume type.

   - **Max read MB/s**

     Sets the read bandwidth limit in MB/s.

   - **Max write MB/s**

     Sets the write bandwidth limit in MB/s.

   - **Enable deduplication**

     Turns deduplication on for this volume type.

     This setting depends on storage class capability.

- **Enable compression**

  Turns compression on for this volume type.

  This setting depends on storage class capability.

- **Enable encryption**

  Turns encryption on for this volume type.

- **Shared**

  Controls whether the volume type is shared.

- **Allowed Accounts**

  From the dropdown, select the accounts that can use the volume type when it is not shared.

6. To confirm creation of the volume type, select **Finish**.

29.1.4 Modifying a volume type

Note

Built In volume types cannot be modified.

To modify an existing volume type:

1. Select **Storage Management > Volume Types**.

   The **Volume Types** screen opens, displaying the list of volume types.

2. In the volume types list, locate the volume type to modify.

   Select the volume type by one of the following methods:

   - Mark the volume type's checkbox.
   - Click the volume type's name.

   The **Modify** option will display in the top toolbar.

3. From the top toolbar, select **Modify**.

   The **Modify Volume Type** dialog opens.

4. In the **Details** step, you can optionally update the following:

   - **Name**

     The volume type name.

   - **Description**

     Optional text that helps identify the purpose of the volume type.

   - **Alias**

     Optional alternate name for AWS compatibility APIs. If no value is provided, the original name is used.

5. Select **Next** to continue to the **Performance** step.

6. In the **Performance** step, you can optionally update the following:

   - **Max read IOPs**

     Sets the read IOPS limit for volumes that use this volume type.

- **Max write IOPs**

  Sets the write IOPS limit for volumes that use this volume type.

- **Max read MB/s**

  Sets the read bandwidth limit in MB/s.

- **Max write MB/s**

  Sets the write bandwidth limit in MB/s.

- **Enable deduplication**

  Turns deduplication on for this volume type.

  This setting depends on storage class capability.

- **Enable compression**

  Turns compression on for this volume type.

  This setting depends on storage class capability.

- **Shared**

  Controls whether the volume type is shared.

- **Allowed Accounts**

  From the dropdown, select the accounts that can use the volume type when it is not shared.

7. To confirm updating the changes to the volume type, select **Finish**.

29.1.5 Deleting a volume type

Note

The current default volume type cannot be deleted.

To delete it, first set another volume type as the default.

To delete a volume type:

1. Select **Storage Management > Volume Types**.

   The **Volume Types** screen opens, displaying the list of volume types.

2. In the volume types list, locate the volume type to delete.

   Select the volume type by one of the following methods:

   - Mark the volume type's checkbox.
   - Click the volume type's name.

   The **Delete** option will display in the top toolbar.

3. From the top toolbar, select **Delete**.

   The **Delete Volume Type** confirmation dialog opens, displaying the name of the volume type to delete.

4. To confirm deleting the selected volume type, select **Delete**.

29.1.6 Disabling provisioning for a volume type

To stop new provisioning through a volume type.

1. Select **Storage Management > Volume Types**.

   The **Volume Types** screen opens, displaying the list of volume types.

   The **Provisioning** column shows whether the volume type is **Enabled** or **Disabled**.

2. In the volume types list, locate the volume type to disable.

   Select the volume type by one of the following methods:

   - Mark the volume type's checkbox.
   - Click the volume type's name.

   The **Disable** option will display in the top toolbar.

3. From the top toolbar, select **Disable**.

   The **Disable Volume Type** confirmation dialog opens, displaying the name of the volume type to disable.

4. To confirm disabling the selected volume type, select **Disable**.

   The selected volume type's **Provisioning** column changes to **Disabled**.

29.1.7 Enable provisioning for a volume type

To allow new provisioning through a volume type.

1. Select **Storage Management > Volume Types**.

   The **Volume Types** screen opens, displaying the list of volume types.

   The **Provisioning** column shows whether the volume type is **Enabled** or **Disabled**.

2. In the volume types list, locate the volume type to enable.

   Select the volume type by one of the following methods:

   - Mark the volume type's checkbox.
   - Click the volume type's name.

   The **Enable** option will display in the top toolbar.

3. From the top toolbar, select **Enable**.

   The **Enable Volume Type** confirmation dialog opens, displaying the name of the volume type to enable.

4. To confirm enabling the selected volume type, select **Enable**.

   The selected volume type's **Provisioning** column changes to **Enabled**.

29.1.8 Setting a default volume type

To choose which volume type is used by default for new volumes.

Note

A volume type that is already the default cannot be set again.

1. Select **Storage Management > Volume Types**.

   The **Volume Types** screen opens, displaying the list of volume types.

   The **Default** column shows which volume type is the current default.

2. In the volume types list, locate the volume type to set as default.

   Select the volume type by one of the following methods:

   - Mark the volume type's checkbox.
   - Click the volume type's name.

   The **Set Default** option will display in the top toolbar.

3. From the top toolbar, select **Set Default**.

   In the **Default** column, the **Default** check mark indication moves to the selected volume type.

   On provisioning, new volumes will default to the selected volume type.

29.2 Recommended best practices

- Only use the default volume types.
- Use the **Alias** field only when you need an alternate AWS API name.
- Use provisioning disablement when you need to stop new use of a volume type without deleting it.
- Review the **Shared** setting carefully. If the volume type is not shared, set the correct accounts in **Allowed Accounts**.
- Use the details panel to confirm health, state, provisioning status, and capping settings after a change.

29.3 Troubleshooting

- **Modify is unavailable**

  The selected volume type might be built in. Built-in volume types cannot be modified.

- **Delete is unavailable**

  The selected volume type might be the current default. Set a new default volume type first, and then delete the old one.

- **Provisioning shows Disabled**

  New provisioning is turned off for that volume type. Use **Enable** to allow provisioning again.

- **Set Default is unavailable**

  The selected volume type is already the default.

- **Allowed Accounts cannot be edited**

  The volume type is set as shared. Clear **Shared** first to make account selection relevant.

- **Dedupe or compression cannot be changed**

  The selected storage class might not support that capability.

- **Encryption cannot be changed during modification**

  The modify flow does not show the encryption field. Review the existing capping values in the details panel.

CHAPTER

THIRTY

BACKUP AND RESTORE TASKS

Backup and Restore Tasks data is associated with backup protection groups.

This page displays Backup to Object Storage (B2OS) tasks only. Local (snapshot) operations which don't involve B2OS don't appear on this Backup and Restore Tasks page.

This page allows an admin to view running B2OS tasks across the system for all tenants, mostly to provide an overview and visibility for the admin to monitor running B2OS tasks from a single pane of glass, and to provide details in case something goes wrong, for example, a B2OS task that is taking too long to finish, too many tasks running against the same VPSA, telco-line, or Object Storage, and more.

In normal use, this page shows task data generated by protection-group backup or restore activity. If no backup protection groups exist, the page will typically have no task data to show. If task records exist but related protection-group objects are unavailable, some rows may still appear, but the protection-group field can be blank.

The page shows task records in a table view. It combines task data with related volume data, backup protection group data, and remote snapshot data.

The page updates automatically while it is open. The data refresh interval is 10 seconds.

30.1 View current tasks

Use this operation to review the tasks that are currently being executed.

To open the page:

1. Select **Storage Management > Backup / Restore Tasks**.

   The **Backup / Restore Tasks** page opens, listing task records in a table view.

   UI fields and prompts on this page include:

   - **Backup Protection Group**

     The related backup protection group for the task.

     Use it to identify which protection group started the task.

   - **VSC Backup Protection Group ID**

     The external identifier of the backup protection group.

     Use it when you need the backend-facing group ID for tracking or support.

   - **Snapshot**

     The related remote snapshot resource for the task, when one is available.

     Use it to identify the snapshot tied to the task.

- **Volume**

  The related volume for the task.

  Use it to identify which volume the task is working on.

- **VSC Snapshot ID**

  The external identifier of the snapshot.

  Use it when you need the backend-facing snapshot ID.

- **VSC Volume ID**

  The external identifier of the volume.

  Use it when you need the backend-facing volume ID.

- **Volume Size**

  The size of the related volume.

  Use it to understand the scope of the task.

- **VPSA Name**

  The name of the VPSA associated with the task.

  Use it to see which storage service instance is involved.

- **VPSA Status**

  The job status reported for the VPSA.

  Use it to compare the storage-side job state with the task state.

- **VPSA Nova ID**

  The internal identifier for the VPSA.

  Use it for correlation and support work.

- **Status**

  The snapshot task status.

  Use it to check whether the task is running, completed, failed, or in another reported state.

- **Snapshot Type**

  The type of snapshot used by the task.

  Use it to distinguish between task types in the table.

- **Capacity**

  The task capacity value in storage units.

  Use it to understand the amount of data involved in the task.

- **Progress**

  The reported progress value for the task.

  Use it to track task completion.

- **Average Sync Rate**

  The average synchronization rate for the task.

  Use it to judge current transfer performance.

- **Estimated Remaining Time**

  The reported remaining time for the task.

  Use it to estimate when the task may finish.

- **Job ID**

  The backend job identifier for the task.

  Use it when you need a specific job reference for logs or support.

30.1.1 View related resource details

Linked resources allow drilldowns from a task entry to the related resource:

- Backup protection group entries
- Snapshot entries
- Volume entries

To open a related resource:

1. Locate the task row you want to inspect.
2. In the row, select the linked resource name in the relevant column:
   - **Backup Protection Group**

     Opens the related protection group details page.

   - **Snapshot**

     Opens the related remote snapshot details page.

   - **Volume**

     Opens the related volume details page.
3. Review the destination page for that resource.

30.1.2 Show or hide table columns

Use the filters to control which task details are visible in the table.

The page provides a column picker in the list view.

To change the visible columns:

1. Open the Backup / Restore Tasks page.
2. In the list view, open the column picker.
3. Select the columns you want to show.
4. Clear the columns you do not want to show.
5. Review the updated table.

Columns marked as non-default in the UI include:

- **VSC Backup Protection Group ID**

  Hidden by default. Show it when you need the external group ID.

- **VSC Snapshot ID**

  Hidden by default. Show it when you need the external snapshot ID.

- **VSC Volume ID**

  Hidden by default. Show it when you need the external volume ID.

- **Volume Size**

  Hidden by default. Show it when you want more size detail in the table.

- **VPSA Nova ID**

  Hidden by default. Show it when you need the VPSA internal ID.

30.2 Recommended best practices

- Keep the page open while you are watching active storage work. The page refreshes automatically every 10 seconds.
- Use **Status**, **Progress**, **Average Sync Rate**, and **Estimated Remaining Time** together. This gives a clearer view than checking one field alone.
- Turn on the non-default ID columns before opening a support case or checking backend logs. These IDs are useful for exact correlation.
- Open the linked **Backup Protection Group**, **Snapshot**, or **Volume** entry when you need more context about a task.
- When the table is empty, confirm that there are no active tasks before you continue troubleshooting elsewhere.

30.3 Troubleshooting

30.3.1 The page shows no rows

If the page shows no task rows, check whether the page message says that there are no tasks being executed.

If that message is shown, the UI is reporting that no active tasks are currently listed.

30.3.2 The page does not finish loading

The page uses a loading spinner until its data is fetched.

If the spinner stays visible, the required data may not have finished loading.

Refresh the page and check again.

30.3.3 A related name is blank

Some related fields are shown only when a matching resource is found.

If **Backup Protection Group**, **Snapshot**, or **Volume** is blank, the task row may not have a resolved linked resource in the current data.

Use the corresponding ID field, if available, to continue your investigation.

30.3.4 You need more detail for a task

Turn on the non-default columns from the column picker.

These columns expose additional identifiers and size values that can help you trace the task.

30.3.5 You need more context about the resource behind a task

Open the linked entry from the **Backup Protection Group**, **Snapshot**, or **Volume** column.

This is the supported way to move from the task list to the related resource details.

CHAPTER

THIRTYONE

NETWORKING AND LOAD BALANCER SERVICE ENGINES

The Service Engines screens show service engine groups by type:

- **Networking** service engines that are deployed in the region.

  The Networking category includes networking service engines such as **DNS** related engines and the **NAT Gateway** service.

- **Load Balancer** service engines that are deployed in the region.

Each service engine type uses the same screen layout and has the same functional controls.

The difference is the service engine group that you select from the left navigation.

These screens can be used to review service engine entries, check version and status information, and manage the enabled state for each listed engine version.

31.1 Service Engines management

Service engines are grouped by type and opened from the group icon in the left navigation pane.

The **Service Engines** screens show the service engines that are deployed in the region.

This screen helps an administrator to see which service engines are present, which version each engine uses, whether a newer version is available, whether the engine is enabled, and the current engine status.

The **Service Engines** screens support these lifecycle operations when they are available for the current engine state:

- Initialization
- Reset after an initialization failure
- Enable
- Disable
- Upgrade

The screen also shows the current state of each engine entry:

- **Engine**

  The service engine name.

- **Version**

  The current engine version.

- **Upgrade Available**

  The target version that is available for upgrade. It indicates whether there is a newer version that can be applied.

- **Updated At**

  The date and time when the engine version entry was last updated.

- **Enabled**

  The toggle to enable or disable the service engine version.

- **Status**

  The current engine state, such as **Active** or **Disabled**, indicates whether the engine version is currently in service.

- **Filter**

  The list filter box in the upper-right area of the screen. Use this field to narrow the list when many engine entries are shown.

31.1.1 Initializing a service

When the service is not initialized, the UI shows an initialization action instead of the normal ready-state list workflow.

To initialize a service:

1. Select **Service Engines** and then the relevant type: **Networking** or **Load Balancer**.
2. Locate the initialization prompt for the service.
3. Select **Initialize**.
4. Wait for the service state to change from the initialization flow to the normal screen state.
5. Review the screen again and confirm that engine entries are listed.

31.1.2 Resetting a service after a failed initialization

When the service is in an error state, the UI shows a reset action.

To initialize a service:

1. Select **Service Engines** and then the relevant type: **Networking** or **Load Balancer**.
2. Locate the failed initialization prompt for the service.
3. Select **Reset**.
4. Wait for the reset and initialization flow to complete.

#. Review the screen again and confirm that the service returns to a usable state.

31.1.3 Enabling an engine version

To place an engine version into service:

1. Select **Service Engines** and then the relevant type: **Networking** or **Load Balancer**.
2. Find the engine row that you want to enable.
3. In the **Enabled** column, turn on the toggle.
4. Wait for the screen to refresh.
5. Confirm that **Status** shows an active state.

31.1.4 Disabling an engine version

To remove an engine version from service:

1. Select **Service Engines** and then the relevant type: **Networking** or **Load Balancer**.
2. Find the engine row that you want to disable.
3. In **Enabled** column, turn off the toggle.
4. Wait for the screen to refresh.
5. Confirm that **Status** shows **Disabled**.

31.1.5 Upgrading an engine version

The UI provides an upgrade action only when the engine version is marked as upgradable.

To upgrade a service engine:

1. Select **Service Engines** and then the relevant type: **Networking** or **Load Balancer**.
2. Select the engine row that shows a value in **Upgrade Available**.
3. From the top toolbar, select **Upgrade** for the engine version.
4. In the confirmation dialog, review the engine name and version.
5. Confirm the upgrade.
6. Wait for the upgrade operation to finish.
7. Review the **Version**, **Upgrade Available**, and **Status** columns to confirm the new state.

31.2 Recommended best practices

- Review the **Status** field before making a change.
- Review **Upgrade Available** before planning maintenance.
- Use the **Enabled** toggle only for the specific engine row you intend to change.
- After any change, refresh your review of **Version**, **Updated At**, and **Status**.
- Use **Filter** to reduce the risk of changing the wrong engine when the list grows.
- Do not treat **Pool** as a planning field for disabled entries, because the screen can show that field only for enabled versions.

31.3 Troubleshooting

31.3.1 An engine shows as Disabled

A disabled engine is not currently enabled for service.

1. Check the **Enabled** toggle for the row.
2. If needed, turn on the toggle.
3. Review **Status** again after the screen refreshes.

31.3.2 No engine entries are listed

The service might not be initialized yet.

1. Check whether the screen shows an initialization prompt.
2. If it does, select **Initialize**.
3. Wait for the service to finish initialization.
4. Return to the screen and review the list again.

31.3.3 Initialization failed

The service can enter an error state during initialization.

1. Check whether the screen shows a reset prompt.
2. Select **Reset**.
3. Wait for the service to return to a usable state.
4. Review the screen again.

31.3.4 Upgrade is not available

The engine version might not be marked as upgradable.

1. Review the **Upgrade Available** field for the row.
2. If no upgrade value is shown, no upgrade is currently available from this screen.
3. Continue to monitor the screen for a future upgrade entry.

31.3.5 Pool is empty

This can occur when the engine version is not enabled.

1. Check the **Enabled** toggle for the row.
2. Check the **Status** field.
3. If the version is disabled, expect **Pool** to be empty on this screen.

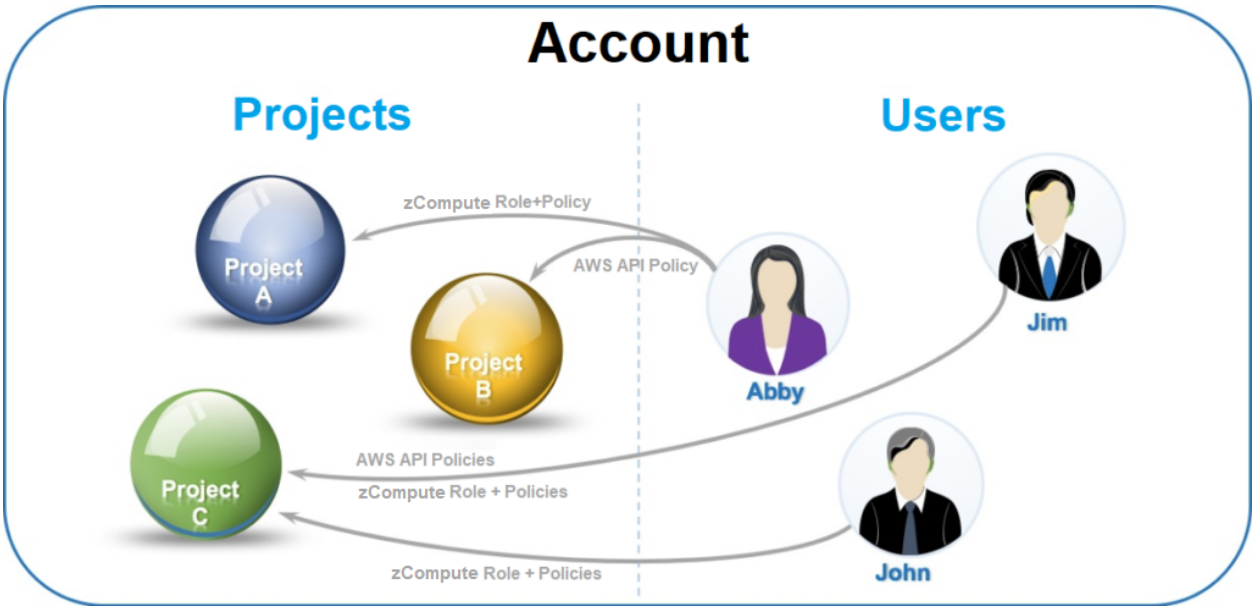
CHAPTER

THIRTYTWO

INTRODUCTION TO IDENTITY AND ACCESS

32.1 Accounts and Projects

The virtual resources of the Zadara Cloud Services region are managed through an administrative hierarchy of accounts and projects as shown in figure below. The Zadara Cloud Services region consists of one or more accounts, each of which contains one or more projects. Virtual resources such as VLANs, instances, volumes, images and snapshots are created per project, account or region. Users or User Groups, which are members of an account, can be assigned different projects within their account through a Zadara Cloud Services role, together with one or more Zadara Cloud Services and AWS API policies. This assignment enables you to provide access to Zadara Cloud Services to numerous users or groups, while dividing and separating the resources that each user will be able to view, create, and manage.



See the video on the basics of zCompute Identity and Access:

32.2 Identity Provider (IdP)

Zadara Cloud Services users and groups can be accessed from an identity provider such as MS Active Directory. This is done by first connecting the Zadara Cloud Services account to the MS Active Directory. Connection with the identity provider must be configured independently for each Zadara Cloud Services account.

See *Connecting an Account to a Microsoft Active Directory Identity Provider* for detailed steps.

Once an account becomes connected to Active Directory, the newly available Active Directory users and groups must be assigned Zadara Cloud Services Project Roles and Policies. This can be done either through CLI commands or via the GUI.

Each of these users will now be able to login to Zadara Cloud Services with their Active Directory name and password.

If an account is not connected to an identity provider, users can be manually created within Zadara Cloud Services as members of this account.

Note

An account with an IdP configured should have no local member users, because keeping local member users:

- Bypasses policies enforced by the IdP
- Defeats some of the purposes of using such an identity provider

Once an IdP is configured, changes cannot be applied to local users.

It is recommended to create and configure a "break-the-glass" local admin user before configuring an IdP.

32.3 Roles and Policies

Access to Zadara Cloud Services functionality is gained by assigning the user to one or more projects within this account. The user is then assigned policies and roles per project which are used as follows:

1. **Policies** define the ZCS and AWS functionality or APIs, which are being permitted for the user. Policies essentially determine **what** functionality is permitted. Multiple policies can be assigned.
2. **Roles** determines which of those policies or APIs can actually be assigned or used, by a user. Both ZCS and AWS roles are used. A user must be assigned a single ZCS role. In addition, a user may be associated with one or more AWS roles. Roles essentially determine the type of user **who** is permitted to perform this functionality.

32.3.1 ZCS Roles

There are three Zadara Cloud Services roles, each allowing different functionality, as follows:

1. **Member** - This role allows the user to use policies and APIs for creating, viewing, modifying and deleting virtual resources belonging to projects to which the user has been assigned. This is the standard role for most users.
2. **MSP or Tenant Admin** - In addition to allowing the use of those policies and APIs which are granted to a Member, the MSP or Tenant Admin role also allows the user to use policies and APIs for creating and managing new projects and users within a specific account, assigning to these users, per project, roles and Zadara Cloud Services and AWS policies. It is recommended that each account have at least one user with this role.
3. **Zadara Ops Admin** - In addition to allowing the use of those policies and APIs which are granted to a Member and an MSP/Tenant Admin, the Zadara Ops Admin role also allows the user to use policies and APIs for viewing, creating, managing and deleting all physical resources, such as nodes (servers), disks, storage pools, physical networks, etc., and administrative entities such as accounts. It is possible to create more than one user per region with Zadara Ops Admin rights.

   Note

   The Zadara Ops Admin user `admin` is only available to a Zadara user.

Two important guidelines concerning the assignment of Zadara Cloud Services roles:

1. Only assign a single role per project.
2. When assigning multiple projects to a user, assign the same role for each project.

Note

Effective user and group permissions are determined by the intersection of the user's and group's assigned roles and policies.

For example, if a user is assigned the Tenant Admin role, and one of their Symp policies is ReadOnlyAdministrator, then the effective permission for that user would be read-only permission on all of the account's resources and aspects.

When Zadara Cloud Services is installed, it comes with a single account containing one project and one user available to Zadara.

- The account is called **cloud_admin**.
- The project inside is called **default**.
- The special Zadara user is called **admin**.

This user, who serves as the System Administrator of the entire region, comes assigned to the 'default' project with the 'Admin' Zadara Cloud Services role, the 'FullAccess' Zadara Cloud Services Policy and the 'AdministratorAccess' AWS API policy. This built-in account, project and user cannot be deleted or modified in any way.

Important

It is highly recommended to create and keep a "break the glass" local account admin user, with its credentials vaulted for future use. (i.e. This is not a personal user who can leave the organization potentially causing the credentials to be lost, but rather a user for emergency "break the glass" use cases).

This admin user can be used for maintenance of the IdP connection if it stops working properly for any reason.

32.3.2 AWS Roles

AWS IAM Roles are policy-based tokens with temporary credentials, allowing a user temporary access to AWS services and actions which the user is normally not permitted to access. These users may be from different projects or even different accounts. These roles can also be embedded into specific instances allowing these instances access to the necessary actions.

Note

The AWS IAM roles are independent of the Zadara Cloud Services roles, which together with Zadara Cloud Services policies, grant access to Zadara Cloud Services services and actions.

The AWS IAM role consists of the following:

1. Permissions which give access to certain Zadara Cloud Services, supported AWS services or actions.
2. Trust policy that defines the relationship between user per project and this role.
    1. This nature of the relationship may be 'allow' which grants permission to the specified users to assume the role, or 'deny' which prevents these users from assuming the role.
    2. This permission may be granted to multiple users of the same projects, different projects within the same account, or even users of different accounts.
3. The maximum session duration that can be requested when assuming this role.

32.3.3 Policies

Policies can be defined and assigned to a user per project for both Zadara Cloud Services and AWS API. While AWS APIs and Zadara Cloud Services functionality may overlap, they are essentially two independent areas of functionality, requiring separate sets of policies. A user can be granted access to both AWS API and Zadara Cloud Services functionality on the same project.

1. The two policy types support the granting of access to one area without granting access to the other. For example, users working with only AWS APIs do not need access to the Zadara Cloud Services API/GUI.

2. The Zadara Cloud Services API is more extensive and all of its functionality is not covered by the AWS APIs.
3. Even APIs that appear to be similar in AWS API and Zadara Cloud Services API, such as “create vm” and “runinstances”, actually permit different actions.

CHAPTER

THIRTYTHREE

ACCOUNTS

33.1 Overview

The virtual resources of the Zadara Cloud Services region are managed through an administrative hierarchy of accounts & projects. Users, who are members of an account, can be assigned different projects within their account. This assignment enables you to provide access to Zadara Cloud Services to numerous users, dividing and separating the resources that each user will be able to view, create, and manage.

33.2 Basic Account Operations

To see account information and perform basic account configurations, navigate to **Identity & Access > Account**.

In the upper half of the account view, the following account information is available:

- Account Id
- Account Name
- MFA Status
- Number of associated projects, groups, and users
- Number of associated compute resources (VM instances, vCPU's, RAM)
- Number of associated networking resources (networks, floating IP's, security groups)
- Amount of associated storage resources (volumes, images, snapshots)

In addition, the following configurations can be made:

- Enforce MFA (Multi-factored Authentication)
- Assign tags which can be used to allow better account visibility, filtering, or reporting.

In the lower half of the account view, 4 tabs are available for the following operations:

1. **Projects** - Projects associated with the account are displayed. To associate the account with a new project, click **Create Project** in the displayed toolbar. An account can be associated with more than one project.
2. **Groups** - Groups associated with the account are displayed. To associate the account with a new group, click **Create Group** in the displayed toolbar.
3. **Users** - Users associated with the account are displayed. Use the toggle buttons in the dispaly to enable/disable a specific user, or activate/deactivate MFA for a specific user once enabled for the account. To associate a new user with the account, click **Create User** in the displayed toolbar.
4. **Limits** - Configure resource limitations for the account as described in the following section.

Note

From the **Identity & Access > Account** view, operations can be per account or per project. To see information or make configurations per project, select a specific project from the list displayed in the project tab. For more information on project-level operations, see *Creating Projects*. If no project is selected from the list, the operations will be per account.

You can always be sure whether you are in the account or project view by the display on top of the toolbar.

- In the **account** view, the display will be:

  Home > Account > *Account_Name*.

- In the **project** view, the display will be:

  Home > Account > *Account_Name* **> Project >** *Project_Name*

33.3 Account Limits

Compute, Services and Storage resources can be limited per account and per project within the account. When these limits are set, Zadara Cloud Services tracks the usage of these resources when they are allocated or freed-up, both at the project level and at the account level.

For a given resource, the sum of the limits **set** for each project within an account may exceed the limit set for the account itself. However the total amount of the resource actually **used** may never exceed either the project limit or account limit set for that resource. For example, there may be a 10-image limit set for account A, and a 5-image limit for each of its three member projects - P1, P2, P3. This sets the total limits of the projects at 15, even though the limit of account A is only 10. Users within account A will be allowed to create images until they either individually reach the limit of 5 in their project, or until a total of 10 images have been created in the account across all projects.

Note

Networking resources can currently be limited only per project.

Limits can be imposed on the following account resources:

1. Compute
   - Number of cores
   - Number of images
   - Number of instances
   - Number of key-pairs
   - RAM
2. Services
   - Number of Kubernetes clusters
   - Number of database instances
   - Number of load balancers
   - Number of registries
3. Storage
   - Number of snapshots

- Number of volumes
- Volume capacity

Note

Storage limits are defined and displayed per storage pool, which are then aggregated to an account limit.

Note

Users with roles Member or Tenant Admin can **view** but not **add** or **modify** limits.

CHAPTER

THIRTYFOUR

PROJECTS

34.1 Creating Projects

If you want to divide your virtual resources among numerous users within each account, you may create different projects within the account. A **Tenant Admin** user can create projects only in their own account.

To create a project:

1. In the zCompute UI, go to the **Identity & Access** > **Accounts** view. Select the account for which you wish to create a project.
2. In the lower half of the account view, select the **Projects** tab and click **Create Project**. The **Create Project** dialog is displayed.
3. Enter the following information:
    1. **Project Name** – enter a name for the new project. The name must be unique within the account. Project names are not case-sensitive.
    2. **Project Description** [Optional] – enter a description of the new project.
    3. **Type** - select the type of project:
        - **VPC** - Virtual Private Cloud project type.
        - **DVS** - Distributed Virtual Switch project type.
    4. **IP Pool** (VPC-type projects) - select one of the IP pools. If there is no available IP pool, request an administrator to create a shared edge network available in your Zadara Cloud Services region.
    5. **Grant Permissions** - leave the checkbox selected (default) to grant yourself permissions on the new project.
4. Click **OK** to create the new project. The new project is created and is displayed in the **Projects** tab of the specific account.

34.2 Enabling or Disabling Projects

To enable or disable a project:

1. In the zCompute UI, go to the **Identity & Access** > **Accounts** view. Select the account containing the project that you wish to enable or disable.
2. In the lower half of the account view, select the **Projects** tab.
3. From the displayed project list, select the desired project.
4. If the project is currently disabled, click **Enable** in the toolbar to enable it.
5. If the project is currently enabled, click **Disable** in the toolabr to disable it.

34.3 Renaming Projects

To rename a project:

1. In the zCompute UI, go to the **Identity & Access** > **Accounts** view. Select the account containing the project that you wish to rename.
2. In the lower half of the account view, select the **Projects** tab.
3. From the displayed project list, select the desired project.
4. In the toolbar, click **Rename**.
5. In the **Rename Project** dialog, enter the desired changes in the project name and description.
6. Click **OK**.

34.4 Deleting Projects

To delete a project:

1. In the zCompute UI, go to the **Identity & Access** > **Accounts** view. Select the account containing the project that you wish to delete.
2. In the lower half of the account view, select the **Projects** tab.
3. From the displayed project list, select the desired project.
4. In the toolbar, click **Delete**.
5. In the **Delete Project** confirmation dialog, click **Delete**. A message confirming the deletion of the project will pop-up in the upper right-hand corner of the screen.

> **Note**
>
> You cannot delete the 'default' project in the cloud_admin account.

34.5 Assigning a User to a Project

To assign a user to a project:

1. In the zCompute UI, go to the **Identity & Access** > **Accounts** view. Select the account containing the project for which you wish to assign a user.
2. In the lower half of the account view, select the **Projects** tab.
3. From the displayed project list, select the desired project.
4. In the toolbar, click **Assign User**.
5. In the **Assign User** dialog, enter the followng:
    1. **User** - select a user from the drop-down list which will display all users associated with the account.
    2. **Project Roles** - select **Member** or **Tenant Admin**.
    3. **Policies**
    4. **AWS API Policies**
    5. Click **OK**.

6. A new user will be created with the selected Zadara Cloud Services role and the default Zadara Cloud Services Policy 'FullAccess'.

34.6 Project Limits

34.6.1 Project Limits Overview

Zadara Cloud Services allows you to set limits on the amount of virtual compute, service, storage and network resources each project can use. It is recommended that each account **Tenant Admin** user set the available resource limits for each of the projects within their account.

34.6.2 Project Limits

To view existing project virtual resources:

1. Navigate to the **Identity & Access** > **Accounts** > view for the specific account containing the project for which the limits should be set.
2. In the lower half of the account view, select the **Projects** tab.
3. From the displayed project list, select the desired project.
4. In the bottom half of the project view, select the **Limits** tab. Existing limits, if any, will be displayed for the following resources:
    1. Compute:
        - Number of cores
        - Number of images
        - Number of instances
        - Number of key-pairs
        - RAM
    2. Services
        - Number of Kubernetes clusters
        - Number of database instances
        - Number of load balancers
        - Number of registries
    3. Storage
        - Number of snapshots
        - Number of volumes
        - Volume capacity

        Note

        Storage limits are defined and displayed per storage pool, which are then aggregated to an Account limit.

    4. Network
        - Floating IPs
        - Networks

- Routers
- Security Groups
- Security Group Rules
- Subnets

To add new project compute, services, storage resource limits:

1. Navigate to the project **Limits** tab as described above.
2. Select the resource category (compute, services, storage) to be limited.
3. Click **Add**.
4. In the **Add Limit** dialog, select the resource to be limited. Note the current usage of the resource will also be displayed.
5. Enter the resource limit to be added. Verify that the limit exceeds the current usage.
6. Click **OK**.

To modify existing project compute, services, storage resource limits:

1. Navigate to the project **Limits** tab as described above.
2. Select the resource category (compute, services, storage) to be modified.
3. On the row with the specific limit to be modified, click on the modify icon (pencil).
4. In the **Edit Limit** dialog, note the current resource usage and limit.
5. Enter the new resource limit. Verify that the new limit exceeds the current usage.
6. Click **OK**.

To remove existing project compute, services, storage resource limits:

1. Navigate to the project **Limits** tab as described above.
2. Select the resource category (compute, services, storage) with limit to be deleted.
3. On the row with the specific limit to be removed, click on the delete icon. The existing limit will be deleted.

34.6.3 Managing Project Virtual Network Resource Limits

To limit project network resources:

1. Navigate to the project **Limits** tab as described above.
2. Click on the **Edit** button on the top-right of the **Network** list.
3. For each virtual network resource, either check the **Unlimited** box, or uncheck it and enter a limit.
4. Click **OK**. The network resource limit list will be updated accordingly.

CHAPTER

THIRTYFIVE

USERS

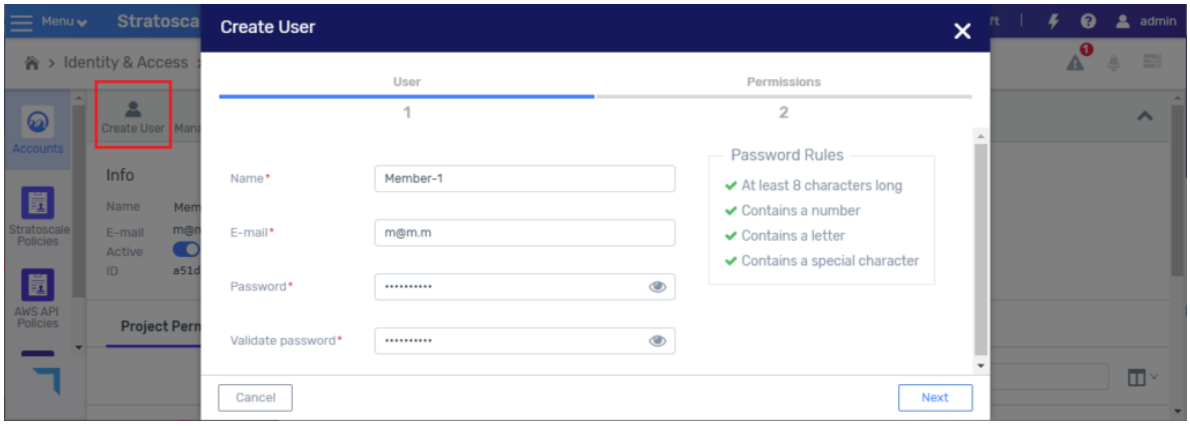
A user can be created within Zadara Cloud Services as a member of a single account. To access and use Zadara Cloud Services, a user must be assigned to at least one project within that account via a Zadara Cloud Services role, and assigned at least one Zadara Cloud Services Policy.

Note

- If a user is assigned to more than one project, the role assigned for each project must be the same.
- To perform these operations, you need to be either the Zadara Admin user who can add users to any account, or a Tenant Admin user who can add users to the tenant account.
- If a user originated from an Identity Provider and not from within Zadara Cloud Services, the only user action which can be performed is the management of their permissions.
- None of these actions can be performed on the system 'admin' user of the cloud_admin account.

35.1 Creating Users

1. Navigate to the **Identity & Access** > **Accounts** view for the specific account for which you wish to create a user.
2. From the bottom half of the view, select the **Users** tab. This will display the currently defined users.
3. From the tab toolbar, click **Create User**.

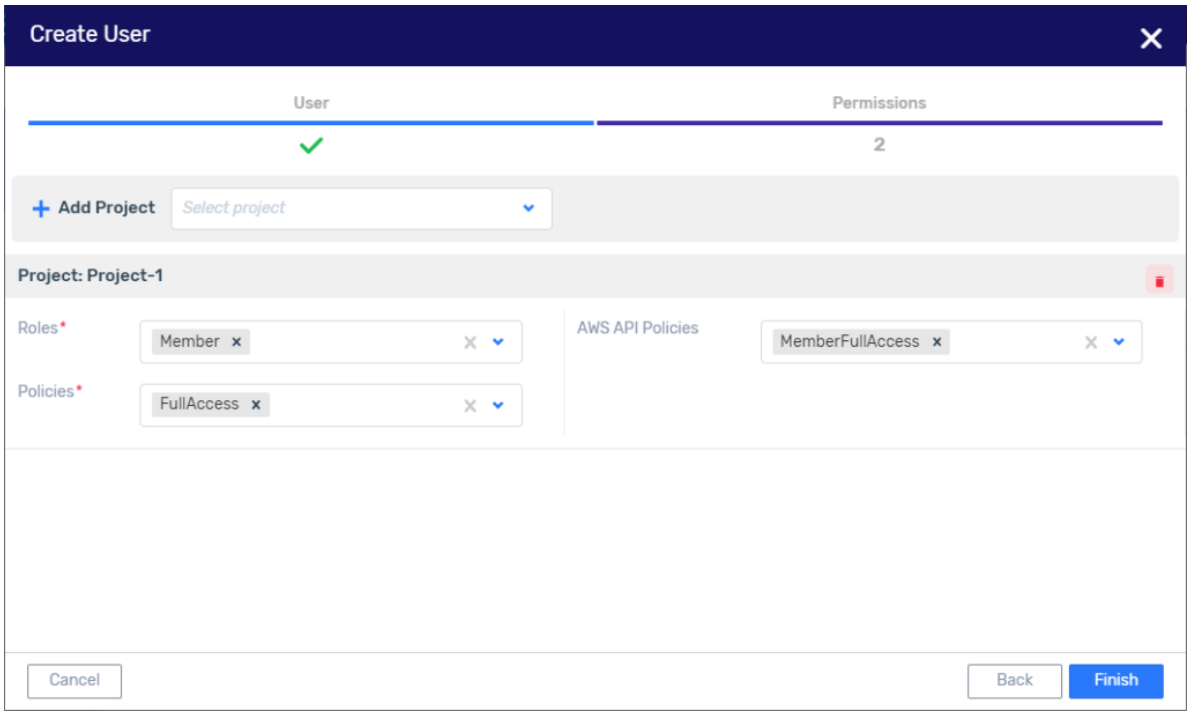


4. In the **Create User** dialog, enter the following information:
    1. **Name** – enter a name. All names in an account must be unique. It is not case-sensitive.
    2. **E-mail** – enter the user's email address.

3. **Password** – enter a password. A valid password must satisfy the following currently defined rules:
   - Be at least 8 characters long.
   - Contain at least 1 number, 1 letter, and 1 special character.

   When entering the password, the display on the right will dynamically indicate when each password rule has been satisfied.
4. **Validate Password** – re-enter the password.

5. Click **Next**. The **Permissions** tab displayed.
6. In the **Add Project** field, select a project from the pull-down list. You can select any project in the user's account. You may also select multiple projects, one at a time.



1. For each project added, enter the following information:
   - **Roles** - select one of the available roles: Member or Tenant Admin. This is required for the first project. Default role: Member.

     Note

     It is highly recommended not to assign multiple roles for the same user in a single project.

   - **Policies** - Select one or more Zadara Cloud Services Policies. The default ZCS policy is FullAccess. This policy grants full access to all of the functionality granted to this role prior to v5.3.4.

     Note

     Effective user and group permissions are determined by the intersection of the user's and group's assigned roles and policies.

     For example, if a user is assigned the Tenant Admin role, and one of their Symp policies is ReadOnlyAdministrator, then the effective permission for that user would be read-only permission on all of the account's resources

and aspects.

- **AWS Policies** - This field is optional and need only be populated if you want to grant access to AWS APIs.

2. To add a user to another project, click **Add Project** and enter the information above for the new project. To remove a project from the list, click the delete icon near the project name.
3. Click **Finish** to create the new user. The new user is created and is displayed in the **Users** tab of the specific account.

35.2 Deleting Users

When deleting a user, this user will no longer have access to Zadara Cloud Services. Deleting a user only removes the user from the system. It does not delete the virtual resources that this user created. These resources will still be accessible to all users with access to the projects for which these resources were created.

To delete a user:

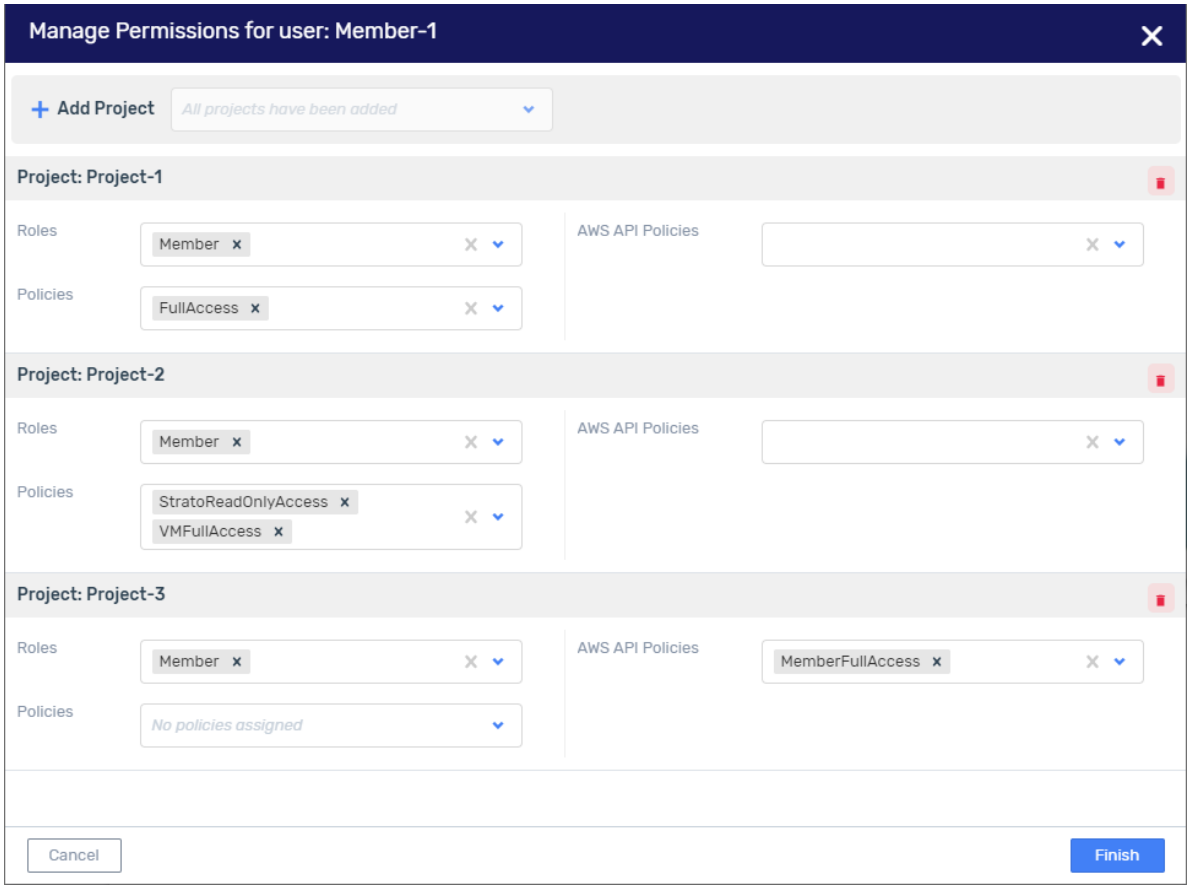
1. Navigate to the **Identity & Access** > **Accounts** view.
2. From the bottom half of the view, select the **Users** tab. This will display the currently defined users.
3. Select the user to be deleted from the displayed list and click **Delete** from the tab toolbar. The **Delete User** confirmation notice appears.
4. Click **OK**. A message confirming the deletion of the user project will pop-up in the upper right-hand corner of the screen.

35.3 Managing Users Permissions

After a user is created, you can modify the user's permissions and add/delete projects assigned to the user.

To manage the permissions of a user:

1. Navigate to the **Identity & Access** > **Accounts** view.
2. From the bottom half of the view, select the **Users** tab. This will display the currently defined users.
3. Select the user whose permissions are to be modified from the displayed list and click **Manage Permissions** from the tab toolbar. The **Manage Permissions for User: <name of selected user>** dialog box is displayed.



4. In the **Add Project** field, select a project from the pull-down list. The following permissions may be modified.
    - **Roles**
    - **Policies**
    - **AWS API policies**
5. To modify permissions for another project, click **Add Project** and enter the information above for the new project. To remove a project from the list, click the delete icon near the project name.
6. Click **Finish** to save the modified permissions.

35.4 Modifying Users

After creating a user, you can modify its name, email address and password expiration status.

1. Navigate to the **Identity & Access** > **Accounts** view.
2. From the bottom half of the view, select the **Users** tab. This will display the currently defined users.
3. Select the user to be modified from the displayed list and click **Modify** from the tab toolbar. The **Modify User** dialog box is displayed.
4. The following may be modified:
    - **Name**
    - **Email Address**
    - **Password Expiration** - Enable/Disable

5. Click **Finish**. A message confirming the updating of the user will pop-up in the upper right-hand corner of the screen.

35.5 Resetting User Passwords

Once a user is created, the password can be reset by the Tenant Admin user of the user's account.

Note

- The password of users in accounts connected to an Identity Provider service, cannot be reset from within Zadara Cloud Services.

To reset a user password:

1. Navigate to the **Identity & Access** > **Accounts** view.
2. From the bottom half of the view, select the **Users** tab. This will display the currently defined users.
3. Select the user whose password is to be modified from the displayed list and click **Set Password** from the tab toolbar. The **Set Password** dialog box is displayed.
4. Enter the following:
    - **Password**
    - **Validate Password** - re-enter the same password.
5. Click **OK**. A message confirming the successful resetting of the password will pop-up in the upper right-hand corner of the screen.

CHAPTER

THIRTYSIX

CONNECTING AN ACCOUNT TO A MICROSOFT ACTIVE DIRECTORY IDENTITY PROVIDER

zCompute accounts' users can be authenticated using an external LDAP-compatible identity provider, such a Microsoft Active Directory domain. Once it's set up, users and groups of the identity provider can be granted permissions to zCompute. To enable this, the zCompute account must be connected to Active Directory. Each zCompute account requires independent configuration for the Identity Provider connection.

Important

As a prerequisite, the zCompute cloud must have a route to the Active Directory domain controllers in order to connect a zCompute account to the Active Directory.

Tenants are advised to consult with their Managed Service Provider.

36.1 Connecting an Account to an Active Directory Identity Provider

Note

Before connecting the zCompute account to selected users in Active Directory, it is recommended to first create a dedicated group for them in Active Directory, and add them to that group.

This allows you to use the filters to select only those users that should be connected to the zCompute account.

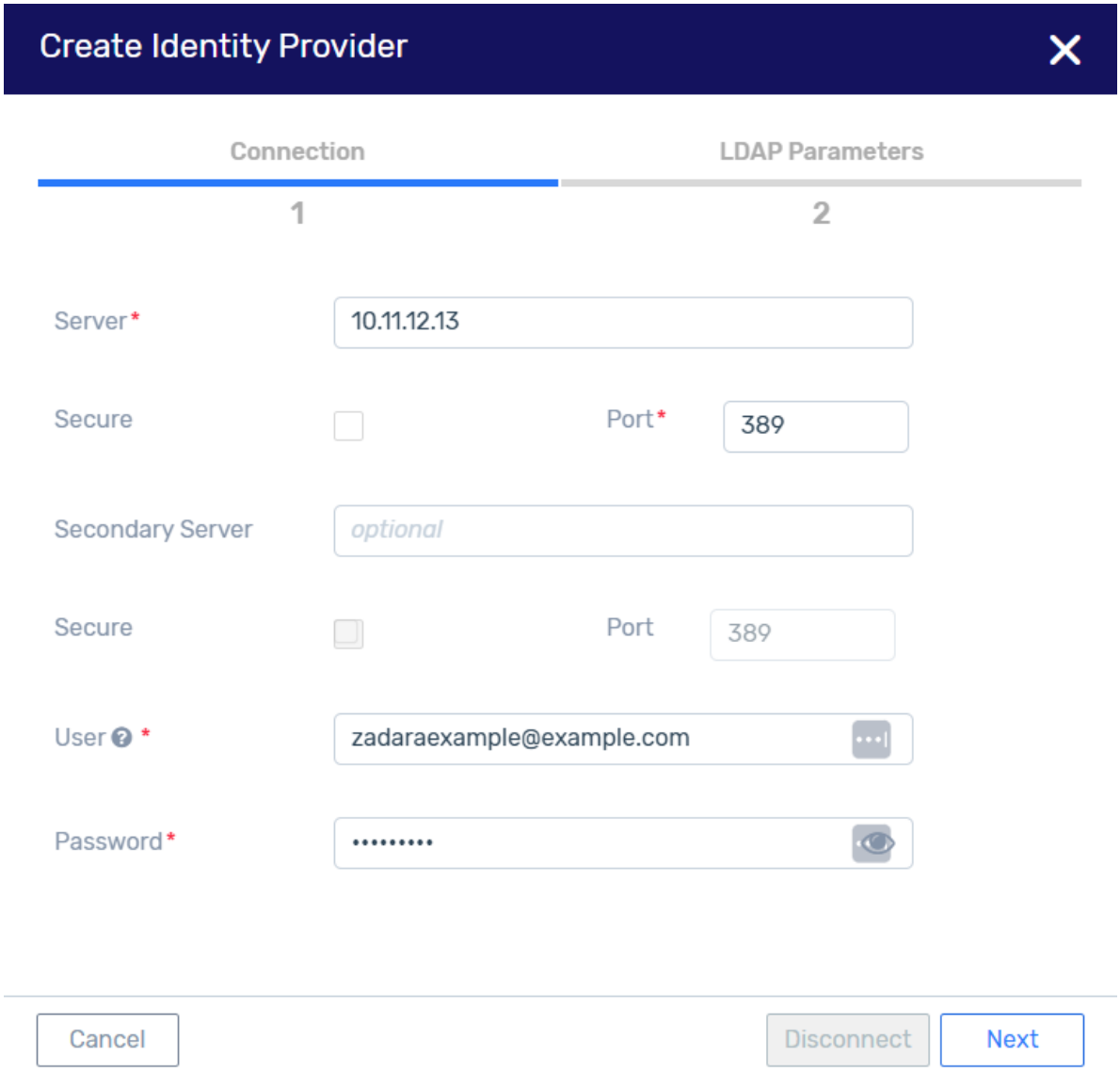
To connect an account to an Active Directory Identity Provider:

1. Navigate to the **Identity & Access** > **Accounts** view, and highlight the row of the account that you wish to connect to the Identity Provider.

   An **Identity Provider** button appears in the toolbar.

2. Click **Identity Provider**.

   The **Create Identity Provider** dialog opens at the **Connection** step.



Note

An existing LDAP connection can be disconnected via the **Disconnect** button.

1. Enter the Active Directory domain controller (DC) details:
   - **Server** - The LDAP server or Active Directory domain controller's IP address.

     For example: `10.11.12.13`

     Note

     If that server has a public DNS name (FQDN), it is possible to use the DNS name instead of the IP address.

     For example: `dc3.mydomain.com`.

   - **Secure** - Check this box if your LDAP server supports the secure LDAP protocol over SSL or TLS.

- **Port** - The reserved port is 389 if the connection is not secure, or 636 if the connection is secure.
- **Secondary Server** (optional) - A backup server if the first one is not available. The Active Directory Domain Controller address, expressed as an IP address or a DNS hostname.

  For example: `10.11.12.14` or `dc4.mydomain.com`.
- **Secure** - Check this box if your secondary LDAP server supports the secure LDAP protocol over SSL or TLS.
- **Port** - The reserved port for the secondary server is 389 if the connection is not secure, 636 if the connection is secure.
- **User** - This is the user principal name (UPN) or distinguished name (DN) of a user through which one can gain access to Active Directory Server.

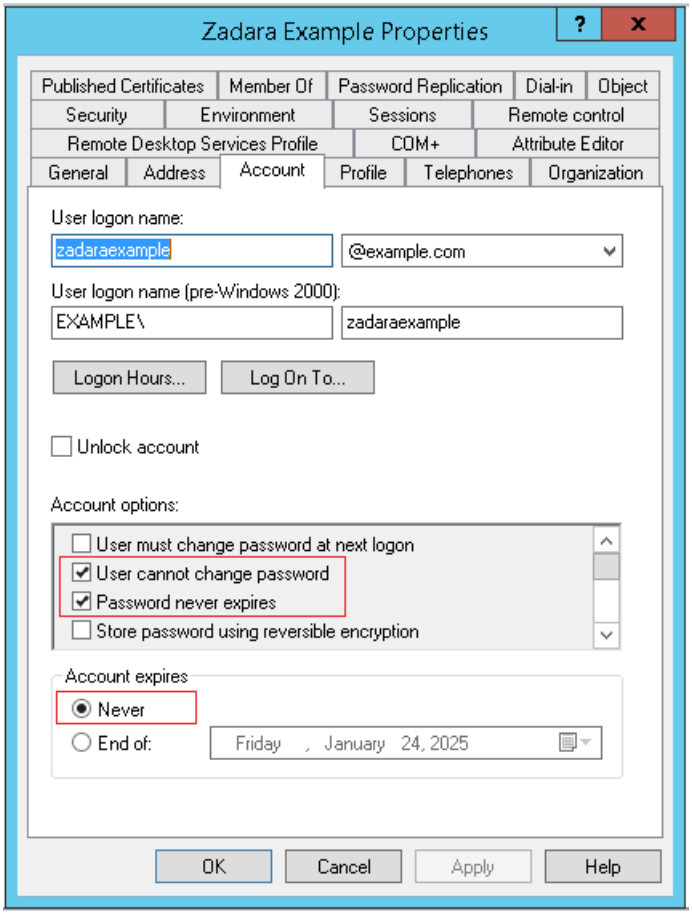
  DN example: `cn=zadaraexample,cn=Users,dc=example,dc=com`

  UPN example: `zadaraexample@example.com`

Note

- This user should not be an administrator of the domain.
- This user establishes the connection to Active Directory, and is used as a “service account” to sync zCompute with the Active Directory domain.

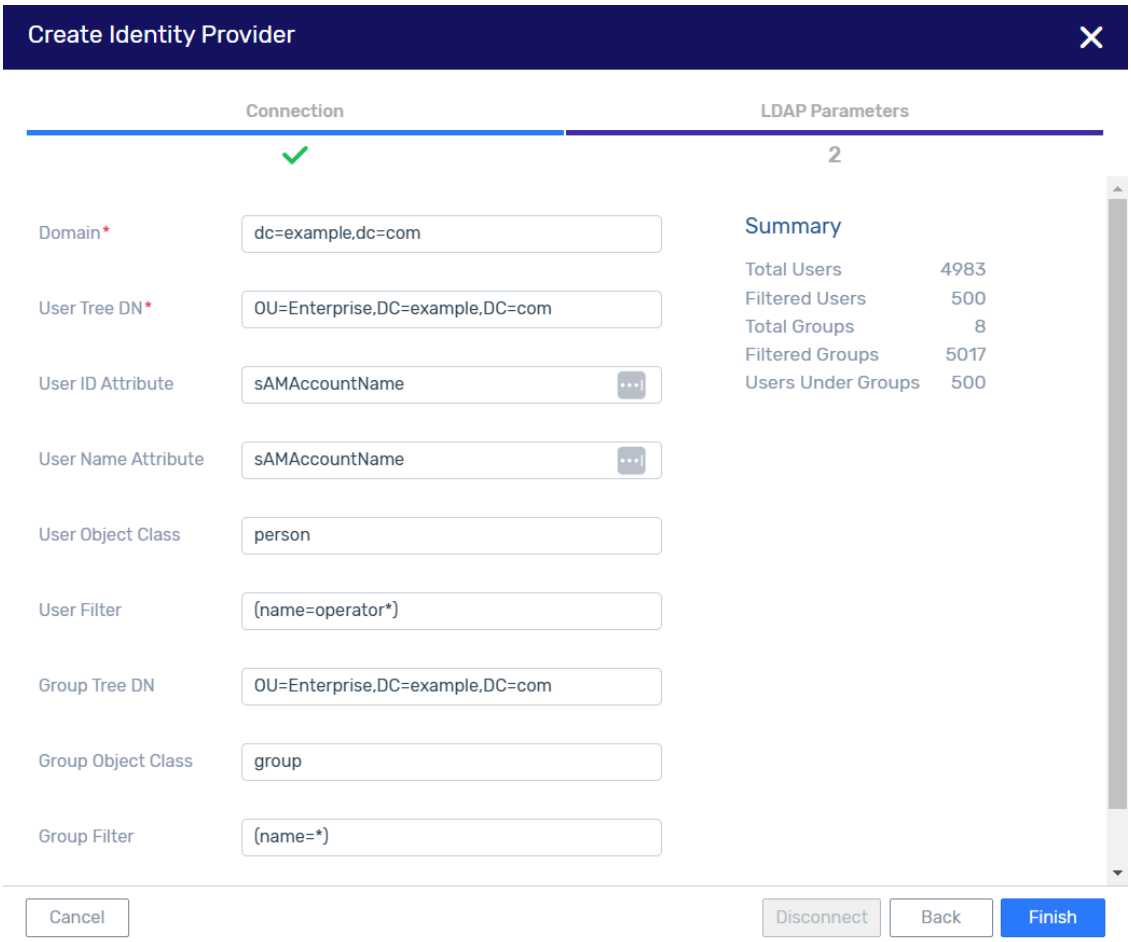
  To guarantee a continuous operation with Active Directory, configure this user’s Active Directory account so that it never expires, and that its password never changes and never expires.



- **Password** - The Active Directory password of the user.
- Click **Next**.

After validating the connection to the Active Directory server, the dialog displays the **LDAP Parameters** step.

2. **LDAP Parameters** - All parameters are expressed in the LDAP syntax.



Note

A **Summary** of LDAP users and groups appears on the right, and updates dynamically as filters are applied.

- **Domain** - The customer's domain in Active Directory.

  For example: `dc=example,dc=com`
- **User Tree DN** - The location in the Active Directory in which the users will be scanned.

  For example: `OU=Enterprise,DC=example,DC=com`
- **User ID Attribute**: The attribute name in Active Directory that represents the User ID.

  For example: `sAMAccountName`.
- **User Name Attribute**: The attribute name in Active Directory that represents the User Name.

  For example: `sAMAccountName`.

- **User Object Class**: The objectClass property of a user object in Active Directory.

  For example: `person`.

- **User Filter** - Filters the users scanned in the User Tree.

  For example:

  - `(name=example-*)` searches the User Tree for any user beginning with `example-`.
  - `(memberOf=cn=grp-exampleldap,cn=Users,dc=example,dc=com)`

    searches for all users who belong to the group `grp-exampleldap` in the Active Directory with domain components matching `dc=example,dc=com`.

Note

The input syntax for the **User Filter** parameter **includes the parentheses**.

- **Group Tree DN** - The location in the Active Directory in which the groups will be scanned.

  For example: `cn=Users,dc=example,dc=com`

Caution

If the **Group Tree DN** is left empty, the UI doesn't display the LDAP users' list.

In this case, it is still possible to grant permissions to individual users, but these users' permissions don't display in the UI.

- **Group Object Class** - Active Directory's group object class for the groups. Default: `group`
- **Group Filter** - Filters the groups scanned in the Group Tree.

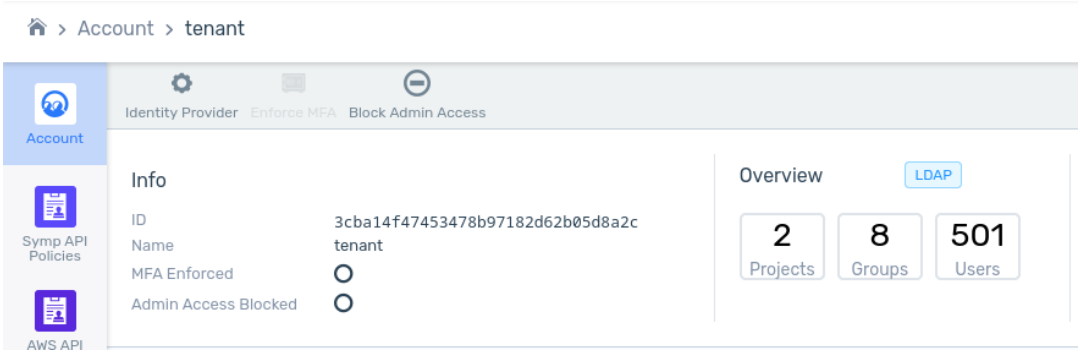
  For example: `(name=grp-*)` searches for all groups with the prefix `grp-`.

Note

The input syntax for the **Group Filter** parameter **includes the parentheses**.

Click **Finish**.

- The selected account is connected to MS Active Directory.
- Users matching the filters described above will appear as LDAP Users of the selected account. On completion of configuring Active Directory as an identity provider for the zCompute account, these users' AD passwords will be used for authentication.
- The groups matching the filters described above will appear as groups of the selected account, containing the users defined in Active Directory.
- The account's top pane **Overview** section displays the **LDAP** badge, and summarizes the number of LDAP **Groups** and LDAP **Users**.

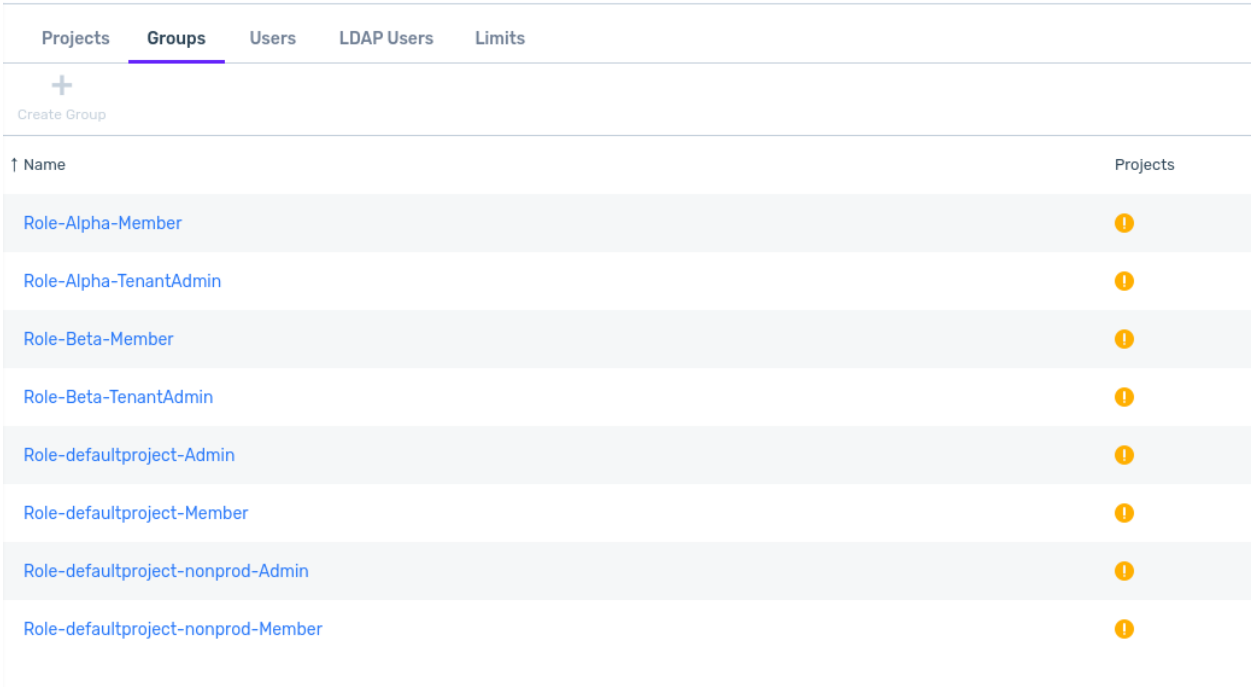


36.2 Viewing LDAP Groups

To view the LDAP Groups:

In the account's lower pane, click the **Groups** tab.

The LDAP Groups list displays.



By default, users in a group can't sign on until the group is assigned project permissions. The Projects column for groups without project permissions displays an alert symbol to indicate this.

To configure project permissions for a group, see *Managing an LDAP Group's Permissions*.

Note

For a zCompute account that is connected to an Active Directory, new users and groups can only be created and managed in the Active Directory.

Important

As a best practice, it is recommended to have a single local (non-LDAP) tenant admin user for “break the glass” scenarios, in the event of LDAP connection loss or malfunction.

36.2.1 Managing an LDAP Group’s Permissions

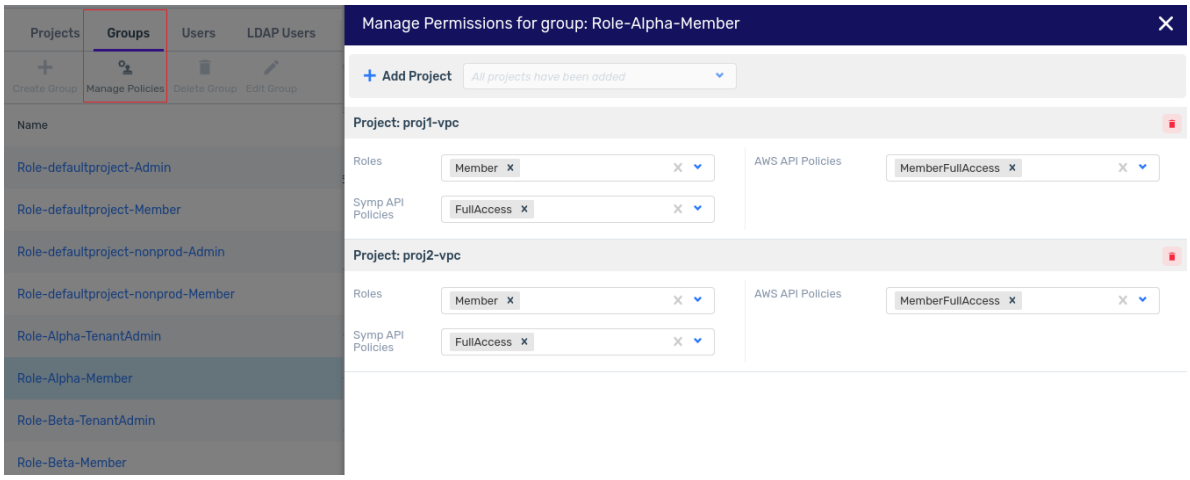
To assign permissions to an LDAP group and its members:

1. In the account’s lower pane, click the **Groups** tab.

   The LDAP Groups list displays.

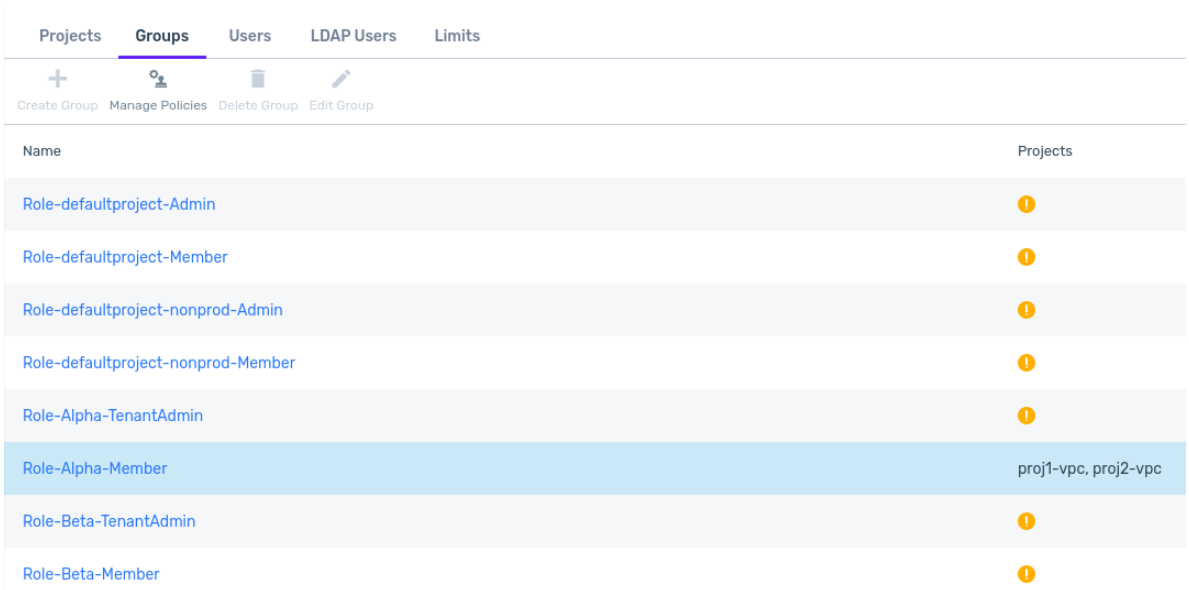
2. Highlight the row of the group to apply permissions.
3. In the lower pane menu bar, click **Manage Policies**.

   The **Manage Permissions for Group** dialog opens.



1. Configure the projects and and their permissions for the group
2. Click **Finish**.

The updated projects appear in the Projects column for the group, in the **Groups** tab.

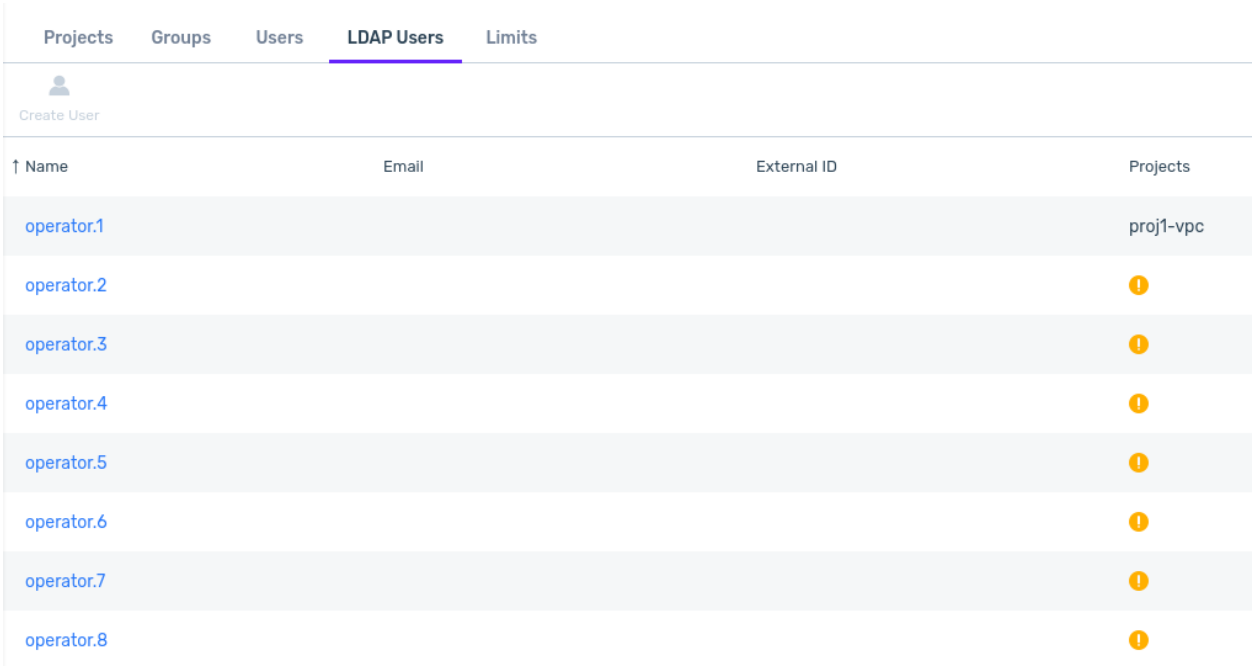


36.3 Viewing LDAP Users

To view the LDAP Users:

In the account's lower pane, click the **LDAP Users** tab.

The LDAP Users list displays.



Note

- LDAP user passwords and email addresses **cannot** be modified within zCompute.
- **Assigning Project Roles and zCompute Policies**

  Before these users will be able to work with zCompute, you must first assign a project to each user, either individually or via the groups of which they are members, together with the group's zCompute role and policies.

Users who are members of groups configured with project permissions have their associated projects listed in the Projects column. For other users, the Projects column displays an alert symbol, indicating that the user cannot sign on until project permissions are assigned. User permissions can be assigned individually or as a member of a group with configured permissions.

36.4 Assigning zCompute Roles and permissions to an Active Directory User - UI

The following UI actions are recommended for assigning the Account and Project 'Admin' roles and permissions to a user added to zCompute via Active Directory:

1. Locate the new user added to zCompute via Active Directory.
    1. On the **Identity & Access** > **Accounts** <account> view for the requested account, click the **LDAP Users** tab.

       The list of users for that account displays.
    2. Click the user to assign zCompute permissions.

       The user's **Project Permissions** tab displays the user's role and policies per project.

2. Assign a project to the user with a project role and zCompute policy.
    1. In the user's upper menu bar, click **Manage Permissions**.
    2. The dialog box **Manage Permissions for User** opens.
    3. From the **+ Add Project** dropdown list, select the project.
    4. In the section that opens for the project, complete the user's permission settings:

       The **Roles**, **AWS API Policies** and **Symp API Policies** have descriptive names that reflect the role or policy scope.

        1. **Roles** - From the dropdown, select the role.

           Multiple roles can be applied.

        2. **AWS API Policies** - From the dropdown, select the policy.

           Multiple AWS API policies can be applied.

        3. **Symp API Policies** - From the dropdown, select the policy.

           Multiple Symp API policies can be applied.

        4. Optionally, repeat these steps to assign roles and policies to the user for additional projects.
        5. Click **Finish** to save the user's roles and policies per project.

    The user's **Project Permissions** tab in the lower pane displays the user's role and policies per project.

36.5 Assigning zCompute Roles and permissions to an Active Directory User - CLI

The following CLI commands are recommended for assigning the Account and Project 'Admin' roles and permissions to a user made available to zCompute via Active Directory:

Note

In this example the assumption is that the zCompute account, `new_account`, which is connected to a domain in Active Directory, already contains the project, `new_project`.

1. Use symp's user list command to locate the user.

```
user list -c id -c name -c domain_id

+----------------------------------+----------+----------------------------------+
| id                               | name     |   domain_id                      |
+==================================+==========+==================================+
| c8a63b29558d4765a6cd78760729a2f7 | new_user | 2d27e2fe6d8a4398b901c4d84c478777 |
+----------------------------------+----------+----------------------------------+
| admin                            | admin    | default                          |
+----------------------------------+----------+----------------------------------+
```

   Note the user's domain ID.

2. Use symp's user list command to verify that the user `new_user` from the Active Directory list, does not already appear in another zCompute account.

Note

A username can appear in more than one zCompute account, but each user's ID is unique even though the username might appear more than once.

```
user list --name new_user -c id -c name -c domain_id

+-----------+----------------------------------+
| id        | c8a63b29558d4765a6cd78760729a2f7 |
| name      | new_user                         |
| domain_id | 2d27e2fe6d8a4398b901c4d84c478777 |
+-----------+----------------------------------+
```

3. Use symp's `domain list` command to list the domains with their IDs and names.

```
domain list -c id -c name

+----------------------------------+---------------+
| id                               | name          |
+==================================+===============+
| 2d27e2fe6d8a4398b901c4d84c478777 | new_account   |
| default                          | cloud_admin   |
+----------------------------------+---------------+
```

4. Use symp's project list command to list the projects with their IDs, names and domains.

```
project list -c id -c name -c domain_name

+----------------------------------+-------------+---------------+
| id                               | name        | domain_name   |
+==================================+=============+===============+
| 4bd79a2fa9574af2a4b9a7a87195f144 | default     | cloud_admin   |
| 1569c28e3a344ee2b3989640499b8eca | new_project | new_account   |
+----------------------------------+-------------+---------------+
```

5. Use symp's role list command to list all roles in zCompute.

```
role list

+--------------------+--------------------+
| id                 | name               |
+====================+====================+
| admin              | admin              |
| tenant_admin       | tenant_admin       |
| _member_           | member             |
+--------------------+--------------------+
```

6. Use symp's project list-roles-on-project command to check if user `new_user` has already been assigned a role in the project `new-project`.

   Syntax: `project list-roles-on-project <project_id> <user_id>`

```
project list-roles-on-project 1569c28e3a344ee2b3989640499b8eca␣
↪c8a63b29558d4765a6cd78760729a2f7

+--------------+--------------+
```

(continues on next page)

(continued from previous page)

```
| value        | tenant_admin |
+--------------+--------------+
```

If the user is already assigned a role other than `admin` in the project `new_project`, use the project revoke-role command to remove the role from `new_user`.

Syntax: `project revoke-role <project_id> <user_id> <role_id>`

```
project revoke-role 1569c28e3a344ee2b3989640499b8eca␣
↪c8a63b29558d4765a6cd78760729a2f7 tenant_admin

+-------+---------+
| value | Success |
+-------+---------+
```

7. Use symp's project grant role command to assign the *admin* role to user `new_user` in the project `new_project`.

   Syntax: `project grant-role <project_id> <user_id> <role_id>`

```
project grant-role 1569c28e3a344ee2b3989640499b8eca␣
↪c8a63b29558d4765a6cd78760729a2f7 admin

+-------+---------+
| value | Success |
+-------+---------+
```

8. Use symp's project list-roles-on-project command to verify that the role of `new_user` in `new_project` is `admin`.

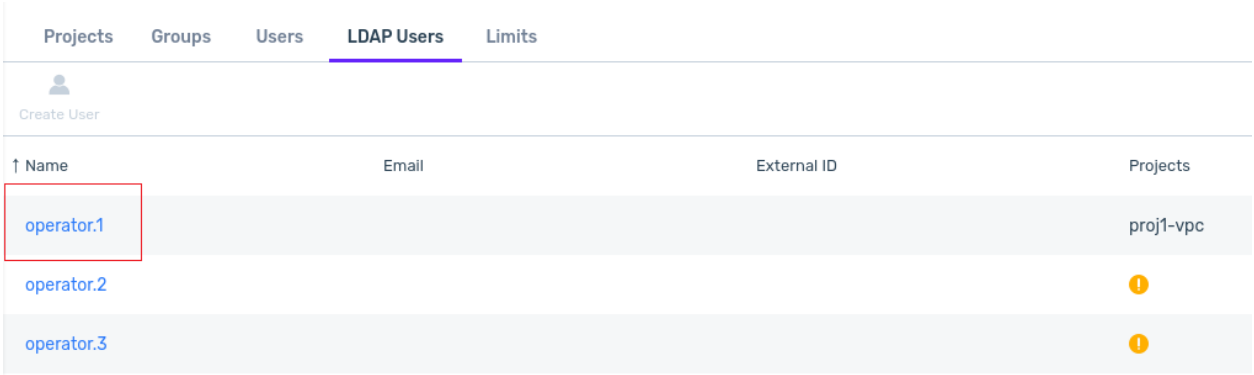
   Syntax: `project list-roles-on-project <project_id> <user_id>`

```
project list-roles-on-project 1569c28e3a344ee2b3989640499b8eca␣
↪c8a63b29558d4765a6cd78760729a2f7

+--------------+--------------+
| value        | admin        |
+--------------+--------------+
```

36.6 LDAP User Sign-on to zCompute

After LDAP users' project and permission assignments are configured in zCompute, the users can sign on to zCompute using their short username, as registered in zCompute, and listed in the account's **LDAP Users** tab.



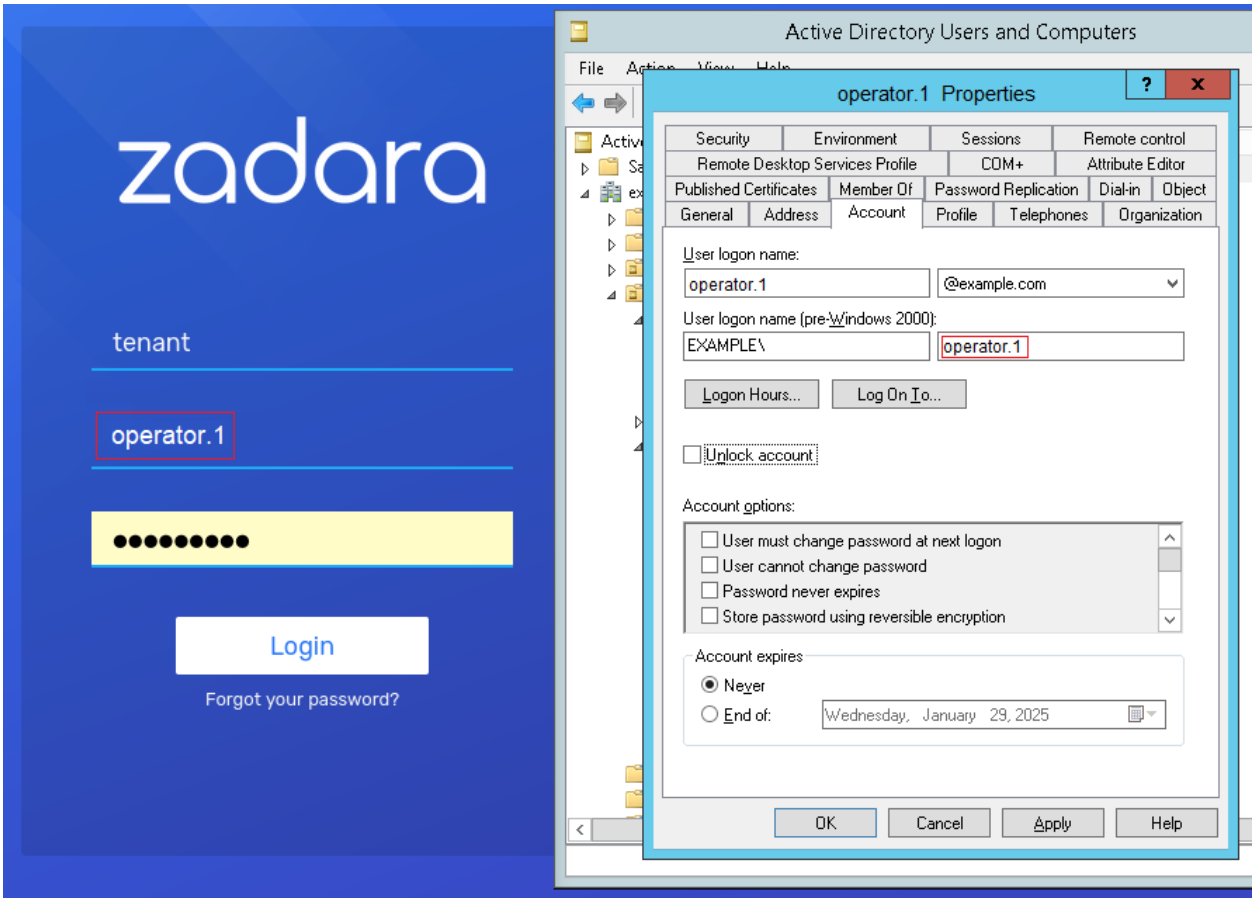
In this example:

- The zCompute account name is `tenant`.

  For this example, this is the account where the administrator loaded LDAP users and groups, and assigned them zCompute projects and permissions.

- The LDAP user's username is `operator.1`.

  This is the short username of the LDAP user, imported into zCompute based on the mapping of the LDAP **User Name Attribute** `sAMAccountName`. The mapping was configured in the **LDAP Parameters** step of the **Create Identity Provider** dialog.



CHAPTER

THIRTYSEVEN

AUTHENTICATION USING ZCOMPUTE APIS

Symp CLI commands can be run by authenticating, without requiring a token.

However, it is advisable to authenticate and then generate a token for use in any automated Symp CLI or API process. This will prevent possible `connection reset by peer` errors when rate limits are exceeded.

37.1 API Endpoints

The list of a cluster's API endpoints is available in the UI.

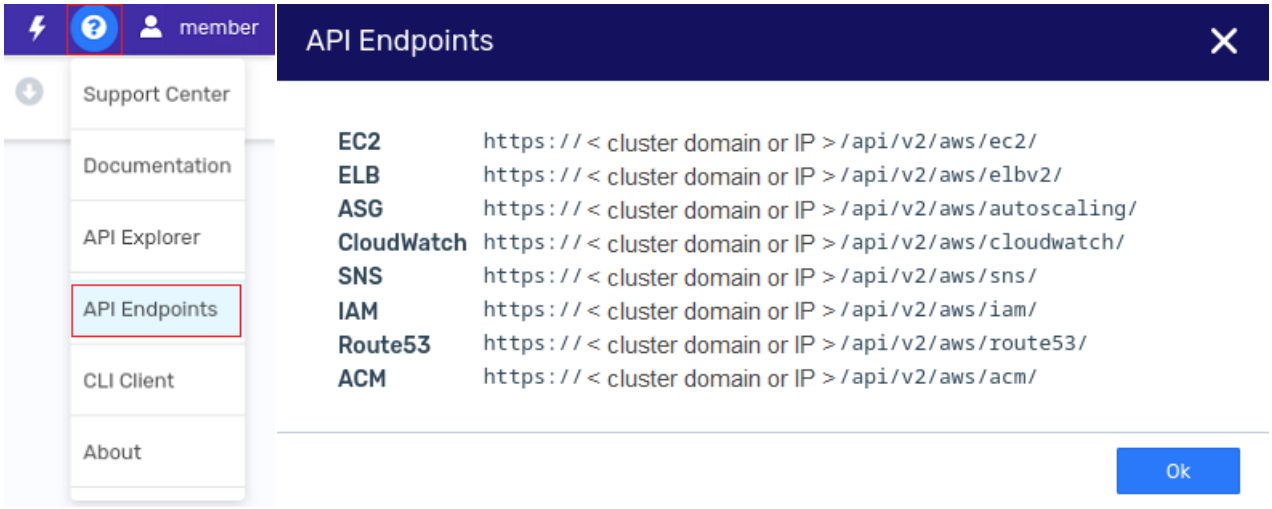
To view the cluster's API endpoints:

1. In the UI, in the upper right corner near the username, click **Help** (the **?** icon).

   The **Help** dropdown menu opens.

2. In the dropdown menu, select **API Endpoints**.

   The **API Endpoints** dialog opens, listing the cluster's API endpoints.



37.2 API Explorer

zCompute provides an interactive Swagger-based API Explorer.

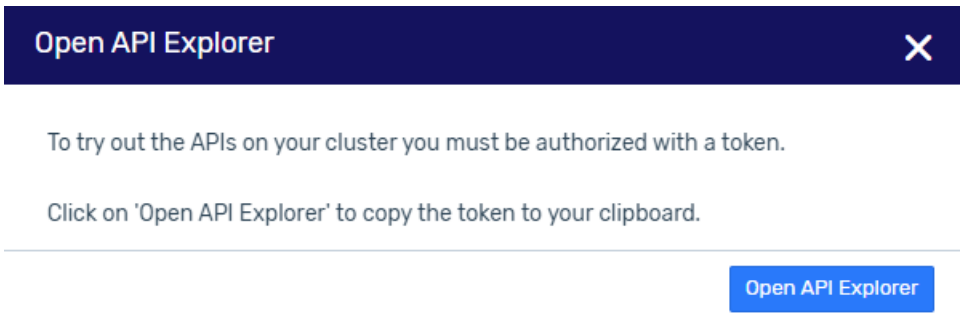
API operations require authorization with a token.

37.2.1 Obtaining a token in the UI

To obtain a token in the UI:

1. In the UI, in the upper right corner near the username, click **Help** (the **?** icon), and in the dropdown menu, select **API Explorer**.

   The **Open API Explorer** dialog opens.



2. Click the **Open API Explorer** button.
   - A token is generated and automatically copied to the clipboard.
   - The **zCompute Symp API Explorer** opens, listing the zCompute APIs.

Note

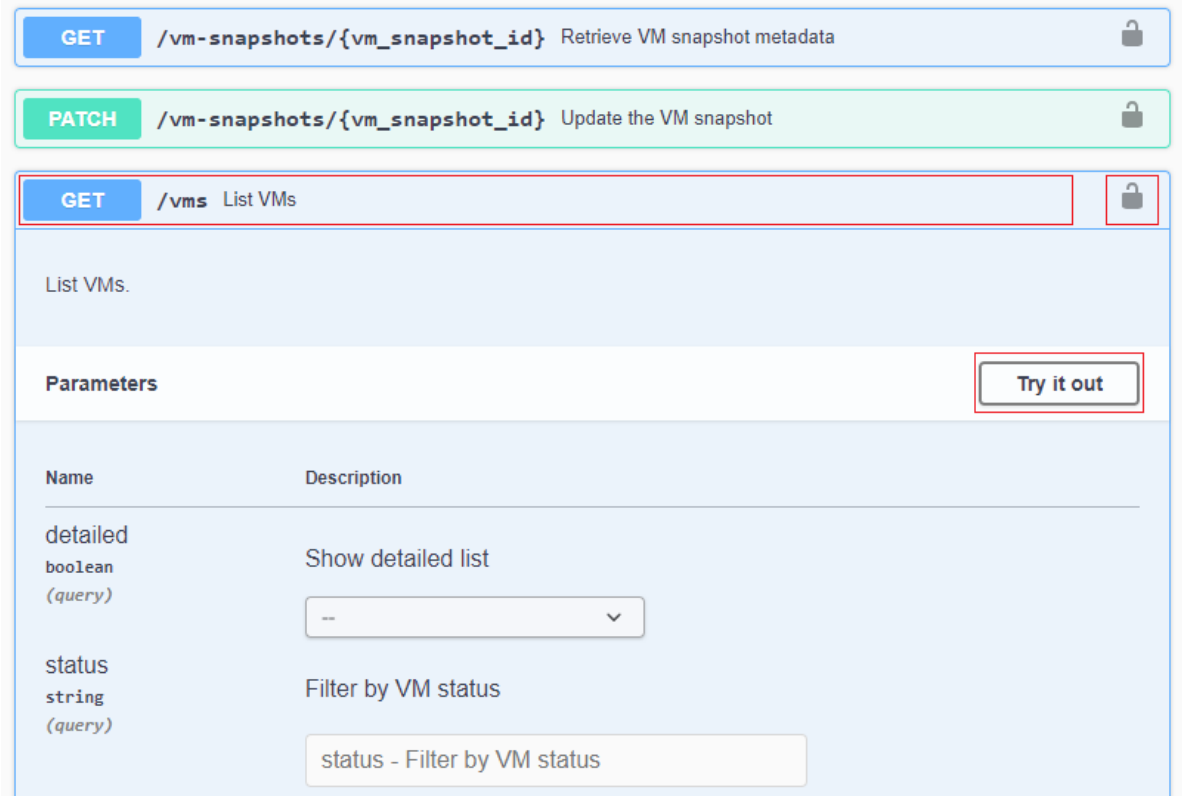
The token in the clipboard can be pasted and used in the API Explorer, or for running API operations using the CLI.

37.2.2 Running API operations in the API Explorer

1. After *Obtaining a token in the UI*, click an API in the **zCompute Symp API Explorer** list.

   A Swagger UI opens, displaying the selected API's REST methods.

2. Scroll down the list to locate the desired operation to run.
3. Click the selected REST method or its description, to expand the view to display its parameters.



4. Click the **Lock** icon on the right, to submit the authorization token.

Note

Alternatively, click the **Authorize** button in the screen header.

The **Available Authorizations** dialog opens.

1. In the **token (apiKey) > Value** field, paste the token from your clipboard.
2. Click **Authorize**.

Note

After the token is authorized and while it is still valid, REST methods for all APIs can be run.

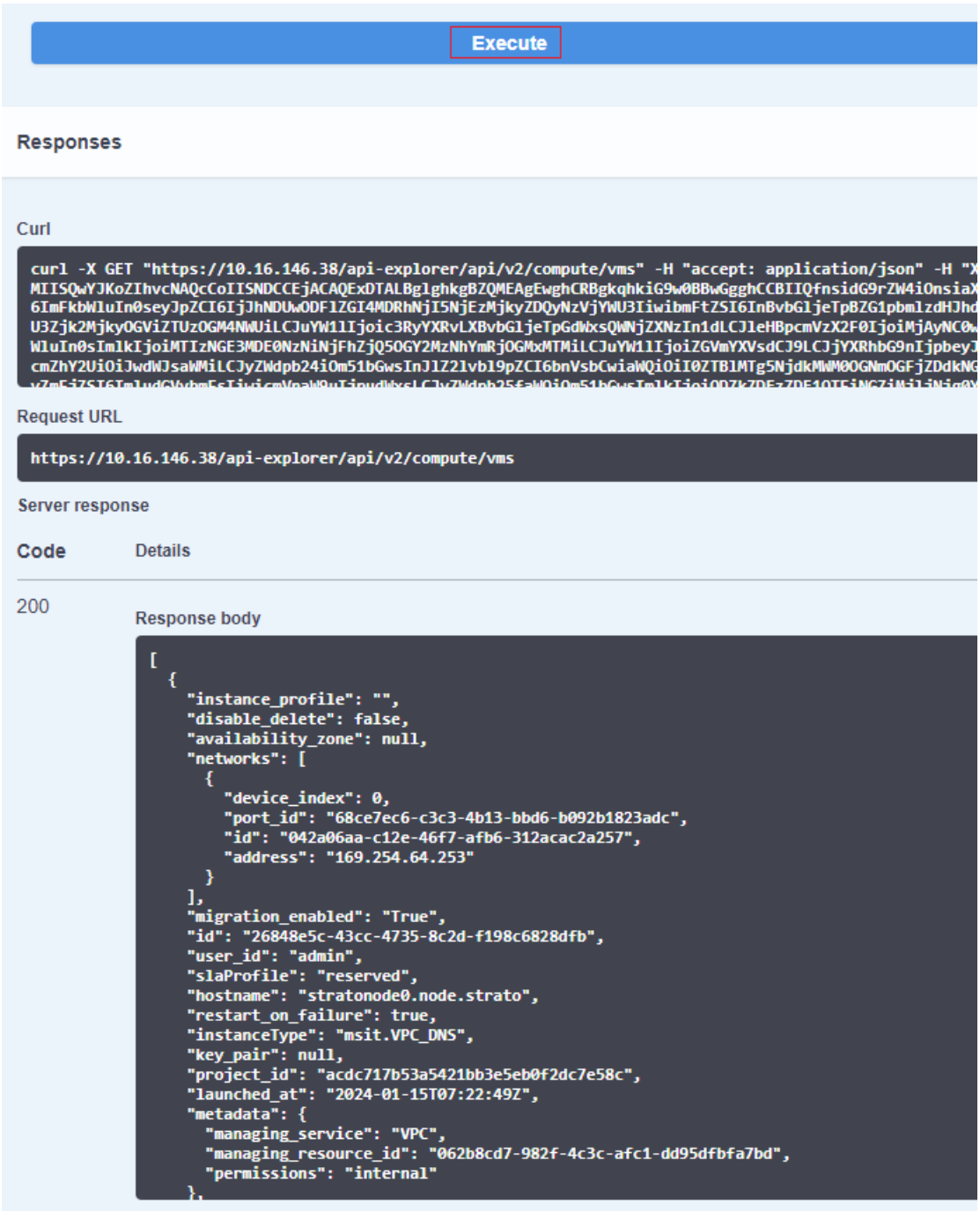
It is not necessary to repeat this authorization step.

3. Click **Close** to close the **Available Authorizations** dialog and return to the selected API REST method screen.

5. Click **Try it out**.
6. Enter parameters for the REST method, and click the **Execute** bar.

   The **Responses** pane expands, displaying:

   - The **Curl** command that was executed.
   - The **Request URL**.
   - The **Server response**, comprising the return code, response body and response headers.



37.3 Generating a token at a bash prompt

Since some commands are limited to the scope of a project, while others have the scope of an entire account (domain), it's necessary to generate a token for the account (domain) scope, and a separate token for the project scope.

Using `curl` on the `/api/v2/identity/auth` endpoint, to generate authentication tokens involves two steps:

First, generating an account (domain) token, and then using the domain token to generate a project token.

1. Generating a domain token:

The `/api/v2/identity/auth` endpoint's authentication token request data structure that defines the **scope** of an entire **account** (domain):

```
{
  "auth": {
    "scope": {
      "domain": {
        "name": "<ACCOUNT_NAME>"
      }
    },
    "identity": {
      "password": {
        "user": {
          "domain": {
            "name": "<ACCOUNT_NAME>"
          },
          "password": "<PASSWORD>",
          "name": "<USER_NAME>"
        }
      },
      "methods": [
        "password"
      ]
    }
  }
}
```

1. The `curl` command at the bash prompt for generating an account (domain) authentication token:

   ```
   curl -D - -H "Content-Type: application/json" -X POST https://<cluster FQDN>/
   ↪api/v2/identity/auth -d \
   '{"auth": {"scope": {"domain": {"name": "<account>"}}, "identity": {"password":
   ↪{"user": {"domain": {"name": "<account>"}, "password": "<password>", "name": "
   ↪<user>"}}, "methods": ["password"]}}}'
   ```

   For example:

   ```
   curl -D - -H "Content-Type: application/json" -X POST https://compute.example.
   ↪com/api/v2/identity/auth -d \
   '{"auth": {"scope": {"domain": {"name": "acc1"}}, "identity": {"password": {
   ↪"user": {"domain": {"name": "acc1"}, "password": "Mypass1!", "name": "user1"}}
   ↪, "methods": ["password"]}}}'
   ```

2. The domain token is returned in the `x-subject-token` header, and can be copy/pasted to assign it to an environment variable for later use.

   For example:

   ```
   export ZC_DOMAIN_TOKEN=MIISUgYJKoZIhvcNAQcCoIISQzCC...<2K+ character string>...
   ↪ngpXjEqQtxlRGuCEIvo46sGMfedc=
   ```

2. Using the previously generated domain token to generate a project token:

   The `/api/v2/identity/auth` endpoint's authentication token request data structure that defines the **scope** as limited to a **project**, and the identity method as a **token**:

```
{
  "auth": {
    "scope": {
      "project": {
        "domain": {
          "name": "<ACCOUNT_NAME>"
        },
        "name": "<PROJECT_NAME>"
      }
    },
    "identity": {
      "token": {
        "id": "'<DOMAIN_TOKEN>'"
      },
      "methods": [
        "token"
      ]
    }
  }
}
```

1. The `curl` command at the bash prompt for generating a **project** authentication token:

   ```
   curl -D - -H "Content-Type: application/json" -X POST https:<cluster FQDN>/api/
   ↪v2/identity/auth -d \
   '{"auth": {"scope": {"project": {"domain": {"name": "<account>"}, "name": "
   ↪<project>"}}, "identity": {"token": {"id": "'<domain token string>'"},
   ↪"methods": ["token"]}}}'
   ```

   For example:

   ```
   curl -D - -H "Content-Type: application/json" -X POST https://compute.example.
   ↪com/api/v2/identity/auth -d \
   '{"auth": {"scope": {"project": {"domain": {"name": "acc1"}, "name": "vpcproj1"}
   ↪}, "identity": {"token": {"id": "'$ZC_DOMAIN_TOKEN'"}, "methods": ["token"]}}}
   ↪'
   ```

2. The project token is returned in the `x-subject-token` header, and can be copy/pasted to assign it to an environment variable for later use.

   For example:

   ```
   export ZC_PROJECT_TOKEN=MIIScwYJKoZIhvcNAQcCoIISZDCC...<2K+ character string>...
   ↪VpADxR3RDrScUbgSwMAaZ8zCSrow=
   ```

3. To use `curl` to call a REST method on an endpoint, provide the token string in the `X-Auth-Token` header:

   ```
   curl -L -X <REST method> "https://<cluster FQDN>/api-explorer/api/v2/<endpoint>/
   ↪" \
   -H "accept: application/json" \
   -H "X-Auth-Token: <token string>"
   ```

   For example, to retrieve the project's tags, use the `GET` method on the `/api-explorer/api/v2/tags/` endpoint, and the project authentication token (assigned in the previous step to the environment variable `$ZC_PROJECT_TOKEN`). This example uses `sed` and `awk` to format the output:

```
curl -L -X GET "https://compute.example.com/api-explorer/api/v2/tags/" \
-H "accept: application/json" \
-H "X-Auth-Token: $ZC_PROJECT_TOKEN" | \
awk 'BEGIN{FS=",";OFS="\n"} FNR==1{$1=$1;print;exit}' | \
sed '/{/{x;p;x;}'
```

The response:

```
  % Total    % Received % Xferd  Average Speed   Time    Time     Time  Current
                                 Dload  Upload   Total   Spent    Left  Speed
100   116  100   116    0     0  11756      0 --:--:-- --:--:-- --:--:-- 12888
100  2302  100  2302    0     0  59137      0 --:--:-- --:--:-- --:--:--     0

[{"description": "Tag2"
 "updated-at": "2023-10-31T08:44:01Z"
  "created-at": "2023-10-31T08:44:01Z"
  "value": ""
  "scope": "public"
  "project-id": "7ff34832b6754f7d91a918302af3e77f"
  "name": "vpctag2"}

  {"description": "Tag 1"
  "updated-at": "2023-10-31T08:10:49Z"
  "created-at": "2023-10-31T08:10:49Z"
  "value": ""
  "scope": "public"
  "project-id": "7ff34832b6754f7d91a918302af3e77f"
  "name": "vpctag1"}]
```

37.3.1 Time-based One-Time Passcode (TOTP) Multi-Factor Authentication (MFA)

If a user has Time-based One-Time Passcode (TOTP) Multi-Factor Authentication (MFA) enabled, you must also include both of the following:

- A `totp` entry in the `methods` array.
- A `totp` object block with details.

For example:

```
{
"auth": {
    "identity": {
    "methods": [
        "password",
        "totp"
    ],
    "password": {
        "user": {
        "domain": {
            "name": "<ACCOUNT_NAME>"
        },
        "name": "<USERNAME>",
        "password": "<PASSWORD>"
```

(continues on next page)

(continued from previous page)

```
        }
    },
    "totp": {
        "user": {
        "domain": {
            "name": "<ACCOUNT_NAME>"
        },
        "name": "<USERNAME>",
        "passcode": "<TOTP_CODE>"
        }
    }
    },
    "scope": {
    "domain": {
        "name": "<ACCOUNT_NAME>"
    }
    }
}
}
```

37.4 Generating a token in Symp CLI

1. To generate an authentication token via the Symp CLI shell, use the Symp `login` command:

   ```
   login <user> <password> <domain> <project> -f value
   ```

   For example:

   ```
   login user1 Mypass1! acc1 vpcproj1 -f value

   MIISUgYJKoZIhvcNAQcCoIISQzCC...<2K+ character string>...
   ↪ngpXjEqQtxlRGuCEIvo46sGMfedc=
   ```

   Alternatively, an example of a single command at the bash prompt to generate an authentication token and assign it to an environment variable:

   ```
   export ZC_PROJECT_TOKEN=`symp -k --url https://10.11.12.13 -d acc1 -u user1 --
   ↪project vpcproj1 -p Mypass1! login user1 Mypass1! acc1 vpcproj1 -f value`

   echo $ZC_PROJECT_TOKEN

   MIISUgYJKoZIhvcNAQcCoIISQzCC...<2K+ character string>...
   ↪ngpXjEqQtxlRGuCEIvo46sGMfedc=
   ```

2. Using the token to sign on to Symp from the bash prompt:

   ```
   symp -k --url https://<cluster IP address> --project-token <token string>
   ```

   For example, generating the token (as in the previous step), and using it to sign on to Symp:

   ```
   export ZC_PROJECT_TOKEN=`symp -k --url https://10.11.12.13 -d acc1 -u user1 --
   ↪project vpcproj1 -p Mypass1! login user1 Mypass1! acc1 vpcproj1 -f value`
   ```

(continues on next page)

(continued from previous page)

```
symp -k --url https://10.11.12.13 --project-token $ZC_PROJECT_TOKEN
```

```
Starting new HTTPS connection (1): 10.16.145.114
Connecting in insecure mode!
Starting new HTTPS connection (1): 10.16.145.114
Starting new HTTPS connection (1): 10.16.145.114

 d88888b  dP    dP 8888ba.88ba   888888ba  dP     dP   .88888.  888888ba  dP    dP
88.     "' Y8.  .8P 88  `8b  `8b  88    `8b 88     88  d8'   `8b 88    `8b Y8.  .8P
`Y88888b.  Y8aa8P  88   88   88 a88aaaa8P' 88aaaaa88a 88     88 88     88  Y8aa8P
      `8b    88    88   88   88  88        88     88  88     88 88     88    88
d8'   .8P    88    88   88   88  88        88     88  Y8.   .8P 88     88    88
 Y88888P     dP    dP   dP   dP  dP        dP     dP   `8888P'  dP     dP    dP

Tap <TAB> twice to get list of available commands.
Type --help to get help with any command
Symphony >
```

37.5 Python authentication example

When using a zCompute API, the user is required to pass a valid token in an `X-Auth-Token` HTTP header.

Users can generate a valid token by logging into the system using the `identity/auth` API.

The following Python script is an example of this process:

```
#!/usr/bin/python
from __future__ import print_function
import copy
import getpass
import os
import sys
import argparse
import requests
try:
    input = raw_input
except NameError:
    pass

IDENTITY_API_SUFFIX = "identity"
VPC_API_SUFFIX = "vpc"
DEFAULT_HEADERS = {
    "Accept": "application/json",
    "Content-Type": "application/json"
}
_debug = False


def _debug_msg(msg, *args, **kwargs):
    global _debug
    if _debug:
```

(continues on next page)

(continued from previous page)

```
        print(msg.format(*args, **kwargs), file=sys.stderr)


def generate_totp_passcode(secret):
    """Generate TOTP passcode.
    :param bytes secret: A base32 encoded secret for the TOTP authentication
    :returns: totp passcode as bytes
    """
    try:
        import mintotp
    except:
        sys.exit("TOTP code generation library is not present. Please install mintotp␣
↪using 'pip install mintotp'")
    return mintotp.totp(secret)


class ZComputeApi(object):

    def __init__(self, api_base_url, account_name, project_name, user, password,
                 insecure=False, totp_code=None, mfa_secret=None, debug=False):
        self._api_base_url = '{}/api/v2'.format(api_base_url)
        self._account_name = account_name
        self._project_name = project_name
        self._user = user
        self._password = password
        self._insecure = insecure
        self._token = None
        self._totp_code = totp_code
        self._mfa_secret = mfa_secret
        self._debug = debug

    def reset_api_params(self, token=None, project_name=None):
        if token:
            self._token = token
        if project_name:
            self._project_name = project_name

    def send_request(self, method='GET', api_path='', api_body=None, headers=None, raise_
↪on_status=True):
        api_url = '{}/{}'.format(self._api_base_url, api_path)
        if headers:
            headers = copy.deepcopy(headers)
            headers.update(DEFAULT_HEADERS)
        else:
            headers = copy.deepcopy(DEFAULT_HEADERS)
        if self._token:
            headers['X-Auth-Token'] = self._token
        try:
            response = requests.request(
                method=method,
                url=api_url,
                headers=headers,
```

(continues on next page)

(continued from previous page)

```
                json=api_body,
                verify=not self._insecure
            )
            if raise_on_status:
                response.raise_for_status()
        except Exception:
            _debug_msg("Failed sending {} {} reuquest".format(method, api_url))
            raise
        return response

    def send_api_request(self, method='GET', api_path='', api_body=None):
        return self.send_request(method=method, api_path=api_path, api_body=api_body).
↪json()

    def _get_project_scope(self):
        return {"project": {"name": self._project_name, "domain": {"name": self._account_
↪name}}}

    def _get_domain_scope(self):
        return {"domain": {"name": self._account_name}}

    def _get_password_auth(self):
        return {
            "methods": [
                "password"
            ],
            "password": {
                "user": {
                    "name": self._user,
                    "password": self._password,
                    "domain": {
                        "name": self._account_name
                    }
                }
            }
        }

    def _add_totp_code(self, identity_auth, totp_code):
        if identity_auth.get('methods'):
            identity_auth['methods'].append('totp')
        else:
            identity_auth['methods'] = ['totp']
        identity_auth['totp'] = {
            "user": {
                "name": self._user,
                "passcode": totp_code,
                "domain": {
                    "name": self._account_name
                }
            }
        }
        return identity_auth
```

(continues on next page)

(continued from previous page)

```
    def get_project_token(self):
        identity_auth = self._get_password_auth()
        if self._mfa_secret:
            self._totp_code = generate_totp_passcode(self._mfa_secret)
        if self._totp_code:
            self._add_totp_code(identity_auth, self._totp_code)
        auth_json = {
            "auth": {
                "identity": identity_auth,
                "scope": self._get_project_scope()
            }
        }
        response = self.send_request('POST', '{}/auth'.format(IDENTITY_API_SUFFIX), auth_
↪json, raise_on_status=False)
        if response.status_code == requests.codes.UNAUTHORIZED and response.json().get(
↪'receipt'):
            os_receipt = response.headers.get('openstack-auth-receipt')
            if self._mfa_secret:
                _debug_msg("Generating MFA secret")
                totp_code = generate_totp_passcode(self._mfa_secret)
            else:
                totp_code = str(input('MFA Code: ')).lower().strip()
            identity_auth = self._add_totp_code({}, totp_code)
            auth_json = {
                "auth": {
                    "identity": identity_auth,
                    "scope": self._get_project_scope()
                }
            }
            response = self.send_request('POST', '{}/auth'.format(IDENTITY_API_SUFFIX),
                                         auth_json,
                                         headers={"openstack-auth-receipt": os_receipt})
        else:
            response.raise_for_status()

        return response.headers['x-subject-token']

    def get_account_token(self):
        auth_json = {
            "auth": {
                "identity": self._get_password_auth(),
                "scope": self._get_domain_scope()
            }
        }
        auth_response = self.send_request('POST', '{}/auth'.format(IDENTITY_API_SUFFIX),␣
↪auth_json)
        return auth_response.headers['x-subject-token']

    def get_user_default_project(self):
        details_api_uri = '{}/users/myself/projects'.format(IDENTITY_API_SUFFIX)
        projects = self.send_api_request('GET', details_api_uri)
```

(continues on next page)

(continued from previous page)

```
        if len(projects) > 1:
            sys.exit("There are {} projects: {}\n"
                     "please select one using --project flag".format(
                         len(projects), ', '.join([p['name'] for p in projects])))
        return projects[0]['name']


def ask_for_value(parameter, current, hide=False):
    if hide:
        value = getpass.getpass("{} []: ".format(parameter))
    else:
        value = input("{} [{}]: ".format(parameter, current or '')).strip()
    if value:
        return value
    return current


def main():
    global _debug
    parser = argparse.ArgumentParser(description="Obtain zCompute API Token using user␣
↪credentials")
    parser.add_argument("cluster", help="Cluster API endpoint IP or host name")
    parser.add_argument("--interactive",
                        help="Will get login credentials interactively",
                        action='store_true',
                        default=False)
    parser.add_argument("--account",
                        help="Account name (will use environment variable ZCOMPUTE_
↪ACCOUNT if not provided)",
                        default=os.environ.get('ZCOMPUTE_ACCOUNT', None),
                        required=False)
    parser.add_argument("--username",
                        help="User name (will use environment variable ZCOMPUTE_USER if␣
↪not provided)",
                        default=os.environ.get('ZCOMPUTE_USER', None),
                        required=False)
    parser.add_argument("--mfa-secret",
                        help="MFA secret to generate totp code MFA enabled login "
                             "(will use environment variable ZCOMPUTE_MFASECRET if not␣
↪provided)",
                        default=os.environ.get('ZCOMPUTE_MFASECRET', None),
                        required=False)
    parser.add_argument("--totp-code",
                        help="TOTP code to use for MFA enabled login "
                             "(will use environment variable ZCOMPUTE_TOTP if not␣
↪provided)",
                        default=os.environ.get('ZCOMPUTE_TOTP', None),
                        required=False)
    parser.add_argument("--project",
                        dest='project_name',
                        help="Project name (will use environment variable ZCOMPUTE_
↪PROJECT if not provided)"
```

(continues on next page)

(continued from previous page)

```
                              " if not provided and only one exists in the account will it
↪",
                        default=os.environ.get('ZCOMPUTE_PROJECT', None),
                        required=False)
    parser.add_argument("--password",
                        help="Password - will use environment variable ZCOMPUTE_PASSWORD␣
↪if not provided",
                        default=os.environ.get('ZCOMPUTE_PASSWORD', None))
    parser.add_argument("--debug", help="Print debug messages", default=False, action=
↪'store_true')
    parser.add_argument("-k", "--insecure", action='store_true', help="Do not verify␣
↪cluster certificate", default=False)
    args = parser.parse_args()
    if args.debug:
        _debug = True
    if args.interactive:
        args.account = ask_for_value('account', args.account)
        args.project_name = ask_for_value('project name', args.project_name)
        args.username = ask_for_value('User name', args.username)
        args.mfa_secret = ask_for_value('MFA secret', args.mfa_secret, hide=True)
        args.totp_code = ask_for_value('TOTP code', args.totp_code)
        args.password = ask_for_value('Password', args.password, hide=True)

    url = 'https://{}'.format(args.cluster)
    if args.cluster.startswith('https://'):
        url = args.cluster
    api = ZComputeApi(url, args.account, args.project_name, args.username, args.password,
                      insecure=args.insecure, totp_code=args.totp_code, mfa_secret=args.
↪mfa_secret, debug=args.debug)
    if not args.password:
        sys.exit("Please provide a password")
    if not args.project_name:
        if args.account is None:
            sys.exit("Please provide the account name to login into")
        if args.username is None:
            sys.exit("Please provide the user name to login with")
        api.reset_api_params(token=api.get_account_token())
        project_name = api.get_user_default_project()
    else:
        if args.project_name is None:
            sys.exit("Please provide the project name to login into")
        project_name = args.project_name
    api.reset_api_params(project_name=project_name)
    print(api.get_project_token())


if __name__ == '__main__':
    main()

|api-endpoints-list|
```

CHAPTER

THIRTYEIGHT

AWS API POLICIES

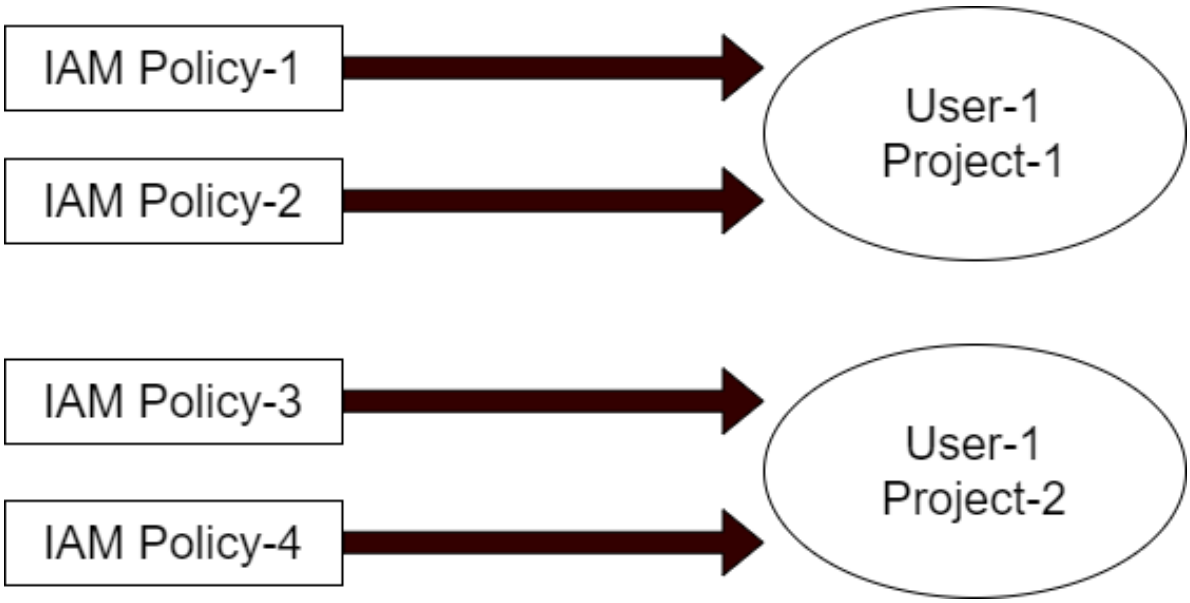
38.1 Introduction

Usage of all Zadara Cloud Services-supported AWS services and actions are governed by their corresponding AWS-managed policies. These policies can be assigned per project to users, groups of users, and STS (Security Token Service) Roles. Zadara Cloud Services usage is governed by Zadara IaaS (Infrastructure as a Service) policies together with Zadara Cloud Services roles. Zadara Cloud Services supports both AWS Managed Policies and Zadara Cloud Services Managed Policies.

38.2 AWS IAM API Policies and AWS Roles Overview

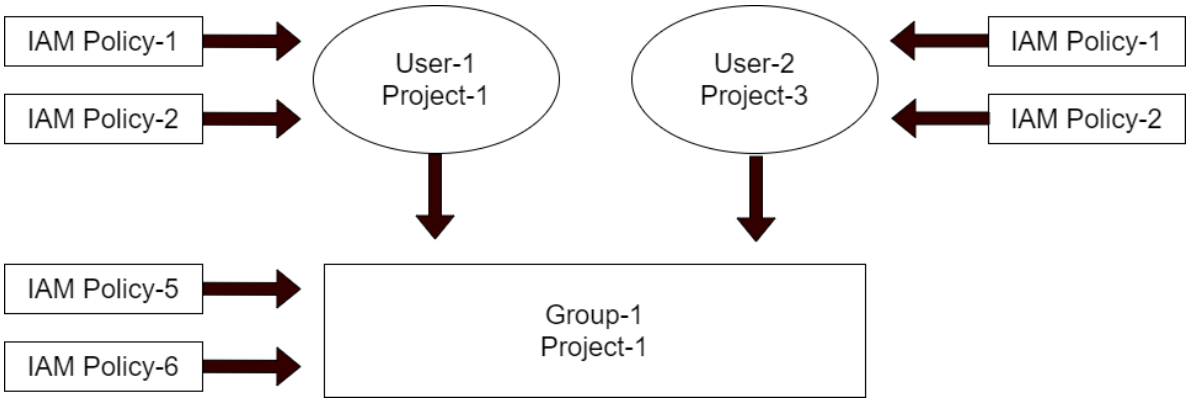
The following section provides different examples in the use of policies and roles.

1. **Policies are attached to users, groups or IAM roles only within the context of a project.**



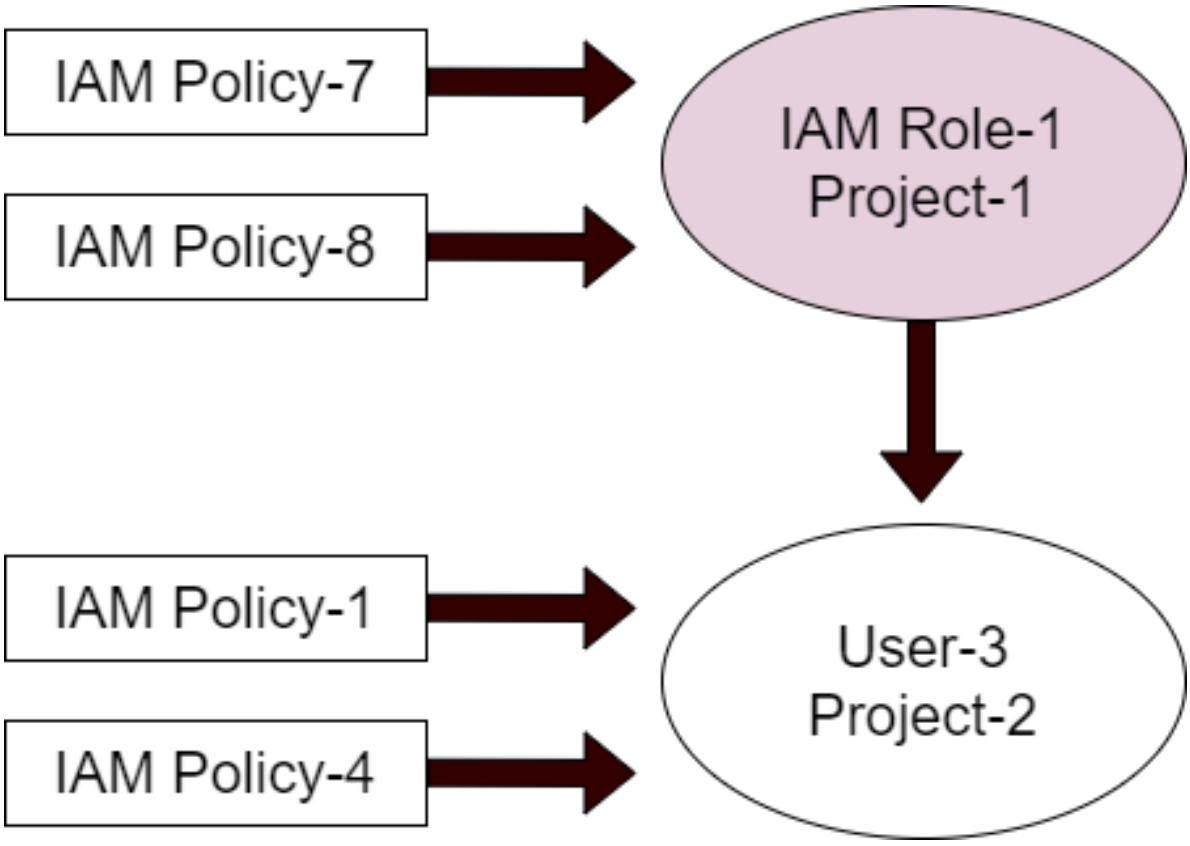
   User-1 will have the permissions defined in IAM Policy 1 and IAM Policy-2 only when working within the context of Project-1. When working within the context of Project-2, User-1 will have the permissions defined in IAM Policy-3 and Policy-4.

2. **Policies attached to users and groups within the context of the same project are aggregated.**



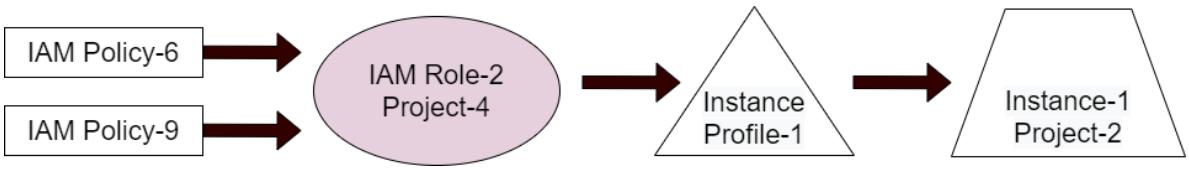
User-1 will have the permissions defined in IAM Policy-1 and IAM Policy-2 in addition to those defined in IAM Policy-5 and IAM Policy-6. On the other hand, User-2 will have the permissions defined in IAM Policy-5 and IAM Policy-6 when working within the context of Project-1. But when working within the context of Project-3, they will both have the permissions defined in IAM Policy-1 and Policy-2.

3. **An IAM role with attached IAM policies which is added to a user, grants the user temporary credentials within the context of the IAM role's project, once assumed.**



When User-3 assumes IAM Role-1, he will have the temporary permissions defined in IAM Policy-7 and IAM Policy-8, when working within the context of Project-1. When working within the context of Project-2, User-3 will have the permanent permissions defined in IAM Policy-1 and IAM Policy-4.

4. **An IAM role with attached IAM policies which is attached to an Instance via an instance profile, grants the instance permanent credentials within the context of the IAM role's project.**



Instance-1 will have the permissions defined in Policy-6 and Policy-9, permanently, when working within the context of Project-4.

38.3 Managed AWS API Policies Supported by Zadara-IaaS

38.3.1 Zadara Cloud Services-managed AWS IAM API policies

Name	Description
AmazonS3BucketManagen	Provides the ability to create buckets and read their data.
EC2AMIDeleteOnly	Provides the ability to delete an EC2 AMI.
EC2AMIDescribeInstan	Provides the ability to describe instances (including their statuses and attributes), and to create and describe images.
EC2AMIFullAccess	Provides full access to all EC2 AMI actions.
EC2AMIReadOnlyAcces	Provides read only access to all EC2 AMI actions.
EC2ManageInstances	Provides read-only access to all EC2 (which includes EC2, VPC, EBS, VM Import/Export) actions, in addition to permission to start and stop EC2 instances.
MemberFullAccess	Provides limited access to IAM policies and full access to all other supported services.
STSAssumeRole	Provides the ability to obtain an IAM role using the “assume-role” action.
STSFullAccess	Provides full access to all STS actions.

ZCS-managed policy definitions

JSON policy definition examples

Examples of JSON policy definitions for ZCS managed IAM policies:

EC2AMIDeleteOnly

JSON policy definition for the EC2AMIDeleteOnly ZCS managed IAM policy:

```
[
   {
      "Action": [
         "ec2:DeregisterImage"
      ],
      "Resource": [
         "*"
      ],
      "Effect": "Allow"
   }
]
```

EC2AMIDescribeInstances

JSON policy definition for the EC2AMIDescribeInstances ZCS managed IAM policy:

```
[
   {
      "Action": [
         "ec2:DescribeImageAttribute",
         "ec2:DescribeImages",
         "ec2:CreateImage",
         "ec2:DescribeInstanceAttribute",
         "ec2:DescribeInstanceStatus",
         "ec2:DescribeInstances"
      ],
      "Resource": [
         "*"
      ],
      "Effect": "Allow"
   }
]
```

EC2AMIFullAccess

JSON policy definition for the EC2AMIFullAccess Zadara IaaS-managed IAM policy:

```
[
   {
      "Action": [
         "ec2:*Tags",
         "ec2:*Image*"
      ],
      "Resource": [
         "*"
      ],
      "Effect": "Allow"
   }
]
```

EC2AMIReadOnlyAccess

JSON policy definition for the EC2AMIReadOnlyAccess ZCS managed IAM policy:

```
[
   {
      "Action": [
         "ec2:Describe*Image*",
         "ec2:DescribeTags"
      ],
      "Resource": [
         "*"
      ],
      "Effect": "Allow"
   }
]
```

EC2ManageInstances

JSON policy definition for the EC2ManageInstances ZCS managed IAM policy:

```
[
      {
      "Action": [
         "ec2:RebootInstances",
         "ec2:Describe*",
         "autoscaling:Describe*",
         "ec2:StartInstances",
         "ec2:DescribeTags",
         "elasticloadbalancing:Describe*",
         "ec2:StopInstances",
         "cloudwatch:GetMetricStatistics",
         "cloudwatch:Describe*",
         "cloudwatch:ListMetrics"
      ],
      "Resource": [
         "*"
      ],
      "Effect": "Allow"
   }
]
```

MemberFullAccess

JSON policy definition for the MemberFullAccess ZCS managed IAM policy:

```
[
   {
      "NotAction": [
         "iam:*"
      ],
      "Resource": [
         "*"
      ],
      "Effect": "Allow"
   },
   {
      "Action": [
         "iam:ChangePassword",
         "iam:GetUser",
         "iam:CreateAccessKey",
         "iam:ListAccessKeys",
         "iam:DeleteAccessKey"
      ],
      "Resource": [
         "arn:aws:iam::*:user/${aws:username}"
      ],
      "Effect": "Allow"
   },
   {
      "Action": [
```

(continues on next page)

(continued from previous page)

```
            "iam:GetRole",
            "iam:ListAttachedRolePolicies",
            "iam:ListRoles",
            "iam:*InstanceProfile*",
            "iam:GetPolicy",
            "iam:GetPolicyVersion"
        ],
        "Resource": [
            "*"
        ],
        "Effect": "Allow"
    },
]
```

STSAssumeRole

JSON policy definition for the STSAssumeRole ZCS managed IAM policy:

```
[
    {
        "NotAction": [
            "iam:*"
        ],
        "Resource": [
            "*"
        ],
        "Effect": "Allow"
    },
    {
        "Action": [
            "iam:ChangePassword",
            "iam:GetUser",
            "iam:CreateAccessKey",
            "iam:ListAccessKeys",
            "iam:DeleteAccessKey"
        ],
        "Resource": [
            "arn:aws:iam::*:user/${aws:username}"
        ],
        "Effect": "Allow"
    },
    {
        "Action": [
            "iam:GetRole",
            "iam:ListAttachedRolePolicies",
            "iam:ListRoles",
            "iam:*InstanceProfile*",
            "iam:GetPolicy",
            "iam:GetPolicyVersion"
        ],
        "Resource": [
            "*"
```

(continues on next page)

(continued from previous page)

```
      ],
      "Effect": "Allow"
   },
]
```

STSFullAccess

JSON policy definition for the STSFullAccess ZCS managed IAM policy:

```
[
   {
      "Action": [
         "sts:*"
      ],
      "Resource": [
         "*"
      ],
      "Effect": "Allow"
   }
]
```

38.3.2 AWS-managed IAM API policies supported by Zadara Cloud Services, with links to their JSON policy documents

Name	Description
AdministratorAccess	Provides full access to AWS services and resources.
AmazonDynamoDBFullAccess	Provides full access to Amazon DynamoDB via the AWS Management Console.
AmazonDynamoDBFullAccesswithDataPipeline	Provides full access to Amazon DynamoDB including Export/Import using AWS
AmazonDynamoDBReadOnlyAccess	Provides read only access to Amazon DynamoDB via the AWS Management Co
AmazonEC2ContainerRegistryFullAccess	Provides administrative access to Amazon ECR resources.
AmazonEC2ContainerRegistryPowerUser	Provides full access to Amazon EC2 Container Registry repositories, but does n
AmazonEC2ContainerRegistryReadOnly	Provides read-only access to Amazon EC2 Container Registry repositories.
AmazonEC2FullAccess	Provides full access to Amazon EC2 via the AWS Management Console.
AmazonEC2ReadOnlyAccess	Provides read only access to Amazon EC2 via the AWS Management Console.
AmazonEC2ReportsAccess	Provides full access to all Amazon EC2 reports via the AWS Management Conso
AmazonEC2RoleforAWSCodeDeploy	Provides EC2 access to S3 bucket to download revision. This role is needed by
AmazonEKSClusterPolicy	This policy provides Kubernetes the permissions it requires to manage resourc
AmazonEKSServicePolicy	This policy allows Amazon Elastic Container Service for Kubernetes to create a
AmazonElasticFileSystemFullAccess	Provides full access to Amazon EFS via the AWS Management Console.
AmazonElasticFileSystemReadOnlyAccess	Provides read only access to Amazon EFS via the AWS Management Console.
AmazonElasticMapReduceEditorsRole	Default policy for the Amazon Elastic MapReduce Editors service role.
AmazonElasticMapReduceforAutoScalingRole	Amazon Elastic MapReduce for Auto Scaling. Role to allow Auto Scaling to add
AmazonElasticMapReduceforEC2Role	Default policy for the Amazon Elastic MapReduce for EC2 service role.
AmazonElasticMapReduceFullAccess	Provides full access to Amazon Elastic MapReduce and underlying services tha
AmazonElasticMapReduceReadOnlyAccess	Provides read only access to Amazon Elastic MapReduce via the AWS Managem
AmazonElasticMapReduceRole	Default policy for the Amazon Elastic MapReduce service role.
AmazonEMRCleanupPolicy	Allows the actions that EMR requires to terminate and delete AWS EC2 resourc
AmazonRDSBetaServiceRolePolicy	Allows Amazon RDS to manage AWS resources on your behalf.
AmazonRDSDataFullAccess	Allows full access to use the RDS data APIs, secret store APIs for RDS database

Name	Description
AmazonRDSEnhancedMonitoringRole	Provides access to Cloudwatch for RDS Enhanced Monitoring.
AmazonRDSFullAccess	Provides full access to Amazon RDS via the AWS Management Console.
AmazonRDSPreviewServiceRolePolicy	Amazon RDS Preview Service Role Policy.
AmazonRDSReadOnlyAccess	Provides read only access to Amazon RDS via the AWS Management Console.
AmazonRDSServiceRolePolicy	Allows Amazon RDS to manage AWS resources on your behalf.
AmazonRoute53DomainsFullAccess	Provides full access to all Route53 Domains actions and Create Hosted Zone to
AmazonRoute53DomainsReadOnlyAccess	Provides access to Route53 Domains list and actions.
AmazonRoute53FullAccess	Provides full access to all Amazon Route 53 via the AWS Management Console.
AmazonRoute53ReadOnlyAccess	Provides read only access to all Amazon Route 53 via the AWS Management Co
AmazonS3FullAccess	Provides full access to all buckets via the AWS Management Console.
AmazonS3ReadOnlyAccess	Provides read only access to all buckets via the AWS Management Console.
AmazonSNSFullAccess	Provides full access to Amazon SNS via the AWS Management Console.
AmazonSNSReadOnlyAccess	Provides read only access to Amazon SNS via the AWS Management Console.
AmazonSNSRole	Default policy for Amazon SNS service role.
AmazonSQSFullAccess	Provides full access to Amazon SQS via the AWS Management Console.
AmazonSQSReadOnlyAccess	Provides read only access to Amazon SQS via the AWS Management Console.
AmazonVPCCrossAccountNetworkInterfaceOperations	Provides access to create network interfaces and attach them to cross-account
AmazonVPCFullAccess	Provides full access to Amazon VPC via the AWS Management Console.
AmazonVPCReadOnlyAccess	Provides read only access to Amazon VPC via the AWS Management Console.
AutoScalingConsoleFullAccess	Provides full access to Auto Scaling via the AWS Management Console.
AutoScalingConsoleReadOnlyAccess	Provides read-only access to Auto Scaling via the AWS Management Console.
AutoScalingFullAccess	Provides full access to Auto Scaling.
AutoScalingNotificationAccessRole	Default policy for the AutoScaling Notification Access service role.
AutoScalingReadOnlyAccess	Provides read-only access to Auto Scaling.
AutoScalingServiceRolePolicy	Enables access to AWS Services and Resources used or managed by Auto Scali
AWSAutoScalingPlansEC2AutoScalingPolicy	Policy granting permissions to AWS Auto Scaling to periodically forecast capac
AWSCertificateManagerFullAccess	Provides full access to AWS Certificate Manager (ACM).
AWSCertificateManagerReadOnly	Provides read only access to AWS Certificate Manager (ACM).
AWSElasticLoadBalancingServiceRolePolicy	Service Linked Role Policy for AWS Elastic Load Balancing Control Plane.
AWSKeyManagementServiceCustomKeyStoresServiceRolePolicy	Enables access to AWS services and resources required for AWS KMS custom ke
AWSKeyManagementServicePowerUser	Provides access to AWS Key Management Service (KMS).
CloudWatchActionsEC2Access	Provides read-only access to CloudWatch alarms and metrics as well as EC2 me
CloudWatchFullAccess	Provides full access to CloudWatch.
CloudWatchReadOnlyAccess	Provides read only access to CloudWatch.
DatabaseAdministrator	Grants full access permissions to AWS services and actions required to set up a
DynamoDBReplicationServiceRolePolicy	Permissions required by DynamoDB for cross-region data replication.
ElasticLoadBalancingFullAccess	Provides full access to Amazon ElasticLoadBalancing, and limited access to oth
ElasticLoadBalancingReadOnly	Provides read only access to Amazon ElasticLoadBalancing and dependent ser
IAMFullAccess	Provides full access to IAM via the AWS Management Console.
IAMReadOnlyAccess	Provides read only access to IAM via the AWS Management Console.
IAMSelfManageServiceSpecificCredentials	Allows an IAM user to manage their own Service Specific Credentials.
IAMUserChangePassword	Provides the ability for an IAM user to change their own password.
IAMUserSSHKeys	Provides the ability for an IAM user to manage their own SSH keys.
NetworkAdministrator	Grants full access permissions to AWS services and actions required to set up a
PowerUserAccess	Provides full access to AWS services and resources, but does not allow manage
RDSCloudHsmAuthorizationRole	Default policy for the Amazon RDS service role.
ReadOnlyAccess	Provides read-only access to AWS services and resources.
SecretsManagerReadWrite	Provides read/write access to AWS Secrets Manager via the AWS Management C
SecurityAudit	The security audit template grants access to read security configuration metad
SystemAdministrator	Grants full access permissions necessary for resources required for application

Name	Description
ViewOnlyAccess	This policy grants permissions to view resources and basic metadata across all

38.4 Working with Managed AWS API Policies

38.4.1 Working with Managed AWS API Policies via the GUI

1. **To retrieve the full list of ZCS-supported managed AWS API policies:**
    1. Navigate to the **Identity & Access > AWS API Policies** view.
2. **To display the JSON policy definition of a specific managed AWS API policy:**
    1. In the **Identity & Access > AWS API Policies** view, click on the desired policy and select the **Policy** tab. The policy definition will be displayed.
3. **To display all the users, groups and roles assigned to a specific managed AWS API policy:**
    1. In the **Identity & Access > AWS API Policies** view, click on the desired policy and select the **Assignments** tab. A list of all of the users, groups and roles assigned to this policy will be displayed.

38.4.2 Working with Managed AWS API Policies via the CLI

1. **To retrieve the entire list of Zadara Cloud Services-supported managed AWS API policies:**
    1. Enter the following command:

```
Zadara Cloud Services @ Account-1/Project-1 > policy list
```

```
The list of all Zadara Cloud Services-supported managed policies will
be displayed, together with their IDs.
```

2. **To display the JSON policy definition of a specific policy:**
    1. First locate the ID of the desired policy from the list of policies with the previous command:

```
Zadara Cloud Services @ Account-1/Project-1 > policy list
```

    2. Using the ID of the desired policy, enter the following command to get its policy definition:

```
Zadara Cloud Services @ Account-1/Project-1 > policy␣
↪get ced7e6aca00340bd84e396c71763c7d8
```

    A variety of details about this policy including its policy definition will be displayed.

3. **To display all of the users, groups and roles assigned to a specific policy:**
    1. First locate the ID of the desired policy from the list of policies with the previous command:

```
Zadara Cloud Services @ Account-1/Project-1 > policy list
```

    2. Using the ID of the desired policy, enter the following command to get its policy definition:

```
Zadara Cloud Services @ Account-1/Project-1 > policy get-
↪entities ced7e6aca00340bd84e396c71763c7d8
```

    All of the users, groups and roles attached to the selected policy, will be displayed.

4. **To display all the assigned policies for all of users, groups and roles:**

    1. Enter the following command:

    ```
    Zadara Cloud Services @ Account-1/Project-1 > policy assignments-list
    ```

    All of the assigned policies for all users, groups and roles, will be displayed.

CHAPTER

THIRTYNINE

AWS IAM ROLES AND INSTANCE PROFILES

39.1 AWS IAM Roles

AWS IAM Roles are policy-based tokens with temporary credentials allowing a user temporary access to AWS services and actions which the user is normally not permitted to access. These users may be from different projects or even different accounts. These roles can also be embedded into specific instances allowing these instances access to the necessary actions.

Important

The AWS IAM roles are independent of the Zadara Cloud Services roles, which together with Zadara Cloud Services policies, grant access to ZCS services and actions.

The AWS IAM role consists of the following:

1. Permissions policy which give access to certain ZCS supported AWS services or actions. This is defined in the **Identity & Access > AWS Roles** view, **Attached Policies** tab.
2. Trust policy that defines the relationship between user per project and this role. This is defined in the **Identity & Access > AWS Roles** view, **Trust Policy** tab.
    1. The nature of the relationship may be 'allow' which grants permission to the specified users to assume the role, or 'deny' which prevents these users from assuming the role.
    2. This permission may be granted to multiple users of the same projects, different projects within the same account, or even users of different accounts.
3. The maximum session duration that can be requested when assuming this role.

39.1.1 Viewing the AWS IAM Role List

1. Navigate to the **Identity & Access > AWS Roles** view. The defined AWS IAM Role list will be displayed.
2. Select a specific role in the display list to view the following detailed information:
    1. **Events** - displays events related to the role.
    2. **Attached Policies** - displays AWS API policies attached to this role.
    3. **Trust Policy** - displays entities allowed/denied to assume this role.

39.1.2 Creating an AWS IAM Role

1. Navigate to the **Identity & Access > AWS Roles** view.
2. In top toolbar, click **Create**.
3. In the **Create Role** dialog, **Details** tab, enter the following:

1. **Name** - name of the role AWS IAM Role to be defined.
2. **Description** - description of role..
3. **Max. Session Duration** - maximum time (in seconds) that a user can assume this role.
    - **Default:** 3600 seconds (1 hour).
    - **Maximum:** 12 hours (43,200 second).
4. **Policies** - Select one or more managed AWS policies which define the permissions of this role.
5. Click **Next**.

4. In the **Allow Assuming** tab, the trust policy is defined for this role. Multiple trust policies can be defined by clicking **Add**. For each trust policy, enter the following information:
    1. **User/Service/Any** - defines for what user or service the role will be applied.
        - If **User** is selected, enter the project name and user name.
        - If **Service** is selected, enter the service type. Currently, the only available service type for the trust policy is **VM**.
        - If **Any** is selected, enter project name. The trust policy will be applied to all users.
    2. Click **Next**.
5. In the **Deny Assuming** tab, users or services who can not assume this role may be defined. This is an optional configuration and by default there are no denied users or services. If defined, the same **user/service/any** options described for the **allow assuming** step above are also available for the **deny assuming** step. **Note:** If the same user and project are defined for both **allow** and **deny**, the 'Deny' prevails.
6. Click **Finish**. The new role will appear in the **Identity & Access > AWS Roles** view.

39.1.3 Deleting an AWS IAM Role

Important

A role cannot be deleted if policies are attached to it.

1. Navigate to the **Identity & Access > AWS Roles** view.
2. From the displayed list, select the role to be deleted.
3. From top toolbar, click **Delete**. The **Delete Role** dialog will open.
4. Click **Delete** to confirm. The role will be removed from the **Identity & Access > AWS Roles** list.

39.1.4 Modifying an AWS IAM Role

1. Navigate to the **Identity & Access > AWS Roles** view.
2. From the displayed list, select the role to be modified.
3. From top toolbar, click **Modify**. The **Modify Role** dialog will be displayed. The following parameters can be modified:
    1. **Max. Session Duration** - maximum time (in seconds) that a user can assume this role.
    2. **Policies** - managed policies which define the permissions of this role.
    3. **Description**
4. Click **OK** to save the changes.

Note

The following can not be modified:

1. The project for which this role is defined.
2. The users who may assume this rule.
3. The users who are prevented from assuming this role.

39.1.5 Applying an AWS IAM Role via the CLI

When an AWS IAM role has been created and a policy attached, the role can be applied via the Zadara Cloud Services CLI. This in turn will generate the credentials needed to access the AWS services and actions defined in the role. In order to assume a role you must know its ID.

1. Find the AWS IAM roles you want to use:
    1. Run role iam-list:

        Zadara Cloud Services @ user1/cloudacc1 > role iam-list

        If you are a Member or Tenant Admin user, this command returns a list of any roles created in the currently logged-in project, together with the ID's of all users who can assume these role(s).
    2. To check if the entity_ids for any of the users is your own, use the following command:

        Zadara Cloud Services @ user1/cloudacc1 > user get-my-details

        This will display your user_id along with your user_name.
    3. Regardless of AWS IAM roles discovered, there may be roles available which are not visible to you. To verify that you have the complete list of roles available to you, it is recommended that a Zadara Admin user does the following:
        - List all of the AWS IAM Roles

        ```
        Zadara Cloud Services  @ user1/cloudacc1 > role iam-list
        ```

        All AWS IAM roles in the account will be displayed with the ID's of all users to which these roles are available.
        - List all of the users

        ```
        Zadara Cloud Services  @ user1/cloudacc1 > user list
        ```

        This will display all of the users in the account with their names and ID's. It will now be possible to discover which roles are available for which users.
2. Once you have the Role-ID, you can assume the AWS IAM role via the 'role assume-role **role_id session_name**' command, where role_id is supplied in the output above, and session_name is selected by the user.

```
Zadara Cloud Services @ user1/cloudacc1 > role assume-role 565197da-2f13-48dc-b232-
↪bdffc756b7f9 Session-1
```

This returns the following information:

```
==================== ====================================
**access_key_id**     23515d96a5da4408821783e0b9aa6ff1

**created_at**        2019-03-10T20:55:42Z
```

(continues on next page)

(continued from previous page)

```
**duration_seconds**  3600

**expires_at**        2019-03-10T21:55:42Z

**external_id**       none

**policy_id**         none

**project_id**        fc268815422e471da6756c7918b03d01

**role_assumer**      5120ef807769455b822095497b55ffac

**role_id**           565197da-2f13-48dc-b232-bdffc756b7f9

**role_name**         Role-2

**secret_access_key** 94d8a61f419b4edbb638abc399e7a420

**session_name**      Session-1

**token**             cd9b0253e0034cdd976c21c317c
===================== ====================================
```

3. Use the access_key_id, secret_access_key and token to access the AWS services and actions.

39.2 Instance Profiles

An instance profile is a container for an AWS IAM role. It can be used to pass role information to an EC2 instance when the instance starts. When an AWS IAM role, embedded in an instance profile, becomes attached to an instance, its credentials become permanent.

The following actions can be performed with instance profiles:

39.2.1 Viewing Instance Profiles

1. In the zCompute UI, go to **Compute > Instance Profiles**.

   The instance profiles list is displayed.

2. Click an instance profile in the list to view its details.

   The detailed instance profile view is displayed, with the option to modify the instance profile by removing its associated role, and the option to delete the instance profile.

39.2.2 Creating an Instance Profile

To create an instance profile:

1. In the zCompute UI, go to **Compute > Instance Profiles**.

   The instance profiles list is displayed.

2. In the top menu bar, click **Create**.

   The **Create Instance Profile** dialog opens.

3. In the **Create Instance Profile** dialog, enter:

- **Name** - a name to identify the instance profile.
- **Role** - optionally select an AWS IAM role from the dropdown.
- Click **Finish**.

The new instance profile is displayed in the instance profiles list.

39.2.3 Modifying an Instance Profile

An instance profile can be modified by changing its AWS IAM role. This involves removing the current AWS IAM role from instance profile, and adding a new role.

Removing an AWS IAM Role

1. In the zCompute UI, go to **Compute > Instance Profiles**.

   The instance profiles list is displayed.

2. To modify an instance profile, click its entry in the list.

   The detailed instance profile view is displayed.

3. In the top menu bar, click **Remove Role**.

   The **Remove Role from Instance Profile** confirmation dialog opens.

4. Click **Delete** to remove the AWS IAM role from the instance profile.

Adding an AWS IAM Role

1. In the zCompute UI, go to **Compute > Instance Profiles**.

   The instance profiles list is displayed.

2. To modify an instance profile, click its entry in the list.

   The detailed instance profile view is displayed.

3. In the top menu bar, click **Add Role**.

   The **Add Role to Instance Profile** dialog opens.

4. Select an AWS IAM role from the dropdown, to add to the instance profile.
5. Click **Ok**.

39.2.4 Working with Instance Profiles via the CLI

Adding an AWS IAM Role into an Instance

Using instance profiles, it is possible to grant permissions to an instance to access specific Zadara Cloud Services-supported AWS services. Although, in the context of a user, role permissions are temporary, in the context of an instance, the application is guaranteed to have credentials as long as the instance profile is attached to the instance, and that instance profile has a role embedded in it.

1. From the Zadara Cloud Services GUI, create an AWS IAM role together with its AWS IAM policies and trust policies.

   Note

   This must be performed by a Zadara Admin or a Tenant Admin user.

2. To create an instance profile, use the Zadara Cloud Services CLI command: 'instance-profile create **name**'.

   Note

   This command may be performed by any user.

   ```
   Zadara Cloud Services @ user1/cloudacc1 > instance-profile create instance-profile-1
   ```

   To view basic details about the newly created instance-profile, use command *instance-profile -1*.

   ```
   ============== ==================================
   **id**         dd56d017-0167-42b7-a130-8706d746493e
   **name**       instance-profile-1
   **created_at** 2019-03-11T02:58:30Z
   **path**       /
   **project_id** 4331358dff9b4c29aa53c982e92801f6
   **roles**      []
   ============== ==================================
   ```

3. Find the AWS IAM roles that you can embed in the instance profile as follows:

   1. Run role iam-list:

      ```
      Zadara Cloud Services @ user1/cloudacc1 > role iam-list
      ```

      If you are a Member or Tenant Admin user, the above command returns a list of any roles created in the currently logged-in project, together with the ID's of all users who can assume these role(s), as shown below.

      Note

      If you are assigned to more than one project, you must login to each project and run 'role iam-list' to get the complete list of roles available to you.

   2. To check if the entity_ids for any of the users is yours run the following command:

      ```
      Zadara Cloud Services @ user1/cloudacc1 > user get-my-details
      ```

      This will display your user_id along with your user_name.

4. Embed the role in the instance profile with the Zadara Cloud Services CLI command 'instance-profile add-role **instance_profile-id role_id**', as follows: (Take the instance_profile_id and the role_id from the steps above.) **Note:** This command can be performed by a Member or Tenant admin if the project of the role is the same as the logged-in project.

   ```
   Zadara Cloud Services @ user1/cloudacc1 > instance-profile add-role **dd56d017-0167-
   ↪42b7-a130-8706d746493e**  **565197da-2f13-48dc-b232-bdffc756b7f9**
   ```

5. Use the GUI to create an instance.

   Note

   If you create this instance from the CLI, you can add the instance-profile on creation. This will remove the need for steps 6 and 8.

6. From the Zadara Cloud Services CLI using the CLI command "vm list" locate the ID of the instance that you just created.

7. Since the instance profile operation is very sensitive to network latency and cluster load, the system may time out before finishing the operation. It is therefore recommended to increase the timeout on the metadata service connection and/or allow retries. These can be configured inside the VM that is connected to the instance-profile by setting the following environment vars:

```
AWS_METADATA_SERVICE_TIMEOUT >1
AWS_METADATA_SERVICE_NUM_ATTEMPTS >1
```

8. Using the Zadara Cloud Services CLI command 'vm update **–instance-profile INSTANCE_PROFILE vm_id**" attach the instance profile to the instance you just created, as follows: (Take the instance_profile_id and the vm_id from the steps above.) **Note:** The role and VM must be defined for the same project. A Member user and Tenant Admin user can perform this command if they are logged in to the same project as that of the role and VM.

```
Zadara Cloud Services @ user1/cloudacc1 > vm update -**-instance-
↪profile** **dd56d017-0167-42b7-a130-8706d746493e d6ca69e7-1533-4740-9881-
↪395d442719f5**
```

Instance Profiles CLI commands

1. To create an instance profile enter the following command:

```
Zadara Cloud Services @ Account-1/Project-1 > instance-profile get <name of␣
↪instance-profile>
```

   A variety of details about this instance profile will be displayed, including it's ID

2. To add an AWS IAM role to an instance profile:
   1. Procure the role-id via the following command

```
Zadara Cloud Services @ Account-1/Project-1 > role iam-list
```

      The list of all IAM roles will be displayed, together with their ID's.

   2. Using the ID of the desired instance profile, returned from the first command above, enter the following command:

```
Zadara Cloud Services @ Account-1/Project-1 > instance-profile add-role
↪<instance_profile_id> <role_id>
```

      Only one role can be added to an instance profile.

3. To remove an AWS IAM role from an instance profile, enter the following command:

```
Zadara Cloud Services @ Account-1/Project-1 > instance-profile remove-role
↪<instance_profile_id> <role_id>
```

4. To retrieve the entire list of Instance profiles in the cluster enter the following command:

```
Zadara Cloud Services @ Account-1/Project-1 > instance-profile list
```

   The list of all instance profiles will be displayed, together with their ID's.

5. To remove an Instance profile:

   Using the ID of the desired instance profile retrieved above, enter the following command:

```
Zadara Cloud Services @ Account-1/Project-1 > instance-profile remove <instance_
↪profile_id>
```

6. To retrieve information about a specific instance profile:

   Using the ID of the desired instance profile retrieved above, enter the following command:

```
Zadara Cloud Services @ Account-1/Project-1 > instance-profile get <instance_
↪profile_id>
```

   A variety of details about this instance profile including its ID, will be displayed.

CHAPTER

FORTY

ZADARA CLOUD SERVICES POLICIES

40.1 Overview

Zadara Cloud Services Policies define the permissions for accessing Zadara Cloud Services functionality. For each project available to a user, one or more Zadara Cloud Services Policies and a single Zadara Cloud Services Role are assigned. If the role of the user does not match the role type of any of the APIs included in the policy, then the user cannot perform these APIs, even if the policy which includes these APIs was assigned to the user.

Whenever a new user is created via the GUI, a FullAccess policy is the suggested default Zadara Cloud Services policy. When a new user is created via the API/CLI, FullAccess is automatically assigned Zadara Cloud Services policy. This policy provides full access to all of the services that each Zadara Cloud Services role granted in the past, prior to v5.3.4.

Note

In the ZCS GUI, Zadara Cloud Services policy is shown in the **Identity & Access > Symp API Policies** view.

40.2 Working with Zadara Cloud Services Policies

Zadara Cloud Services policies can be assigned while creating a user on the Permissions tab of the Create User wizard, or on the Manage Permissions dialog for an existing user. For more information on assigning policies to users, see *Creating Users*.

Note

If a Zadara Cloud Services policy is assigned to a user while the user is in the middle of a session, it will not take effect until the user first logs out of the current session and then logs in again to a new session.

40.2.1 Assigning Permissions for Creating a VM

As an example of how to use policies to allow certain functionality, the following section explains how to use Zadara Cloud Services policies to grant the permissions needed by a user to create a VM.

1. A Member user with only VMFullAccess rights, grants permission to create a VM and select an Instance Type, but there are no permissions to view the following entities:
    1. **Image**
    2. **Storage Pool**
    3. **Subnet**
2. A Member user with legacy permissions providing full access to all Zadara Cloud Services member functionality, can create a VM. All of the fields in the Create VM dialog, including Image, Storage Pool, and Subnet are available to these users.

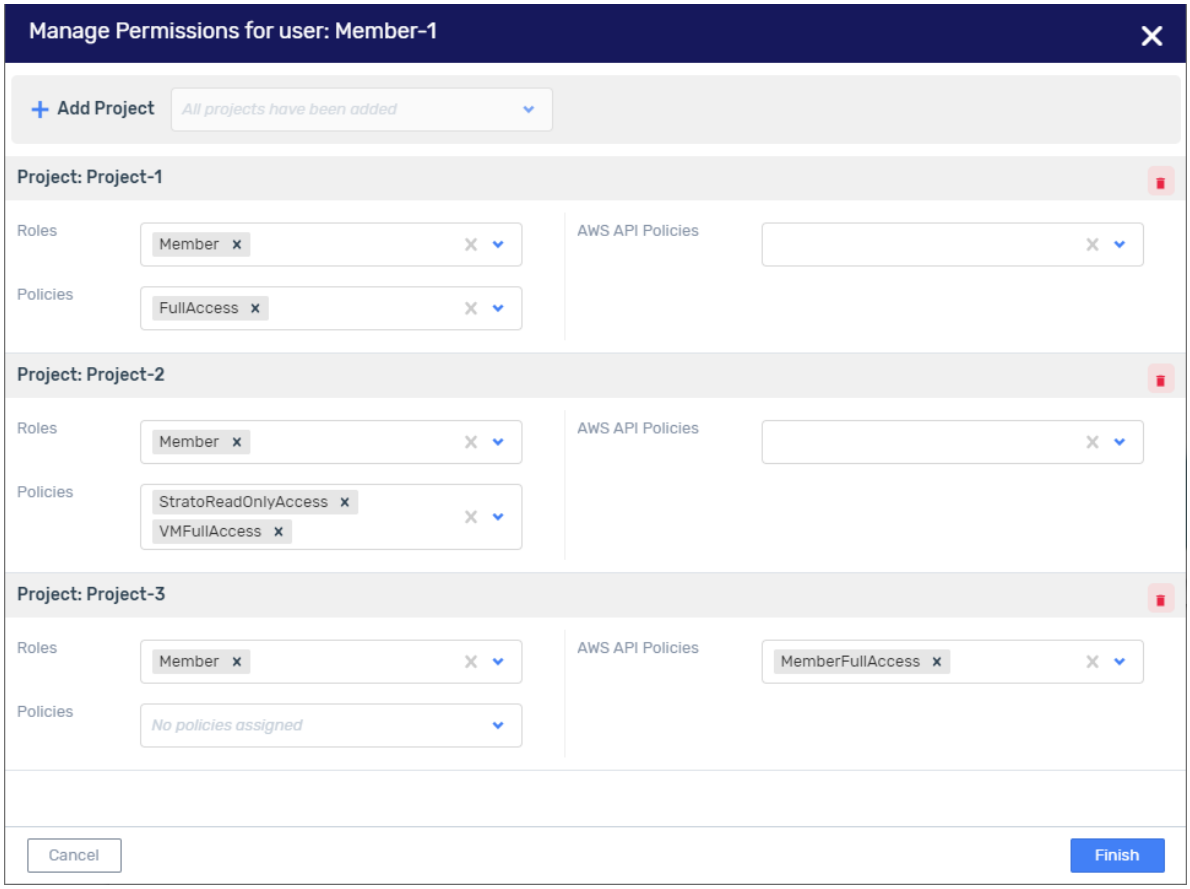
3. A Member user assigned the VMFullAccess Zadara Cloud Services policy together with the ZadaraReadOnlyAccess policy can also create a VM.
    1. The VMFullAccess Zadara Cloud Services policy permits the creation of a VM and the selection of any Instant Type.
    2. The ZadaraReadOnlyAccess Zadara Cloud Services policy permits the viewing and selecting of any entity available to the legacy Member user, including the following:
        - **Images**
        - **Storage Pools**
        - **Subnets**
4. A Member user assigned the VMFullAccess Zadara Cloud Services policy together with the following read-only Zadara Cloud Services policies can also create a VM:
    1. The VMFullAccess Zadara Cloud Services policy permits the creation of a VM and the selection of an Instant Type.
    2. The ImagesReadOnlyAccess Zadara Cloud Services policy permits the viewing and selecting of Images.
    3. The StorageReadOnlyAccess Zadara Cloud Services policy permits the viewing and selecting of Storage Pools.
    4. The VPCReadOnlyAccess Zadara Cloud Services policy permits the viewing and selecting of Subnets.

40.2.2 Assigning Permissions for Multiple Projects to a Single User

The following section explains how to use Zadara Cloud Services policies to assign multiple projects to a single user. This can be done when creating the user or by managing permissions of an existing user. For example, to assign permissions for only Zadara Cloud Services functionality (but not AWS API functionality) for each of three different projects, the configurations below should be made.

Projects:

- Project-1 - Full legacy Member role permissions
- Project-2 - Permissions to create VMs and many other actions concerning VMs
- Project-3 - To be determined



In the **Manage Permissions for user** view, the following information should be configured:

1. **Roles** - Verify that the **Member** role has been assigned to each project.

   Note

   It is strongly recommended not to assign different roles to the same user for different projects.

2. **Policies** - Each Project should be assigned the following Zadara Cloud Services policies:

   1. Project-1 - Select the `FullAccesspolicy` to grant Member legacy permissions.
   2. Project-2 - Select the `VMFullAccessandStratoReadOnlyAccess` to grant permissions to create a VM, including viewing Images, Storage Pools and Subnets, access to which is necessary when creating a VM.
   3. Project-3 - Leave empty until User-1's responsibilities for this project have been determined.

      Note

      As long as at least one project has both assigned a Zadara Cloud Services Role and Policy, the user's permissions for other projects need not be completely defined.

3. **AWS API policies** - Verify the MemberFullAccesspolicy is removed from each of the projects.
4. Click **Finish**.

40.2.3 List of Managed Zadara Cloud Services Policies

Zadara Cloud Services Policies basically consist of two policies per service, one for **FullAccess** of all the service's APIs, and one for **ReadOnly access** of all the service's APIs.

1. **FullAccess** – a policy which provides full access to all Zadara Cloud Services APIs based on the user's role.
2. **ReadOnlyAccess** - a policy which provides read-only access to all Zadara Cloud Services APIs based on the user's role.

Exceptions to this rule are:

1. The Identity Service which has two additional policies:
    1. **IdentityBasicUsage** - which provides access to those identity entities available to a Member user.
    2. **IdentitySTSAssumeRole** - which provides the ability to obtain an IAM role using the 'assume-role' API.
2. The Snapshot service which has three sets of FullAccess and ReadOnlyAccess policies, one set for each of the following:
    1. Snapshot
    2. RemoteSnapshot
    3. RemoteVMSnapshot
3. The VM service which has two sets of FullAccess and ReadOnlyAccess policies, one set for each of the following:
    1. VM
    2. VMSnapshot

The table below summarizes all Zadara Cloud Services Policies with their descriptions and the roles for which they are enforced. To see the actual content of a policy go to the **Identity & Access > Symp API Policies** view, and click on the requested policy.

	Zadara Cloud Services Policy	Description	Available to:
1.	AlarmsFullAccess	Provides full access to all alarms APIs	All ZCS roles
2.	AlarmsReadOnlyAccess	Provides read-only access to all alarms APIs	All ZCS roles
3.	AutoScalingFullAccess	Provides full access to all autoscaling-group APIs	All ZCS roles
4.	AutoScalingReadOnlyAccess	Provides read-only access to all autoscaling-group APIs	All ZCS roles
5.	CRSFullAccess	Provides full access to all container registry APIs	All ZCS roles
6.	CRSReadOnlyAccess	Provides read-only access to all container registry APIs	All ZCS roles
7.	CertificateManagerFullAccess	Provides full access to all certificate APIs	All ZCS roles
8.	CertificateManagerReadOnlyAccess	Provides read-only access to all certificate APIs	All ZCS roles
9.	CloudWatchFullAccess	Provides full access to all cloudwatch APIs	All ZCS roles

continues on next page

Table 1 – continued from previous page

	Zadara Cloud Services Policy	Description	Available to:
10.	CloudWatchReadOnlyAccess	Provides read-only access to all cloudwatch APIs	All ZCS roles
11.	ConversionsFullAccess	Provides full access to all conversion APIs	All ZCS roles
12.	ConversionsReadOnlyAccess	Provides read-only access to all conversion APIs	All ZCS roles
13.	DBCFullAccess	Provides full access to all DB cluster APIs	All ZCS roles
14.	DBCReadOnlyAccess	Provides read-only access to all DB cluster APIs	All ZCS roles
15.	DBSFullAccess	Provides full access to all DB APIs	All ZCS roles
16.	DBSReadOnlyAccess	Provides read-only access to all DB APIs	All ZCS roles
17.	EngineFullAccess	Provides full access to all engine APIs	All ZCS roles
18.	EngineReadOnlyAccess	Provides read-only access to all engine APIs	All ZCS roles
19.	ExternalEndpointFullAccess	Provides full access to all external-endpoint APIs	All ZCS roles
20.	ExternalEndpointReadOnlyAccess	Provides read-only access to all external-endpoint APIs	All ZCS roles
21.	EventsAccess	Provides access to all events APIs	All ZCS roles
22.	FullAccess	Provides full access to all Zadara Cloud Services APIs based on user's scope	All ZCS roles
23.	GuestnetToolFullAccess	Provides full access to all guestnet-admin-tool APIs	All ZCS roles
24.	GuestnetToolReadOnlyAccess	Provides read-only access to all guestnet-admin-tool APIs	All ZCS roles
25.	HotUpgradeFullAccess	Provides full access to all hot-upgrade APIs	Admin
26.	HotUpgradeReadOnlyAccess	Provides read-only access to all hot-upgrade APIs	Admin
27.	IdentityBasicUsage	Provides access to basic identity entities operations.	All ZCS roles
28.	IdentityFullAccess	Provides full access to identity APIs	All ZCS roles

continues on next page

Table 1 – continued from previous page

	Zadara Cloud Services Policy	Description	Available to:
29.	IdentityReadOnlyAccess	Provides read-only access to identity APIs	All ZCS roles
30.	IdentitySTSAssumeRole	Provides the ability to obtain an IAM role using the 'assume-role' API	All ZCS roles
31.	ImagesFullAccess	Provides full access to all image APIs	All ZCS roles
32.	ImagesReadOnlyAccess	Provides read-only access to all image APIs	All ZCS roles
33.	InspectorFullAccess	Provides full access to all inspector APIs	Admin
34.	InspectorReadOnlyAccess	Provides read-only access to all inspector APIs	Admin
35.	KubernetesFullAccess	Provides full access to all Kubernetes APIs	All ZCS roles
36.	KubernetesReadOnlyAccess	Provides read-only access to all Kubernetes APIs	All ZCS roles
37.	LbaasFullAccess	Provides full access to all LBaaS APIs	All ZCS roles
38.	LbaasReadOnlyAccess	Provides read-only access to all LBaaS APIs	All ZCS roles
39.	MapReduceFullAccess	Provides full access to all map-reduce APIs	All ZCS roles
40.	MapReduceReadOnlyAccess	Provides read-only access to all map-reduce APIs	All ZCS roles
41.	MetricsAccess	Provides access to all metrics APIs	All ZCS roles
42.	NFSFullAccess	Provides full access to all NFS APIs	All ZCS roles
43.	NFSReadOnlyAccess	Provides read-only access to all NFS APIs	All ZCS roles
44.	NodesFullAccess	Provides full access to all node APIs	Admin
45.	NodesReadOnlyAccess	Provides read-only access to all node APIs	Admin
46.	ObjectStoresFullAccess	Provides full access to all object-store APIs	All ZCS roles
47.	ObjectStoresReadOnlyAccess	Provides read-only access to all object-store APIs	All ZCS roles

continues on next page

Table 1 – continued from previous page

	Zadara Cloud Services Policy	Description	Available to:
48.	ProtectionFullAccess	Provides full access to all protection APIs	All ZCS roles
49.	ProtectionReadOnlyAccess	Provides read-only access to all protection APIs	All ZCS roles
50.	QuotasFullAccess	Provides full access to all quota APIs	All ZCS roles
51.	QuotasReadOnlyAccess	Provides read-only access to all quota APIs	All ZCS roles
52.	RemoteSnapshotFullAccess	Provides full access to all remote-snapshot APIs	All ZCS roles
53.	RemoteSnapshotReadOnlyAccess	Provides read-only access to all remote-snapshotAPIs	All ZCS roles
54.	RemoteVMSnapshotFullAccess	Provides full access to remote VM Snapshot APIs	All ZCS roles
55.	RemoteVMSnapshotReadOnlyAccess	Provides read-only access to remote VM Snapshot APIs	All ZCS roles
56.	Route53FullAccess	Provides full access to all Route53 APIs	All ZCS roles
57.	Route53ReadOnlyAccess	Provides read-only access to all Route53 APIs	All ZCS roles
58.	SnapshotFullAccess	Provides full access to all local compute-snapshot APIs	All ZCS roles
59.	SnapshotReadOnlyAccess	Provides read-only access to all local compute-snapshot APIs	All ZCS roles
60.	StorageFullAccess	Provides full access to all storage APIs	Admin
61.	StorageReadOnlyAccess	Provides read-only access to all storage APIs	All ZCS roles
62.	StratoReadOnlyAccess	Provides read-only access to all Zadara Cloud Services APIs based on user's scope	All ZCS roles
63.	VMFullAccess	Provides full access to VM APIs	All ZCS roles
64.	VMReadOnlyAccess	Provides read-only access to VM APIs	All ZCS roles
65.	VMSnapshotFullAccess	Provides full access to VM Snapshot APIs	All ZCS roles
66.	VMSnapshotReadOnlyAccess	Provides read-only access to VM Snapshot APIs	All ZCS roles

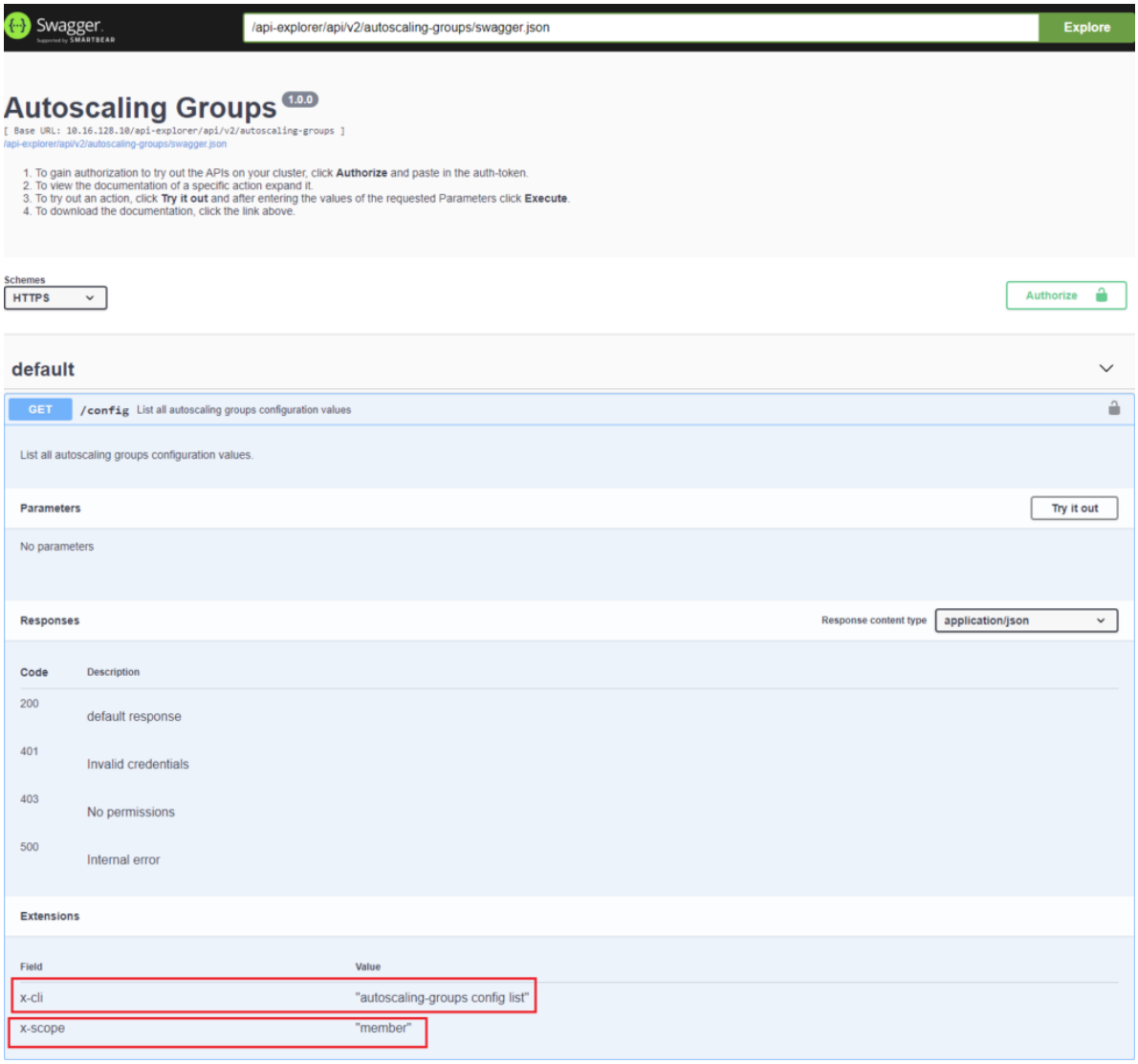
continues on next page

Table 1 – continued from previous page

	Zadara Cloud Services Policy	Description	Available to:
67.	VPCFullAccess	Provides full access to all VPC APIs	All ZCS roles
68.	VPCReadOnlyAccess	Provides read-only access to all VPC APIs	All ZCS roles
69.	VolumesFullAccess	Provides full access to all volume APIs	All ZCS roles
70.	VolumesReadOnlyAccess	Provides read-only access to all volume APIs	All ZCS roles

Notes:

1. It is currently possible to assign a Zadara Cloud Services policy to a user, for whose role, the APIs in this policy will not be permitted. For example, it is possible to assign the NodesFullAccess policy to Member users, even though the APIs in this policy will not be permitted for Members. In a future release, this invalid assignment will be prevented.
2. A Zadara Cloud Services policy which is permitted for a Member or Tenant Admin role, may nevertheless include some APIs/CLIs which are not permitted for that role. For example, the VMFullAccess policy whose APIs/CLIs are permitted for all roles, including the Member role, contains the API, 'vm live-migrate' which migrates VMs from one node to another. This action can be performed only by a Zadara Admin user. To determine the role for which specific APIs of a Zadara Cloud Services policy are permitted, access the API Explorer for the service covered by the policy and open the API, as shown below.



3. In the Extensions section at the bottom of the window, the x-cli field displays the name of the CLI while the x-scope field displays the role with the least amount of permissions which may use this CLI. (e.g. ‘member’ has less permissions than ‘tenant_admin’). If x-scope = ‘all’ permissions are not required to perform the CLI.

CHAPTER

FORTYONE

AWS IDENTITY-AND-ACCESS MANAGEMENT (IAM)

AWS API Reference	Ignored Param	Optional Parameters	Required Parameters
AddRoleToInstanceProfile	[]	[]	InstanceProfileName RoleNam
AddUserToGroup	[]	[]	GroupName UserName
AttachGroupPolicy	[]	[]	GroupName PolicyArn
AttachRolePolicy	[]	[]	PolicyArn RoleName
AttachUserPolicy	[]	[]	PolicyArn UserName
ChangePassword	[]	[]	NewPassword OldPassword
CreateAccessKey	[]	UserName	[]
CreateGroup	[]	Path	GroupName
CreateInstanceProfile	[]	Path	InstanceProfileName
CreateLoginProfile	[]	PasswordResetRequired	UserName Password
CreatePolicy	[]	Description Path	PolicyDocument PolicyName
CreateRole	[]	Description MaxSessionDuration Path	AssumeRolePolicyDocument
CreateUser	[]	Path	UserName
DeleteAccessKey	[]	UserName	AccessKeyId
DeleteGroup	[]	[]	GroupName
DeleteInstanceProfile	[]	[]	InstanceProfileName
DeleteLoginProfile	[]	[]	UserName
DeletePolicy	[]	[]	PolicyArn
DeleteRole	[]	[]	RoleName
DeleteUser	[]	[]	UserName
DetachGroupPolicy	[]	[]	GroupName PolicyArn
DetachRolePolicy	[]	[]	RoleName PolicyArn
DetachUserPolicy	[]	[]	PolicyArn UserName
GetGroup	[]	Marker MaxItems	GroupName
GetInstanceProfile	[]	[]	InstanceProfileName
GetLoginProfile	[]	[]	UserName
GetPolicy	[]	[]	PolicyArn
GetPolicyVersion	[]	[]	PolicyArn VersionId
GetRole	[]	[]	RoleName
GetUser	[]	UserName	[]
ListAccessKeys	[]	Marker MaxItems UserName	[]
ListAttachedGroupPolicies	[]	Marker MaxItems PathPrefix	GroupName
ListAttachedRolePolicies	[]	Marker MaxItems PathPrefix	RoleName
ListAttachedUserPolicies	[]	Marker MaxItems PathPrefix	UserName
ListEntitiesForPolicy	[]	EntityFilter Marker MaxItems PathPrefix	PolicyArn
ListGroups	[]	Marker MaxItems PathPrefix	[]
ListGroupsForUser	[]	Marker MaxItems	UserName
ListInstanceProfiles	[]	Marker MaxItems PathPrefix	[]

Table 1 – continued from previous page

AWS API Reference	Ignored Param	Optional Parameters	Required Parameters
ListInstanceProfilesForRole	[]	[]	RoleName
ListMFADevices	[]	Marker MaxItems UserName	[]
ListPolicies	[]	Marker MaxItems OnlyAttached PathPrefix Scope	[]
ListRoles	[]	Marker MaxItems PathPrefix	[]
ListUsers	[]	Marker MaxItems PathPrefix	[]
RemoveRoleFromInstanceProfile	[]	[]	InstanceProfileName RoleNam
RemoveUserFromGroup	[]	[]	GroupName UserName
UpdateAccessKey	[]	UserName	AccessKeyId Status
UpdateAssumeRolePolicy	[]	[]	RoleName PolicyDocument
UpdateGroup	[]	NewGroupName NewPath	GroupName
UpdateLoginProfile	PasswordResetRequired	Password PasswordResetRequired	UserName
UpdateRole	[]	Description MaxSessionDuration	RoleName
UpdateRoleDescription	[]	[]	RoleName Description
UpdateUser	[]	NewPath NewUserName	UserName

41.1 AWS-STS

AWS API Reference	Ignored Param	Optional Parameters	Required Parameters	Unsupported Params
AssumeRole	[]	DurationSeconds	RoleArn RoleSessionName	ExternalId Policy SerialNumber TokenCode
GetCallerIdentity	[]	[]	[]	[]

CHAPTER

FORTYTWO

AWS CONSOLE

From within the zCompute UI, you can use the AWS Console to run AWS CLI commands.

Supported AWS services include EC2 and S3.

42.1 Using the AWS Console

To use the AWS Console:

1. Select **Consoles > AWS**.

   The AWS shell appears.

2. At the `aws>` prompt, run AWS commands.

   Example:

```
Running AWS Shell
aws> ec2 describe-images
{
    "Images": [
        {
            "Name": "dbs_manager_mysql_5_5_002_default_v1",
            "ImageId": "ami-1d14fdb4da924c5b87d962d487747bdc",
            "State": "available",
            "Architecture": "",
            "ImageLocation": "None (dbs_manager_mysql_5_5_002_default_v1)",
            "RootDeviceType": "ebs",
            "OwnerId": "5870a3972fdf41dd85fdce48f7c7d373",
            "CreationDate": "2018-02-22T19:11:56",
            "Public": true,
            "ImageType": "machine"
        },
        {
            "Name": "persistent/rootfs-centos7-workload.qcow2",
            "ImageId": "ami-4f735f3b79a24ffd8c7286fc5b504f12",
            "State": "available",
            "Architecture": "",
            "ImageLocation": "None (persistent/rootfs-centos7-workload.qcow2)",
            "RootDeviceType": "ebs",
            "OwnerId": "5870a3972fdf41dd85fdce48f7c7d373",
            "CreationDate": "2018-02-22T15:07:48",
            "Public": false,
```

(continues on next page)

(continued from previous page)


```
            "ImageType": "machine"
        },
        {
            "Name": "cirros-0.3.1-x86_64-disk.qcow2",
            "ImageId": "ami-22e8f0f0290b4b5dbd9180fa02a77a92",
            "State": "available",
            "Architecture": "",
            "ImageLocation": "None (cirros-0.3.1-x86_64-disk.qcow2)",
            "RootDeviceType": "ebs",
            "OwnerId": "5870a3972fdf41dd85fdce48f7c7d373",
            "CreationDate": "2018-02-22T15:07:48",
            "Public": false,
            "ImageType": "machine"
        },
        {
            "Name": "persistent/trusty-server.qcow2",
            "ImageId": "ami-b3c8012c79f44de2a63ee34b2adeeb40",
            "State": "available",
            "Architecture": "",
            "ImageLocation": "None (persistent/trusty-server.qcow2)",
            "RootDeviceType": "ebs",
            "OwnerId": "5870a3972fdf41dd85fdce48f7c7d373",
            "CreationDate": "2018-02-22T15:07:48",
            "Public": false,
            "ImageType": "machine"
        }
    ]
}
aws>
```


3. Exit the console by pressing F10.

42.2 Copying console output

To copy output from the AWS Console:

1. In your Chrome browser, highlight the console output that you want to copy.
2. Press **Enter** to copy the selected output to the clipboard.

CHAPTER

FORTYTHREE

SYMP CONSOLE

The Zadara zCompute Symp client is a unified CLI (Command Line Interface) to manage and operate your zCompute services.

From within the zCompute UI, you can use the Symp Console to run zCompute Symp CLI commands.

43.1 Using the Symp Console

To use the Symp Console:

1. Select **Consoles > Symp**.

   The Symp shell appears.

2. At the `Symphony >` prompt, run Symp commands.

   Example:

   Listing the projects in default output format:

```
Symphony > project list
+----------------------------------+--------------+--------------------+----------
↪------------------------+-------------+---------+----------+--------+-----------
↪----------------------+
| id                               | name         | description        | domain_
↪id                      | domain_name | enabled | is_domain | is_vpc | parent_
↪id                      |
+==================================+==============+====================+==========================
| 12d782aa3b8543e48110235d0d09ae73 | proj3-vpc    |                    |␣
↪abd8a5a9b33b4d9e8e3a8219ec2950fb | tenant      | true    | false     | true   |␣
↪kld8a5a9b33b4d9e8e3a8219ec2950fb |
| 34e07f1cc0f444a0960563ad8d5bee4a | dvsproj1     |                    |␣
↪cdc4aa4db5fe494eaa15c0a85189136b | account1    | true    | false     | false  |␣
↪mnc4aa4db5fe494eaa15c0a85189136b |
| 56464d96fc0744efb6d99b9f6e57fe1a | proj2-vpc    |                    |␣
↪efd8a5a9b33b4d9e8e3a8219ec2950fb | tenant2     | true    | false     | true   |␣
↪opd8a5a9b33b4d9e8e3a8219ec2950fb |
| 785dd3b474cc4f48bcd24039914a5197 | proj1-vpc    |                    |␣
↪ghd8a5a9b33b4d9e8e3a8219ec2950fb | tenant3     | true    | false     | true   |␣
↪qrd8a5a9b33b4d9e8e3a8219ec2950fb |
| 90872b1cb68546ed9ec7a2fc30ffecee | vpcproj1     |                    |␣
↪ijc4aa4db5fe494eaa15c0a85189136b | account2    | true    | false     | true   |␣
↪stc4aa4db5fe494eaa15c0a85189136b |
+----------------------------------+--------------+--------------------+----------
```

(continues on next page)

(continued from previous page)

```
↪------------------------+-------------+---------+-----------+--------+------------
↪----------------------+
```

Listing the projects in json format:

```
Symphony > project list -f json
[
{
    "is_domain": false,
    "name": "proj3-vpc",
    "enabled": true,
    "domain_name": "tenant",
    "is_vpc": true,
    "domain_id": "abd8a5a9b33b4d9e8e3a8219ec2950fb",
    "parent_id": "kld8a5a9b33b4d9e8e3a8219ec2950fb",
    "id": "12d782aa3b8543e48110235d0d09ae73",
    "description": ""
},
{
    "is_domain": false,
    "name": "dvsproj1",
    "enabled": true,
    "domain_name": "account1",
    "is_vpc": false,
    "domain_id": "cdc4aa4db5fe494eaa15c0a85189136b",
    "parent_id": "mnc4aa4db5fe494eaa15c0a85189136b",
    "id": "34e07f1cc0f444a0960563ad8d5bee4a",
    "description": ""
},
{
    "is_domain": false,
    "name": "proj2-vpc",
    "enabled": true,
    "domain_name": "tenant2",
    "is_vpc": true,
    "domain_id": "efd8a5a9b33b4d9e8e3a8219ec2950fb",
    "parent_id": "opd8a5a9b33b4d9e8e3a8219ec2950fb",
    "id": "56464d96fc0744efb6d99b9f6e57fe1a",
    "description": ""
},
{
    "is_domain": false,
    "name": "proj1-vpc",
    "enabled": true,
    "domain_name": "tenant3",
    "is_vpc": true,
    "domain_id": "efd8a5a9b33b4d9e8e3a8219ec2950fb",
    "parent_id": "qrd8a5a9b33b4d9e8e3a8219ec2950fb",
    "id": "785dd3b474cc4f48bcd24039914a5197",
    "description": ""
},
{
```

(continues on next page)

(continued from previous page)


```
    "is_domain": false,
    "name": "vpcproj1",
    "enabled": true,
    "domain_name": "account2",
    "is_vpc": true,
    "domain_id": "ghc4aa4db5fe494eaa15c0a85189136b",
    "parent_id": "stc4aa4db5fe494eaa15c0a85189136b",
    "id": "90872b1cb68546ed9ec7a2fc30ffecee",
    "description": ""
}
]
```


3. To exit the console, enter `exit` at the prompt.

For further information on Symp, see Symp CLI Client and Symp CLI Reference Guide in the zCompute User Guide.

43.2 Copying console output

To copy output from the Symp Console:

1. In your Chrome browser, highlight the console output that you want to copy.
2. Press **Enter** to copy the selected output to the clipboard.