NetApp

以AI專用架構，高效分析資料價值

Full Stack AI Data and Experiment Management Across a Hybrid Cloud

Lawrence Hsu/許宏俊
資深技術顧問
2020/08/21

大綱

- Machine Learning WorkFlows(AI 模型)
- AI/Container Computing 及Storage 特性
- NetApp 產品特色
- AI 架構流程展示
- VDI防疫包淺談

NetApp

Machine Learning WorkFlows(AI 模型)

人工智慧的演進

ARTIFICIAL INTELLIGENCE

Techniques that enable computers to mimic human intelligence.

1950’s

MACHINE LEARNING

Ability to learn without being explicitly programmed

1980’s

DEEP LEARNING

Learning based on Deep Neural Network

2010’s

NetApp

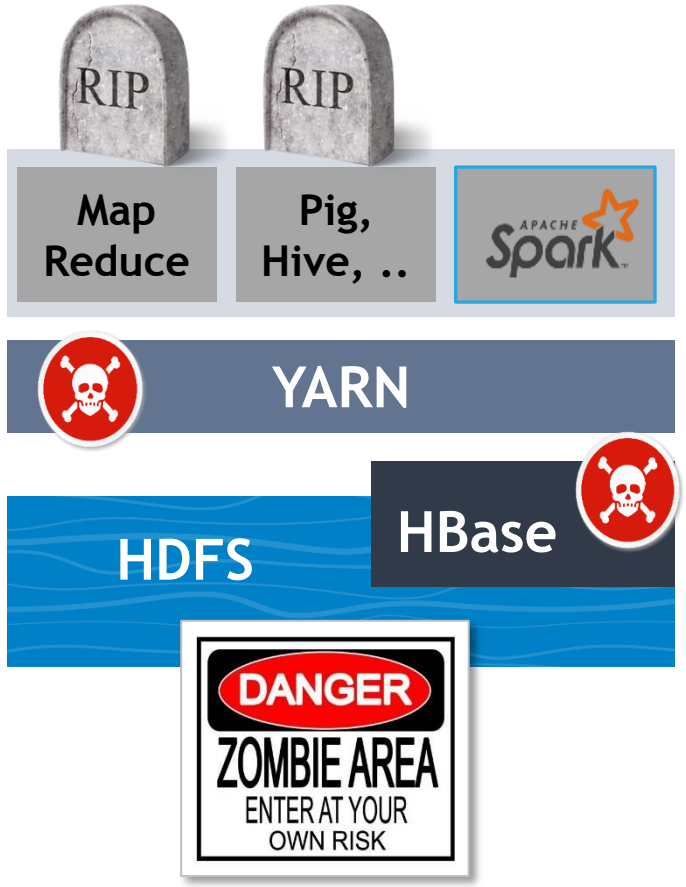
Goodbye Hadoop, hello Auto ML

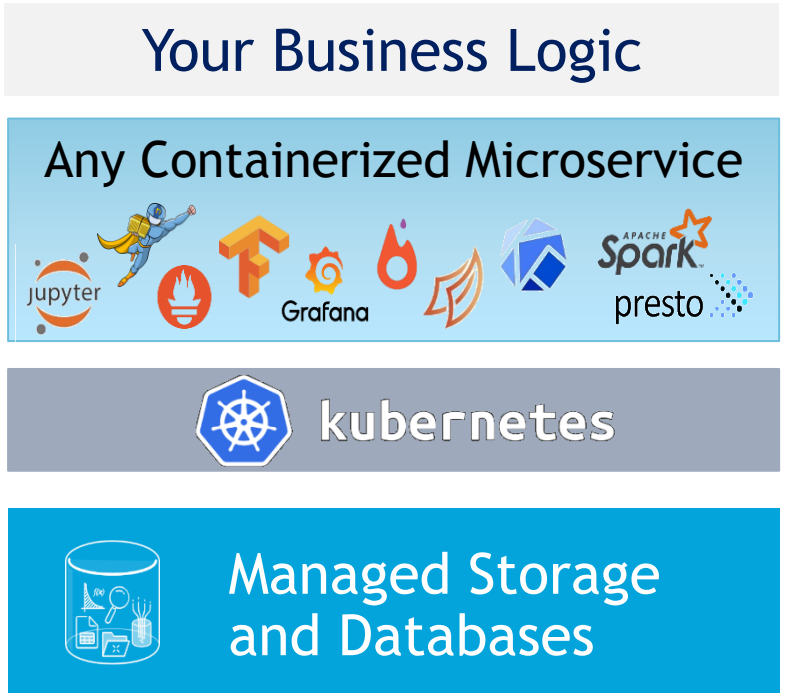
AutoML成主流AI服務標配

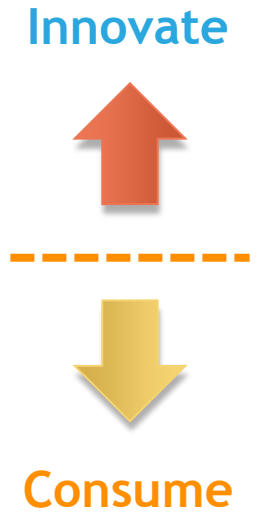
Middleware

Orchestration

Data

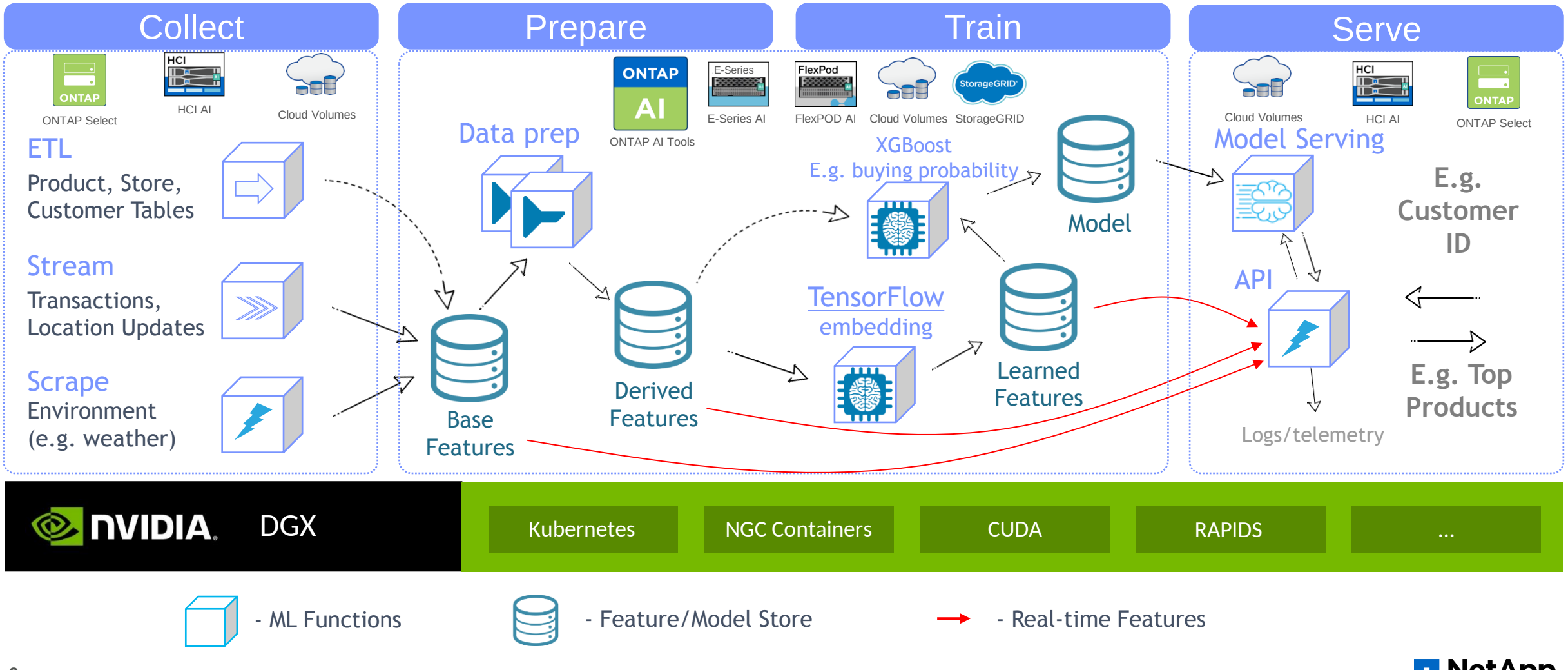






NetApp

AI/ML的流程分析



面臨的挑戰

- 資料是新一代的原始碼
  - 需要能夠保存版本化的AI模型基準
- 數據工作區
  - 需要有能力對海量數據集進行實驗，而不必擔心弄亂重要資料來源
- 可追溯性和合規性
  - 需要具備將每個生產模型追溯到用於訓練的確切數據集的能力
- 不同的團隊在不同的位置工作
  - 需要具備匯總和統一AI計算孤島的能力
- 許多異構數據來源
  - 需要具備將不同數據源集成到統一AI數據管道中的能力

AutoML 組成元件範例

Example for an ONTAP AI environment

GPU compute node(s)
For high-performance compute workloads

Storage node(s)
For high-performance data storage

Management node(s)
Used for cluster management

Provisioning node
Orchestrates the initial setup of the cluster

Deploy Kubeflow

Deploy additional services

Deploy basic services

Deploy TRIDENT

Deploy kubernetes on compute node(s)

Deploy kubernetes on management node(s)

Provision compute node(s)

Prepare management node(s)

Prepare provisioning node

Initial infrastructure provisioning

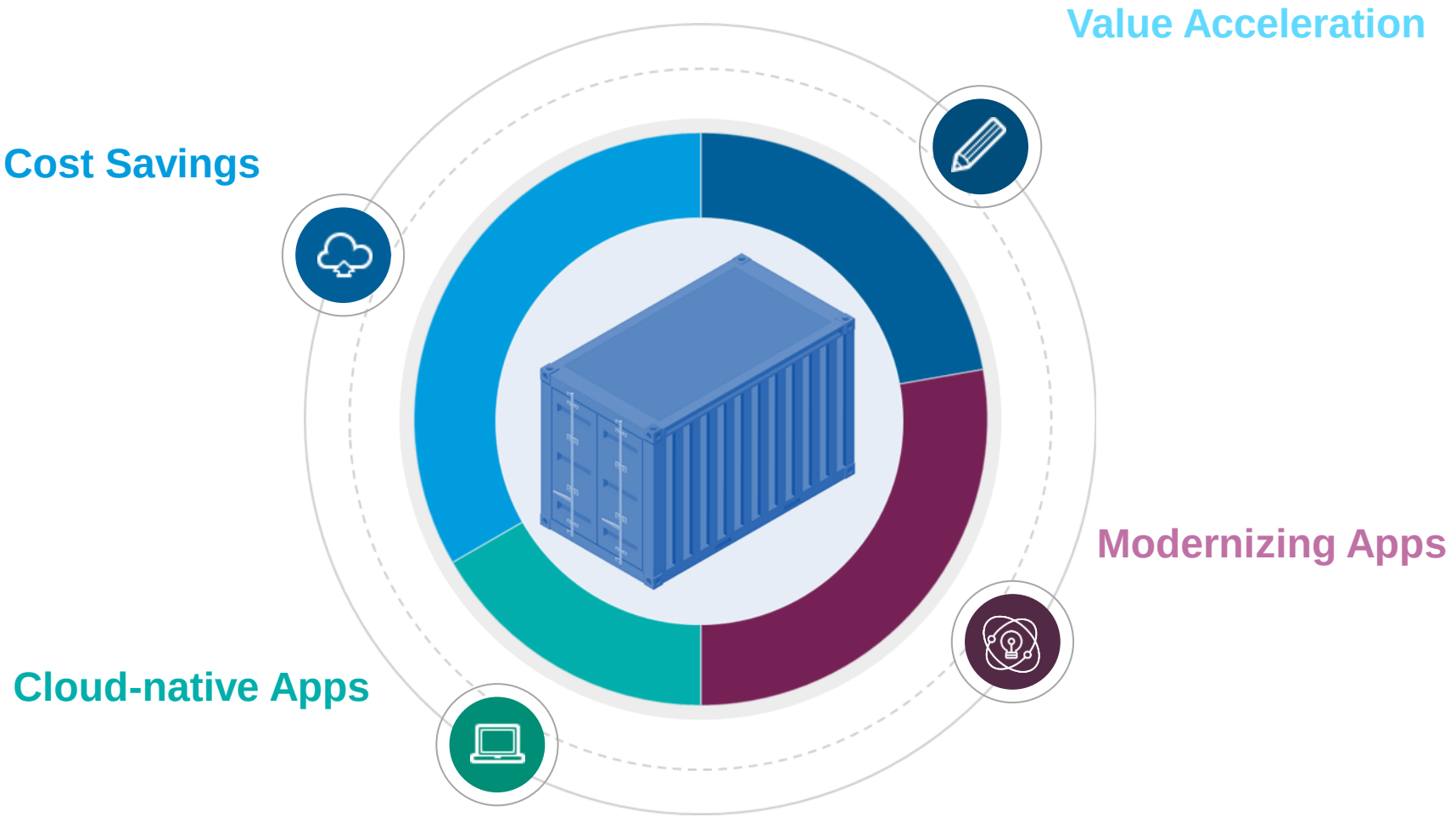
DeepOps

Picture source: NVIDIA POD reference architecture (with NetApp HCI added)

NetApp

AI/Container Computing 及Storage 特性

為何要用Containers?



NetApp

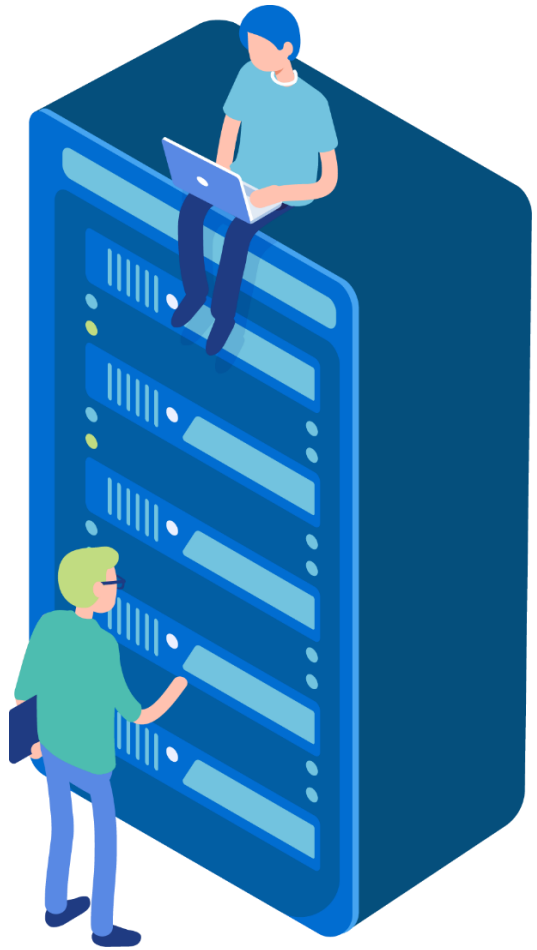
為何需要永久性儲存?

Containers are ephemeral — data is not

There are no real stateless applications

Consuming storage should be easy

Data persistence requirements in the enterprise have not changed



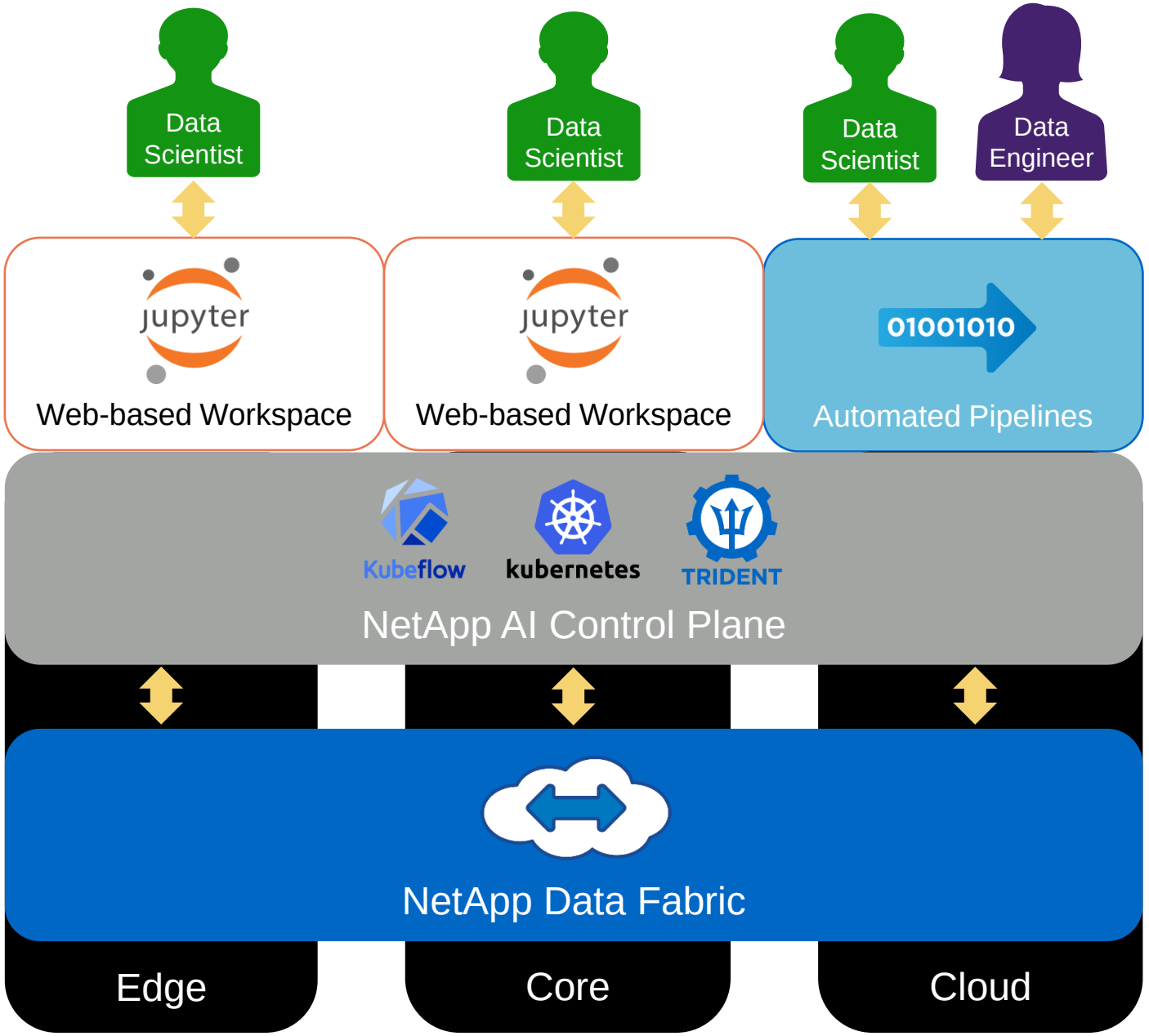
NetApp

NetApp 產品特色

NetApp AI Control Plane

Full-stack AI data and workload management

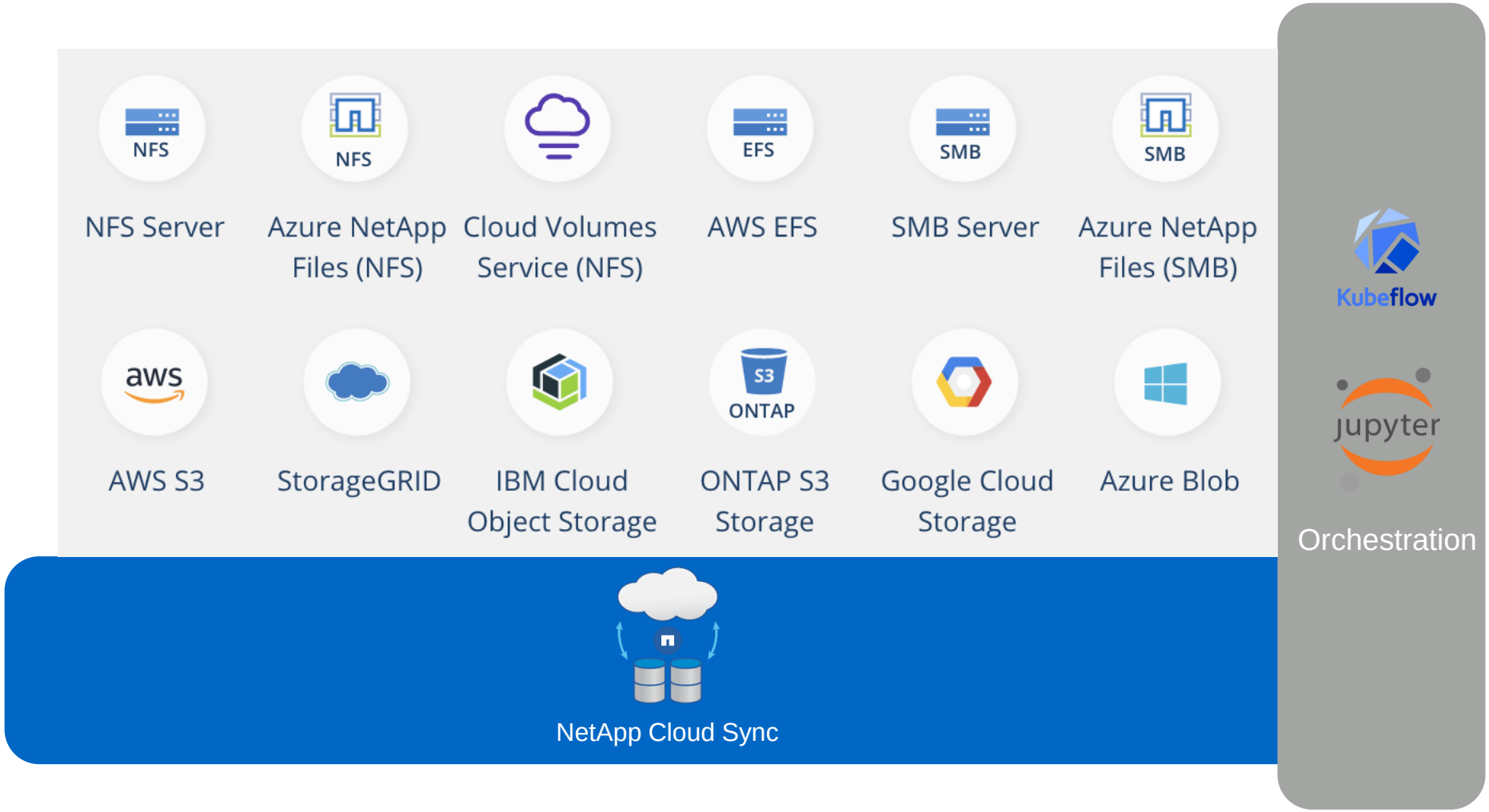
1) On-demand Jupyter Workspaces
2) 自動化數據準備，培訓和部署工作流程
3) 工作空間和工作負載可以跨越邊緣，核心和雲
   - Choice of any compute and/or cloud
   - Cross-site Data Scientist collaboration
4) 內建版本控制
   - Full dataset to model traceability
   - Seamlessly switch between model versions for dev/test, A/B testing, etc.



NetApp

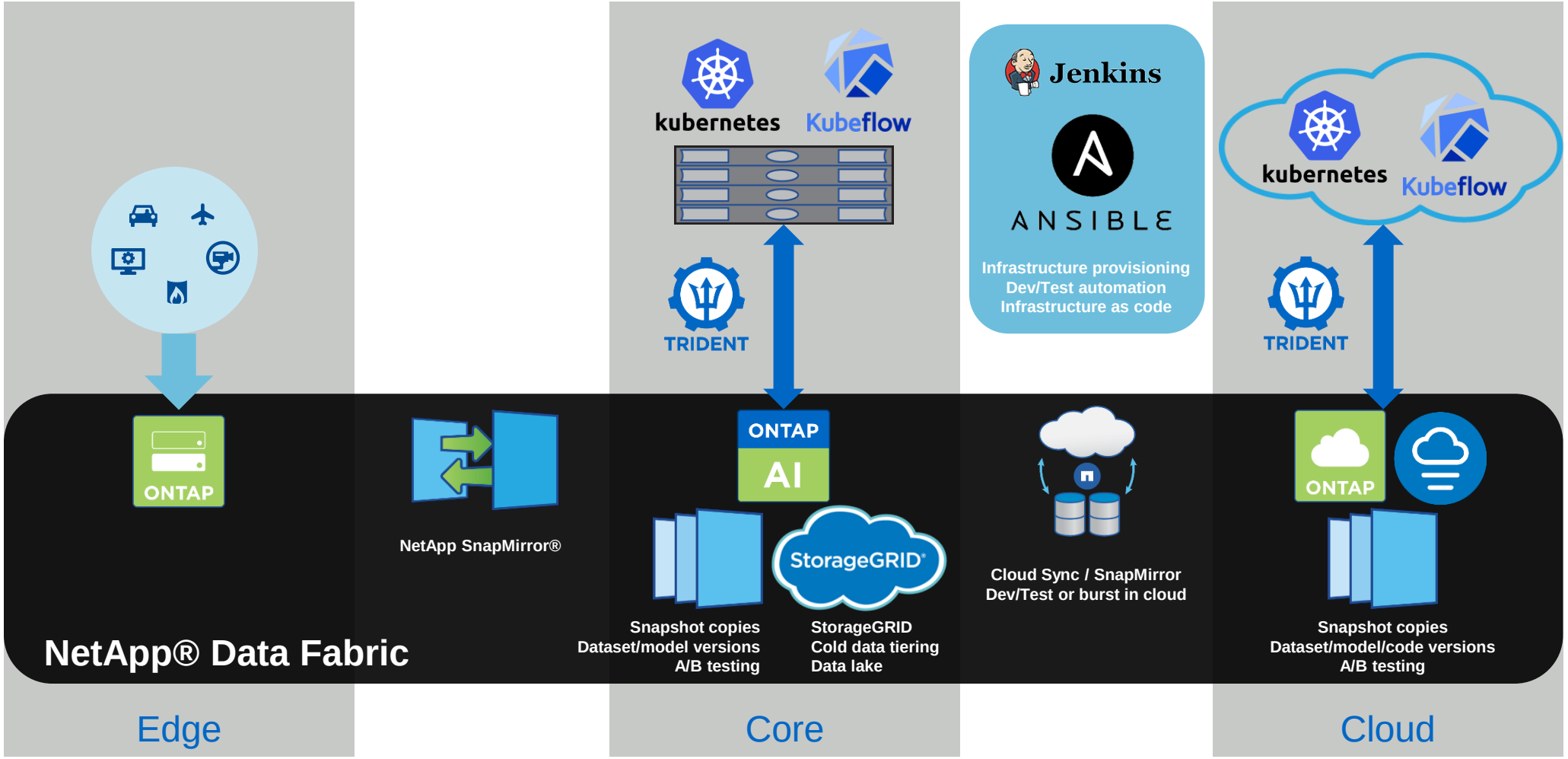
統一數據管道

Integrate heterogeneous data sources; centrally orchestrated



靈活的架構

Data and workloads are available whenever and wherever they are needed



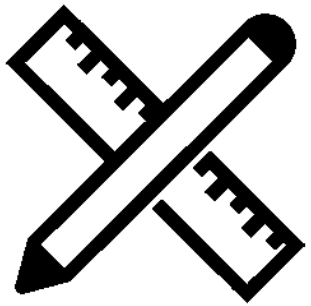
NetApp

合作加速AI數據發展

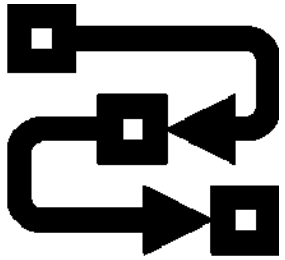
在本地擴展AI面臨的挑戰

Addressing design, deployment and operations bottlenecks

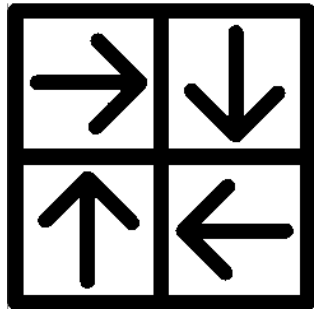
DESIGN GUESSWORK



DEPLOYMENT COMPLEXITY



MULTIPLE POINTS OF SUPPORT



NetApp AI Family

Start small, scale big

HALF RACK

1x DGX-1
1x AFF A220
1x 5624Q

HALF RACK

2x DGX-1
1x AFF A300
2x 5624Q

HALF RACK

4x DGX-1
1x AFF A700
2x 5624Q

FULL RACK

9x DGX-1
1x AFF A800
2x 3232C

FULL RACK

3x DGX-2
1x AFF A800
2x 3232C

FLEXPOD AI

1x UCS 480 ML M5
1x UCS 5108
1x AFF A800
2x 9336C
2x 6454 FI

NetApp

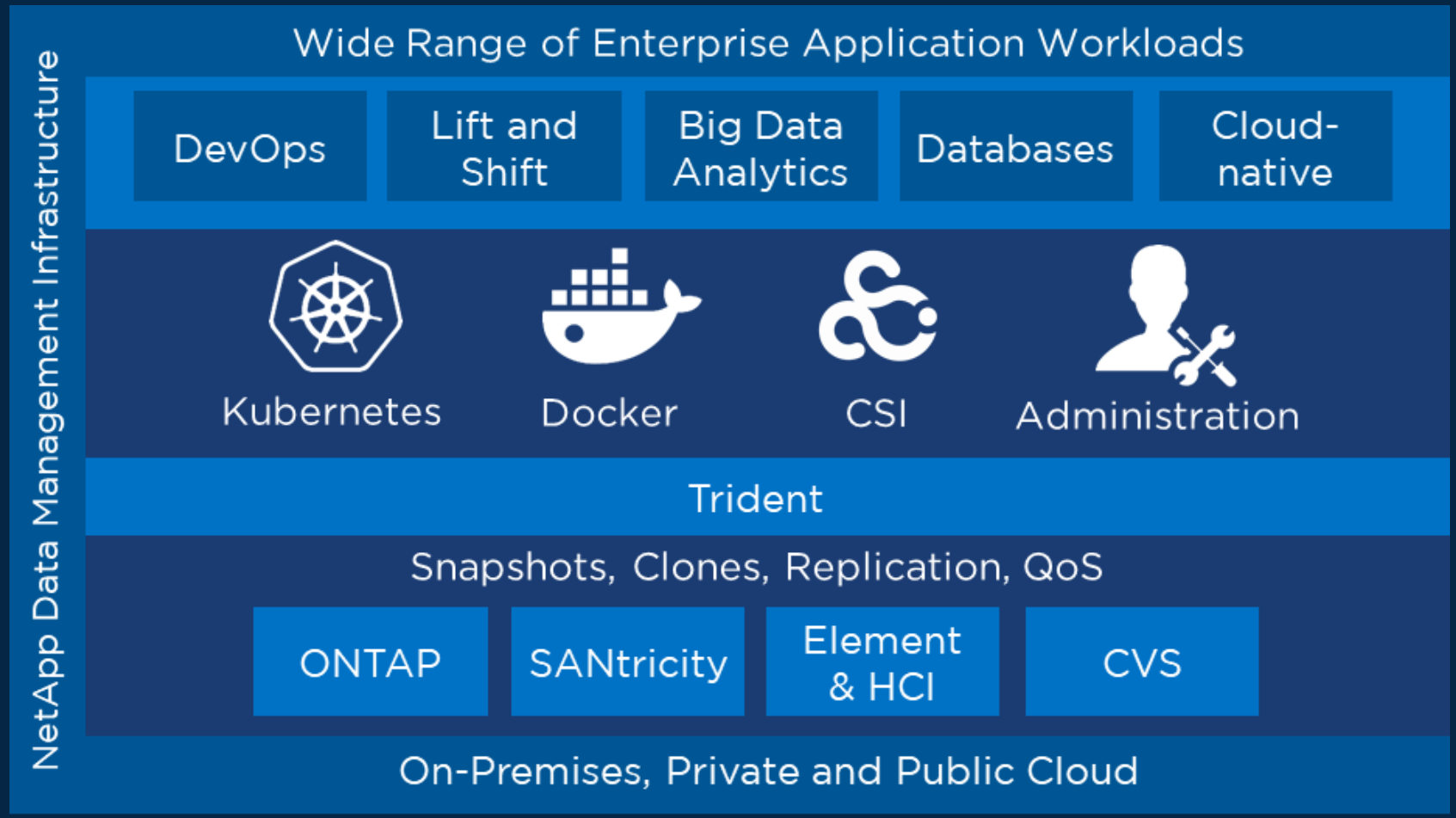
NetApp對Container 環境的支援

Value to Kubernetes Admins

- Native integration with Kubernetes and CSI
- Manage persistence via standard K8s interface

Value to Storage Admins

- Automated, self-provisioning
- Control without micromanaging
- Persistent storage with data management capabilities

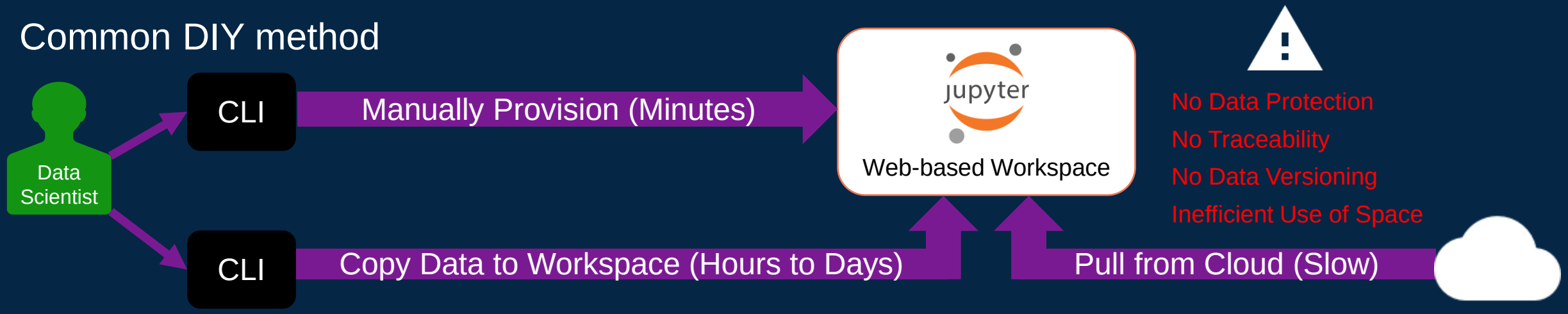


NetApp

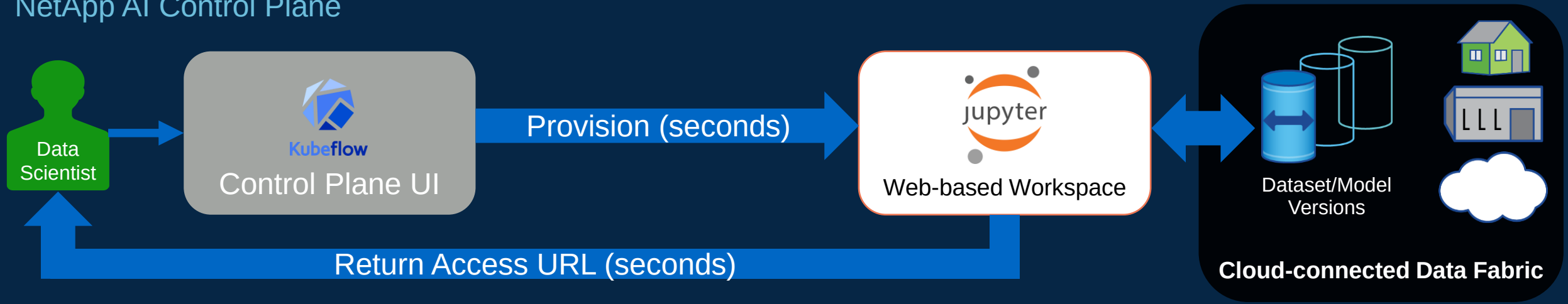
AI 架構流程展示

數據科學家/開發人員工作區創建

Common DIY method



NetApp AI Control Plane



NetApp

Resources

- **Videos**:
  - **Overview:** NetApp AI Control Plane video
  - **Demo:** Provision a Jupyter Workspace with NetApp AI Control Plane
- **Blog Posts:**
  - **Overview:** **How to Effectively Integrate Data Management into Your MLOps Processes with NetApp AI Control Plane**
  - **Technical How-to:** **AI/ML/DL Dataset and Model Versioning from Within a Jupyter Notebook**
- **Solution Brief**: NetApp AI Control Plane solution brief
- **Technical Paper**: NetApp AI Control Plane technical report
- **On-demand Webcast**: AI Control Plane: Full Stack AI Data Management for Hybrid Cloud

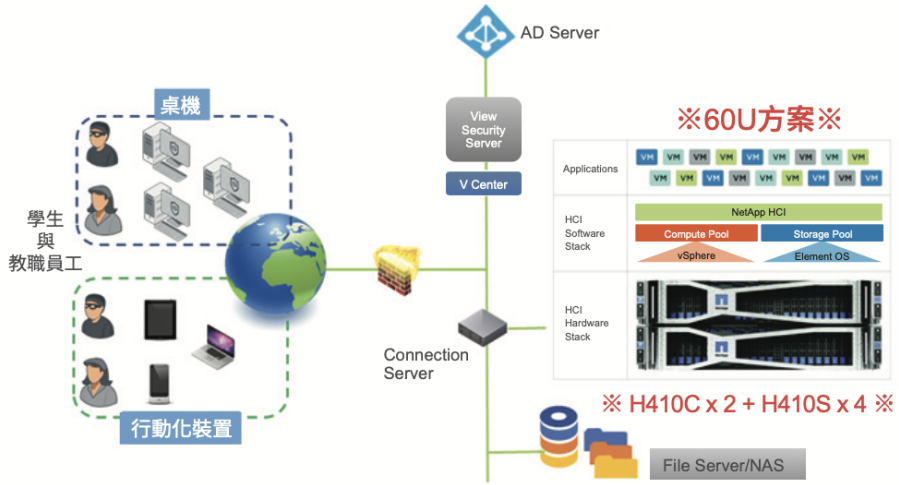
netapp.com/ai

NetApp

VDI防疫包淺談

泰瑩科技與 NetApp HCI 建構 VDI + 伺服器虛擬化
創建無阻礙的資料學習中心

NetApp HCI 開展超融合系統新視界，幫助各大專院校
簡化教學部署，教學暢行無阻



透過 NetApp HCI 儲存、網路、運算等三大元件
建構完美遠程登入學習系統

- NetApp HCI+VDI 提供多元數位教學，可自主學習、彈性修課、課業不中斷。
- 遠距教學零時差，於課堂時間登入 HCI VDI 系統同步進入學校網路環境，藉由 All Flash 避免開機風暴，提升虛擬桌面存取速度，並加快學習進度。
- NetApp HCI 具備 Cloud Ready，能夠無縫串連 AWS、Azure、Google Cloud Platform 等多個主流公有平台，為日後引進公有雲做準備。
- 以 NetApp HCI 全快閃儲存架構結合 VMware 虛擬化平臺，靈活調節伺服器虛擬化環境的資源與效能。
- 利用 NetApp HCI 雲端解決方案來幫助管理者加快軟體安裝及部署作業，同時透過壓縮和重複資料刪除技術，節省儲存成本。

聯絡窗口 | 請洽業務處 劉先生 0933-743-666 / Mail : andy.liu@tai-win.com.tw

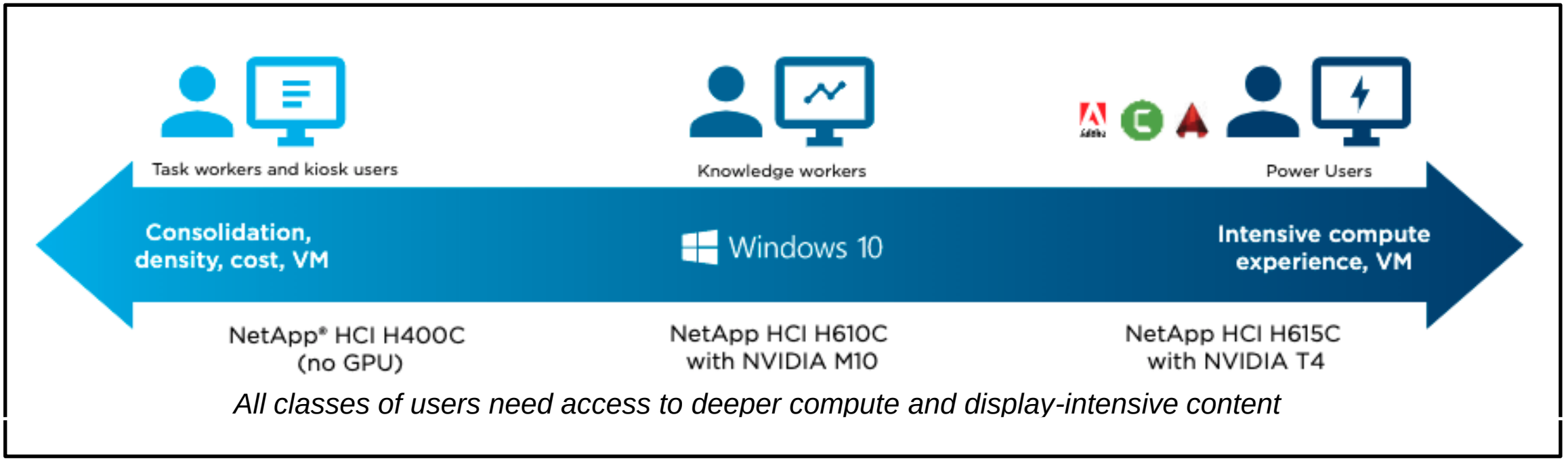
NetApp®

泰瑩科技股份有限公司 TAI WIN TECHNOLOGY CORPORATION

NetApp

NetApp HCI for VDI

Addressing unpredictable and widely varying user workloads in a complicated VDI environment can require painful re-architecting. Introducing new workloads might rob resources from existing workloads, leaving users unhappy.



A better approach to infrastructure for VDI is needed.

NetApp

NetApp HCI and VDI

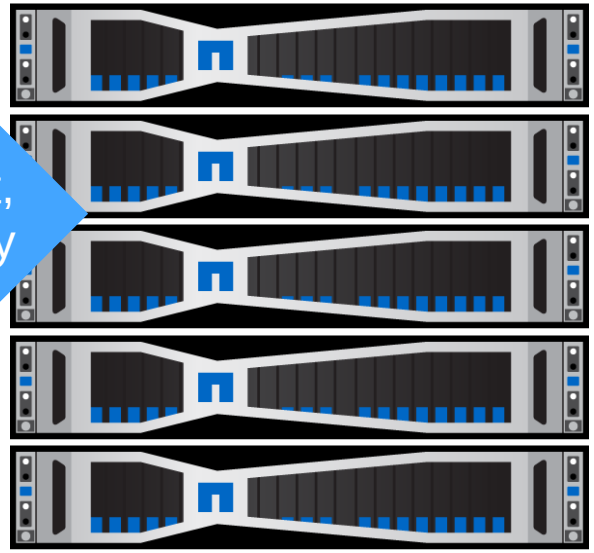
Predictable repeatable scale of compute or storage

Why NetApp HCI for VDI ?

Small incremental scale versus POD-based approach → Scale on your terms

Granular performance controls for workload isolation → Dictate performance

Simplified management beyond initial deployment → Simple management, automation ready



User workloads are impossible to predict

Users' behaviors, different day different behaviors → Control the unknown

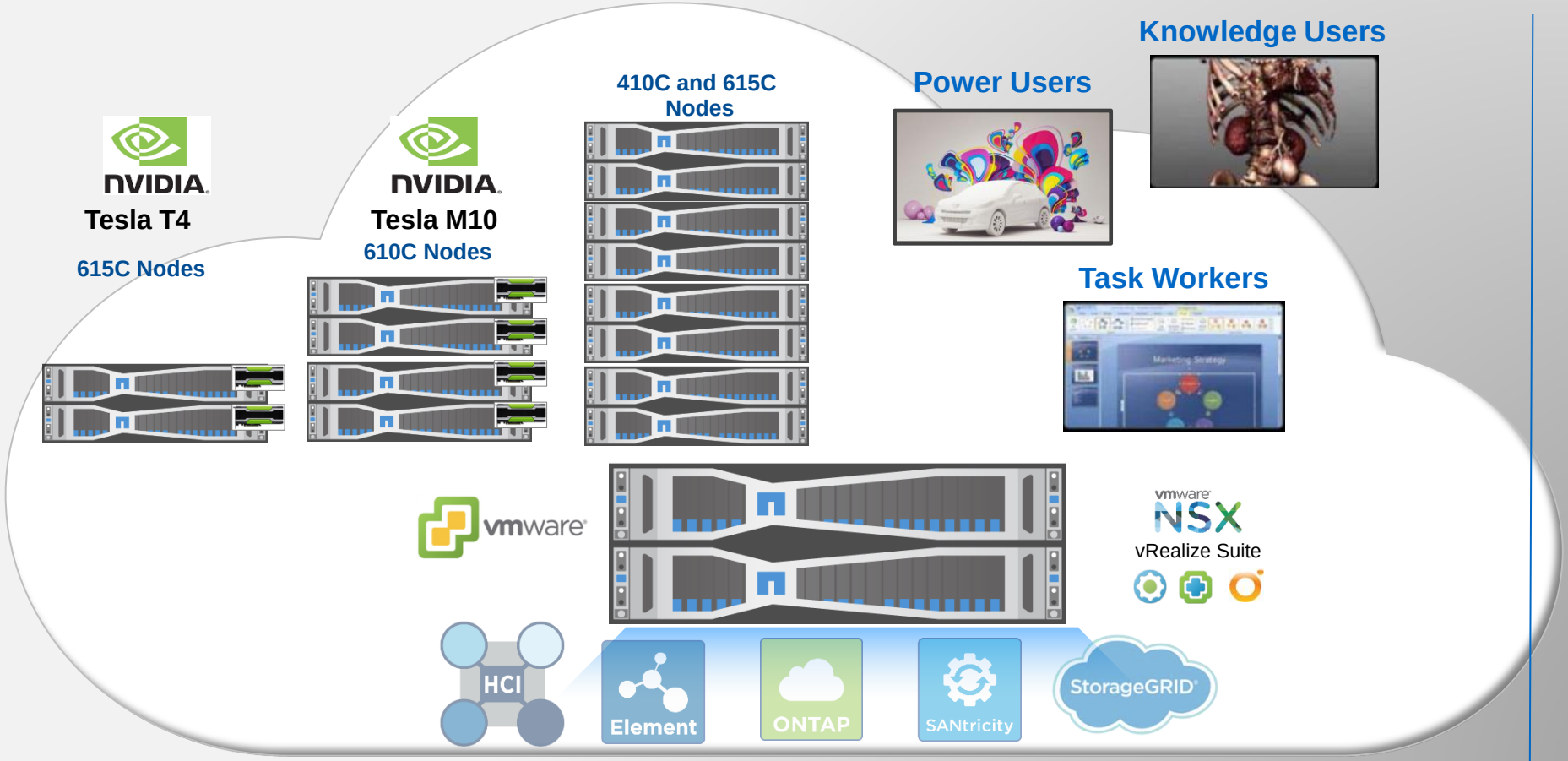
Desktop and apps initial and changing requirements → Meet challenges head on

Assessment helps to get close but doesn't always predict → Monitor and plan

NetApp

A VDI Solution Built to Meet Your Needs

Start Small - Expand - Extend

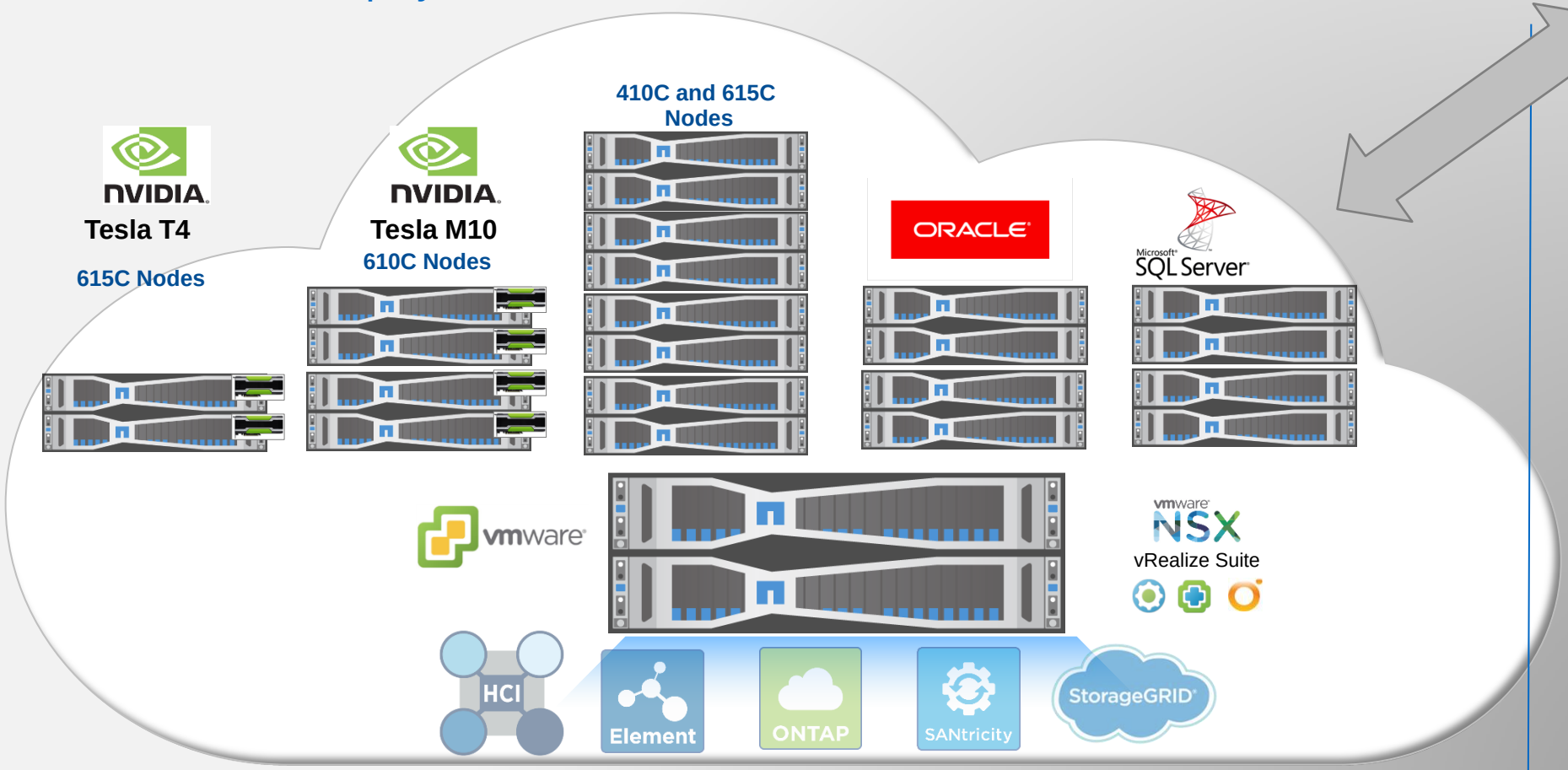


NetApp HCI

- Start Small and Grow
- Add workloads
- Control and guarantee performance
- Extend existing application
- Meet new expectations and business requirements without creating a new environment

An HCI Solution Built for the Data Center

Consolidate – Simplify – Deliver



NetApp HCI

- Start Small and Grow
- Add workloads
- Control and guarantee performance
- Extend existing application
- Meet new expectations and business requirements without creating a new environment

Elastic compute and storage environment that will grow with you and not create silos.

Seven Reasons Why NetApp HCI

1. **Independent Scale:** Maximize resources and minimize hypervisor tax. Grow IT based on business needs, not architectural deficiencies.

2. **Workload Consolidation:** No more silos, No more unnatural workload constraints. More VMs per dollar. More secure data. Better QoS, dedicated performance and guaranteed service levels.

3. **Open Hybrid Multicloud:** No cloud lock-in. Consistent IT consumption across public cloud, private cloud and on-premises.

4. **Guaranteed Efficiency:** The industry's most effective storage efficiency guarantee with no impact to system performance.

5. **End-to-End Automation:** Zero learning curve. Use VMware deployment and management tools that you already know. And don't waste time and money on additional proprietary software.

6. **Simplicity:** Easily support your virtualized environment as is. NetApp HCI is transparent and does not require changes to policies or procedures.

7. **Proactive Protection:** Monitor, troubleshoot and optimize your entire infrastructure with NetApp Cloud Insights. Prevent issues early and accelerate resolution with NetApp Active IQ.

NetApp