

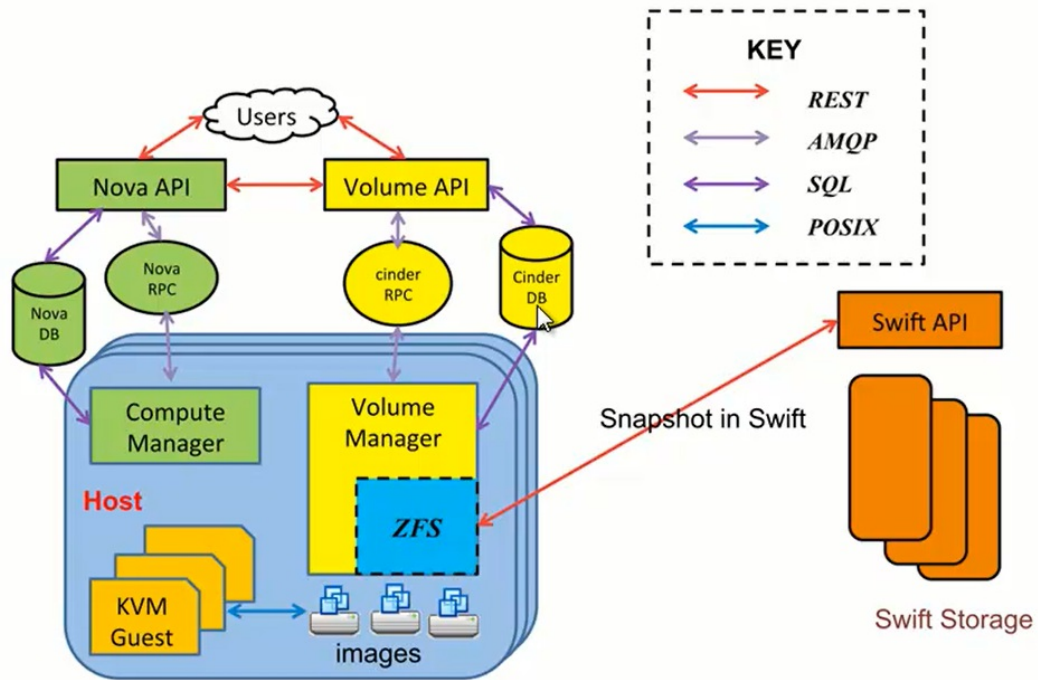
关于 Cinder

- ▶ F版本以前，没有Cinder，对应的组件为nova-volume
- ▶ 提供了Rest API
- ▶ Cinder的目标
 - 减少Nova的复杂性，降低Nova的负载
 - 支持多种后端存储
 - 增加和其他组件之间的交互
- ▶ 参与Cinder项目的公司与组织
 - Rackspace
 - HP
 - Ceph
 - NetApp
 - Citrix
 - ...

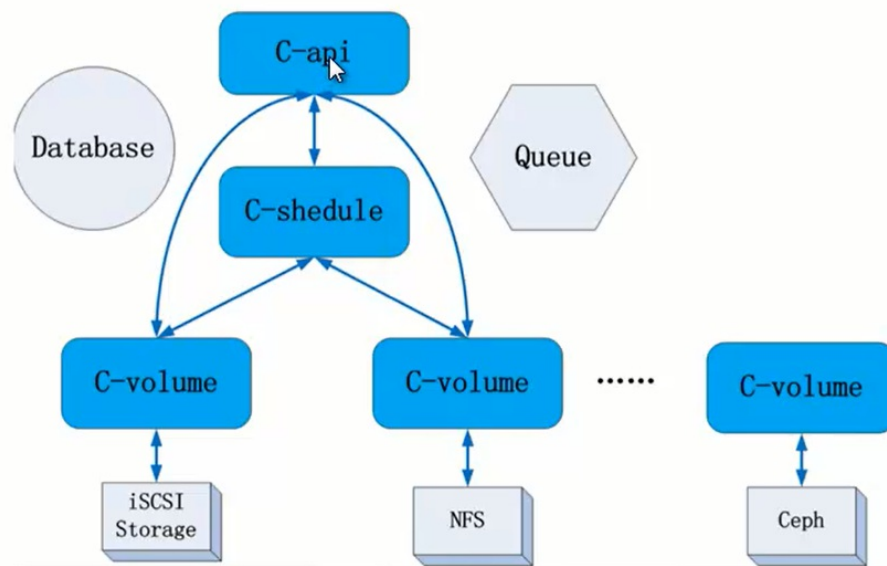
OpenStack & Storage

	Cinder / Block Storage	Swift / Object Storage
Objectives	<ul style="list-style-type: none"> • Storage for running VM disk volumes on a host • Ideal for performance sensitive apps • Enables Amazon EBS-like service 	<ul style="list-style-type: none"> • Ideal for cost effective, scale-out storage • Fully distributed, API-accessible • Well suited for backup, archiving, data retention • Enables Dropbox-like service
Use Cases	<ul style="list-style-type: none"> • Production Applications • Traditional IT Systems • Database Driven Apps • Messaging / Collaboration • Dev / Test Systems 	<ul style="list-style-type: none"> • VM Templates • ISO Images • Disk Volume Snapshots • Backup / Archive • Image / Video Repository
Workloads	<ul style="list-style-type: none"> • High Change Content • Smaller, Random R/W • Higher / "Bursty" IO 	<ul style="list-style-type: none"> • Typically More Static Content • Larger, Sequential R/W • Lower IOPS

Cinder 的位置



Cinder的架构



很明显volume-manage是对底层不同存储类型的屏蔽和封装

Cinder各组件的功能

- ▶ API service:
 - 负责接受和处理Rest请求，并将请求放入RabbitMQ/Qpid队列。
- ▶ Scheduler service
 - 处理任务队列的任务，并根据预定策略选择合适的Volume Service节点来执行任务。
- ▶ Volume service:
 - 该服务运行在存储节点上，管理存储空间

Cinder API

- ▶ Volume create/delete/list/show
- ▶ Create from image, snapshot, volume
- ▶ Snapshot create/delete/list/show
- ▶ Backup create/restore/delete/list/show
- ▶ Volume attach/detach (called by Nova)
- ▶ Volume types
- ▶ Quotas

备份是全拷贝，快照是什么呢？

Volume Types

- ▶ Admins can create tiers of storage
 - Specify requirements from storage
- ▶ Users can specify a tier when creating a volume

Create Volume

Volume Name
web-volume

Description
Additional information here...

Type
tier1
tier2
tier3

Description:
Volumes are block devices that can be attached to instances.

Volume Quotas
Total Gigabytes (350 GB) 650 GB Available
Number of Volumes (3) 7 Available

Cancel Create Volume

volume types, 举例比如后端不同类型的存储类型（比如后端一个是ceph另一个是iscsi类型），这样可以新建两个不同的类型volumetypes。这只是其中一种举例，可以指定其他方面。

Cinder Scheduler

- ▶ Chooses back-end to place a new volume on
- ▶ Configurable plugins for scheduler
 - Simple
 - Chance
 - Filter
- ▶ Most common is the filter scheduler
- ▶ Has plug-able filters & weights

Filter Scheduler: Example Flow

- ▶ Drivers continuously report capabilities and state
- ▶ Scheduler starts with list of all back-ends
- ▶ Filters out unsuitable back-ends
 - Insufficient free space
 - Insufficient capabilities
- ▶ Sorts according to weights (e.g., available space)
- ▶ Returns best candidate

Cinder Volume

- ▶ Manager contains generic code
 - e.g., High-level flow, DB & quota updates
- ▶ Drivers contain back-end specific code
 - Linux LVM
 - Storage controllers from various vendors
 - Distributed file systems
- ▶ Admin can run multiple cinder-volume instances
 - Each able to manage multiple back-ends
- ▶ Each back-end is generally configured to interact with one storage pool
- ▶ Multi-threading

Example Flow: Attach Volume

- ▶ Nova calls Cinder via its API, passing connection information
 - e.g., host name, iSCSI initiator name, FC WWPNS
- ▶ cinder-api passes message to cinder-volume
- ▶ Manager does initial error checking and calls volume driver
- ▶ Volume driver does any necessary preparation to allow the connection
 - e.g., give the nova host permissions to access the volume
- ▶ Volume driver returns connection information, which is passed to Nova
 - e.g., iSCSI iqN and portal, FC WWPNS
- ▶ Nova creates the connection to the storage using the returned information
- ▶ Nova passes the volume device/file to the hypervisor

Cinder 存储解决方案

- ▶ 使用本地存储，即Cinder LFS
 - 直接使用逻辑卷
- ▶ 使用动态存储系统，即Cinder ZFS
 - 这种解决方案提供了多种存储特性，如：
 - Cache
 - Copy on Write
 - Snapshot & Clone
 - Changeable block size
- ▶ 使用分布式存储系统，即Cinder DFS
 - 常见的分布式存储系统有：
 - Ceph
 - Sheepdog
 - Glusterfs

Cinder本身没有做存储的实现，只是提供driver接口封装，都是第三方来支持的。第一种是直接使用本地存储节点（装了volume服务节点）的磁盘。

常用的Backend Storage产品

- ▶ Netapp
- ▶ EMC
- ▶ Huawei
- ▶ Ceph
 - 基于Linux
 - 可扩展
 - 高负载
 - 高可靠性
- ▶ Sheepdog
 - 专为QEMU/KVM设计的分布式存储系统
 - 可扩展

解决方案的选择

- ▶ LFS, ZFS还是DFS?
 - 根据云的规模考虑存储方案
 - 根据预算考虑存储方案
 - 公有云的特征
 - 私有云的特征

私有云块存储需求

- ▶ 集中存储
- ▶ 高性能、高可靠
- ▶ 在预算范围内实现性能最大化, 同时保障数据可靠
- ▶ 无可挑剔的用户体验

公有云块存储需求

- ▶ 对存储进行分级
- ▶ 异构的存储后端
- ▶ 高容量、高可用、成本可控
- ▶ QoS保障
- ▶ 良好的存储网络设计

公有云对存储进行分级，不一定所有的客户都需要高性能的存储。因为分级了所以后端肯定异构存储类型（NFS、iscsi、Ceph等）。QoS带宽肯定需要控制