學校私有雲之架設實務分享

李忠憲

大綱

1. 甚麼是私有雲？
2. 建置私有雲要如何規劃？
3. OpenStack 軟體功能介紹
4. 各種 OpenStack 硬體配置方案
5. 安裝流程
6. 虛擬機映像檔之取得
7. 控制面板操作介面
8. 如何維護節點？

從虛擬化到私有雲

- 虛擬化 - 在一臺實體機器上執行多臺虛擬機，為保護虛擬機映像檔的安全，通常會使用獨立執行的NAS或磁碟陣列作為儲存空間，配合異地備份，可以達到更高的安全性。
- 很多人以為做到上面的機制就算是一朵「雲」，但私有雲其實是一種架構，而非一項軟體技術。首先私有雲是由多個節點組成，節點依照任務來劃分，包含：控制節點、計算節點（跑虛擬機的實體機器）、網路節點、儲存節點（BigData）、認證節點...等。
- 以上的各種節點可以隨意組合，在實驗階段可以通通裝在同一臺電腦上（單一節點），但這樣就不是雲了，要稱為雲至少要三個節點以上。
- 透過雲計算，可以達到虛擬機的動態配置、動態遷移、即時修復。同時可以把多臺虛擬機叢集為一臺機器提供平行運算的服務。

雲的分類

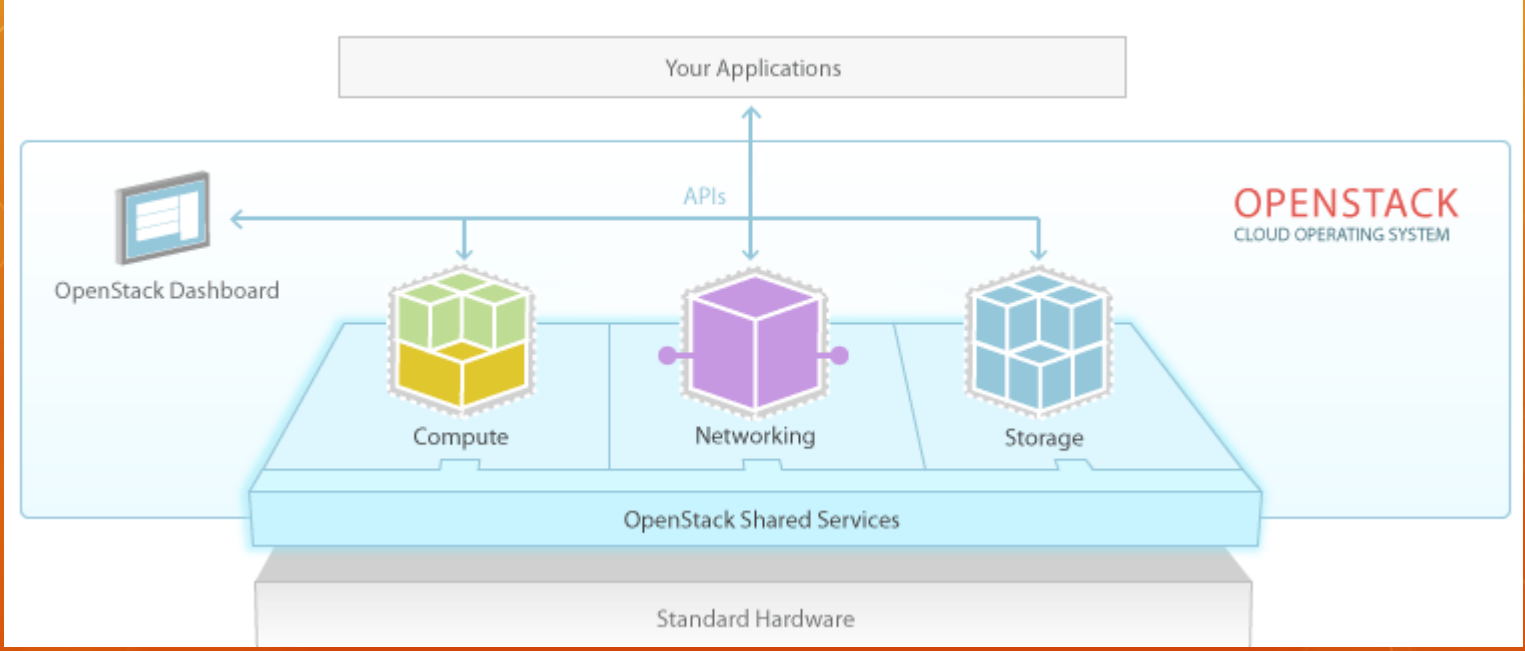
- 以設備歸屬的角度來區分 - 公共雲、私有雲、混合雲
- 以系統建置的角度區分 - 實體雲、虛擬雲（向別人租用雲端平臺）
- 以產業分工來區分：
  - 雲端設備（IaaS） - 基礎設備與雲端建置
  - 雲端平臺（PaaS） - 提供虛擬主機和程式開發平台。例如：台大觔斗雲
  - 雲端服務（SaaS） - 提供網際網路應用程式。例如：優學網

現有的IaaS技術

- 公共雲：
  - Amazon EC2
  - Google 雲端運算
- 私有雲：
  - Eucalyptus（開源Amazon雲端運算技術）
  - Hadoop（開源Google雲端運算技術）
  - OpenStack（開源軟體，內建於各種Linux作業系統中，例如：ubuntu、centos、redhat、fedora……等）
  - Rackspace Cloud（OpenStack的開發廠商之一，基於OpenStack）
  - Windows System Center（$$）
  - Vmware vCloud（$$）
  - Xen Cloud Platform（$$）

本次研習將介紹OpenStack解決方案

OpenStack 概念圖



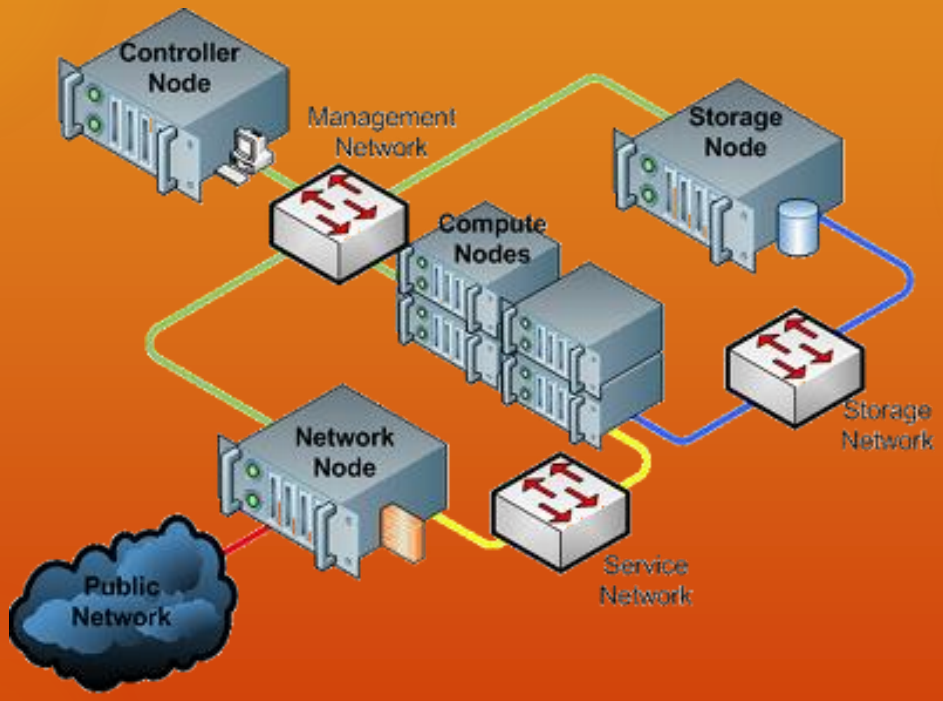
規劃私有雲要考慮甚麼？ Part 1

- 有幾臺伺服器需要虛擬化？各需要多少顆CPU？多少RAM？
  - 伺服器通常都需要2顆CPU才順暢，工作站則只需1顆
  - 不同版本作業系統需要的RAM也不同
- 我的硬體夠用嗎？
  - 基本要求：CPU需支援虛擬指令集（1顆CPU可虛擬成8顆CPU，也就是說一個計算節點最多只能跑8臺虛擬機）
  - 計算節點至少需要32GB RAM（需保留4GB給雲端系統使用，其餘可用來跑虛擬機）。
  - 需要幾臺實體機？採購新主機還是要加RAM？
- 虛擬機映像檔要儲存在哪裡？
  - 專用 Storge：NAS、外掛磁碟陣列、網路磁碟陣列
  - 多臺實體機組成的儲存節點
  - 有幾臺實體機可以作為儲存節點？儲存節點應包含：storage proxy一臺、storage container兩臺以上

規劃私有雲要考慮甚麼？ Part 2

- 採用哪一種網路架構？
  - 共享模式：虛擬機透過 NAT 技術與實體機共享 IP（對外提供服務時需要 port forwarding， throughput 不好）
  - 橋接模式：虛擬機透過路由技術取得學校內部 IP（可使用校內DHCP配發 IP 做到遷移免設定，throughput 稍好）
  - 交換器模式：虛擬機透過第二片網卡跑 VLAN，由網路節點自動配置 IP（ throughput 最好，但每台實體機都需要兩片網卡）
  - 需要一臺有三張網卡的實體機做為網路節點，需要一臺已經切割好VLAN的交換器
- 還有多餘的機器可用？
  - 可考慮建置兩套控制節點及網路節點（透過 hearbeat 做即時監控和服務轉移）
  - 這部分功能 OpenStack 未提供，需仰賴第三方插件或自行開發

OpenStack 網路架構圖



控制節點各種軟體功能描述

- 同一個私有雲的所有服務，都必須透過 keystone 伺服器認證取得授權碼（token）才能提供服務。
- 所有服務所需要的資料表，統一建置於 MySQL 中
- 所有節點透過 RabbitMQ 來互相溝通
- Glance服務用來管理虛擬機映像檔，將映像檔註冊於資料庫中，並進行資源分配
- Quantum服務用來管理網路組態，於資料庫中建立新的網路組態，並指配給虛擬機
- Nova Service 函式庫（包含：api、cert、consoleauth、schduler、novncproxy），用來控制計算節點以便配置虛擬機
- Cinder服務用來提供磁碟掛載服務（透過 iscsi initiator 和 isci target）
- Horizon服務提供網頁管理介面

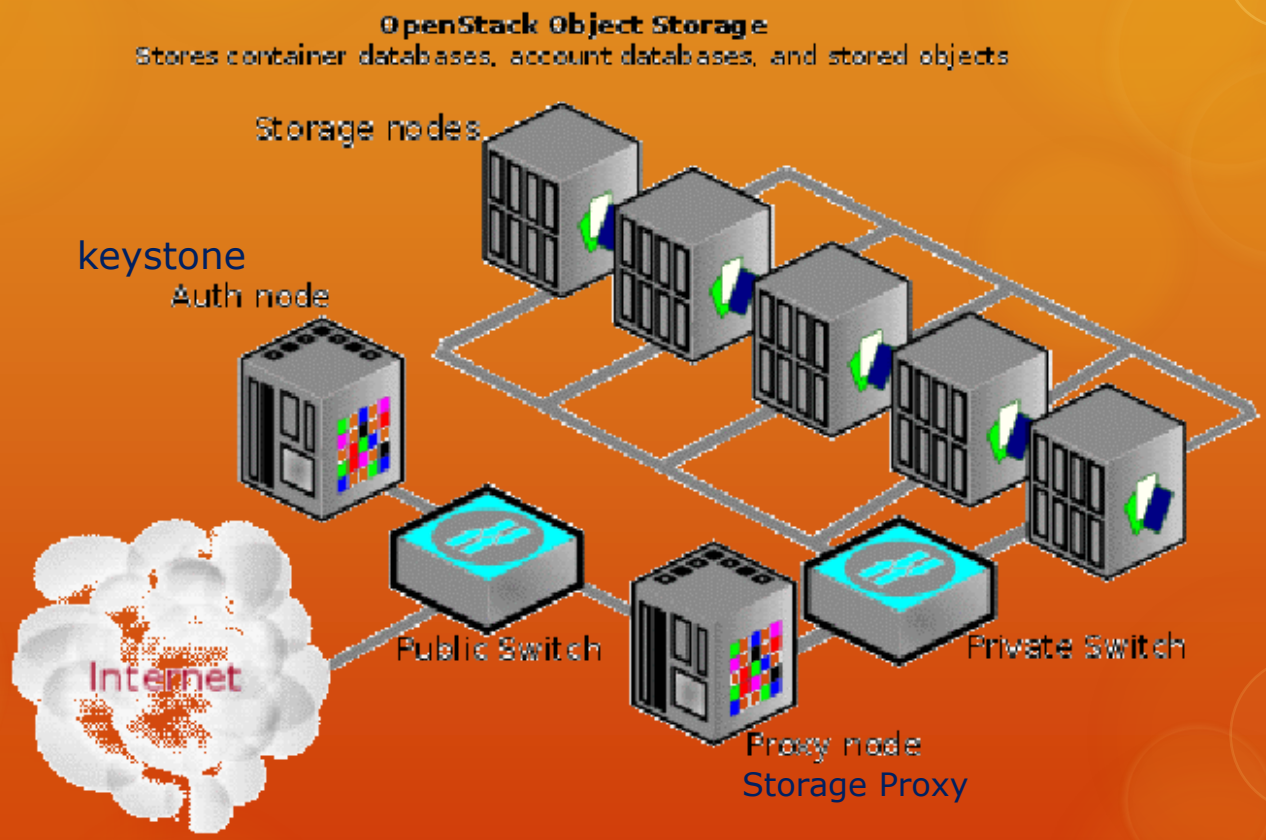
計算節點各種軟體功能描述

- 透過 KVM 進行虛擬化
- OpenVSwitch把網卡模擬成交換器
- Quantum-Agent接受Quantum服務的指揮，進行封包繞送
- Nova-Compute 服務用來管理虛擬機

網路節點各種軟體功能描述

- OpenVSwitch把網卡模擬成交換器
- Quantum-Agent接受Quantum服務的指揮，進行封包繞送

儲存節點架構圖


OpenStack Object Storage
Stores container databases, account databases, and stored objects
Storage nodes
keystone
Auth node
Internet
Public Switch
Private Switch
Proxy node
Storage Proxy

儲存節點各種軟體功能描述

- Swift-Proxy提供儲存服務的統一窗口，接受存取要求後，進行資料蒐集或散佈儲存。

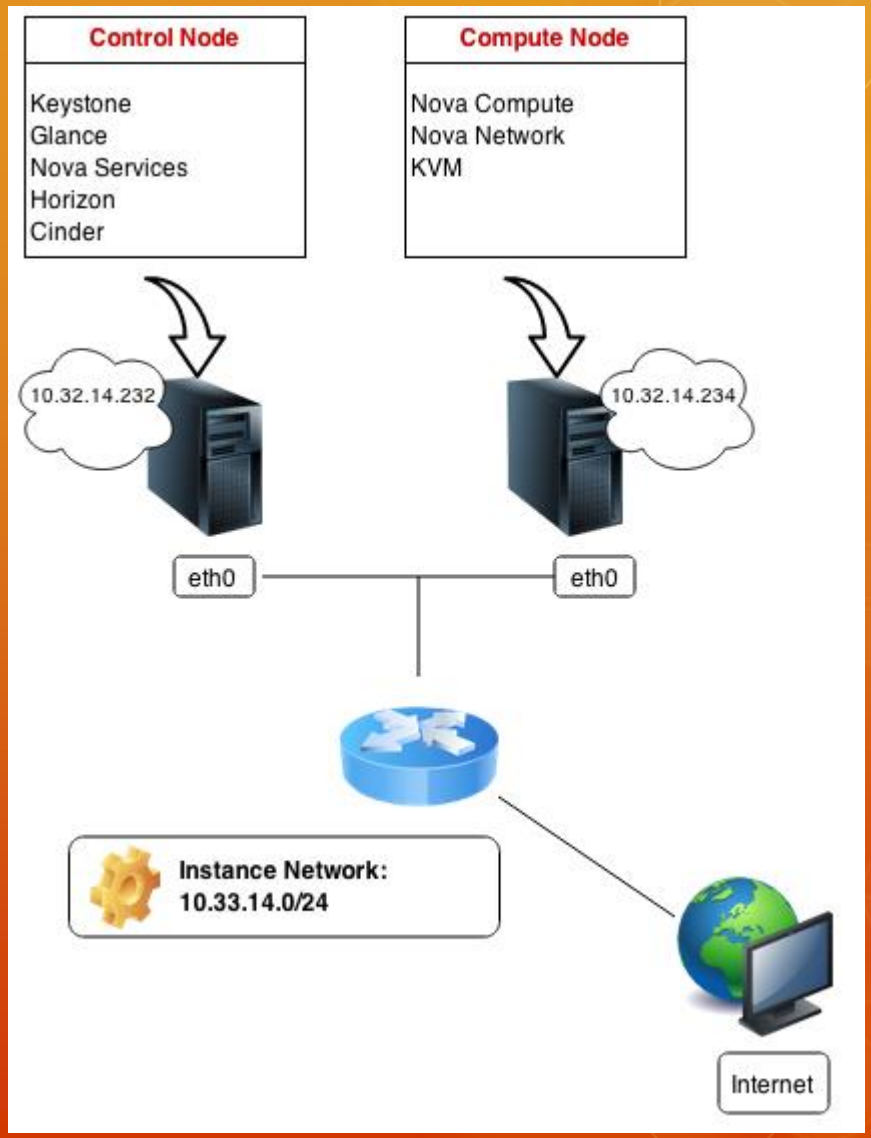
- Swift-Account用來管控連線與檔案存取權限。
- Swift-Container用來管理儲存容器，儲存容器可視為一個 Block Device（磁碟區塊）
- Swift-Object用來管理檔案

使用 NAS 而不使用 Storage Node

- 將 NAS 切割出來的磁區透過 iscsi initiator 掛載到控制節點上，並格式化為 lvm2
- 在控制面板（Dashboard）上建立新的儲存空間時，會從 lvm2 磁區切割出一塊指定大小的分割區，並由控制節點透過 iscsi target 提供給虛擬機掛載
- 也就是控制節點必須同時跑 iscsi initiator（client） 和 isci target （Server）

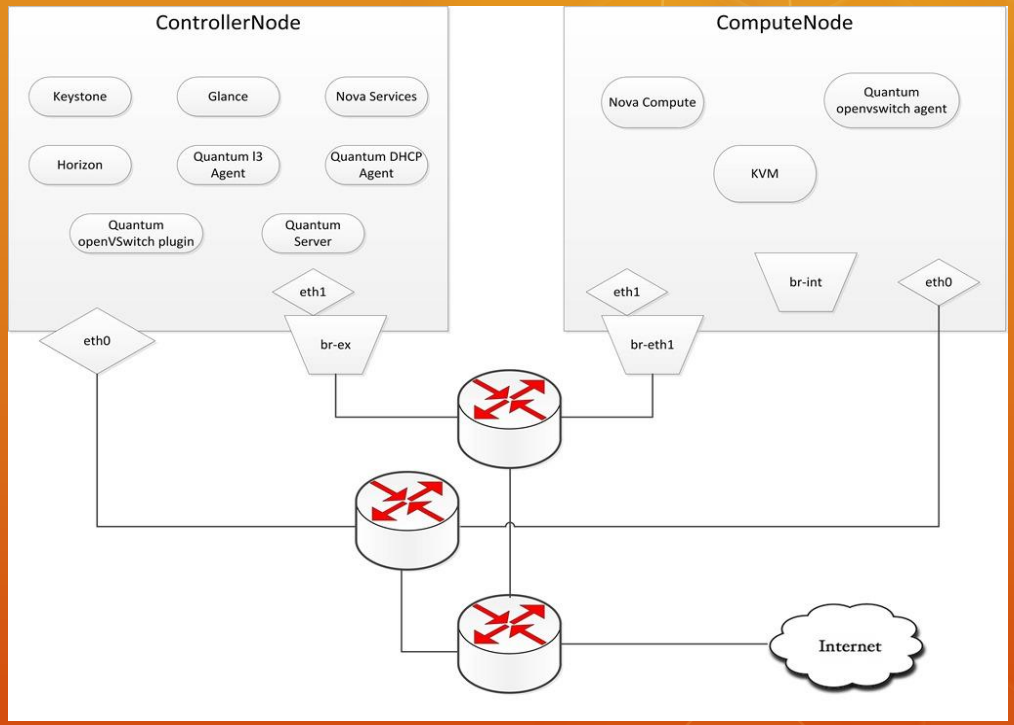
基礎配置

- 每台實體機都只有一片網卡
- 沒有額外的網路交換器可以使用
- 不使用網路節點



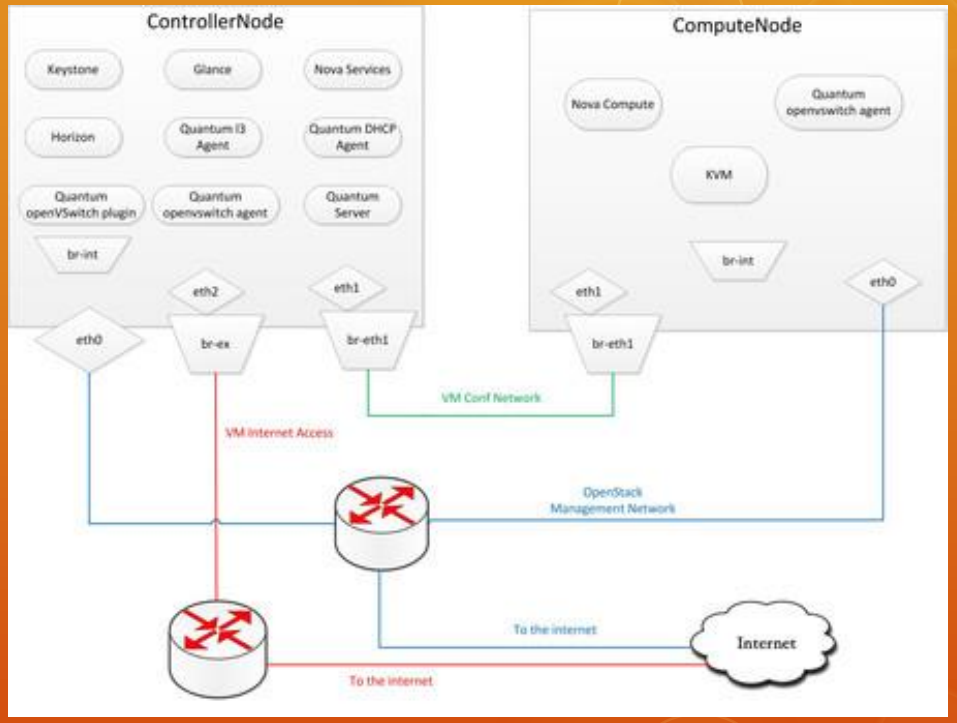
進階配置一

- 每台實體機都有兩片網卡
- 使用一臺已切好3個vlan的網路交換器
- 網路節點直接裝在控制節點上



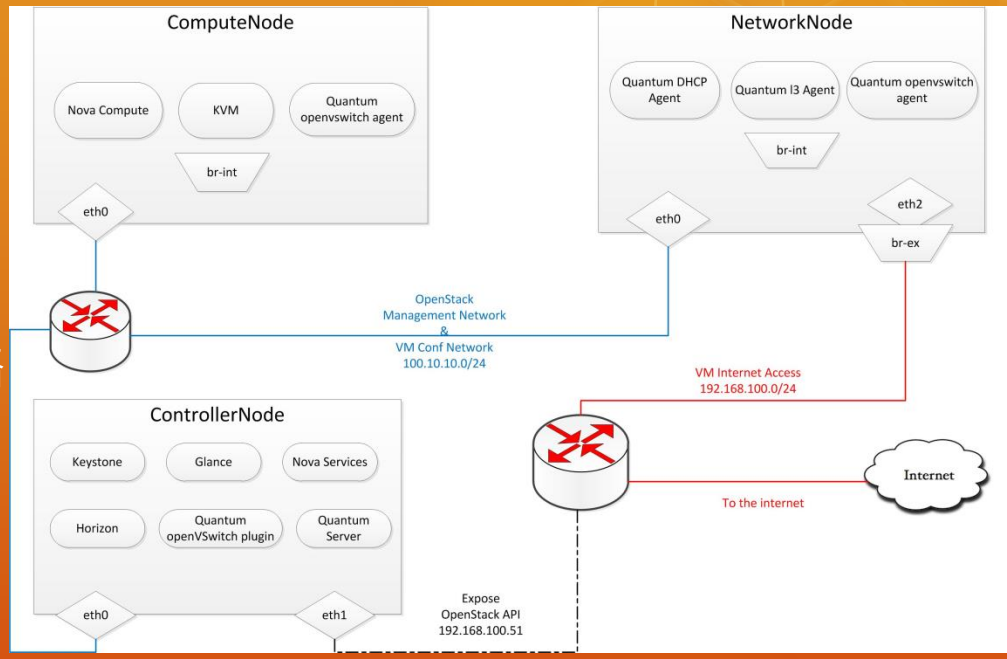
進階配置二

- 每台計算節點實體機都有兩片網卡
- 一台三片網卡的實體機做為控制節點
- 使用一台切割3個Vlan的網路交換器
- 網路節點直接安裝在控制節點上



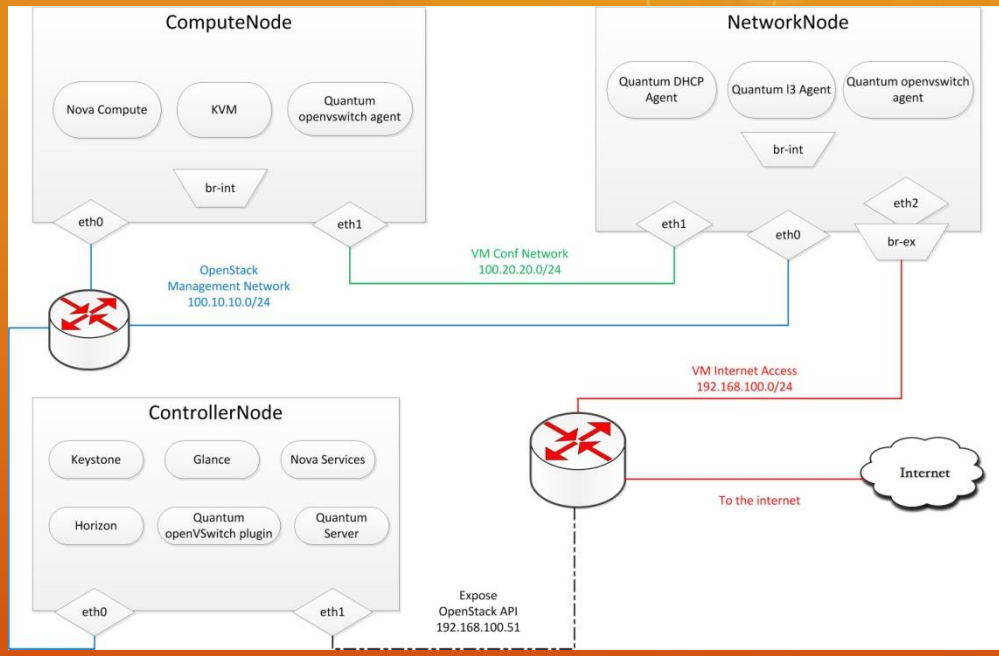
完整配置一

- 每台計算節點實體機都只有一片網卡
- 兩台擁有兩片網卡的實體機做為控制節點和網路節點
- 使用一台切割3個Vlan的網路交換器

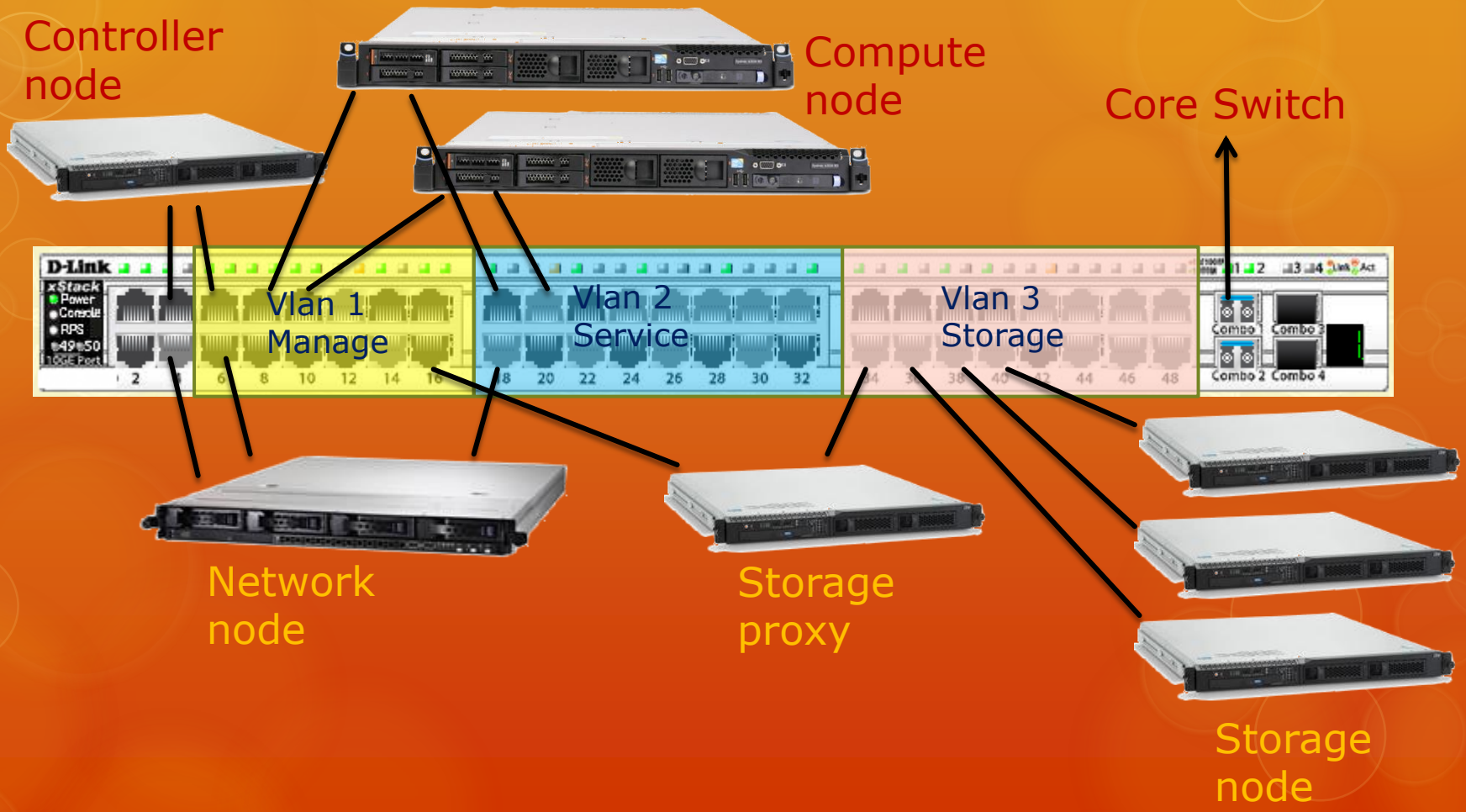


完整配置二

- 每台實體機都有兩片網卡
- 一台擁有三片網卡的實體機做為網路節點
- 使用一台切割3個Vlan的網路交換器



國語實小的方案


Controller node
Compute node
Core Switch
Vlan 1
Manage
Vlan 2
Service
Vlan 3
Storage
Network node
Storage proxy
Storage node

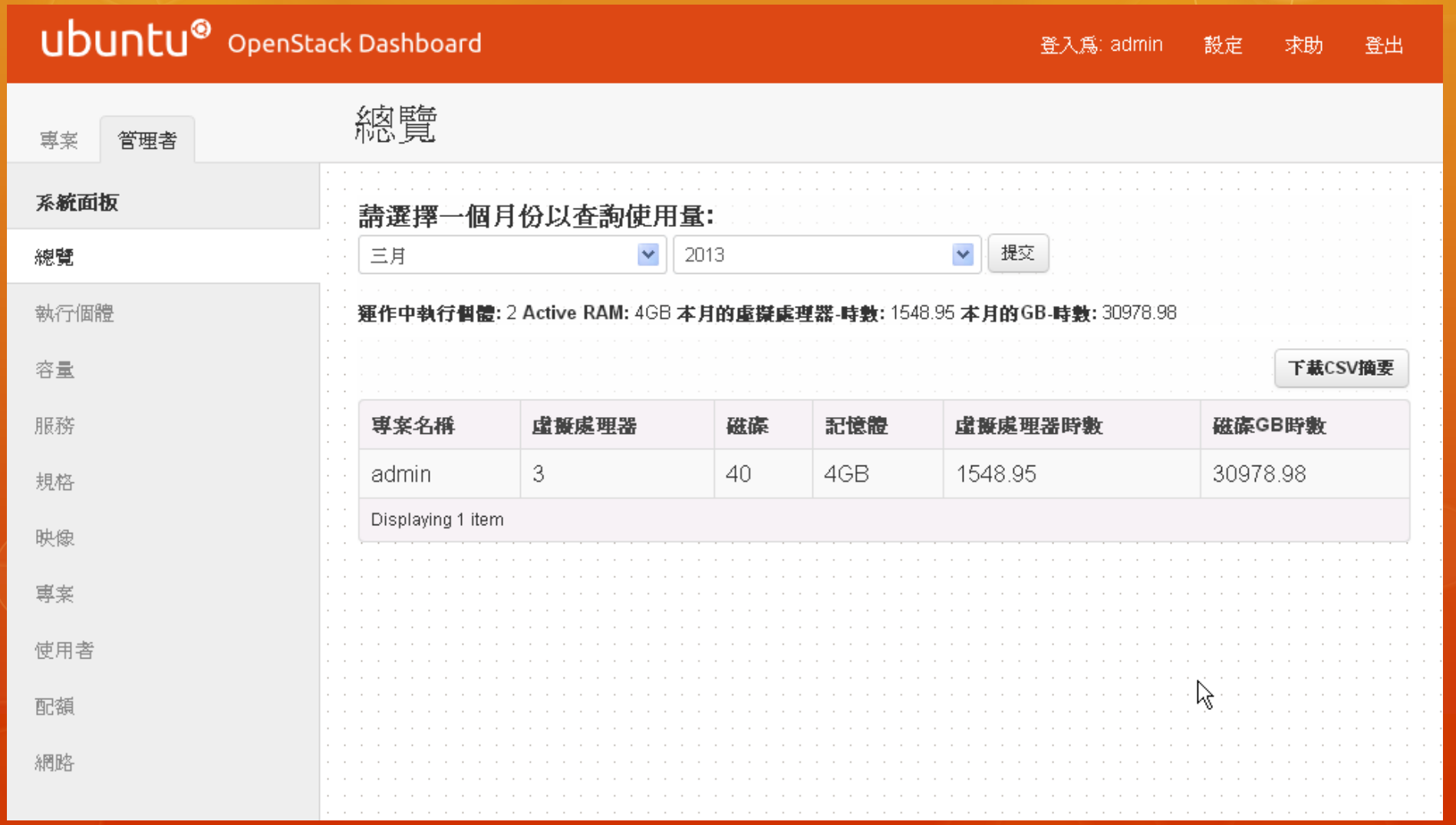
安裝流程

- 請參考 https://github.com/mseknibilel/OpenStack-Folsom-Install-guide

Image 哪裡來？

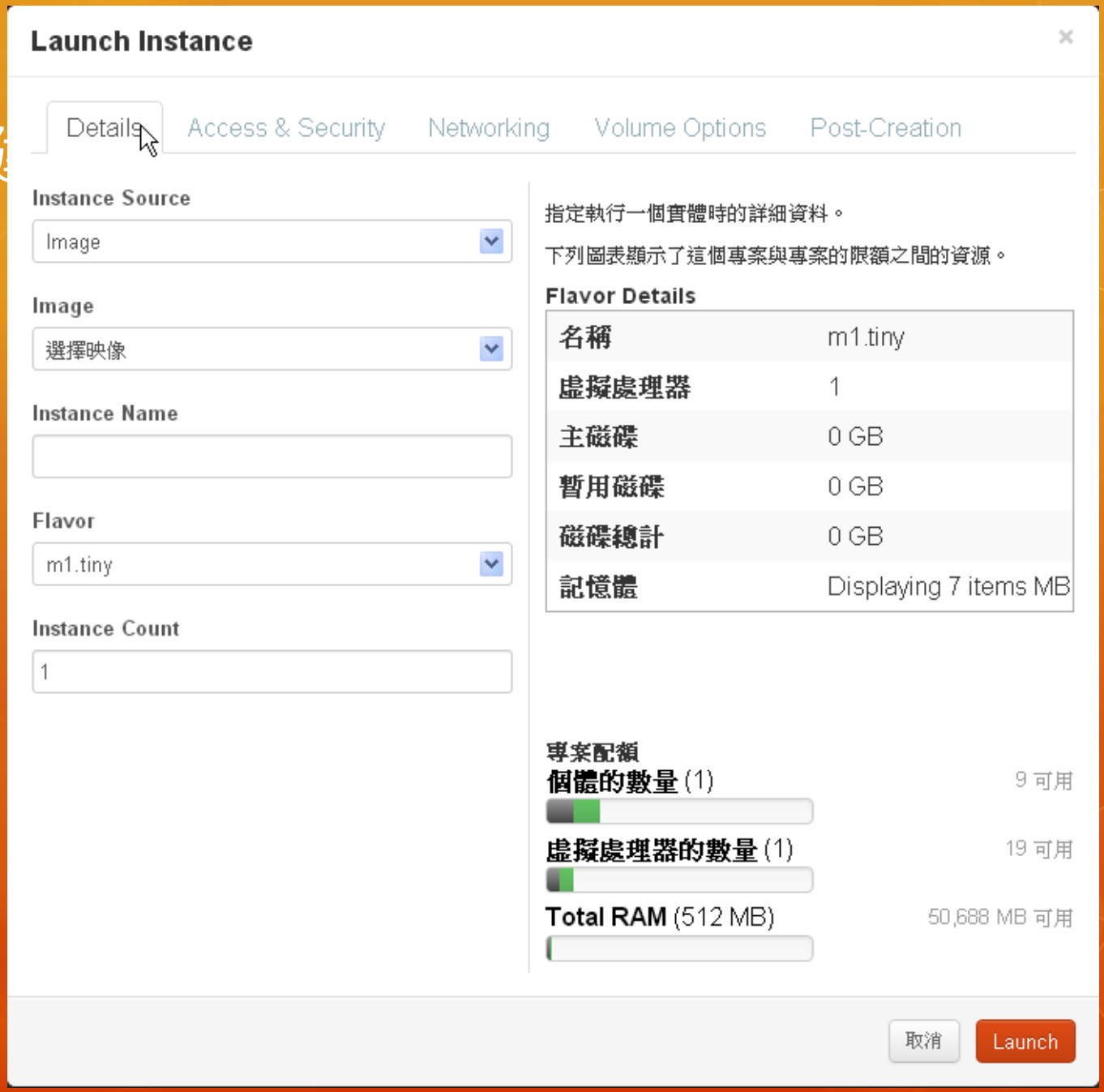
- 光碟映像 ISO->在建立虛擬機後，自動掛載進行安裝程序
  - 例如：Windows server、DRBL-livecd
- 下載現成的虛擬機映像檔（Qcow、vmdk）
  - ubuntu - http://uec-images.ubuntu.com/
  - fedora - http://berrange.fedorapeople.org/images
  - suse - http://susestudio.com/
  - RedHat、Centos - https://github.com/rackerjoe/oz-image-build
  - Windows系列 – 礙於版權規定沒有做好的可以下載

控制面板功能介紹


ubuntu OpenStack Dashboard
登入為: admin
設定
求助
登出
專案
管理者
總覽
系統面板
總覽
執行個體
容量
服務
規格
映像
專案
使用者
配額
網路
請選擇一個月份以查詢使用量:
三月
2013
提交
運作中執行個體: 2 Active RAM: 4GB 本月的虛擬處理器-時數: 1548.95 本月的GB-時數: 30978.98
下載CSV摘要
專案名稱
虛擬處理器
磁碟
記憶體
虛擬處理器時數
磁碟GB時數
admin
3
40
4GB
1548.95
30978.98
Displaying 1 item

執行個體


執行個體
啓動執行個體
Terminate Instances
執行個體名稱
IP位址
大小
金鑰
狀態
工作
電源狀態
動作
drbl test
ext_net
172.22.1.83
net_one
192.168.1.12
m1.tiny | 512MB RAM | 1 VCPU | 0 磁碟
-
Active
None
Running
建立快照
配給浮動IP
編輯執行個體
VNC界面
檢視記錄檔
暫停 執行個體
休眠 執行個體
重啓 執行個體
終止執行 執行個體
Displaying 1 item


Launch Instance
Details
Access & Security
Networking
Volume Options
Post-Creation
Instance Source
Image
Image
選擇映像
Instance Name
Flavor
m1.tiny
Instance Count
1
指定執行一個實體時的詳細資料。
下列圖表顯示了這個專案與專案的限額之間的資源。
Flavor Details
名稱 m1.tiny
虛擬處理器 1
主磁碟 0 GB
暫用磁碟 0 GB
磁碟總計 0 GB
記憶體 Displaying 7 items MB
專案配額
個體的數量 (1) 9 可用
虛擬處理器的數量 (1) 19 可用
Total RAM (512 MB) 50,688 MB 可用
取消
Launch

虛擬磁碟管理


容量
建立容量
Delete Volumes
名稱
敘述
大小
狀態
掛載到
動作
vol10G
10GB
Available
編輯掛載
vol20g
20GB
Available
編輯掛載
Displaying 2 items
建立快照
刪除 容量

建立虛擬磁碟機


建立容量
儲存區名稱
敘述
Additional information here...
容量 (GB)
敘述:
容量是可以掛載到執行個體的分塊磁碟裝置
儲存區限額
總 GB 數 (30 GB)
970 GB 可用
儲存區數量 (2)
8 可用
取消
建立容量

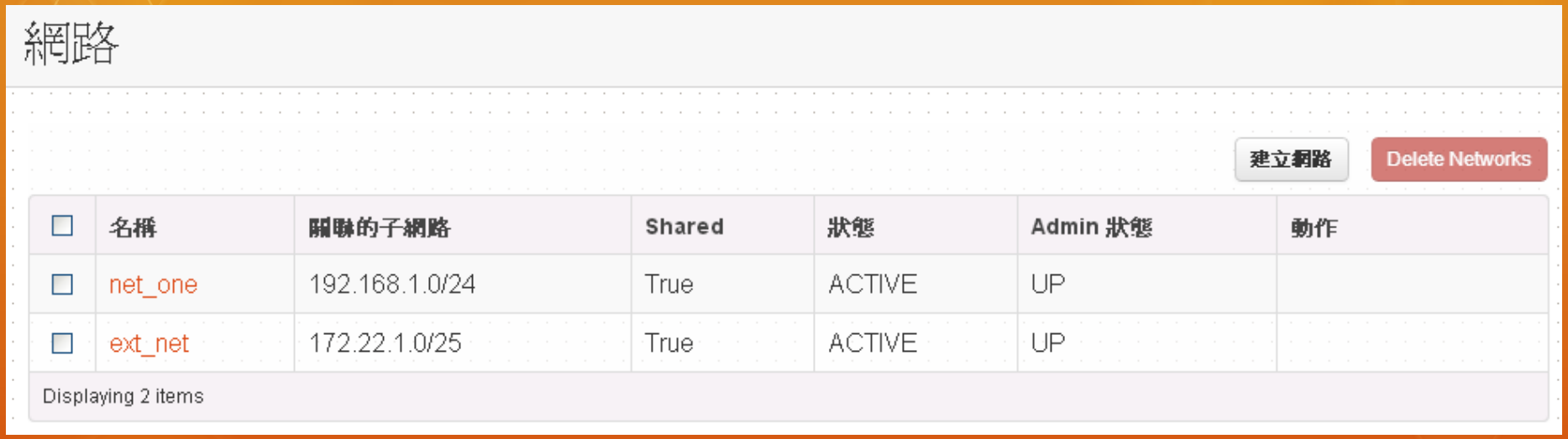
映像及快照管理


映像 & 快照
建立映像
Delete Images
映像名稱
類別
狀態
公開
格式
動作
ubuntu1210_x64
映像
Active
True
ISO
啓動
CentOS 6.0 X86_64
映像
Active
True
QCOW2
啓動
DRBL liveCD
映像
Active
True
ISO
啓動
fedora 16 x86_64
映像
Active
True
QCOW2
啓動
編輯
刪除 映像
Displaying 4 items

建立映像


建立一個映像
名稱
映像位置
http://example.com/image.iso
格式
AKI - Amazon Kernel Image
AMI - Amazon Machine Image
ARI - Amazon Ramdisk Image
ISO - 光碟映像
QCOW2 - QEMU 模擬器
Raw
VDI
VHD
VMDK
公開
詳述：
指定要上傳到映像服務的映像。
Currently only images available via an HTTP URL are supported. The image location must be accessible to the Image Service. Compressed image binaries are supported (.zip and .tar.gz.)
請注意： The Image Location field MUST be a valid and direct URL to the image binary. URLs that redirect or serve error pages will result in unusable images.
取消
建立映像

網路設定


網路
建立網路
Delete Networks
名稱
關聯的子網路
Shared
狀態
Admin 狀態
動作
net_one
192.168.1.0/24
True
ACTIVE
UP
ext_net
172.22.1.0/25
True
ACTIVE
UP
Displaying 2 items

指派IP給執行中的個體

服務運行狀態

服務

Filter

Filter

名稱	服務	主機	已啓用
nova	compute	192.168.1.1	Enabled
quantum	network	192.168.1.1	Enabled
glance	image	192.168.1.1	Enabled
cinder	volume	192.168.1.1	Enabled
ec2	ec2	192.168.1.1	Enabled
keystone	identity (native 後端)	192.168.1.1	Enabled

Displaying 6 items

如何維護節點

- 使用 DRBL 將做好的 compute node 備份成磁碟映像
- DRBL 可以直接使用虛擬機來做，磁碟映像直接儲存於 storage node
- 無論是新增 compute node 或是復原 compute node，只要直接用 DRBL 派送系統就可以了
- 派送好的系統，仍然需要手動修改網路組態，或者是使用 DHCP Server 派送固定IP
- Network node 、Storage node 做法同上
- Controller node 這台比較麻煩，必須採用叢集系統來達到高可用性（HA），請參考 Pacemaker