國家實驗研究院
NATIONAL INSTITUTES OF APPLIED RESEARCH

國網中心雲平台現況、整體規劃與實作

高效能計算與雲端技術組
報告人:潘怡倫、莊家雋

2025/12/01

LLM雲訓練主機

台灣杉系列主機

NARLabs

台灣杉一號 TAIWANIA 1

2017/11

630 nodes / 25200 CPU cores / 256 P100 GPUs
157 TB memory
3.4 PB storage
Omni-Path HFI 100 Gbps

計算力 1.7 PFLOPs
TOP500 #95
Green500 #31

台灣杉二號 TAIWANIA 2

10th 能源效率 20th 計算能量

2018/11

252 nodes / 9072 CPU cores / 2016 V100 GPUs
193.5 TB memory
10 PB storage
InfiniBand EDR 100 Gbps

計算力 9 PFLOPs
TOP500 #20
Green500 #10

台灣杉三號 TAIWANIA 3

2020/11

900 nodes / 50400 CPU cores
172.8 TB memory
9.4 PB storage
InfiniBand HDR 100 Gbps

計算力 2.7 PFLOPs
TOP500 #181
Green500 #69

創進一號(台灣杉四號)高速運算主機

✔ 加速科研創新突破，確保運算資源穩定供應

- 建構新一代以CPU為主的運算平台，支援重點領域(氣候、環境、生醫、材料、能源等超大型課題)
-

TOP500 #222
(2023/11)

Green500 #92
(2023/11)

架構	x86-64	ARM
電腦節點數量	552 台	40 台
CPU處理器	Intel Xeon Platinum 8480+ 56 Cores 2.0GHz *2	Nvidia Grace CPU Superchip 144 Cores
效能 (Rmax)	3.5 petaFLOPS (實測)	0.2 petaFLOPS (預計)
內部高效能網路	InfiniBand NDR 200 Gb/s	
平行檔案系統	IBM Spectrum Scale (GPFS) 9.2 PB	

- ✔ 與目前台灣杉三號使用相同的SLURM排程系統與InfiniBand高效能網路
- ✔ 主機用戶可維持原有的使用習慣，不必重新學習job script語法
- ✔ 113年2月進行友善測試，113年7月正式提供服務

114年新一代高速運算主機建置

AI雲端運算主機，以GPU架構為主，可提供HPC服務與雲端服務(VCS、CCS)

滿足生成式大型語言模型開發需求

提供AI共用運算服務

支援大型科學運算應用

整體算力80PF以上

儲存資源20PB以上

主機能耗PUE 1.35以下

預計114年12月建置完成，115年4月正式提供服務

台灣杉二號雲運算服務

TWCC 臺灣AI雲 TAIWAN COMPUTING CLOUD

https://www.twcc.ai/

快速部署

容器運算服務

Container Compute Service

可快速部署GPU處理器的人工智慧工作環境，提高29%的工作效率

服務項目包含開發型容器、任務型容器。配備 8 個 NVIDIA® Tesla V100 GPU，加速人工智慧訓練、推論與高效能運算，支援 5120 個CUDA核心與 640 個 Tensor 核心，並支援 NVLink 進行GPU之間的資料傳輸。

有效統御

高速運算服務

High-performance Computing

部署跨節點、多顆 GPU 的分散式高速平行運算環境，效能提高 30% 以上

服務項目包含 HPC 高速運算任務、台灣杉二號。配備 NVIDIA® Tesla V100 GPU，加速人工智慧訓練、推論與高效能運算

智算兼備

虛擬運算服務

Virtual Compute Service

短時間內即能建立安全穩固、彈性應用的虛擬運算服務 (VCS) 個體

提供Linux (Ubuntu、CentOS) 作業系統、Windows作業系統。，智算兼備、節省成本最佳的方案。

配備 Intel® Xeon® Gold 61 系列處理器，記憶體處理速度可達 2666 MHz。

大數據匯集

雲端儲存服務

Virtual Compute Service

安全高效率的多樣儲存選擇，適合各式運算情境，多層備份機制，安心儲存資料

HFS高速檔案系統為容器運算服務、高速運算服務搭配使用之儲存方案，註冊帳號即能免費獲得200 GB的儲存空間！另有雲端物件儲存 (COS)及區塊儲存 (BSS) 儲存服務，能與各式運算服務完美搭配運作。

台灣杉二號單節點可運行的最大模型參數

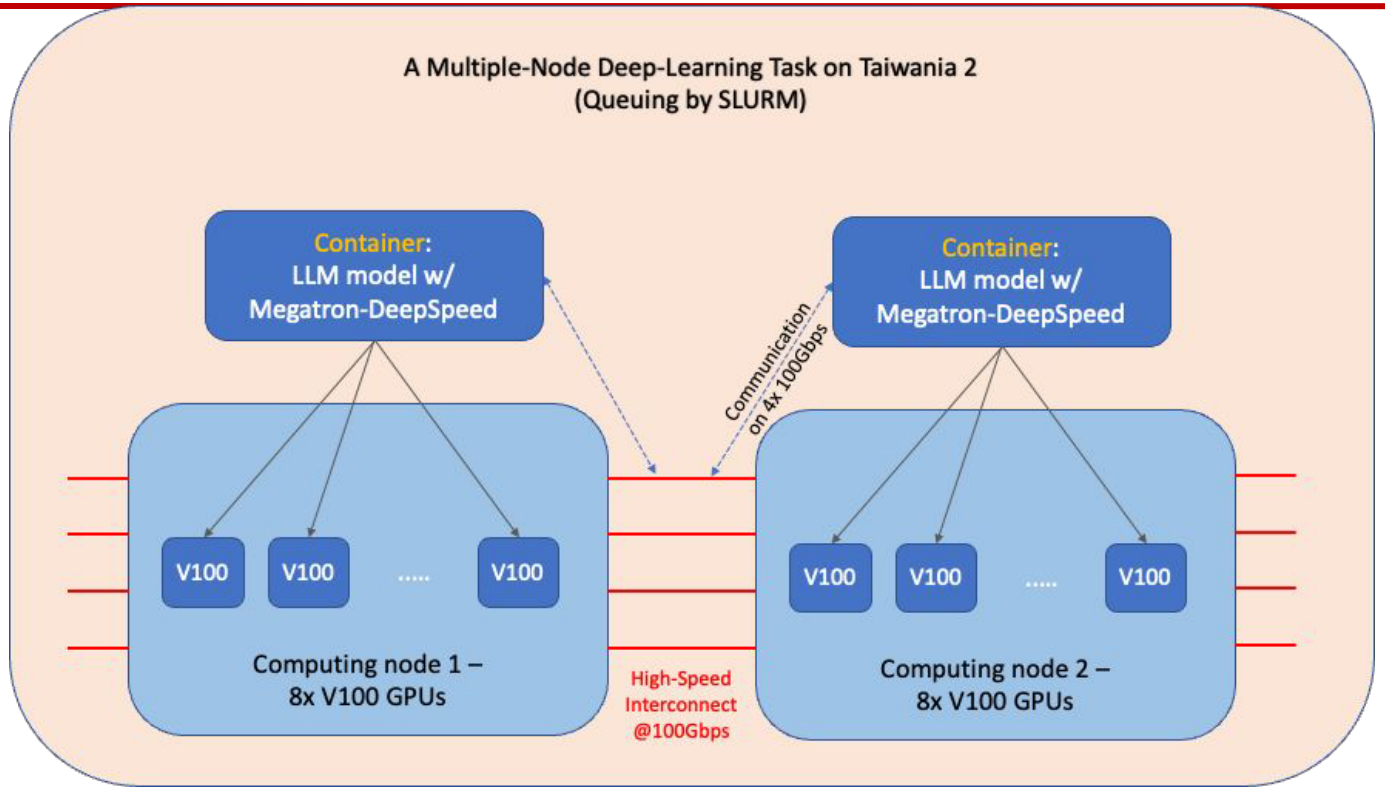
- 模型記憶體需求估算：1B(10億)參數模型在半精度(FP16)模式需使用2GB記憶體 #1 #2
- 在進行GPT模型訓練時將會使用到(系統會配置)5~6倍的GPU記憶體空間
- 模型訓練期間必須保留約25% GPU記憶體空間以防止記憶體不足(Out of Memory/OOM)
- 多尺度GPT模型的記憶體需求與V100-32GB記憶體的比較表：

GPT3/BLOOM Model Size	355M	1.3B	2.7B	6.7B	V100 @50%	13B	V100 @100%	20B	175B
MEM (unit: GB)	0.71	2.6	5.4	13.4	16	26	32	40	350
Fraction of a V100 MEM V100 w/ 32GB Memory	2.2%	8.1%	16.9%	41.9%	50%	81.3%	100%	125%	~1100%
GPU MEM. Allocation (<75% V100 MEM.)	~11%	~40.5%	~84.5%	**OOM**		~405%		~625%	> 6400%
Running on 1 V100 GPU	**V**	**V**	**V**	**X**		**X, 4 V100s**		**X, 7 V100s**	**X, 64 V100s**
Running on 1 node	**V**	**V**	**V**	**V**		**V**		**V**	**X, ~8 nodes**

- 2.7B以上的語言模型訓練需開啟模型平行計算(Model Parallelism)始可運算，但會增加資料交換量會導致運算效率降低
- 單節點可運行的最大模型約是20B模型，但有可能會發生OOM
- 175B模型訓練需要約8台V100節點(64~72 V100s)，改用H100-80GB預估需3台H100節點(24 H100s)運行模型訓練

#1 https://towardsdatascience.com/language-model-scaling-laws-and-gpt-3-5cdc034e67bb
#2 Brown, Tom, et al. "Language models are few-shot learners." Advances in neural information processing systems 33 (2020): 1877–1901.

台灣杉二號跨節點模型訓練運算



跨節點調度 GPU，實現高速分散式平行運算

透過 Slurm 資源調度軟體，操作強大的超級電腦 (HPC)，搭配 MPI 程式，能實現跨節點的分散式平行運算，將高負載的工作量平均分派，提升處理效率！

大頻寬網路串連節點，資料傳輸快速

採用100 Gbs高速網路串連 GPU主機群，有極高的吞吐量與極低的延遲，解決傳統技術的瓶頸，巨量資料傳輸效率不妥協！

GPU Direct 與 RDMA架構，極致加速

透過NVLink 與超高速網路計算 (InfiniBand)架構的整合，使 RDMA (Remote Direct Memory Access)技術可大幅提升跨節點的大數據資料傳輸效率，並顯著提高整體運算效率。

- 開發任務導向容器技術，可大幅簡化多GPU間快速傳輸運算與運用InfiniBand做跨節點或跨平台的平行計算等問題
- 自動化訓練排程、動態部署訓練算力、效能調適容器環境

國網中心雲平台發展沿革

國網中心雲端發展沿革

前瞻計畫:
- VMware平台:VCF建立
- 2025年董事會精進、大考中心轉移

晶創計畫:
- 2025年底建置完成
- 2026年第一季、第二季試營運，2026第四季上線營運

ISO 9001 ISO 27001 BS 10012/PIMS ISO 50001

2017 Peta

台灣杉一號 TAIWANIA 1

2018

台灣杉二號 TAIWANIA 2

2019

TWCC 臺灣AI雲 TAIWAN COMPUTING CLOUD

2020 AI

2021

2022

2023 10 Hyper Kylin

2023

2024 HPC Cloud

2024

1993 2005 2008

前瞻計畫:
- Hyper Kylin: 2025年下線
- 機敏混合雲上線:2024年上線營運, 2025通過ISO

2018

HPC

HPC

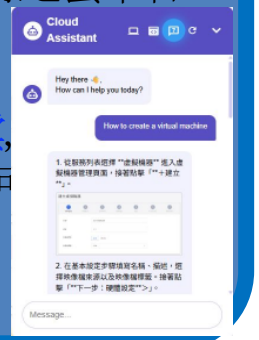
可信賴計畫:
- 下線後的Hyper Kylin肢解形成Staging環境與訓練成員的試煉廠
- 可信賴雲上線:2024年第四季上線營運, 2026預計進行ISO認證

10

雲平台

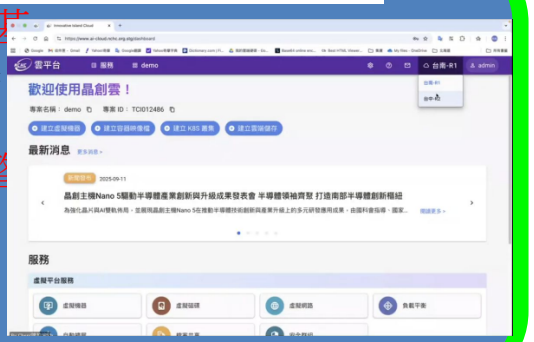
可信賴雲 – 台灣第一組 TRE環境

- 第一組NCHC OpenStack擁有原始碼並開源之雲平台專案提供虛擬機與容器服務
- 雲團隊自主開發QA與執行小助理
- 可整合其他雲平台 – 如前瞻機敏混合雲, 態配置機算資源, 提升服務可用度系統可 99.95%以上



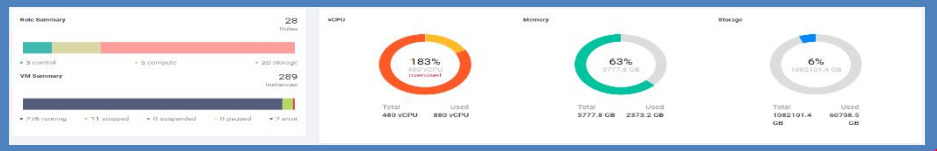
晶創雲平台 IIC

- 以NCHC雲平台原始碼為基底提供服務虛擬機與容器化服務
- 實作Multi-Region機制與整合HPC計算
- 規劃2026 – 2027可承接TWCC使用者



機敏混合雲建構高隱私實驗環境

- 以商用版OpenStack – Cube OS+ CMP提供服務
- 實作機敏環境, 整體雲服務資源使用率超過6成, 運行近300 VMs

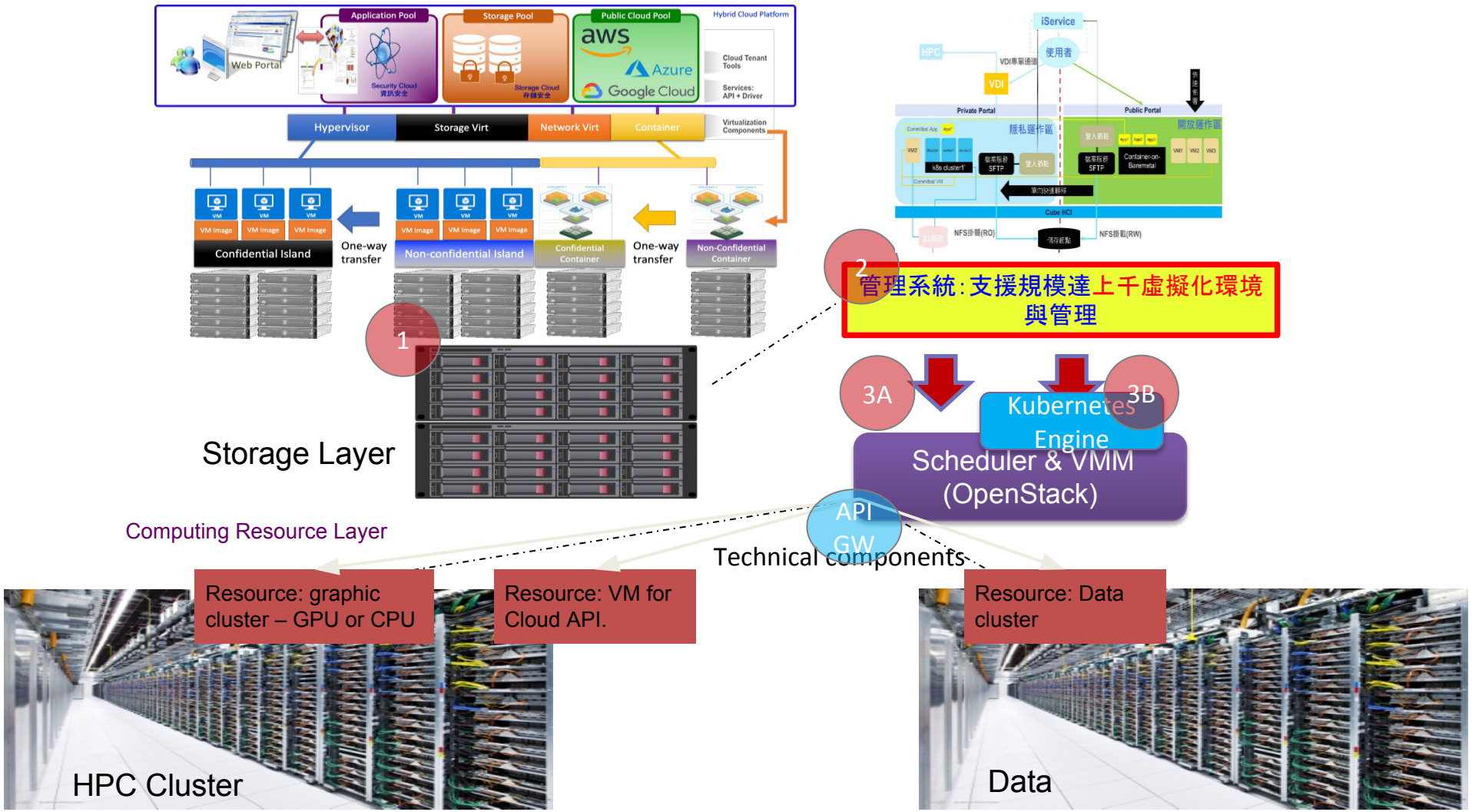


11

可信賴雲平台

- NCHC雲端服務平台
  - 2019年TWCC台灣AI雲正式啟用
  - 2022年Hyper Kylin奇靈雲 2023 10月正式營運
  - 2023年先導型機敏混合雲(簡稱機敏混合雲)建置中
  - 2024年可信賴雲平台
  - 2025年晶創雲平台
- 綜觀上述平台所提供之服務
  - 信任研究環境 (Trusted Research Environment, TRE)
  - 隱私工作區與資源跨專案移轉
  - 分為運算、儲存、網路與使用者API四大部分進行分析
  - 規劃整合HPC高速運算與HFS(HFS高速儲存為TWCC特有項目)

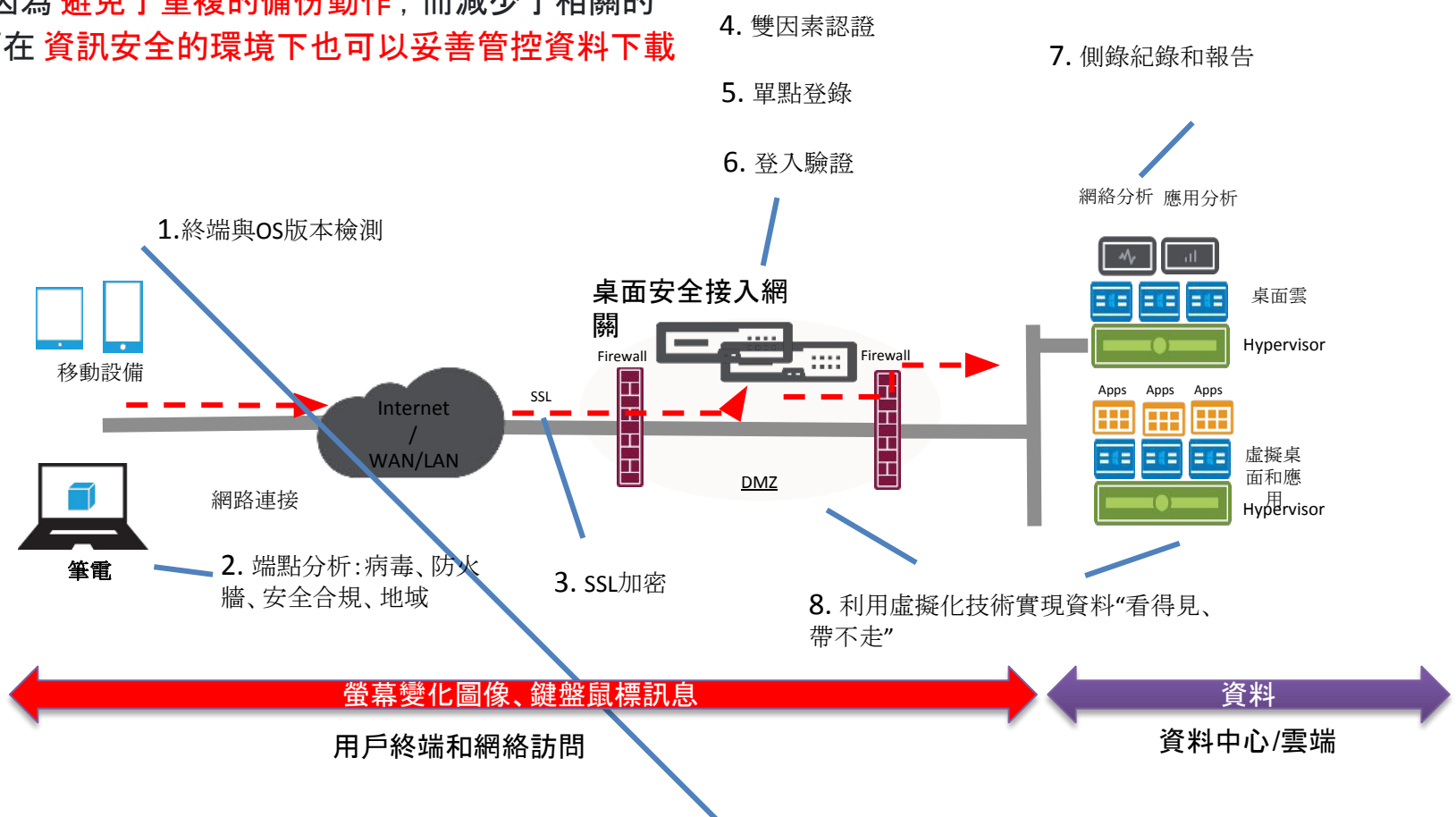
可信賴雲平台


Application Pool
Storage Pool
Public Cloud Pool
Hybrid Cloud Platform
Web Portal
Security Cloud
Storage Cloud
aws
Azure
Google Cloud
Cloud Tenant Tools
Services: API + Driver
Hypervisor
Storage Virt
Network Virt
Container
Virtualization Components
VM
VM Image
Confidential Island
One-way transfer
Non-confidential Island
Confidential Container
One-way transfer
Non-Confidential Container
1
Storage Layer
iService
使用者
HPC
VDI
Private Portal
Public Portal
隱私運作區
開放運作區
Container-on-Baremetal
Cube HCI
2
管理系統:支援規模達上千虛擬化環境與管理
3A
3B
Kubernetes Engine
Scheduler & VMM (OpenStack)
API GW
Technical components
Computing Resource Layer
Resource: graphic cluster – GPU or CPU
Resource: VM for Cloud API.
Resource: Data cluster
HPC Cluster
Data

「信任研究環境」(Trusted Research Environment, TRE)

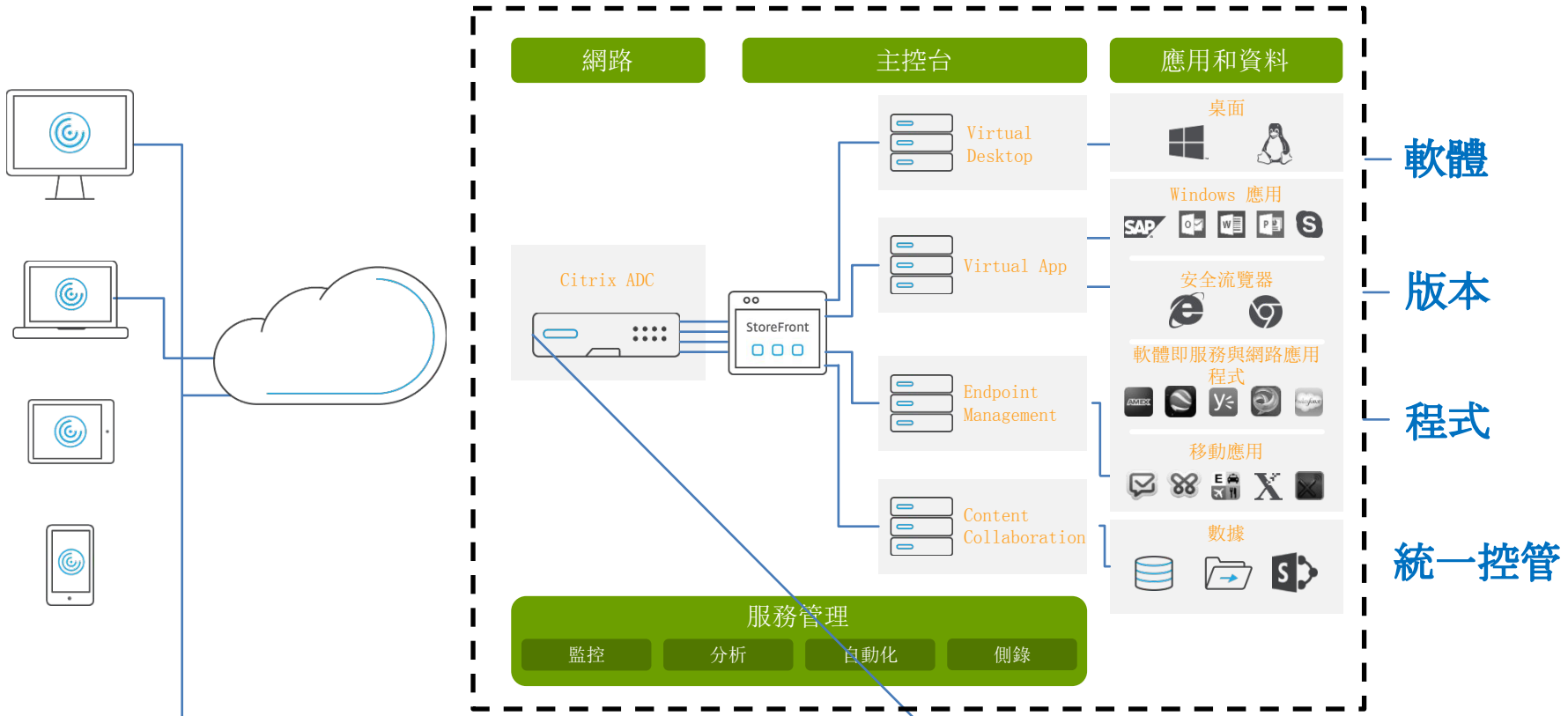
- 將數據分析帶到資訊安全環境執行 (bring analysis to the data)，讓研究人員可以去存取被存放在資訊安全環境的資料集，這種作法有時也被稱作「不落地」

- 如此較能夠確保資料的存取是在安全的資訊環境，也能較能確保個人資料的保護；同時，也因為 避免了重複的備份動作，而減少了相關的傳輸與儲存的費用，而在 資訊安全的環境下也可以妥善管控資料下載



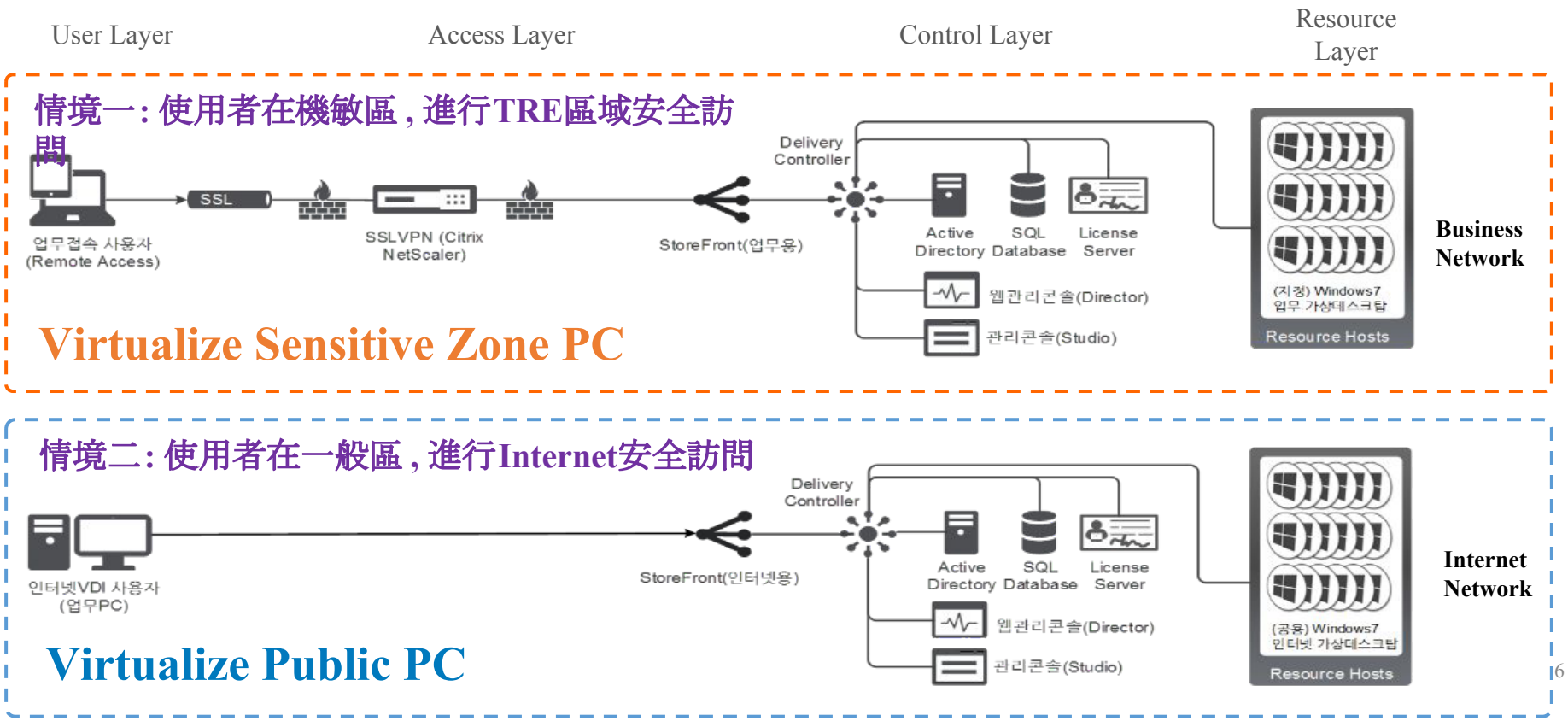
「信任研究環境」(Trusted Research Environment, TRE)

❑ 集中化管理



「信任研究環境」(Trusted Research Environment, TRE)

❑ 雙網隔離架構



願景與目標

願景

- **自主開發雲平台：** 打造Cloud Native技術研發平台並搭配資料串流
- **自主雲平台符合生成式 AI模型研發所需之大量節點調度與雲端管理技術**
- **提供可介接公有雲端服務之動態資源策略**
- **提供滿足 HPC、BigData and AI-related三大族群用 戶之功能服務**

主要目標 – 雲自主開發範圍

- 機敏混合雲平台
- 可信賴雲平台服務
  - 確保可以正常虛擬機器服務與容器服務，使用戶能夠順利地上架雲端。
  - 確保開發上可以Build完整雲平台環境與開發新的功能，以達到可彈性調整服務上雲，確保平台策略符合用戶需求。
  - 可自訂雲端程序，有效監控主機之資源與應用程式
- 晶創雲平台服務
  - 確保可以正常虛擬機器服務與容器服務，使用戶能夠順利地上架雲端。
  - 整合機敏混合雲平台與可信賴雲平台，達到multi-region的實作。

□確認需求面功能，著重在客製化開發與Portal、API的開發

雲自主維運

- 提升各雲平台計算資源的協同運作，其中包括Vmware平台、機敏混合雲與可信賴雲平台
- Cloud Maintenance Part重點:確認平台服務穩定，著重在面對使用者意見回饋與底層平台的穩定

客戶服務

- 提供用戶各種軟體使用問題上的諮詢
- 幫助用戶能夠順利地使用軟體進行研究工作。
- 透過參與研究計畫、舉辦使用者會議等方式，與學界交流和培養人才並推廣中心的計算資源，達到互利的作用。

相關分析比較

雲平台比較項目	機敏混合雲 - 2022建置	可信賴雲 - 2024建置	晶創 - 2025建置
架構	OpenStack – 商用	OpenStack - NCHC	OpenStack - NCHC
資源量	GPU: 32張A30 vCPU: 2560 core Memory: 15 TB Storage: 560 TB	GPU: 30張H100、20張H200 vCPU: 7680 core Memory: 60 TB Storage: SSD 6.6 PB、HDD 7.6 PB	GPU: 70張 H200 vCPU: 19200 core Memory: 150 TB Storage: 15 PB
優勢	機敏區 應用服務APP K8S服務 串連公有雲(AWS)	機敏區 應用服務APP K8S服務 國網可自行開發	Multi-Region/HPC 應用服務APP K8S服務 國網可自行開發
主要服務對象	業界、公部門與國網中心單位	生醫(聯邦計算)、科學計算(安正團隊)、國安國土(志泓/奕良團隊)	使用者對象就是TWCC等相關使用者都以這部份轉移為主
服務情形	業界 • 大橡科技 • SAS公司 公部門 • 內政部 • 疾管署 國網中心 • LLM服務 • 活動與課程支援	國研院 • 國網中心 – 如:生醫團隊 (健康大數據永續平台-建置國家級之友善生醫資料分析與分享平台) • 國儀中心 - (董事會)國儀中心可信賴雲POC • 等等等	目標對象 - TWCC使用者

平台收費比較-服務規格 (詳盡資料請參考計價委員會提報)

虛擬運算服務規格計價 -TWCC			
VM size	vCPU	記憶體	NTD/小時
v.xsuper	4	32 GB	7.10
v.2xsuper	8	64 GB	14.20
v.4xsuper	16	128 GB	28.40

虛擬運算服務規格計價 -機敏混合雲			
VM size	vCPU	記憶體	NTD/小時
Memory.small	**4**	**32 GB**	**6.68**
Memory.medium	8	64 GB	13.37
Memory.large	16	128 GB	26.74

虛擬運算服務規格計價 -機敏混合雲			
VM size	vCPU	記憶體	NTD/小時
Memory.small	**4**	**32 GB**	**7.60**
Memory.medium	8	64 GB	15.20
Memory.large	16	128 GB	30.40

三大公有雲計價

可信賴雲		機敏雲		TWCC		Azure		Amazon		GCP	
Flavor	元/時	Flavor	元/時	Flavor	元/時	Flavor	元/時	Flavor	元/時	Flavor	元/時
Memory.small 4核+32GB	7.6	Memory. small 4核 +32GB	6.68	v.xsuper	7.10	E4s v6	11.13	r7gd.xlarge	8.95	c2d-highmem-4	8.58
H100(80GB) 96核+512GB +120GB	192.0	A30(24G B) 60核 +300GB +120GB	51.54	-	-	NC40ads H100 v5 (8→1) 40核+320GB +3.5TB	213.37	P5.48xlarge (8→1) 24核+256GB +3.75TB	260.05	a3-ultragpu-8g (8→1) 28核+369GB +1.5TB	320.56
H200(140GB) 96核+512GB +120GB	245.0	L4(24GB) (性能較 A30優)	-	-	-	ND96isr_H200_v5 (8→1) 12核+237.5GB +3.5TB	405.03	P5en.48xlarge (8→1) 24核+256GB +3.75TB	299.06	-	-

19

平台功能項目-運算

https://github.com/orgs/Zillaforge/repositories

NARLabs

運算				
	TWCC	機敏混合雲	可信賴雲	晶創雲
虛擬運算	v	v	v	v
虛擬運算 - 資源動態調整	v	v	v	v
虛擬運算 - 隱私工作區	x	v*	v*	x* (模組化設定故可設置)
虛擬運算 - 跨專案移轉	x	v*	v*	x* (模組化設定故可設置)
開發/任務型容器 (container)	v	應用服務APP	應用服務APP, in k8s	應用服務APP, in k8s
Kubernetes as a Service	x	v*	v*	v*
HPC高速運算任務 (Slurm)	v	x	v	v
多雲架構整合 (Multi-Region)	x	x	x	v
程式碼	Binary	NA	v	v
公有雲連動	x	v*	x	x

20

功能總表綜合比較-運算

		TWCC	機敏混合雲	可信賴雲	晶創雲
虛擬運算	**便利性**	第一次須先透過SSH登入建立密碼，才可使用VM Console	部署VM時，UI即可建立密碼，用於登入VM Console，便利性較高		
	隱私性	同專案之管理者可檢視與操控一般使用者的VM Console	同專案之管理者無權限檢視與操控一般使用者的VM Console，隱私性較高		
虛擬運算-資源動態調整		v	v	v	v
虛擬運算-隱私運作區		x	**v***	**v* (模組設定:Y)**	**x* (模組設定:X)**
虛擬運算-跨專案移轉		x	**v***	**v***	
開發/任務型容器 (container)		**v***	v (應用服務APP)	v (應用服務APP)	v (in k8s)
Kubernetes as a service		x	v	v	v
HPC高速運算任務 (Slurm)		**v***	x	v	v
公有雲連動		x	**v***	x	x

21

功能總表綜合比較-儲存

	TWCC	機敏混合雲	可信賴雲	晶創雲
雲端檔案服務 CFS	分享權限設定為計畫成員共享	分享權限設定為 VM 共享		
虛擬機器匯出(下載)	需申請	v	v	v
虛擬機器快照/備份	透過TWCC CLI設定定時建立虛擬機映像檔	透過**UI介面**設定定期自動備份虛擬機		
高速檔案系統 (HFS)	**v***	x	x	x

22

功能總表綜合比較-網路

	TWCC	機敏混合雲	可信賴雲	晶創雲
虛擬網路管理	可建立虛擬網路及設定防火牆	可建立虛擬網路及設定防火牆	可建立虛擬網路及設定防火牆	可建立虛擬網路及設定防火牆
負載平衡	無支援UDP協定	支援多種協定及Header設定		
負載平衡-存取控管	透過安全性群組設定較繁瑣	透過來源白名單設定較直覺		
Auto scaling	v	v	v	v

23

功能總表綜合比較-API

	TWCC	機敏混合雲	可信賴雲	晶創雲
使用者API	透過API金鑰存取	透過API金鑰存取	透過API金鑰存取	透過API金鑰存取

24

總結

- 晶創雲平台 – 以可信賴雲平台管理程式碼為基底開發
  - 提供近似TWCC雲服務之要件(確定業務需求和用例，以確保雲端平台能夠滿足相應的功能和性能要求。)
    - 一站式Portal – VCS/VPS
    - 通用型雲端服務 - 雲端架構，包括應用程式層、資料庫層、儲存層、網路層等。
    - 資源配置：HPC Job Submission
    - 自動化和自動擴展：使用自動化工具和技術來管理和擴展資源，以應對變化的工作負載。
    - 高隱私區：於隱私區提供高監控的環境，而一般通用區達到VCS/VPS的虛擬化服務提供。
  - 整合與擴大資源
    - Cloud Services: 透過應用程式服務App的方式包裝SaaS服務
    - 提供HPCaaS服務
  - 形成應用程式市集，且連動機敏混合雲、可信賴雲，打造三位一體之雲端服務平台
    - Multi-Region:
    - 程式與維運掌握

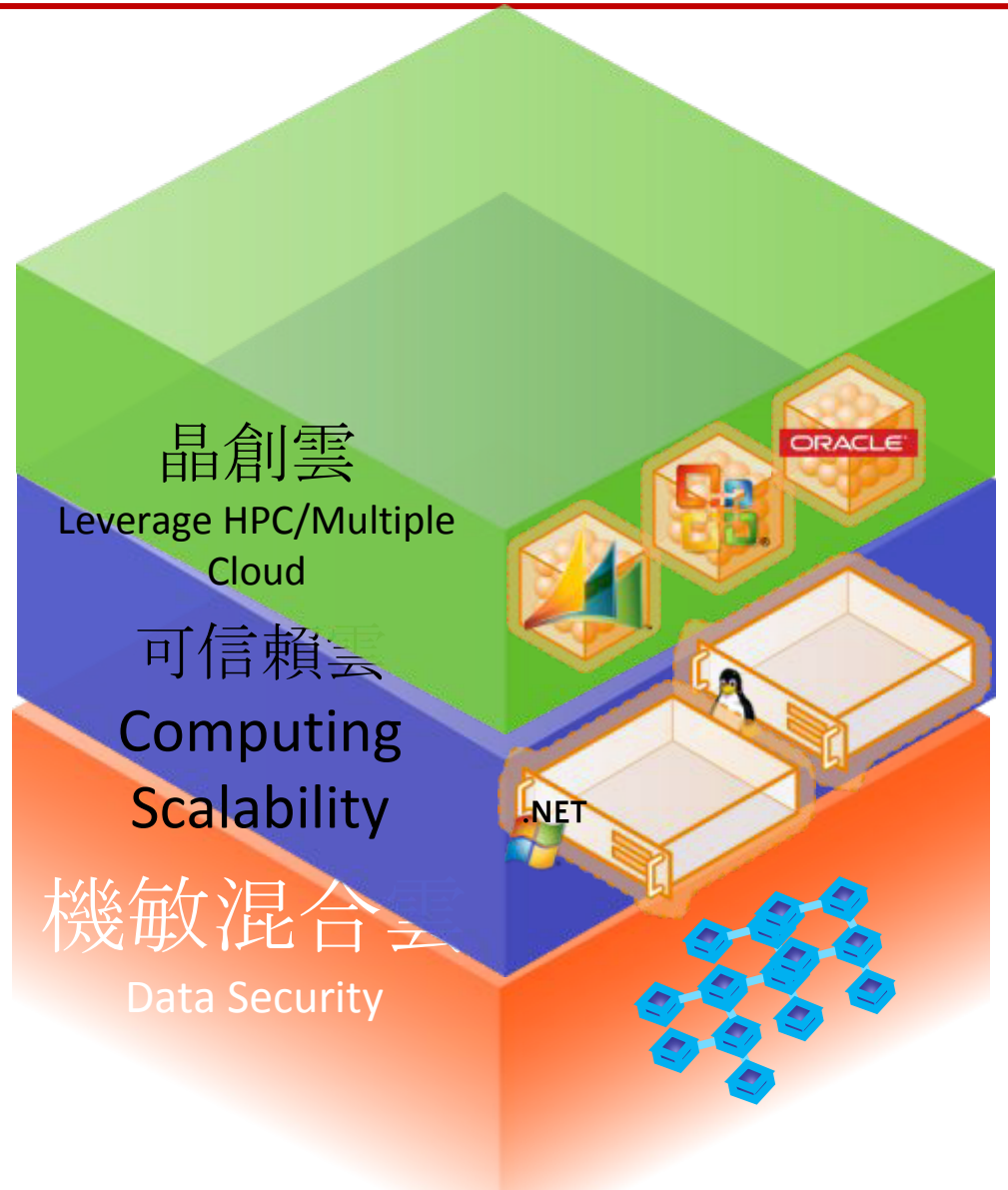
25

打造三位一體之雲端服務平台

長期

中期

近期



26

雲平台管理軟體方向 – 開發、維運、移植與推廣

可信賴雲/晶創雲：可『自主』持續發展的計算資源生態，體現在以下幾個方向：

開發；一站式雲平台服務與技術支援

解決軟體安裝與環境配置的技術門檻，同時享用雲端計算(VM /K8S)與HPC計算之環境，並透過Multi-region**技術**，利用**計算資源監測與用量分析**，提高**資源利用與**研究效率。

轉移TWCC使用者

移植：雲端服務使用

提供多組雲平台(Openstack為基底)，透過計算資源監測與用量分析，合理調度計算資源，提高計算資源使用率。

維運：高可用性虛擬服務 - Multi-Region

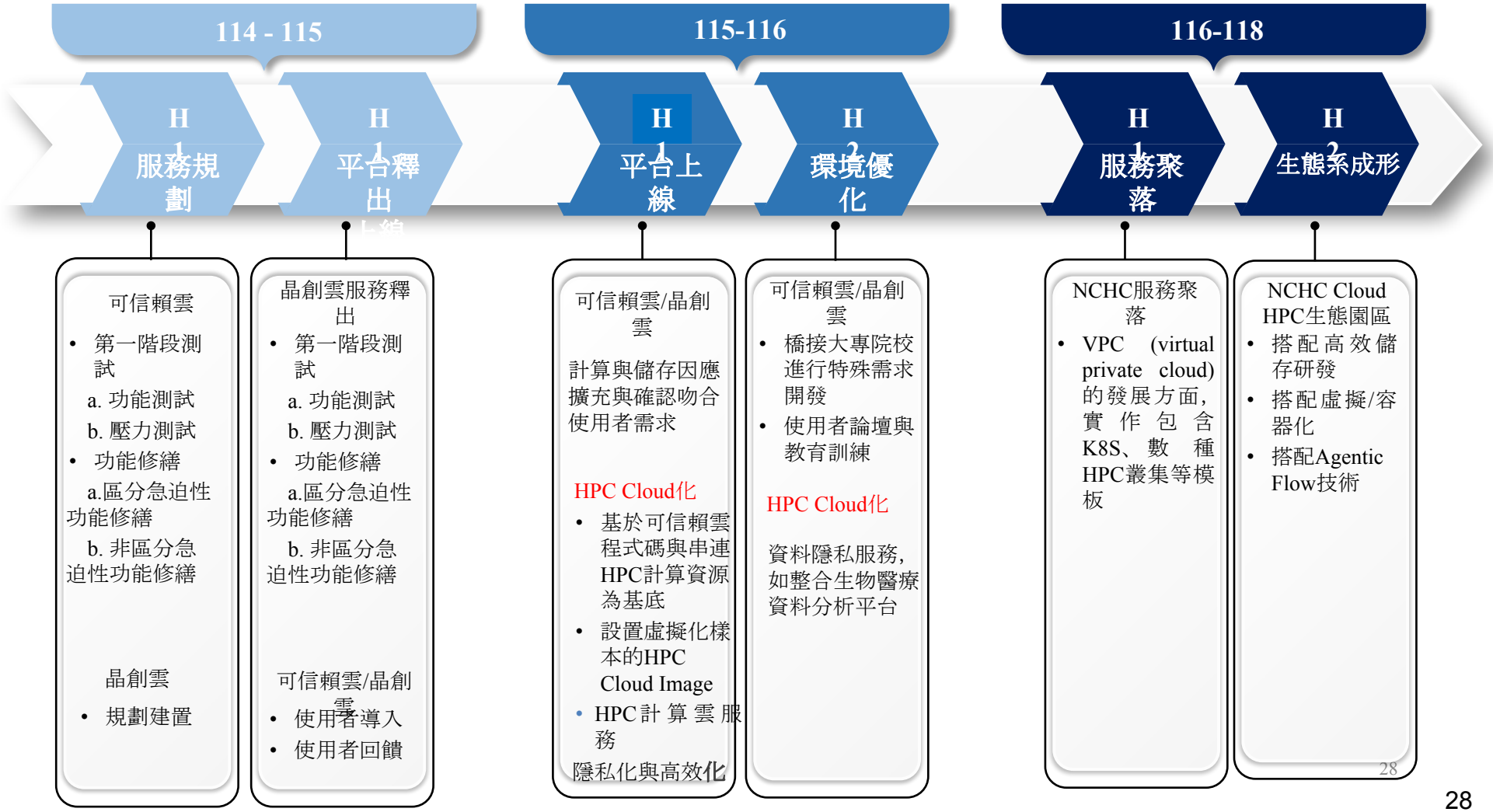
分散式儲存與增加跨域系統備援能力，強化災難備援復原，降低資料損失。

推廣：加速我國人才養成（開放原始碼）

提供教育訓練與使用手冊，降低降低學習門檻，並透過論壇與學校資訊相關科系合作，提升國內雲服務之研發能量。

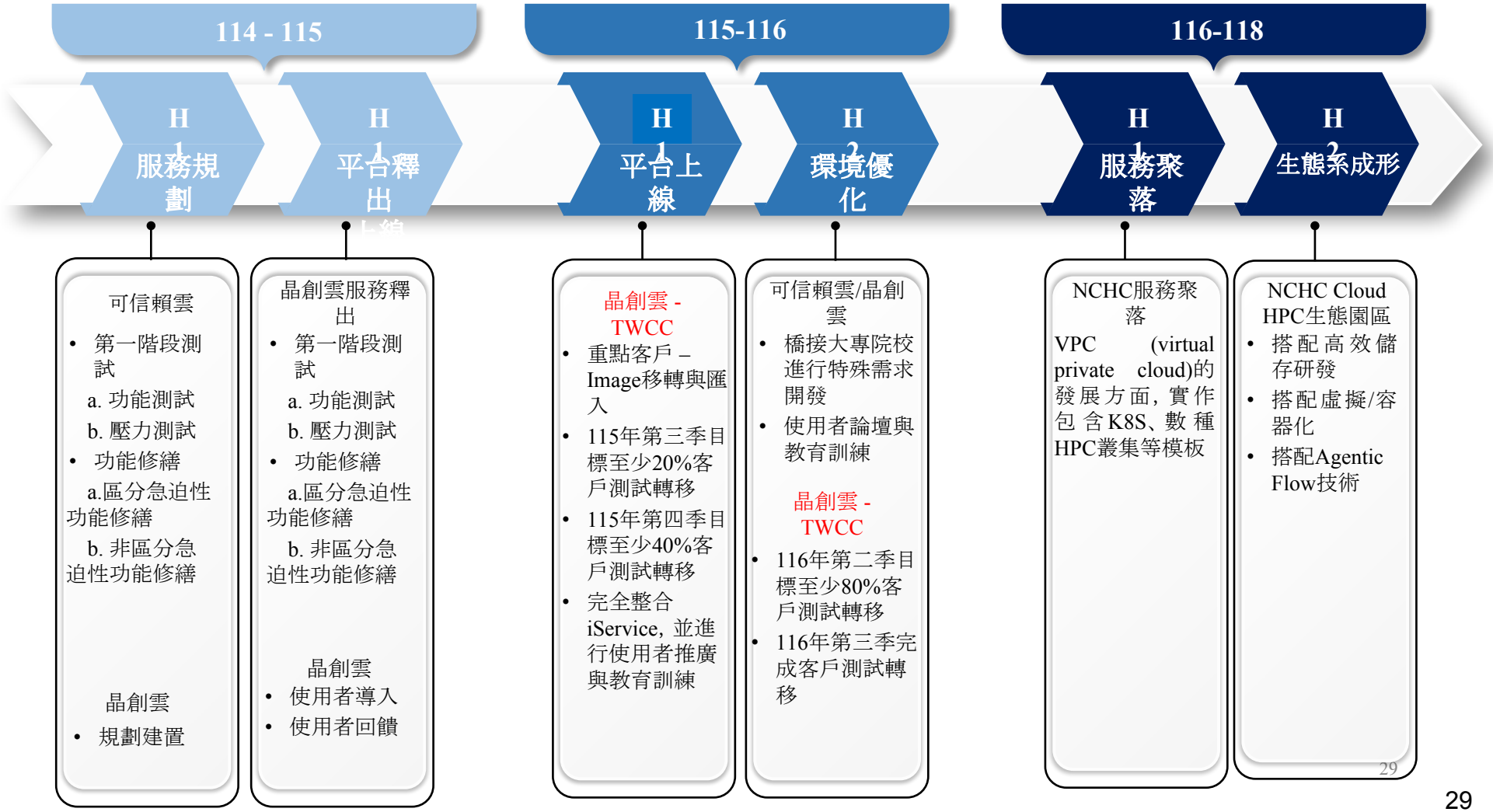
雲平台管理軟體方向 – 開發、維運、移植與推廣

目標對象: HPC、BigData and AI-related三大族群使用者。



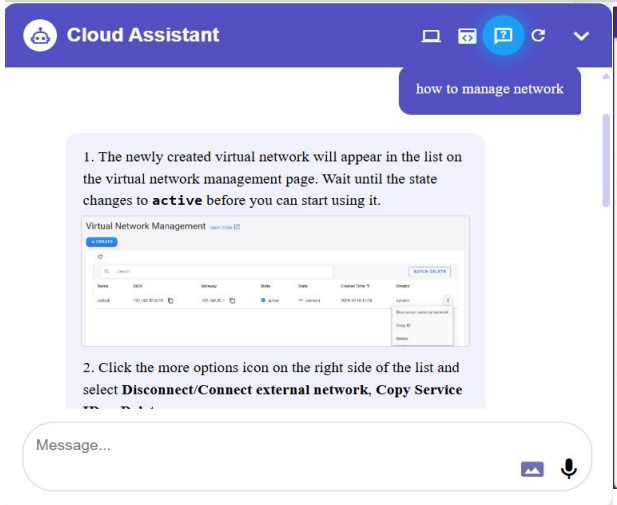
雲平台管理軟體方向 – 開發、維運、移植與推廣

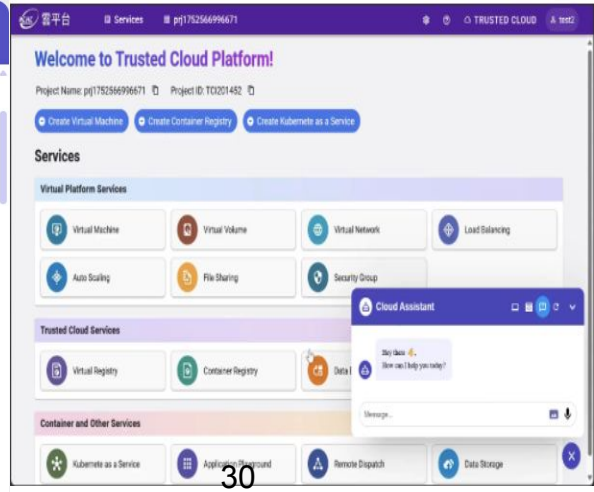
目標對象: HPC、BigData and AI-related三大族群使用者。



雲平台管理軟體規劃

- 自主可信賴雲平台 - 軟體開發
  - HPC Cloud設計
  - MarketPlace設計 – 以QA小助理與執行小助力拋磚引玉
    - 服務上架與下架
    - 開發前端功能元件
- 自主晶創雲
  - iService 整合
    - 帳務機制
  - 系統監控Dashboard
    - 使用國網晶創25資源進行HPC遠端派送功能驗證 已確認
  - Slurm RESTful API對外服務 台中可信賴雲已測通 台南環境待確認
  - SMTP與DNS設定





雲服務 – 自主開發

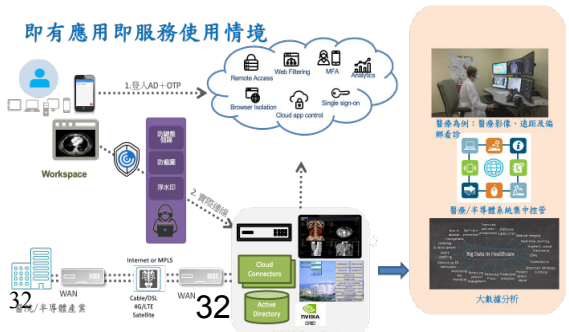
制定自主開發流程	整合性功能:以整合 LLM 服務為例
開發時程:	評估期:中小型開發 3 – 5工作天, 大型開發: 5 – 7工作天 開發期:中小型開發 3 – 5工作天, 大型開發: 5 – 7工作天
服務上線SOP	程式碼檢測、封箱測試 文件/安排友善使用者教育訓練 開放測試 使用者線上文件與影音教材
Beta Test	邀請內部與外部使用者, 並進行意見回覆。

https://github.com/orgs/Zillaforge/repositories

31

完成類公有雲之虛擬機器與容器服務

- 雲平台具體實作
  - 提供台灣第一個「信任研究環境」(Trusted Research Environment, TRE)上線服務
  - 114年教育訓練累計提供行政院公共工程委員會與國網中心等 14個部會參與
  - 可整合其他雲平台 – 如前瞻機敏混合雲，達到可動態配置機算資源，提升服務可用度
- Openstack架構與高資安防禦機制雲平台
  - 提供產業高可用率服(SLA>99.96)
  - 提供生醫團隊進行使用
  - 主要提供雲服務之要件
    - 一站式Portal – 虛擬機/容器服務
    - 高隱私區:於隱私區提供高監控的環境，而一般通用區達到VCS/VPS的虛擬化服務提供。

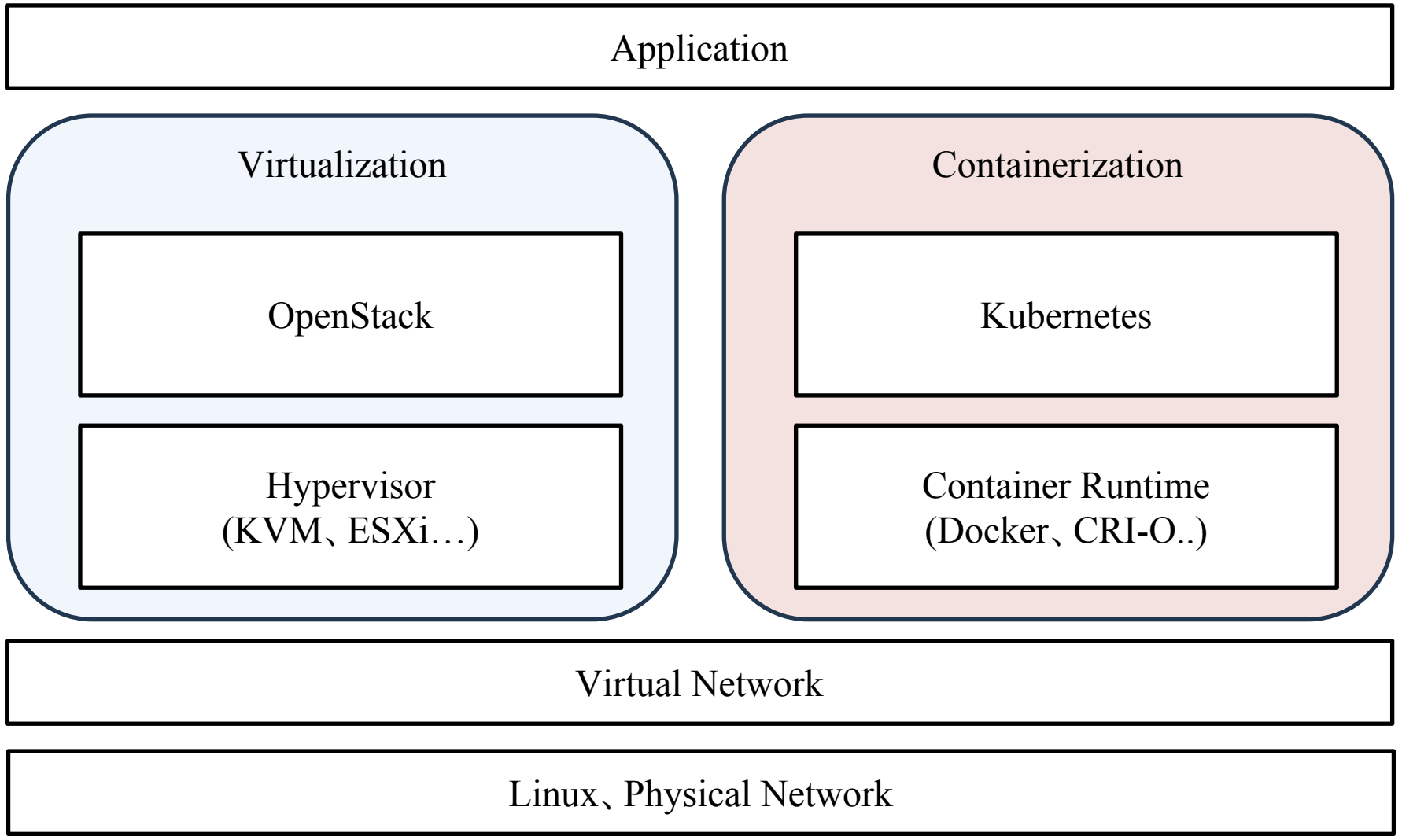


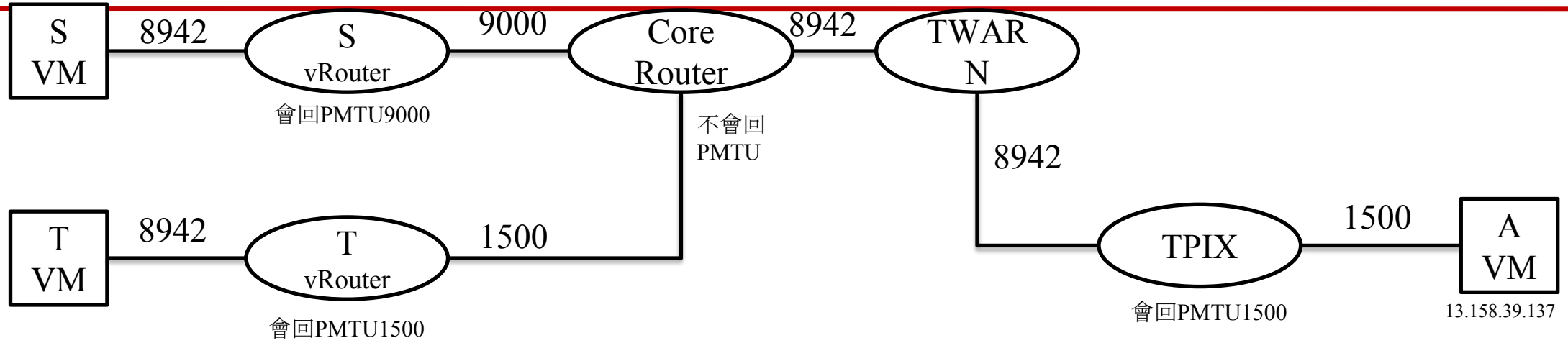
32

NARLabs
國家實驗研究院

國網中心
雲端平台架構概述

你該知道什麼？


Application
Virtualization
OpenStack
Hypervisor
(KVM、ESXi…)
Containerization
Kubernetes
Container Runtime
(Docker、CRI-O..)
Virtual Network
Linux、Physical Network



1. VM-T 連線至VM-S上的 HTTP Server，回傳 2000byte 的資料，會失敗
2. VM-S連到VM-T 上的HTTP Server，就可以正常回傳。
    1. 10分鐘內，VM-T 連線至VM-S也可以正常回傳。
3. VM-A，和VM-T，不論誰是HTTP server，都可以正常回傳

等等，IP封包不是應該會切割嗎?

NARLabs

ip.host == 140.110.160.47 and tcp.port == 41758

No.	Time	Source	Destination	Protocol	Length	Info
8239	29.727806	140.110.160.47	140.110.139.103	TCP	74	41758 → 5123 [SYN] Seq=0 Win=62314 Len=0 MSS=8902 SACK_PERM TSval=3985862156 TSecr=0 WS=128
8241	29.729097	140.110.139.103	140.110.160.47	TCP	74	5123 → 41758 [SYN, ACK] Seq=0 Ack=1 Win=62230 Len=0 MSS=8902 SACK_PERM TSval=563779633 TSecr=3985862156 WS=128
8242	29.729813	140.110.160.47	140.110.139.103	TCP	66	41758 → 5123 [ACK] Seq=1 Ack=1 Win=62336 Len=0 TSval=3985862158 TSecr=563779633
8243	29.729837	140.110.160.47	140.110.139.103	HTTP	159	GET /large-json HTTP/1.1
8245	29.732152	140.110.139.103	140.110.160.47	TCP	66	5123 → 41758 [ACK] Seq=1 Ack=94 Win=62208 Len=0 TSval=563779636 TSecr=3985862158
8246	29.732893	140.110.139.103	140.110.160.47	TCP	233	5123 → 41758 [PSH, ACK] Seq=1 Ack=94 Win=62208 Len=167 TSval=563779636 TSecr=3985862158
8247	29.732950	140.110.160.47	140.110.139.103	TCP	66	41758 → 5123 [ACK] Seq=94 Ack=168 Win=62208 Len=0 TSval=3985862162 TSecr=563779636
8251	29.742708	140.110.139.103	140.110.160.47	TCP	66	[TCP Previous segment not captured] 5123 → 41758 [FIN, ACK] Seq=3039 Ack=94 Win=62208 Len=0 TSval=563779646 TSecr=3985862162
8252	29.743445	140.110.160.47	140.110.139.103	TCP	78	[TCP Dup ACK 8247#1] 41758 → 5123 [ACK] Seq=94 Ack=168 Win=62208 Len=0 TSval=3985862172 TSecr=563779636 SLE=3039 SRE=3040
9248	34.051247	140.110.160.47	140.110.139.103	TCP	78	41758 → 5123 [FIN, ACK] Seq=94 Ack=168 Win=62208 Len=0 TSval=3985866480 TSecr=563779636 SLE=3039 SRE=3040
9250	34.052691	140.110.139.103	140.110.160.47	TCP	66	5123 → 41758 [ACK] Seq=3040 Ack=95 Win=62208 Len=0 TSval=563783956 TSecr=3985866480

```
> Frame 8239: 74 bytes on wire (592 bits), 74 bytes captured (592 bits) on interface \Device\NPF_{4558A056-B8CE-4D5B-A967
> Ethernet II, Src: JuniperNetwo_82:fd:42 (20:93:39:82:fd:42), Dst: PaloAltoNetw_f2:80:10 (b4:0c:25:f2:80:10)
v Internet Protocol Version 4, Src: 140.110.160.47, Dst: 140.110.139.103
    0100 .... = Version: 4
    .... 0101 = Header Length: 20 bytes (5)
  > Differentiated Services Field: 0x00 (DSCP: CS0, ECN: Not-ECT)
    Total Length: 60
    Identification: 0xbdd4 (48596)
  v 010. .... = Flags: 0x2, Don't fragment
      0... .... = Reserved bit: Not set
      .1.. .... = Don't fragment: Set
      ..0. .... = More fragments: Not set
    ...0 0000 0000 0000 = Fragment Offset: 0
    Time to Live: 62
    Protocol: TCP (6)
    Header Checksum: 0x3a74 [validation disabled]
    [Header checksum status: Unverified]
    Source Address: 140.110.160.47
    Destination Address: 140.110.139.103
    [Stream index: 87]
> Transmission Control Protocol, Src Port: 41758, Dst Port: 5123,
```

```
0000  b4 0c 25 f2 80 10 20 93  39 82 fd 42 08 00 45 00   ··%··· · 9··B··E·
0010  00 3c bd d4 40 00 3e 06  3a 74 8c 6e a0 2f 8c 6e   ·<··@·>· :t·n·/·n
0020  8b 67 a3 1e 14 03 03 fe  7b f7 00 00 00 00 a0 02   ·g····· {·······
0030  f3 6a 60 58 00 00 02 04  22 c6 04 02 08 0a ed 93   ·j`X···· "·······
0040  6e 0c 00 00 00 00 01 03  03 07                     n······ ··
```

IP_MTU_DISCOVER (since Linux 2.2)

Set or receive the Path MTU Discovery setting for a socket. When enabled, Linux will perform Path MTU Discovery as defined in RFC 1191 on **SOCK_STREAM** sockets. For non-**SOCK_STREAM** sockets, **IP_PMTUDISC_DO** forces the don't-fragment flag to be set on all outgoing packets. It is the user's responsibility to packetize the data in MTU-sized chunks and to do the retransmits if necessary. The kernel will reject (with **EMSGSIZE**) datagrams that are bigger than the known path MTU. **IP_PMTUDISC_WANT** will fragment a datagram if needed according to the path MTU, or will set the don't-fragment flag otherwise.


NCHC 雲平台
服務
NCHC-可信賴雲
可信賴雲
莊家雋
歡迎使用可信賴雲平台！
專案名稱： NCHC-可信賴雲開發測試計畫-II
專案 ID： GOV113097
建立虛擬機器
建立容器映像檔
服務
虛擬平台服務
虛擬機器
虛擬磁碟
負載平衡
自動擴展
安全群組


NCHC 雲平台
服務
可信賴雲
莊家雋
主頁 > 虛擬機器 > 建立虛擬機器
建立虛擬機器
基本設定
硬體設定
虛擬網路
儲存資訊
認證
初始化指令
檢閱+建立
查看資源使用量 / 配額
型號
GPU (張)
CPU (Cores)
記憶體 (GB)
開機磁碟 (GB)
Basic.small 0 2 8 120
Memory.medium 0 8 64 120
檢閱+建立
<上一步
下一步：虛擬網路>
取消

整體系統架構

事件通知插件　Qumulo 使用量代理

底層系統與管理

- 系統管理網站
- API閘道　監控認證插件
- 收量服務
- 告警服務
- 簡單通知服務
- 資料庫
  - PostgreSQL
  - MariaDB
  - Redis
- 監控系統
  - 系統整合測試單元
  - Prometheus
  - Grafana
- 日誌管理
  - 日誌追蹤服務
  - ELK
- OpenStack
  - Horizon
- Ceph
  - Dashboard
- 內部DNS服務

國網中心雲平台

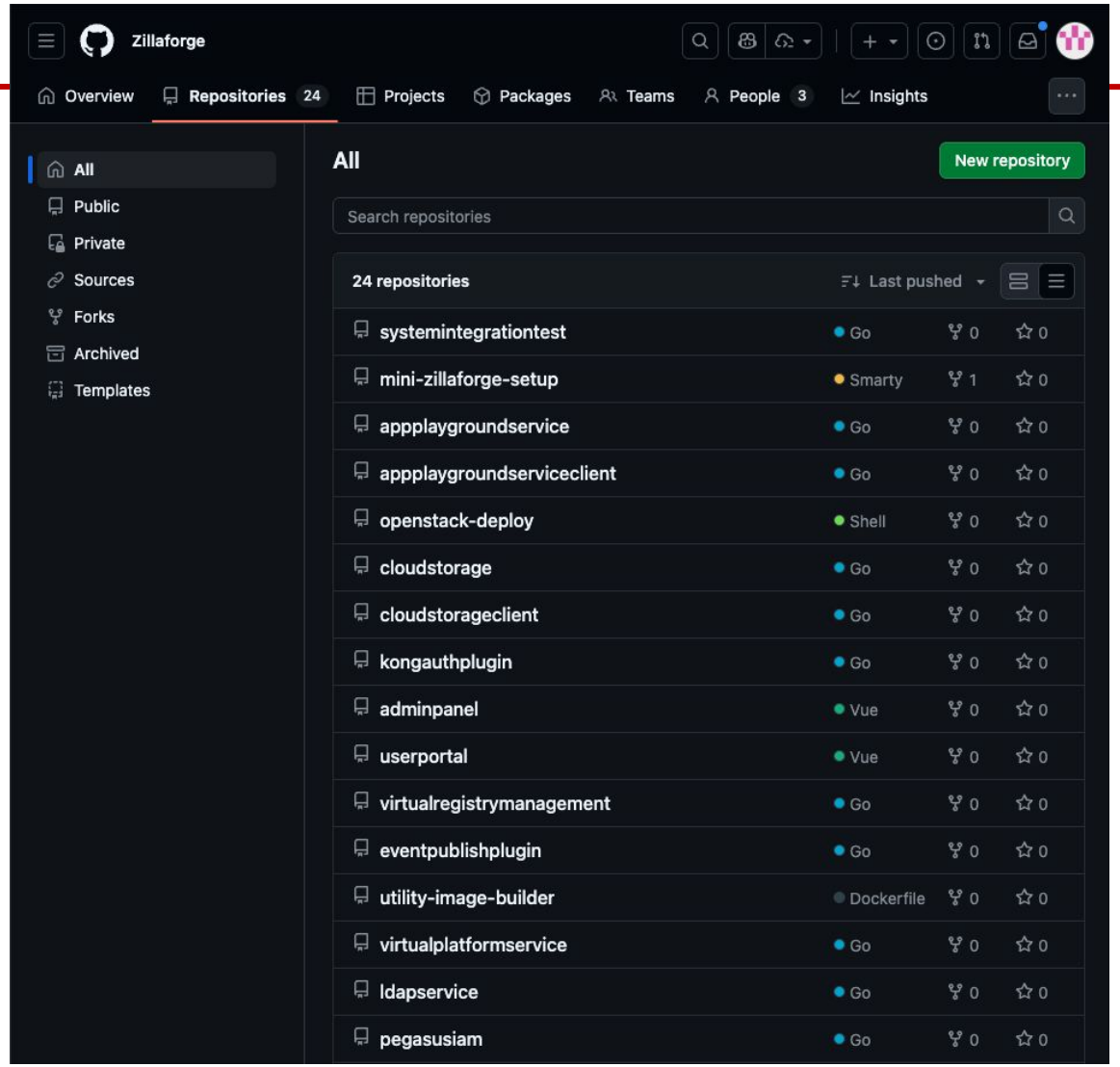
- 國網中心 VDI
- 私有使用者介面
- 私有管理者介面
- 公共管理者介面
- 公共使用者介面
- 認證插件　私有API閘道　回應轉換插件
- 認證插件　公共API閘道　回應轉換插件
- LDAP服務
- 國網中心認證插件　身分認證　國網中心事件代理
- K8S集群服務
- 應用程式服務
- 遠端派送服務
- 虛擬平台服務
- 容器映像檔管理
- 雲端儲存
- 資源轉移服務
- 檔案儲存
- 虛擬映像檔管理

底層

- VM
- Kubernetes
- 系統共用平台
- 一般區 儲存系統
- 機敏區 儲存系統
- 底層虛擬化系統
- 網路交換器 / 防火牆 / 資安設備
- 系統節點
- 管理節點
- CPU 與 GPU 計算節點

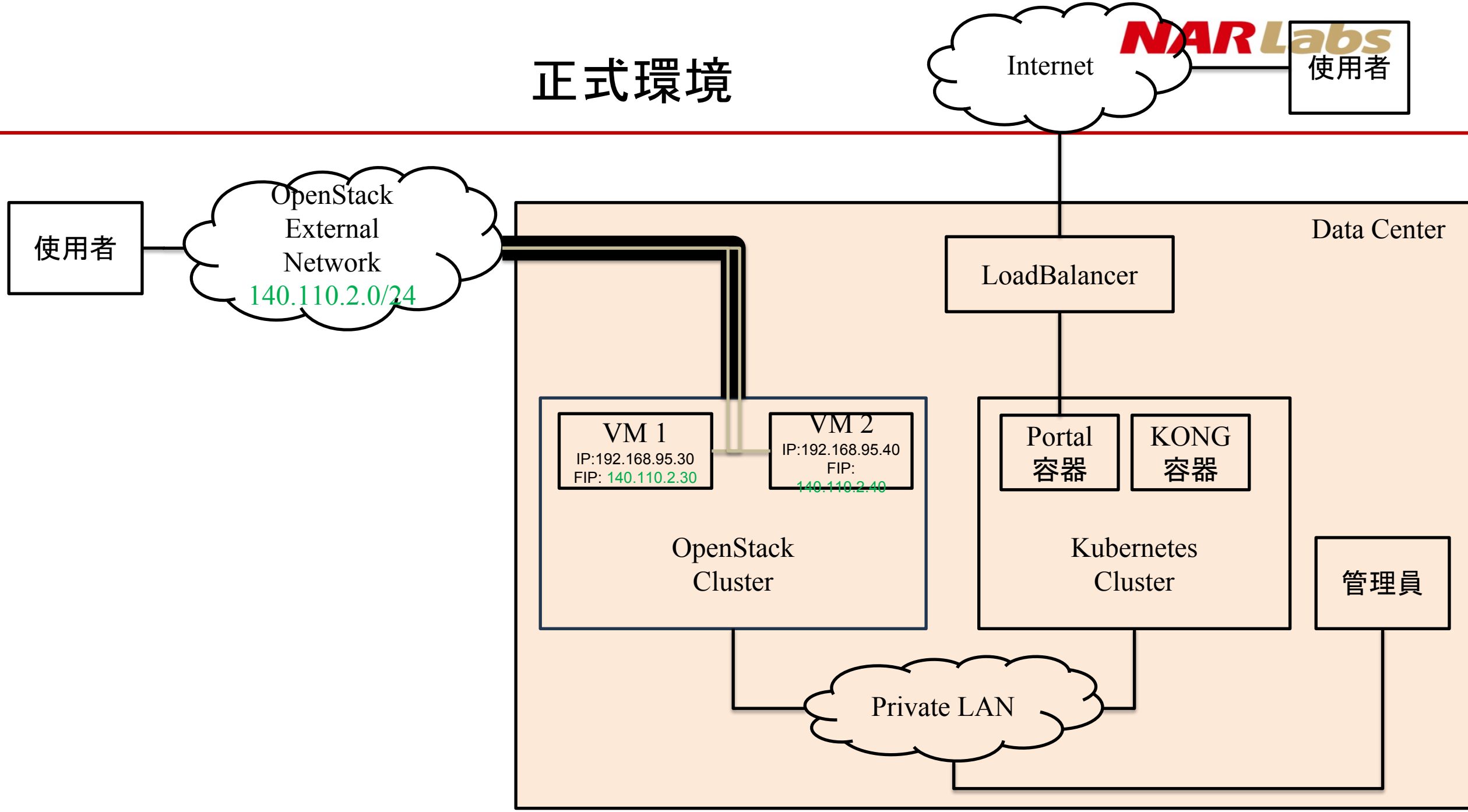
NARLabs

https://github.com/orgs/Zillaforge/repositories

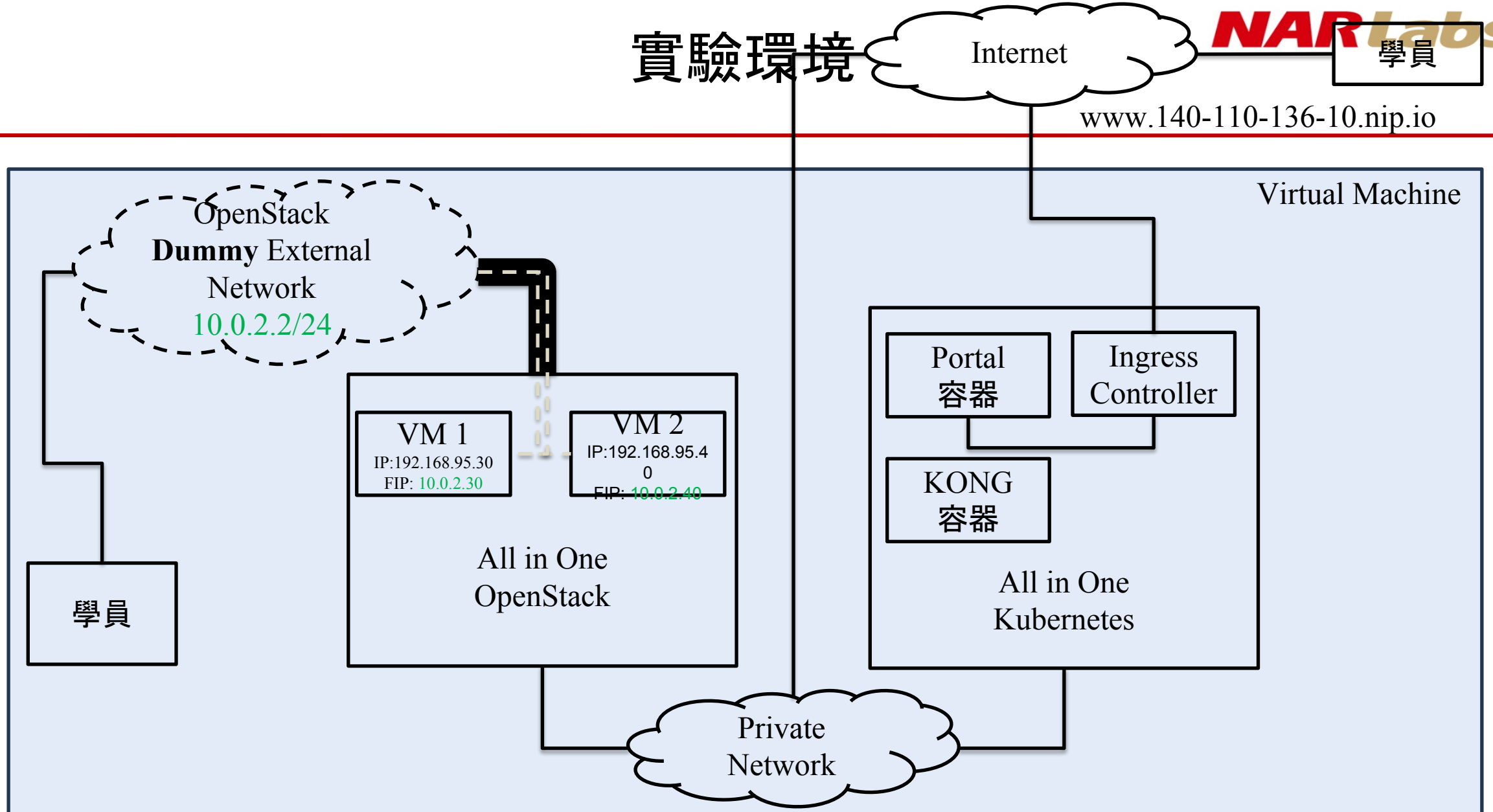


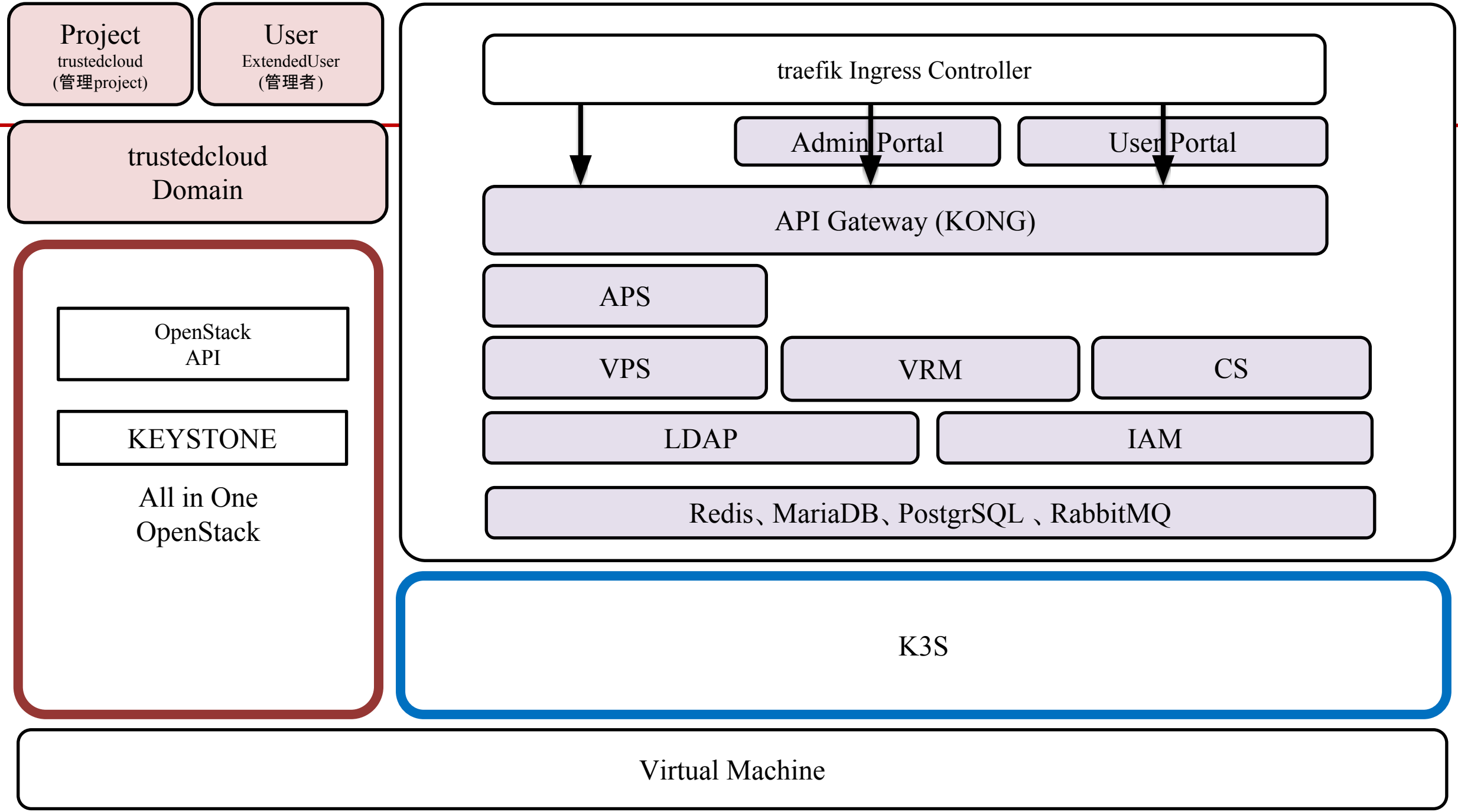
- 部署測試環境:
  - Mini-zillaforge-setup
  - Openstack-deploy
- Portal
  - adminapnel
  - userportal
- 核心功能
  - xxxxx & xxxxClient

正式環境


Internet
使用者
使用者
OpenStack
External
Network
140.110.2.0/24
Data Center
LoadBalancer
VM 1
IP:192.168.95.30
FIP: 140.110.2.30
VM 2
IP:192.168.95.40
FIP:
140.110.2.40
OpenStack
Cluster
Portal
容器
KONG
容器
Kubernetes
Cluster
管理員
Private LAN

實驗環境


Internet
學員
www.140-110-136-10.nip.io
Virtual Machine
OpenStack
Dummy External
Network
10.0.2.2/24
VM 1
IP:192.168.95.30
FIP: 10.0.2.30
VM 2
IP:192.168.95.40
FIP: 10.0.2.40
All in One
OpenStack
學員
Portal
容器
Ingress
Controller
KONG
容器
All in One
Kubernetes
Private
Network


Project
trustedcloud
(管理project)
User
ExtendedUser
(管理者)
trustedcloud
Domain
OpenStack
API
KEYSTONE
All in One
OpenStack
traefik Ingress Controller
Admin Portal
User Portal
API Gateway (KONG)
APS
VPS
VRM
CS
LDAP
IAM
Redis、MariaDB、PostgrSQL 、RabbitMQ
K3S
Virtual Machine

NARLabs

前端網站

User Portal

Admin Panel

Kong Plugin

KongAuthPlugin

KongResponse TransformerPlugin

輔助元件

SystemIntegrationTest

核心元件 gRPC Server

LDAP Service

appplaygroundservice

EventPublishPlugin

PegasusIAM

Cloud Storage

virtualregistrymanagement

virtualplatformservice

核心元件 gRPC Client

appplaygroundservice Client

EventPublishPlugin Client

PegasusIAM Client

Cloud Storage Client

Virtualregistrymanagement Client

Virtualplatformservice Client

公用工具庫

toolkits

Metering toolkits

PegasusMsgQueue Client

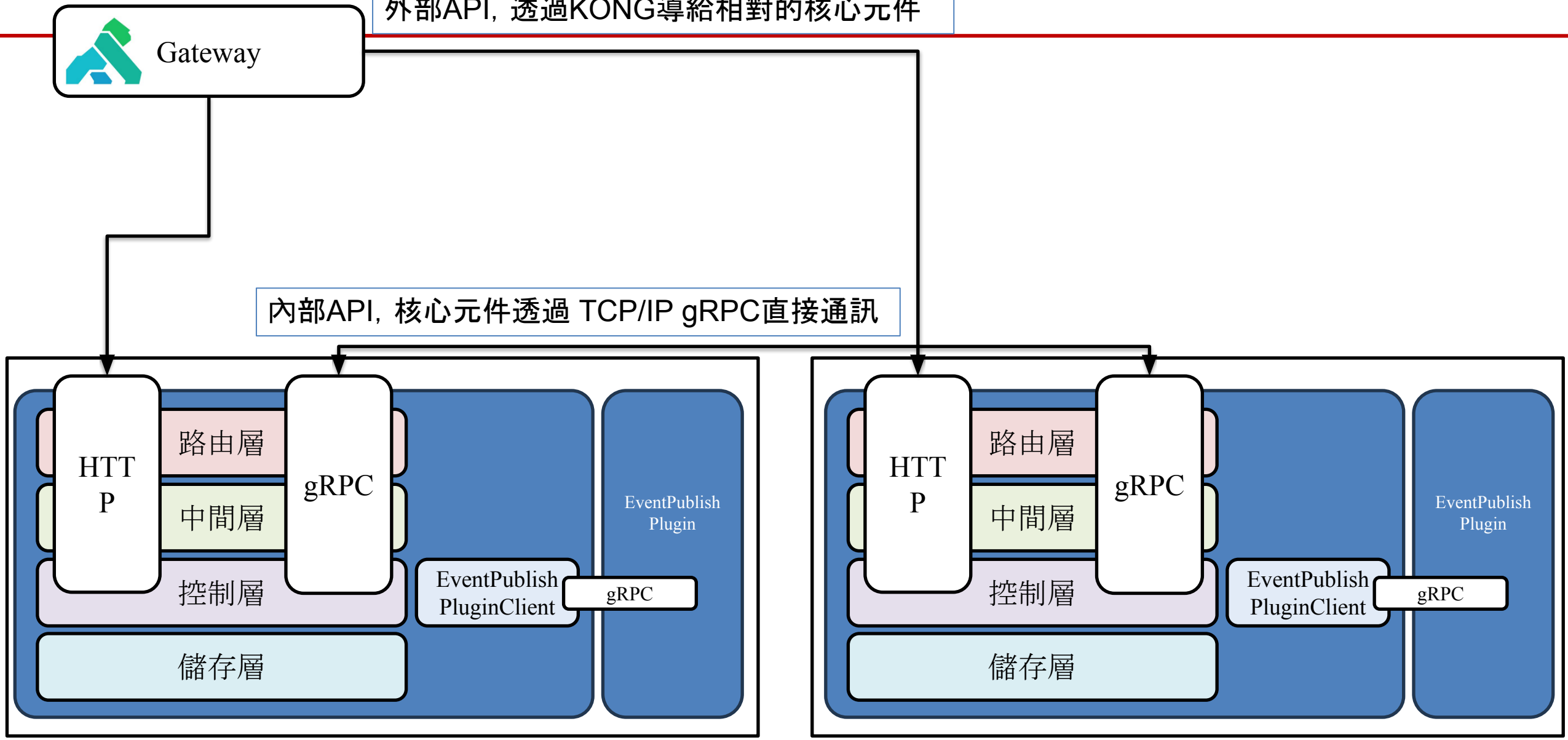
建置工具

Mini-zillaforge-setup

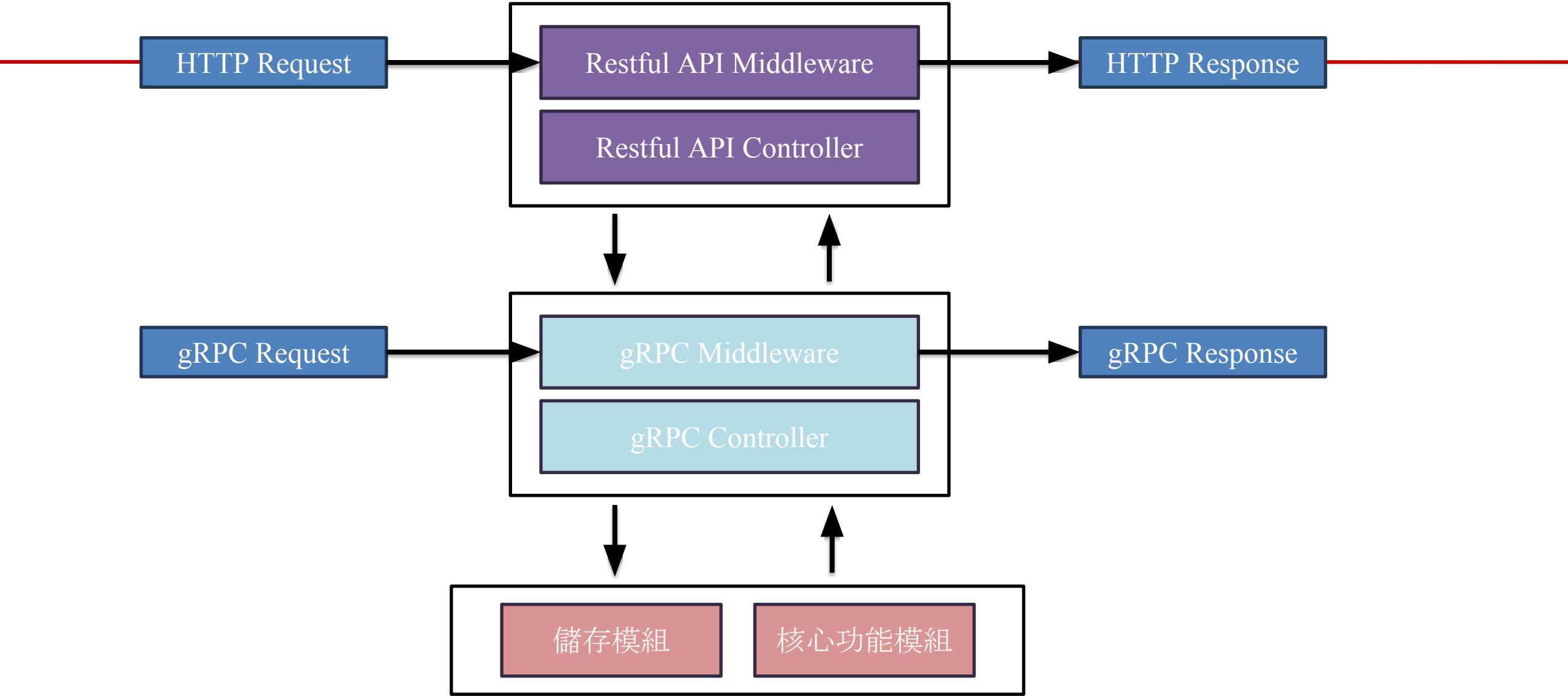
Openstack-deploy

utility-image-builder

外部API，透過KONG導給相對的核心元件



NARLabs


HTTP Request
Restful API Middleware
Restful API Controller
HTTP Response
gRPC Request
gRPC Middleware
gRPC Controller
gRPC Response
儲存模組
核心功能模組

容器化建置 (Containerized Build)

- 核心元件編譯與打包容器映像檔都是透過Docker完成
  - 在不同的環境皆可執行
  - 保證編譯的結果一致
- 在Windows上，建議使用WSL2與透過WSL2安裝Docker

後端元件:

```
# 準備建置用的環境Image
$ git clone https://github.com/Zillaforge/utility-image-builder
$ cd utility-image-builder
$ make release-image-golang

# Clone 核心元件程式碼
$ git clone https://github.com/Zillaforge/xxxxxx
$ make RELEASE_MODE=prod release-image
```

前端網頁:

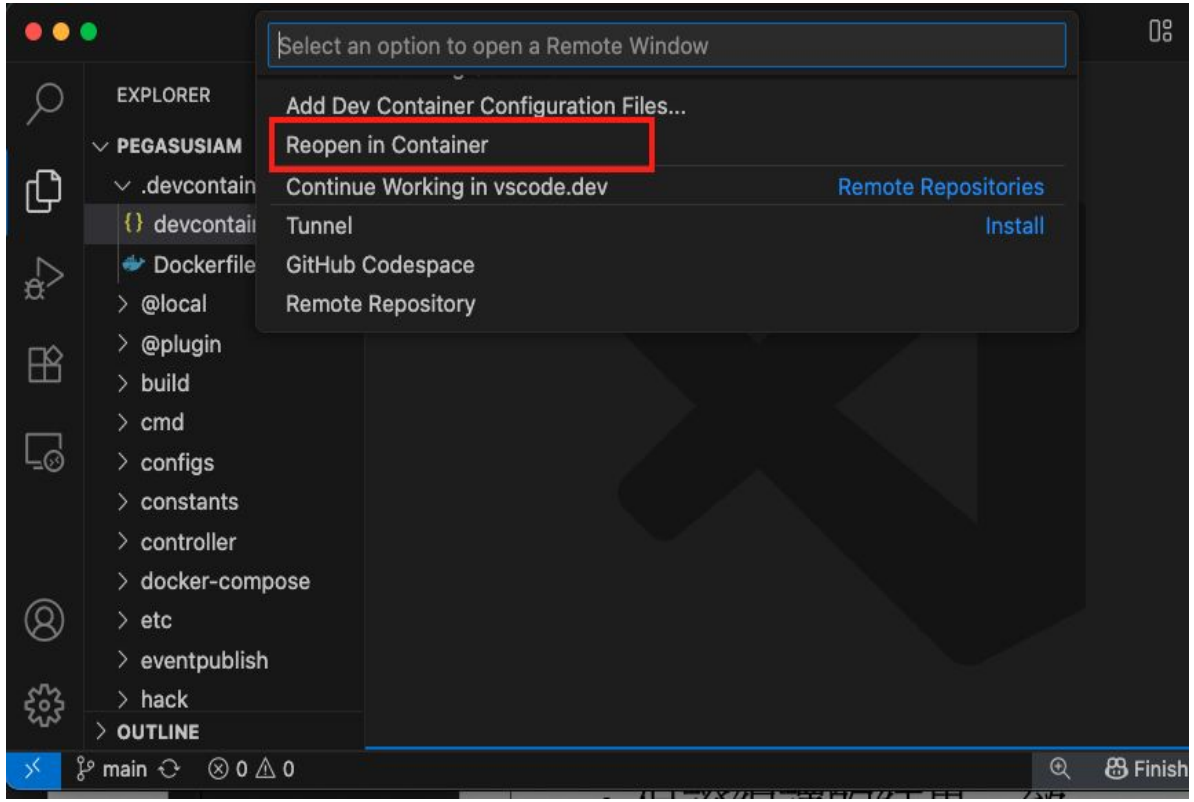
```
# Clone Portal (User & Admin) 程式碼

$ make release-image-public
```

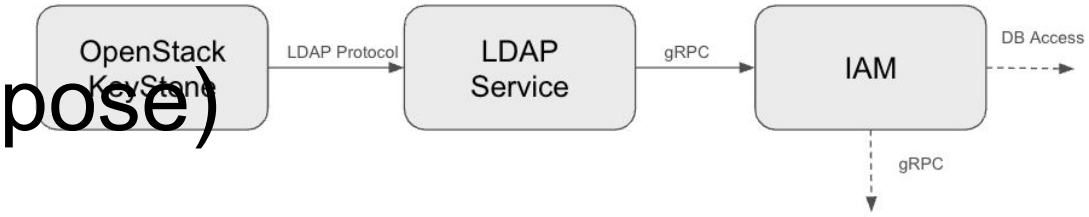
https://blog.miniasp.com/post/2025/06/14/How-to-remove-Docker-Desktop-and-install-Docker-Engine-on-Windows-with-WSL-2

容器化開發 (VSCode DevContainer)

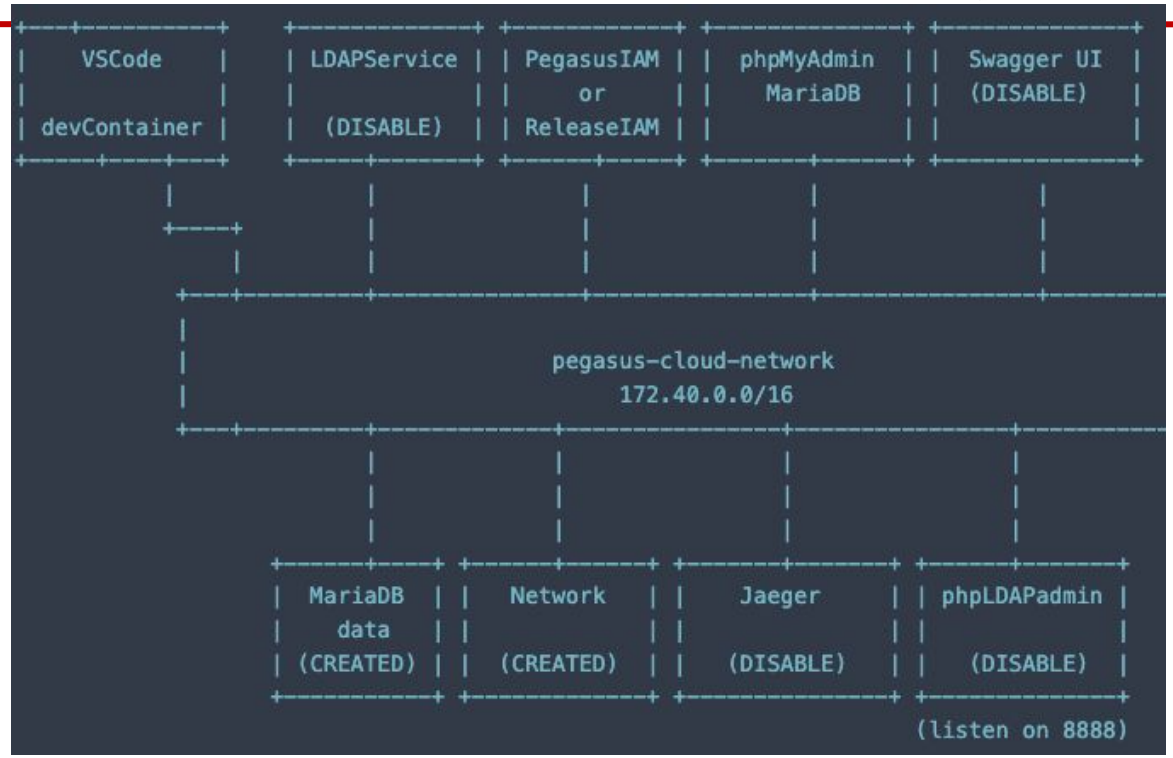
- 透過devcontainer
  - 完整隔離開發環境
  - 確保所有人有相同的開發環境
- 開發核心元件
  - 使用相同的Alpine Golang 開發環境
- 開發前端網端開發
  - 使用相同的Alpine NodeJS開發環境



容器化測試 (docker-compose)



- 透過docker-compose建立容器版測式環境
- 建立獨立的容器網路與Volume
- 建立必要的系統
  - Database、utility…
- 建立要測試以及相依的核心元件
  - 例如 : 測試LDAPSerivice, 要先啟動IAM




```
tree docker-compose
docker-compose
├── README.md
├── etc
│   └── PegasusIAM.yaml
├── persistent
│   ├── docker-compose.network.yaml
│   └── docker-compose.volume.mariadb.yaml
├── service
│   ├── docker-compose.iam.yaml
│   └── docker-compose.ldapservice.yaml
└── system
    ├── docker-compose.jaeger.yaml
    ├── docker-compose.ldap-ui.yaml
    ├── docker-compose.mariadb.yaml
    └── docker-compose.swagger.yaml
```

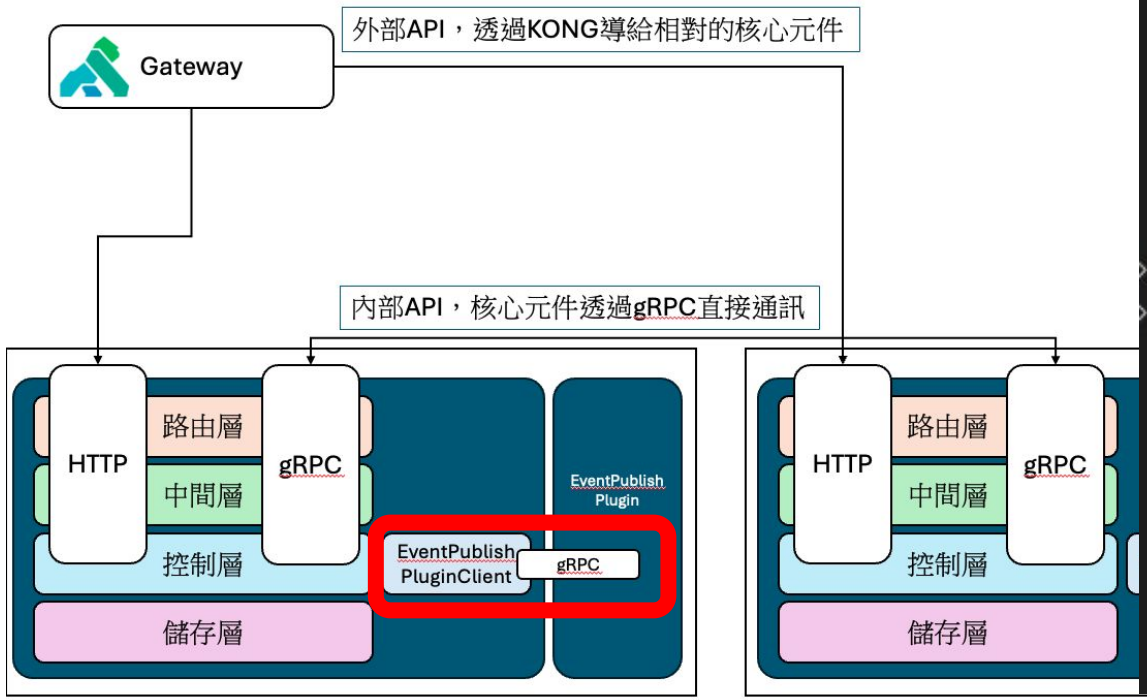
```
$ make start-dev-persistent
$ make start-dev-system
$ make start-dev-service
```

核心元件列表

簡稱	全名	功能
EPP	Event Publish Plugin	幫元件存取Redis
IAM	PegasusIAM	帳號管理
LDAP	LDAP Service	幫IAM提供LDAP介面
CS	Cloud Storage	提供S3服務
VRM	Virtual Registry Management	管理VM Image
VPS	Virtual Platform Service	虛擬化功能API
	User Portal	一般使用者操作雲平台
	Admin Portal	平台管理者對平台進行設定

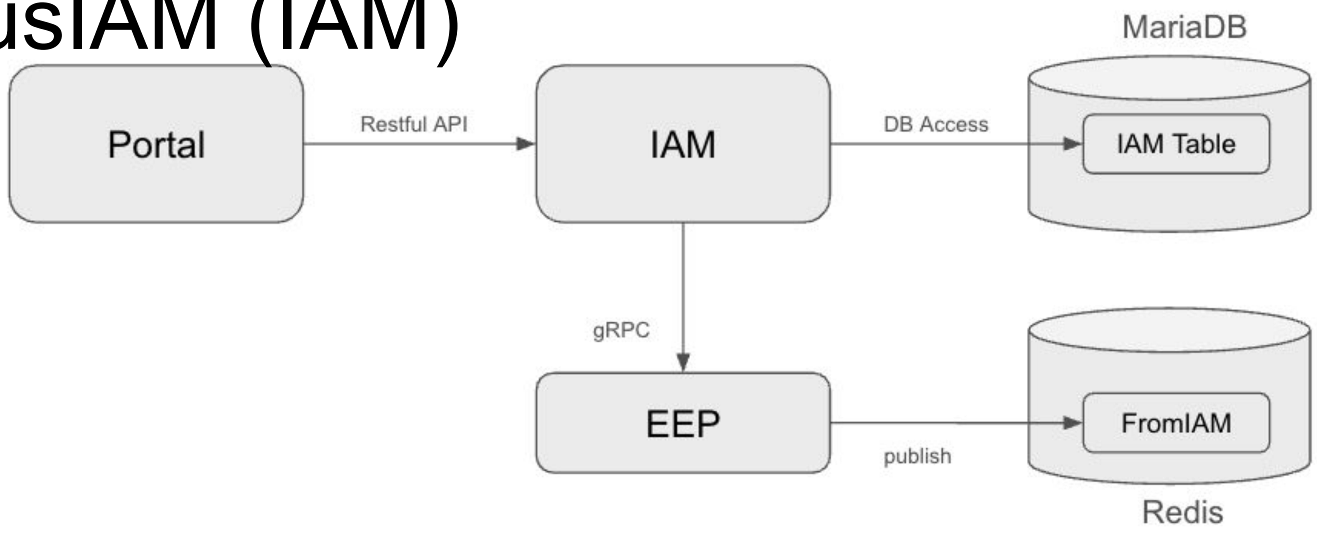
Event Publish Plugin (EPP)

- EPP不單獨部署，而是和每個核心元件在相同的Pod裡
- 核心元件透過unix socket 與EPP gRPC 互動，而不是透過TCP/IP




```
eventpublish:
  eventpublishPlugin:
    plugin:
      name: EventPublishPlugin
      instance: VirtualRegistryManagement-EventPublishPlugin-TrustClou
      version: 0.1.2
      socket_path: "/run/eventpublishplugin.sock"
      logger:…
      tracer:…
      service: my_redis_sentinel
      services:
      - name: my_redis_sentinel
        kind: redis_sentinel
        hosts:
        - redis-sentinel-service:26379
        password: password
        master_group_name: mymaster
```

PegasusIAM (IAM)



- 二個預設帳號
  - admin@ci.asus.com
  - system@ci.asus.com
- 一個預設專案
  - administrator
- IAM與OpenStack上的專案、帳號關係是一對一對應
  - 透過extra 欄位裡的資訊做對應

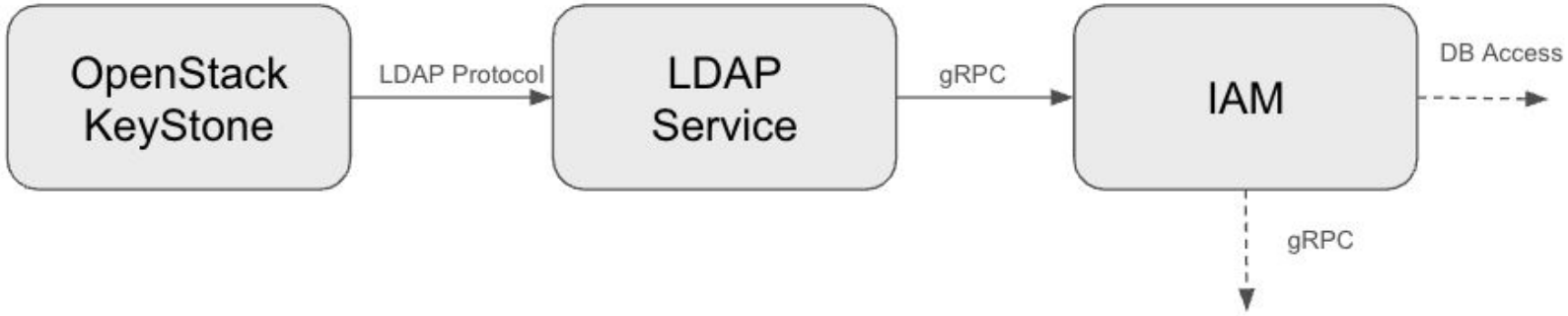
```
MariaDB [iam]> select account, namespace,extra from user;
+-------------------+-------------+-------------------------------------------------------------------------------------------+
| account           | namespace   | extra                                                                                     |
+-------------------+-------------+-------------------------------------------------------------------------------------------+
| system@ci.asus.com | ci.asus.com | {"tw-tc-ad1": {"opsk": {"uuid": "3265c76e39adc88974adec942f2f942d056711a2d4d509480f95fa03c16b07d4"}}} |
| admin@ci.asus.com  | ci.asus.com | {"tw-tc-ad1": {"opsk": {"uuid": "6737f08dd656d4c2530ca8d26d3e840a9075731b7c690c2857bc74ecc67057fa"}}} |
+-------------------+-------------+-------------------------------------------------------------------------------------------+
```

```
MariaDB [iam]> select display_name, extra from project;
+-----------------+------------------------------------------------------------------------------------------------------+
| display_name    | extra                                                                                                |
+-----------------+------------------------------------------------------------------------------------------------------+
| administrator   | {"iservice":{"projectSysCode":"trustedcloud"},"tw-tc-ad1":{"opsk":{"uuid":"470f8268da1e4e8988358a41cfb67fec"}}} |
```

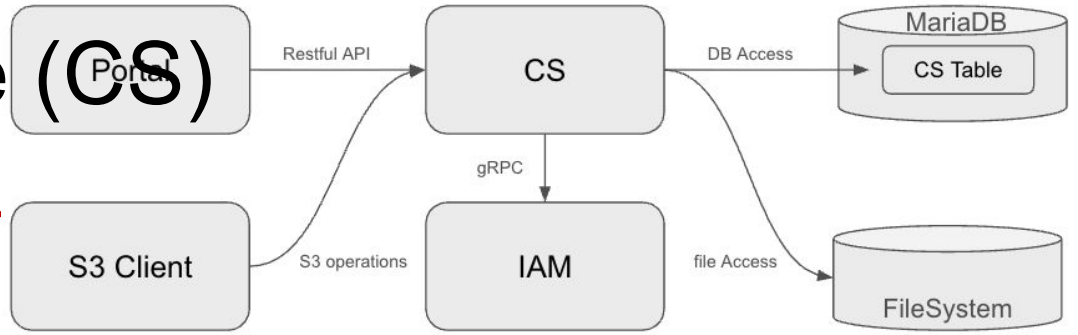
LDAP Service (LDAP)

- IAM提供HTTP 與 gRPC 通訊做帳號與專案的CRUD等操作
- 但應用程式大多支援LDAP做帳號管理

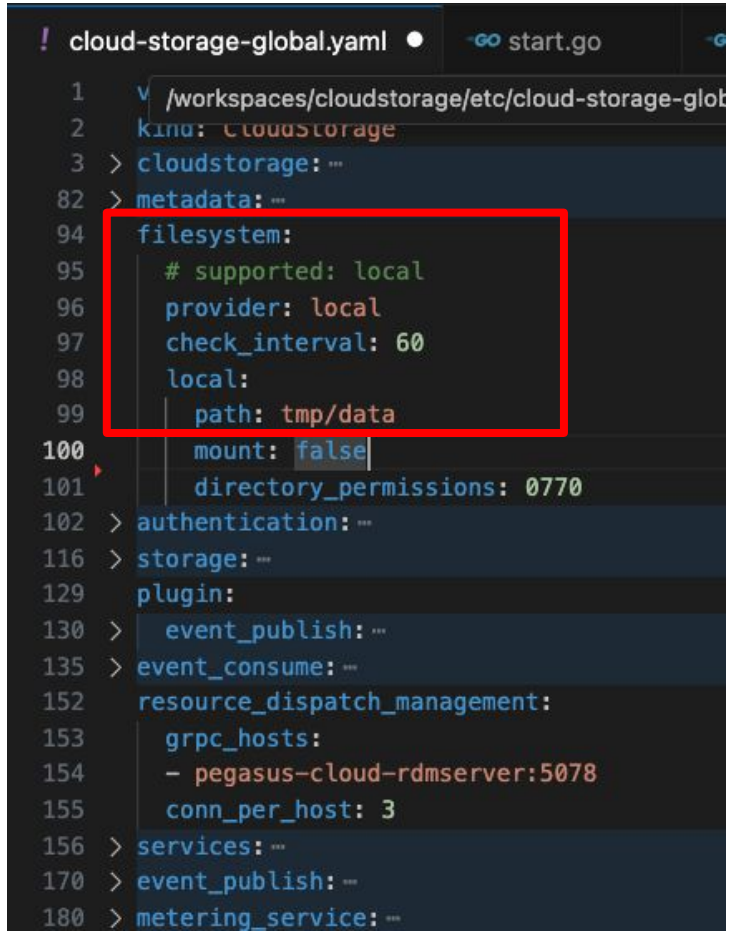
- 透過 LDAP Service 橋接 應用程式的LDAP Client與IAM
  - 不是既有的OpenSource LDAP 專案
  - 提供四種認證機制



Cloud Storage (CS)

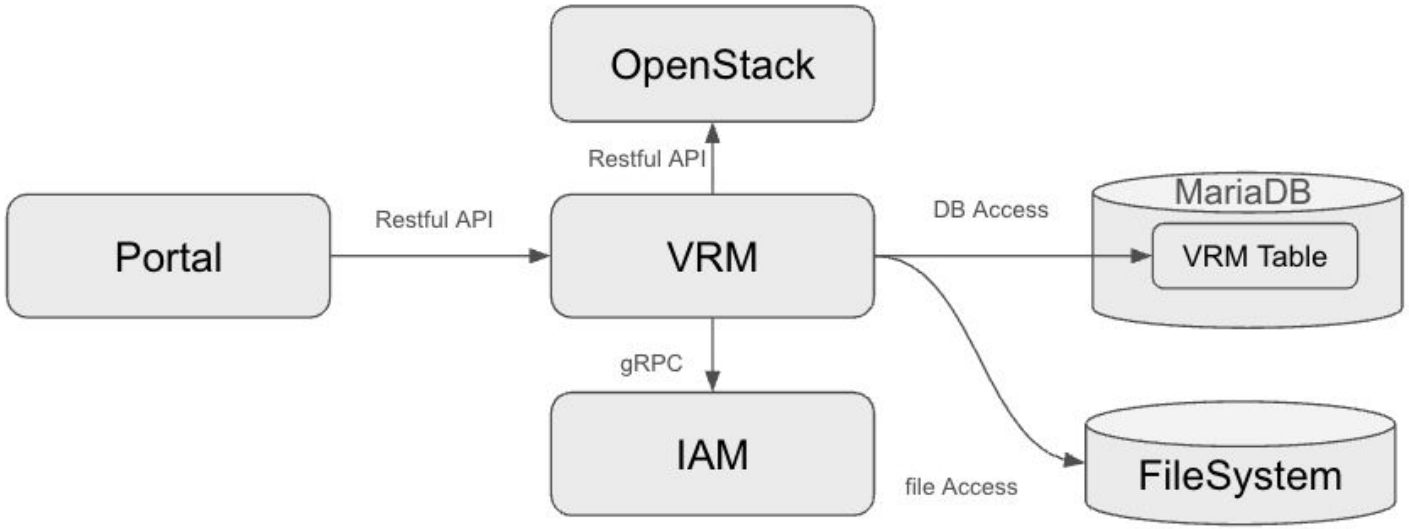


- 客制化開發S3相容的服務
  - 不是既有的 S3 Compatible Storage 專案
- 支援Path-Style 與 Host Style
- 資料是放在local filesystem
  - 可以是一個mount上的檔案系統
- 也會做為和其他服務資料交換的中間儲存
  - VRM 上傳VM Image，VPS轉存到OpenStack Glance



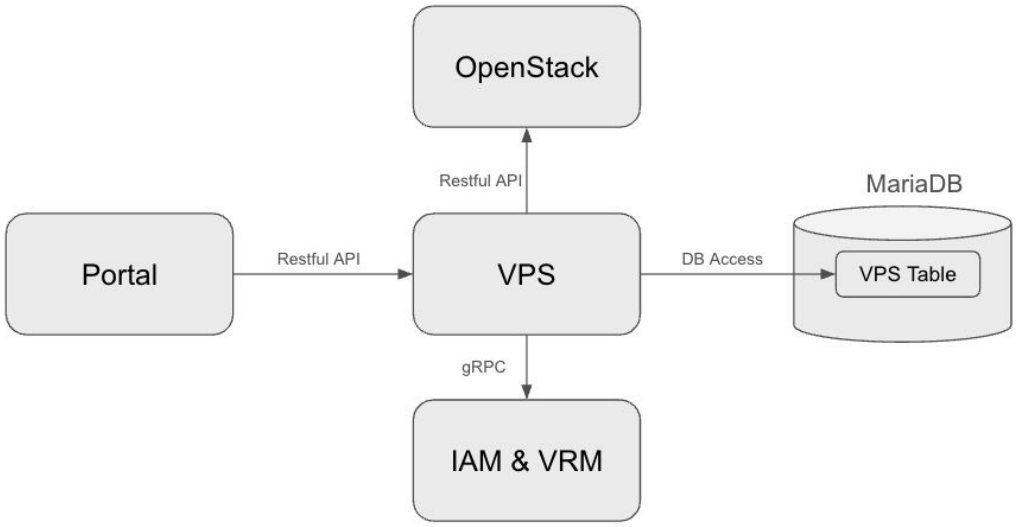
Virtual Registry Management (VRM)

- 管理雲平台上VM Image
  - 與OpenStack Glance上的對應關係



Virtual Platform Servic

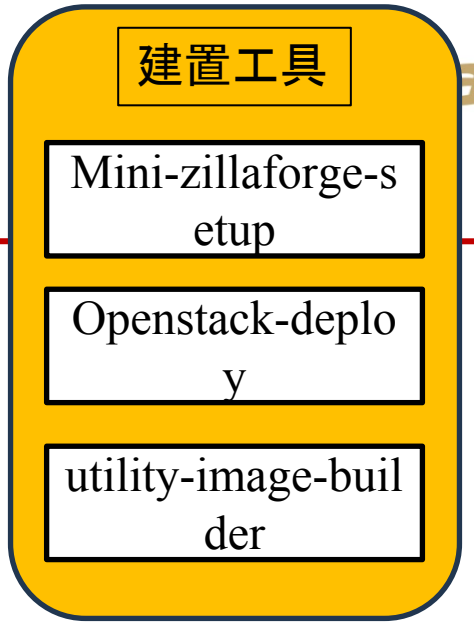
- 和OpenStack溝通，在OpenStack建立相對應的資源
- VPS 要在新Domain與Project啟用後才安裝



trustedcloud.yaml

```
1   version: VERSION_REPLACED
2   kind: VirtualPlatformService
3   virtualplatformservice:
53  storage:
64  authentication:
66  event_consume:
71  services:
87    openstack:
88      auth_url: "keystone_url" #!!!!!!openstack
89      domain: "trustedcloud"
90      admin_project: "trustedcloud"
91      admin: "test@trusted-cloud.nchc.org.tw"
92      admin_pwd: "password123"
93      ldap_enabled: true
94      pwd_method: "otp" # ["reverse", "otp"]
95      pwd_otp_secret: "pwd_otp_secret"
96      network:
97        cidr_blacklist: []
98      micro_version:
99        nova: "2.91" # since 2.92, nova API not allow to generate keypair
100       cinder: "3.70"
101       manila: "2.63"
102     project_name_iam_mapping: "Extra.iservice.projectSysCode" ###############
103   #  share_type: "default_share_type"
104     default_nova_az: "default"
105     gnocchi_granularity: 300
106     gnocchi_granularity_cpu_alarm: 60
107     enable_snapshot_actual_size: false #!!!
108     enable_gpu_quota: false   #!!!
109   manila:
112   vgpu_monitor:
115   roles:
118   timeout: # seconds
121     system_volume_type: "__DEFAULT__"
122     volume_type_whitelist: ["__DEFAULT__"]
123   opsk_role_mapping:
130 namespace:  #!!!
135   network_bonding_mode: true
136 job_sync_interval:
143 scheduler_duration: # every X hours, 0<X<=24 && 24%X==0
146   available_district: "tw-tc-ad1"
147 vm_customize:
155 vrm:
160 memcache:
```

實驗環境安裝



- 8 core CPU 16GB RAM 50G HDD 的 主機
- Ubuntu 2404 (Kolla-Ansible要求)
- 單一網路介面卡 (deploy script 要求)
- 若在VM上，要啟用nested virtualization

Cloud Provider	能否安裝雲平台	能否建立虛擬機	備註
可信賴雲平台	可	可	
機敏雲平台	可	可	
Azure	可	可*	僅部份flavor支援 nested virtualization
TWCC	可	否	不支援nested virtualization
AWS	可	否	不支援nested virtualization
GCP	否*	可	在OS無法取得Netmask, 無法設定VIP

安裝步驟

```
$ git clone https://github.com/Zillaforge/mini-zillaforge-setup.git


$ cd mini-zillaforge-setup
$ git submodule update --init --recursive

# 1. Prerequisites
$ ./prerequisite.sh
$ source ~/.bashrc


# 2. Install services
$ ./install.sh

# 3. Configure integrations
$ ./post-configuration.sh
```

- 準備基礎環境
  - prerequisite.sh
- 安裝
  - install.sh
- 後續設定
  - post-configuration.sh

NARLabs

2025/6/12

- GCP 全球當機 11:46 ~ 18:27