VMware vCloud Director for Service Providers

Architecture Overview

TECHNICAL WHITE PAPER

vmware®

Table of Contents

Scope of Document

This white paper is intended to provide an overview of vCloud Director's modules and components. It explains how vCD operates and leverages VMware virtualization to provide a platform that enables service providers to cater IaaS services for its end customers.

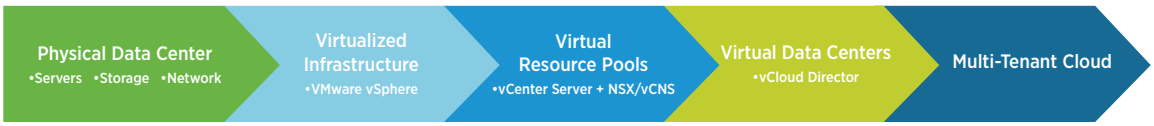
About VMware vCloud Director

VMware vCloud Director is a platform that enables creation of software defined virtual data centers. It enables service providers to take their data centers and convert them into elastic pools of compute resources that can be offered to end customers with various allocation and consumption models. It does this by converting all the physical data center assets, like compute, storage and network, into large pools of virtual resources (via vCenter server and NSX) and partitioning these resources further to provide individual/modular virtual data centers that can be allocated to individual tenants.

vCloud Director leverages VMware vCenter and VMware vSphere to help convert physical compute and storage resources into virtualized resource pools and leverages NSX/vCNS to create corresponding virtual networks.

Apart from providing virtual data centers that can be consumed by multiple tenants for compute and networking, vCloud Director also adds application catalogs and management services. These services enable faster application deployments.

vCloud Director Platform



Platform for Infrastructure Cloud

Architecture Overview

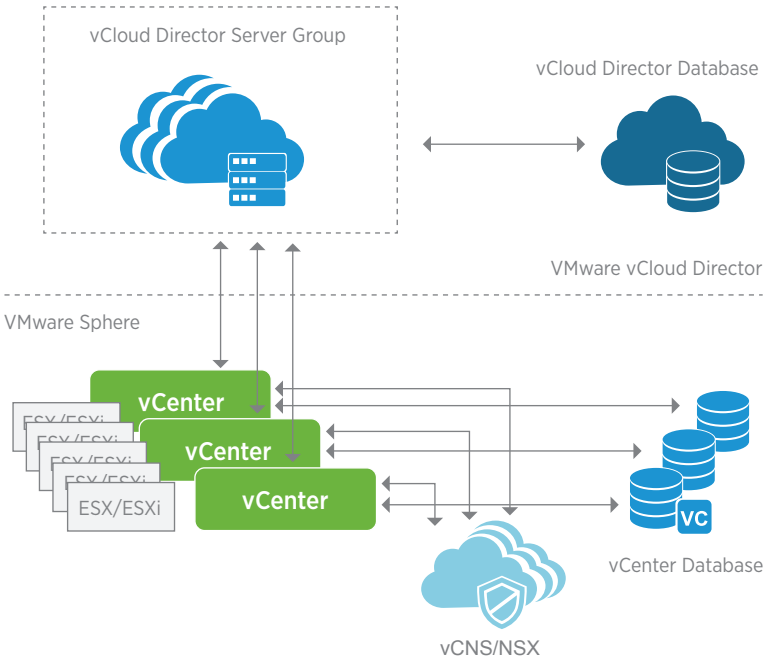
VMware vCloud Director is a robust, multi-tenant infrastructure as service platform designed to consume or accumulate compute, network and storage resources from data centers and in turn convert these resources into individual elastic units that can be provided as a service to and consumed by multiple tenants. The sections below explain how it accumulates resources for use and then how individual tenants can consume these resources.

Data Center resource accumulation by vCloud Director
vCloud Director relies on VMware vSphere and VMware vCenter to provide compute resources and on vCNS/NSX to provide networking resources. It segregates all the resources from vSphere to create a common pool called the Provider vDC. vCD does this by mapping a vCenter cluster to a Provider vDC. All resources available in the cluster get allocated to the Provider vDC. The Provider vDC then creates an abstraction layer from which resources can be derived or further distributed to be allocated as individual units of compute. This individual unit of compute is called the Org vDC. A service provider can then allocate Org vDC's to its end users for consumption.

An Org vDC is derived from a Provider vDC. The Provider vDC maintains this by creating resource pools for each Org vDC in the vCenter. An Org vDC has compute capacity coming from resource pools in vSphere, storage from vSphere data stores and network from vCNS/NSX. Provider vDC allocates resources to Org vDC's based on a predetermined allocation model.

vCloud Director maintains a database of all the resources and inventory from vSphere by periodically synchronizing with vSphere Inventory.

vCloud Director Architecture



An Org vDC is the unit of compute that can be consumed by users of the cloud. An Org vDC is a container for all the virtual machines that are being used in the cloud by a group of users. Virtual machines running within an Org vDC can be clubbed together in virtual appliances. These profiles map to a data store that is made available to the Provider vDC they belong to. Storage profiles have SLA's tied to them. Virtual machines can be allocated to a specific storage profile depending on the SLA needed by them.

A Network is a complete network segment that provides networking services within an Org vDC. Networks are independent segments of virtual networks that connect to Edge gateways and can be routed to external/ public networks.

Enabling Infrastructure Services by vCD

From a resource consumption perspective, it will serve as the compute unit for a set of common business users or service levels. Multiple Org vDC's can be grouped together and belong to a single organization. An enterprise consuming the cloud from a service provider will have one organization that has multiple and each Org vDC can be constructed to have it mapped to specific service profiles such as gold, silver and bronze, or mapped to business groups such as HR, finance or marketing.

Org vDC's are mapped to one or more Org vDC networks. Org vDC networks provide networking services to the virtual machines residing in the Org vDC. They are unique to a specific organization and cannot be shared across organization boundaries. They can be shared by Org vDC's of the same organization. Apart from the Org vDC network, a virtual machine belonging to a virtual appliance can create vApp networks to create further network segment. A vApp network has its gateway associated to an Org vDC Network.

There are three types of networks that a virtual machine or a vApp Network can be connected to:

Isolated Network: a network that is completely isolated and non-routable, suitable for virtual machines that need high security and do not need access to external networks/Internet.

Routed Org vDC Network: virtual machines connected to routed networks can send/receive external network traffic by utilizing Network Address Translation (NAT), and filter traffic by defining firewall rules. Traffic flowing across the routed network can be controlled or limited using firewall, NAT and VPN tunnels.

External Network: This is a direct network connection to the infrastructure vCloud Director is based on and is accessible by all virtual machines in any given organization.

From an administration perspective, service providers can create and can assign tenant admin roles to allocated users for that organization. A tenant admin has admin rights, they can add/remove users, allocate resources and design network services for the organization. Each organization has a unique URL created on top of vCD's base URL. Authorized users can login via their organizations unique URL.

Tenant admins can also enable service catalogs for users of the cloud. These catalogs can either have virtual machine/multi-machine virtual appliance templates or ISO images or files stored in them. Users can take advantage of these templates to provision virtual machines faster.

Constructs of vCloud Director

CONSTRUCT	DESCRIPTION
Organization	An organization is the unit of multi-tenancy that represents a single logical security boundary. An organization contains users, virtual data centers and networks.
Provider Virtual Datacenter	A Provider Virtual Datacenter (VDC) is a grouping of compute and storage resources from a single vCenter Server instance. A provider VDC consists of a pool of physical compute resources and one or more data stores. Multiple organizations can share provider VDC resources.
Organization Virtual Datacenter	An Organization Virtual Datacenter (Org vDC) is a subgrouping of compute and storage resources allocated from a provider VDC and assigned to a single organization. An organization VDC is provisioned resources using vCloud Director resource allocation models. These are represented in vSphere by "resource pools," defined in Table 2.
Resource Allocation Models	Resource allocation models define how resources are provisioned to an organization's VDC from the provider VDC. They also define how resources can be used when deploying virtual applications (vApps) within the organization VDC.
vApp	A vApp is a container for a distributed software solution and is the standard unit of deployment in vCloud Director. A vCloud Director vApp is very different from a vSphere vApp in the manner it is instantiated and consumed in vCloud Director. It enables power operations to be defined and specifically ordered. It consists of one or more virtual machines and can be imported or exported as an OVF package. A vCloud vApp can have additional vCloud-specific constructs such as vApp networks.
Org Networks	An Org Network provides networking services to virtual machines or virtual appliances deployed inside of an Org vDC Network confined to organization boundaries.
Tenant Admin	Tenant Admins have admin rights within an Organization. They can create/import tenant users, add/modify catalogs, etc.
Tenant User	The consumers of the cloud, Tenant Users can add/delete virtual machines/virtual appliances, attach org networks to virtual machines, take/delete snapshots, etc.

Cloud Characteristics

The fundamental characteristics of a cloud platform are designed into vCloud Director. It is developed to optimize better resource consumption, provide modular services that can be extended to multiple customers and yet maintain isolation between resources of each customer on the cloud. Listed below are a few of vCD's characteristics that make this happen.

Elastic Resource Pooling

The fundamental principal that vCD operates to create cloud services is by pooling together all the resources of a data center (Provider vDC) and then allocating them to organizations/customers as needed. Because of this abstraction layer, vCD is able to source resources from its pool for customers when needed and put back resources in the pool when they are done. At the same time, if more hosts/network/storage is added to a cluster, Provider vDC picks up the additional resources and expands its pool. This model helps service providers to better manage and operate their cloud infrastructure as well as provide a strong service platform to its customers.

How it's Done (Allocation Models)
Allocation of resources to customers or vDC's from the provider vDC is governed by an 'Allocation Model'. vCD has three types of models by which it allocates resources to Org vDC's. An Org vDC essentially maps to a resource pool within vSphere, and all the three models have policies that vary in terms of reservations set on the underlying resource pools and the upper limit on the resources that an Org vDC can consume. When a virtual machine is created by a customer in the Org vDC its resource allocation policies are derived from the corresponding resource pool the VM is placed in. Defined below are the three different allocation models in vCloud Director.

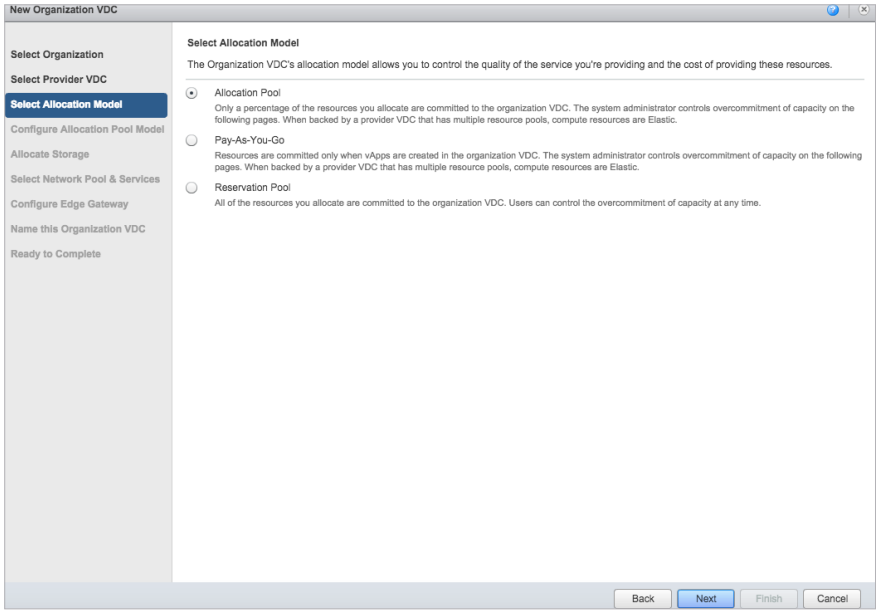
ALLOCATION MODELS	RESOURCE POOL SETTING	VIRTUAL MACHINE SETTING
ALLOCATION POOL	A % of Resource are Guaranteed and a Max Resource Limit is set on the Resource Pool	Resource Guarantee and Limit is inherited from Resource Pool
PAY-AS-YOU-GO	No Resource Guarantee or Limit set at the Reservation Pool	Resource Limits set at the virtual machine level
RESERVATION POOL	Resource Guaranteed and Resource Limits are equal, all the resources are dedicated	No resource setting is defined at the virtual machine, however a user can change limits and reservation per VM.

Because allocation models are defined by limit and resource allocation, they become the basis on which a service provider can create service definitions for its customers. Customers can choose how they consume resources based on their needs. For example, a customer who needs a fixed set of resources can always go with an Org vDC that has a predefined limit set, or the allocation model. A customer can go for pay-as-you-go model when they do not have data on how much of the resources they will consume in the cloud.

The different allocation models help service providers cater service for its customers based on their needs yet the service providers do not need to keep separate provider resources to cater to their business needs. This helps service providers to better manage their data centers.

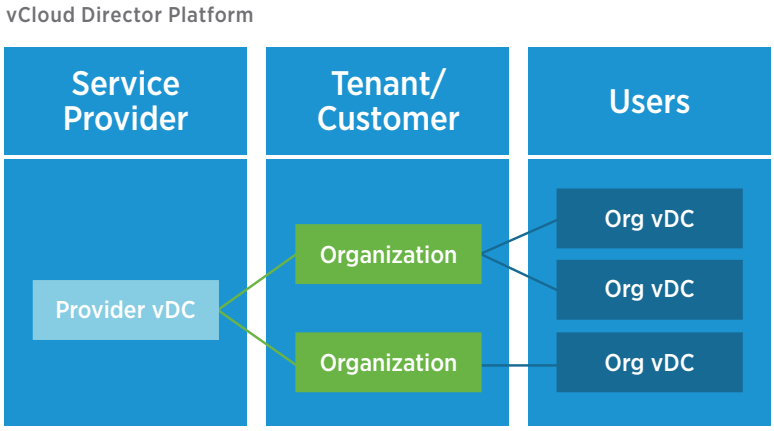
Key Benefits

- **Optimized Data Center Management:** Elastic resource pooling allows service providers to better manage their data centers. The Provider vDC creates an abstraction layer between resource and consumption, helping service providers to add/delete/manage resources without disrupting services.
- **Lower Capex Costs:** Consolidation of data center resources and allocation pools enable over-subscription of resources, which helps service providers avoid oversizing data centers and reduce Capital Expenditure costs.
- **Scale of Resources:** Because of the pooled abstract layer, more resources can always be pumped to scale Provider vDCs. Since the Provider vDC synchronizes with vSphere inventory, additional resources added are automatically picked up by the Provider vDC.



Multi-Tenancy

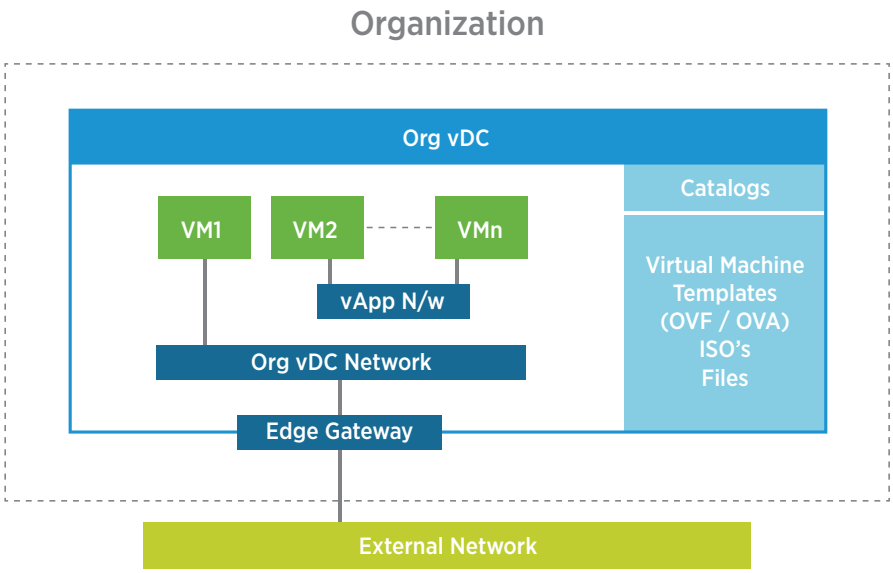
Multi-Tenancy is one of the intrinsic characteristic of an IaaS cloud. vCloud Director has specific modules and constructs built around this basic characteristic. An organization in vCD represents the container in which all of a tenant's resources reside. vCD enables service providers to create isolated compute, storage and network containers that can be mapped to individual tenants of the cloud. vCD does this by 'slicing' resources from the Provider vDC into individual Org vDCs and mapping one or more Org vDCs to organizations.



How it's Done (Organizations)

An organization is the container for a tenant and forms logical boundaries between tenants. Each organization gets units of resources defined by the Org vDCs it has. The resources that get defined at the Org vDC are compute, storage and network.

When a Org vDC is created, a Provider vDC is used to allocate resources to the Org vDC. The Provider vDC maps these resources into different containers in vSphere, which isolates the units of compute. For example, the compute resources CPU and memory are defined by creating a resource pool in vSphere. Storage is mapped by assigning storage profiles to vDC. These storage profiles in turn map to data stores in vSphere. Org Networks created are derived from network pools and map to port groups on the distributed vSwitch that is part of the provider cluster.



Key Benefits

- **Uniform and Standard Service Definitions:** Service providers get a single pool of consolidated resource and offer services to tenants by deriving resources from this pool. Services definitions created once can be applied to multiple tenants. Standard methods of deriving and offering services help maintain uniform environments across tenants.
- **Lower Operational Costs:** Standard and predefined ways of applying services to tenants help reduce the time needed to maintain or create tenant environments. Rather than maintaining manual individual environments for customers, vCD provides defined methods to operating isolated environments for all tenants using the same infrastructure resources.

Customer Self Services

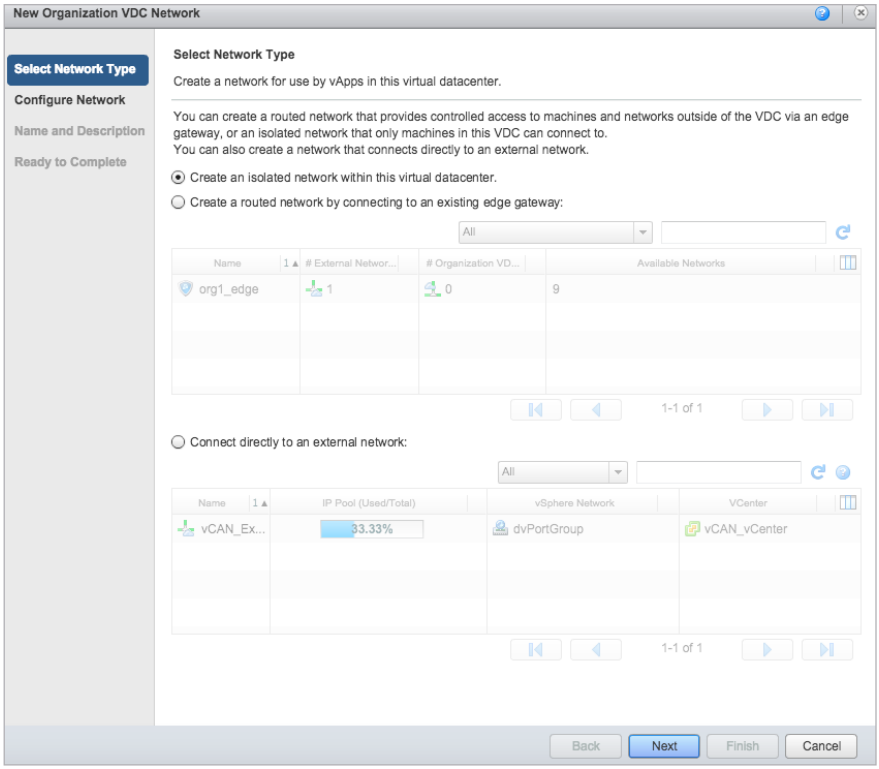
vCloud Director offers a model that helps service providers delegate some of the day-to-day IT operations to its customers. This gives customers more flexibility and control over their cloud environments. A few of the tasks that a customer can do while running the cloud are:

Network Configurations

A tenant user who has appropriate access rights has the ability to create isolated or routed organization networks. These networks can be further assigned to virtual machines. This helps use cases where a tenant is instantiating a new application in the cloud and wants to create a new network designed or crafted for that

specific application. The user does not need to go back to the service provider in order to request more networks, making application provisioning faster.

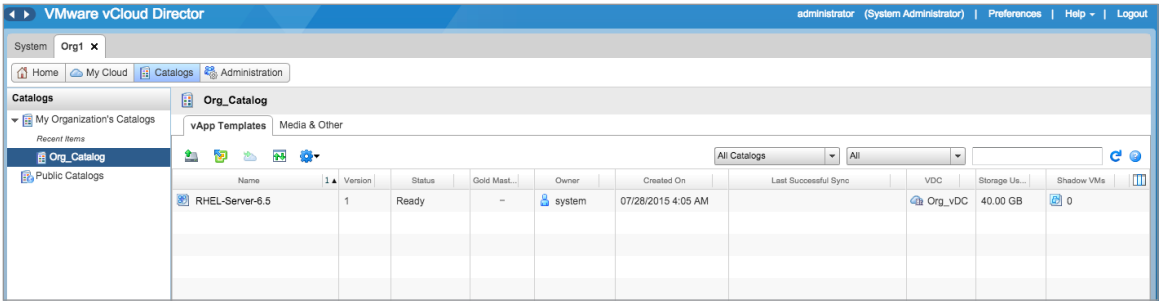
At the same time, if an organization network that is already deployed and is attached to different virtual machines or vApp Networks needs to be modified, a tenant user with appropriate access rights can do this without the need to go back to the service provider.



Application Catalogs

Each Org vDC has a store for publishing virtual machine or virtual appliance templates. These templates can be OVF or OVA files that help users deploy their virtual instances or applications in the cloud faster. Apart from OVA or OVF Files, ISO images can also be uploaded to the catalogs. Catalogs can be used to store multiple file types that users can later use to mount on their virtual machines. Authorized tenant users can upload and maintain these catalogs without the need to have a service provider update/publish catalogs.

In addition, service providers can publish images to public catalogs that are visible to all tenants of the cloud. Public catalogs give service providers the opportunity to easily add additional application service capabilities for their tenants.



Key Benefits

- **Faster Provisioning in the Cloud:** Users can use pre-built templates to provision virtual machines, reducing the time to install and configure operating systems and applications. Guest customization and scripts help pass configuration parameters within the guest operating system, providing users with granular control of their virtual machines.
- **Model to add Application Services:** Service providers can leverage public catalogs and metadata to create additional add-on XaaS offerings to their tenants.

Monitoring and Automation

Service providers using vCloud Director can leverage the following VMware products for monitoring and automation.

Monitor and Analyze Cloud Infrastructure in Real Time

VMware vRealize Operations Manager and VMware vRealize Log Insight offer a "single pane of glass" view to monitor the state of their infrastructure. It can monitor utilization, give performance reports and run analytics. vRealize Operations connects to vSphere environments via vCenter Server and provides hierarchical information on all the constructs in the data center, from vCenter servers and ESXi hosts to virtual machines and their storage and network tiers.

Service providers can for example:

- Monitor the collective health of resource pools mapped to tenants
- Determine when to add additional resources to scale Provider vDC
- Set alerts for when resources are depleting and reach a certain threshold
- Get a bird's-eye view of all the inventory in their infrastructure

vRealize Log Insight collects application and system logs via Syslog and provides analytics capabilities within a visual dashboard. Log Insight imports logs from systems, applications and operating systems via SMTP or web. It then parses and indexes these logs.

Common queries can be run on these indexed logs to determine patterns that can provide insight into system behavior and state. Logs can help capture alerts that are difficult to be captured via common operations monitoring. It can help capture issues that are missed by operations alerts because they are between application subsystems and interdependent. Indexed set of logs can also help run post incident/root-cause analysis.

Dashboards in Log Insight are based on the queries run against logs. A dashboard starts filling for each pattern match that was defined in the query.

Service providers can generate dashboards for various use cases, including:

- Unsuccessful login attempts by tenant/system users
- Database connectivity alerts
- The number and source of rejected connections on an Edge firewall

Automation Using vRealize Orchestrator Plugin for vCloud Director

vRealize Orchestrator (vRO) helps define workflows that can be used to automate common tasks that are run against vCD. vRO has a plugin developed for vCloud Director. This plugin helps vRO understand the user, admin and extension API classes for vCD and also provides a communication channel to talk to vCD. The plugin has out-of-the-box workflows already defined and can help develop custom workflows as well. Service providers can use vRO out-of-the-box workflows or script tasks to do common tasks such as:

- List all enabled/disabled organizations
- Add/edit Provider vDC, organizations and Org vDCs
- Configure networking services on Org vDC networks

vCloud Director API and Extensibility

vCloud Director has an extensive set of REST full API's that can be used to perform a set of operations on the cloud programmatically. These API's are accessible via REST Clients over HTTP. Each object and component in vCloud Director has a corresponding unique ID by which it can be a referenced. The state and properties of each component is stored in the form of XML Elements.

vCloud Director API's can be used to run operations on your cloud without going through the UI, they can be plugged to Automation Frameworks that execute operations on the Cloud. For e.g, when a new tenant is registered, a REST API call to create Organization in vCD can be plugged into an invoice or Billing system. . vCD REST API queries can get help generate infrastructure lists, like, listing all the Organizations belonging to a Provider vDC or all Catalogs in an Organization etc.

SDK's written in Java, .NET and PHP are available to work with vCD's REST API's. SDK's understand vCD's API Reference Schema, how to parse REST calls to create, retrieve, update, delete operations, they make help make working with REST calls easier.

Please refer to the *vCloud Director API Programming Guide* and *vCloud Director SDK for Java /.NET /PHP Developers guide* for further information.

vCloud Director Extension Services

vCloud Director has an Extension API Framework that helps plug additional add-on services to the core set of API's. The API extensibility frameworks helps Service Providers to add their own services by leveraging Infrastructure services from vCloud Director. vCD API also offers vSphere Platform Extension services to manage or retrieve information of the underlying vSphere environments, these API calls can help determine the current inventory of vSphere resources currently being used by the Provider vDC or make changes to vSphere Environments. For e.g, vCloud Extension API's can be used to List External Networks connected to the Provider vDC, Disable/Enable DataStores , update/modify vCenter Server settings etc.

Extension API service allows Service Providers to configure Blocking Tasks, these tasks define events or operations that will not be executed until a Service Provider administrator approves them or a pre-set timer expires. Service Provider can take advantage of Blocking tasks to define operations performed by tenant users that need additional approval or workflows before they can be allowed or executed. A Blocking task can be configured for Creating a new vApp or adding additional Org Networks etc.

Refe the section *Configuring and Using Blocking Tasks* in the *vCloud API Programming Guide* for additional details.

The extensibility framework allows Service Provider to customize and define services beyond what vCloud Director has to offer out of box.

vmware

VMware, Inc. 3401 Hillview Avenue Palo Alto CA 94304 USA Tel 877-486-9273 Fax 650-427-5001 www.vmware.com

08/15