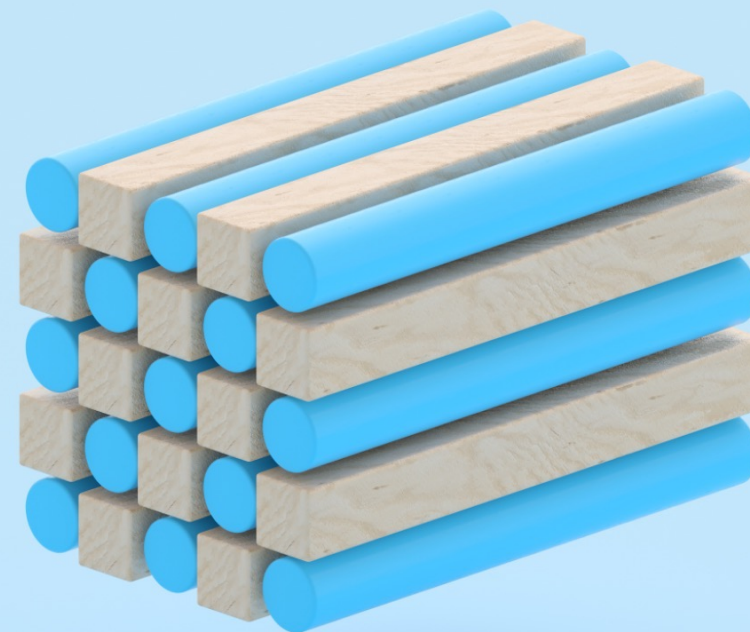


# 擁抱 AI 攜手現代化資料管理

*TANET 2021*

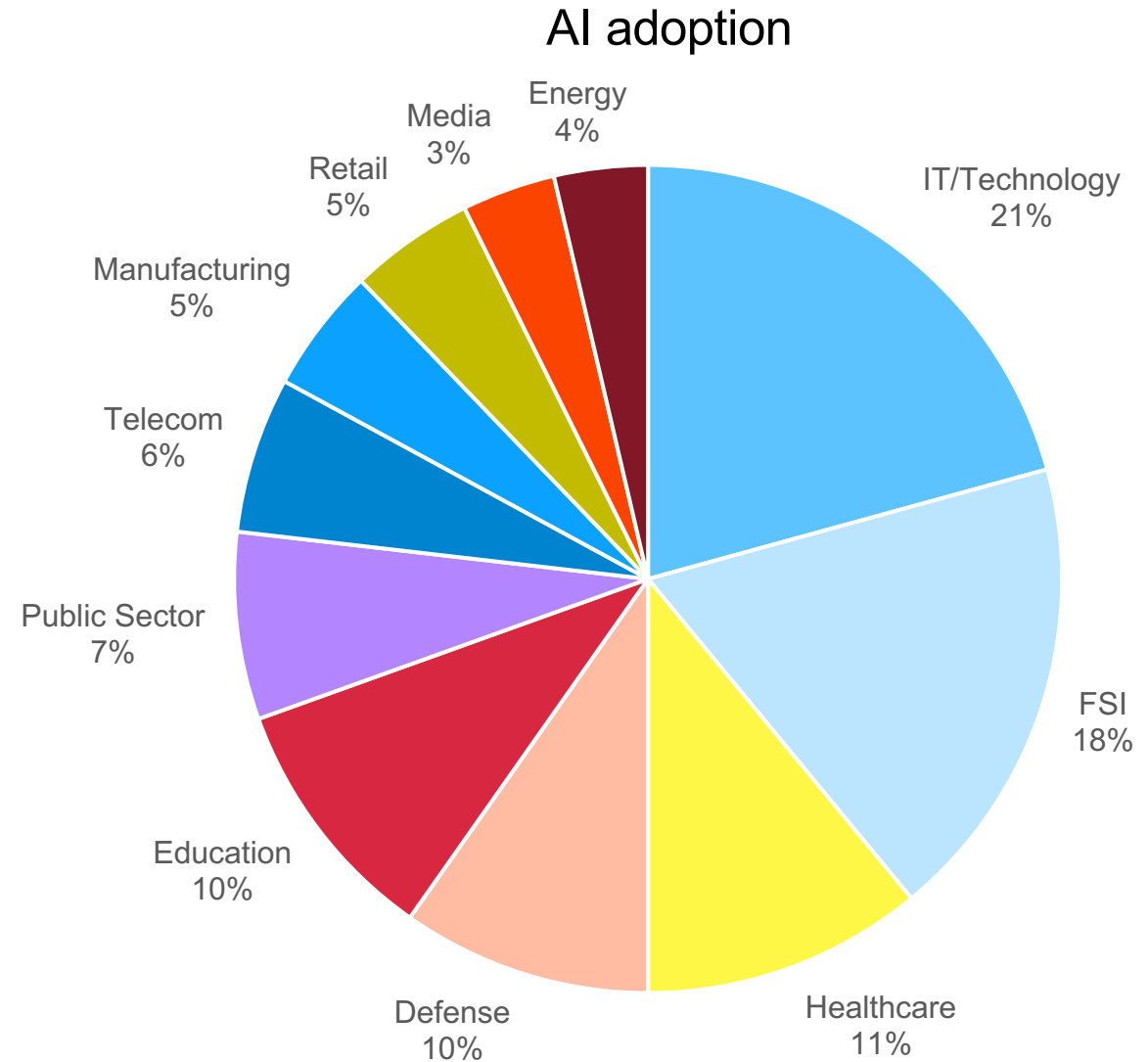
Zero One  
徐榮君 / Sam Syu  
December 11 2021



# AI adoption by industry

## The top drivers of AI adoption in organizations include

- Cyber security
- Automation
- Changing business needs



<https://www.oreilly.com/radar/ai-adoption-in-the-enterprise-2021/>

**IDC: Storage infrastructure  
market for AI workloads  
will grow and reach**

**by 2023**

**It's just “day 1” for data**



**\$10  
Billion**

# Data is critical to AI, but it presents manageability and infrastructure challenges

51%

## Data silos

Eliminate silos with multiworkload quality of service (QoS)

37%

## Tech complexity

Simplify deployment and management of the edge-to-core-to-cloud pipeline

35%

## Data access

Access data at the edge, the core, and the cloud

35%

## Data preparation

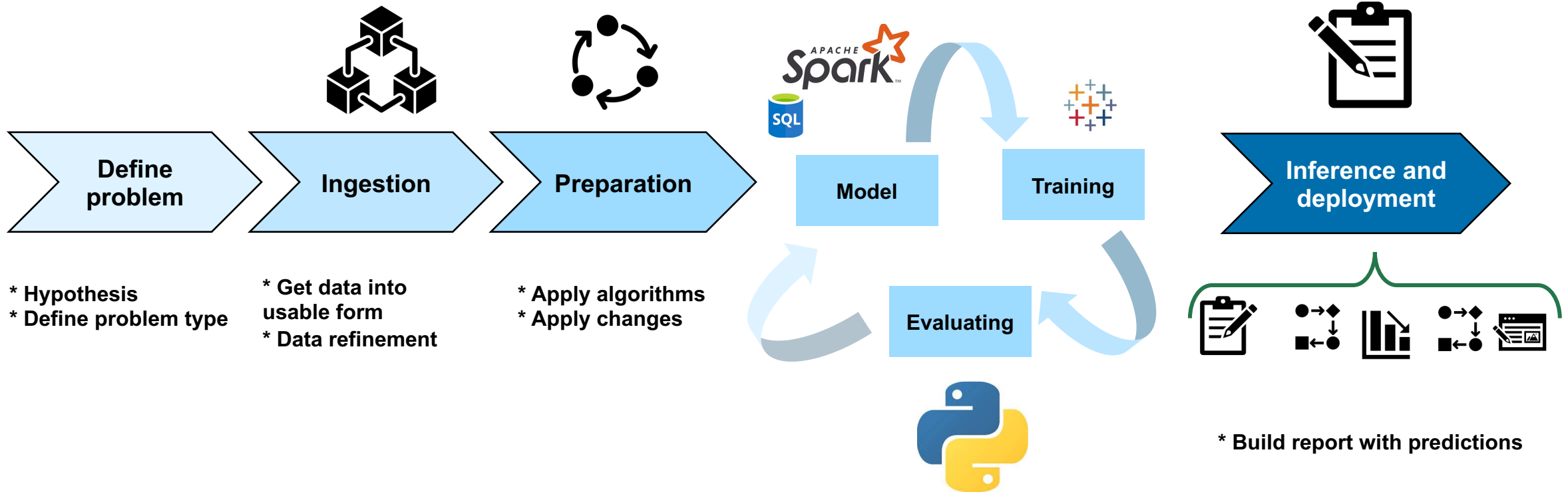
Get the performance and simplicity that you need to power your AI data preparation, training, and inference

“Artificial Intelligence (AI) has **moved well beyond prototyping** and into the phase of execution and implementation.”

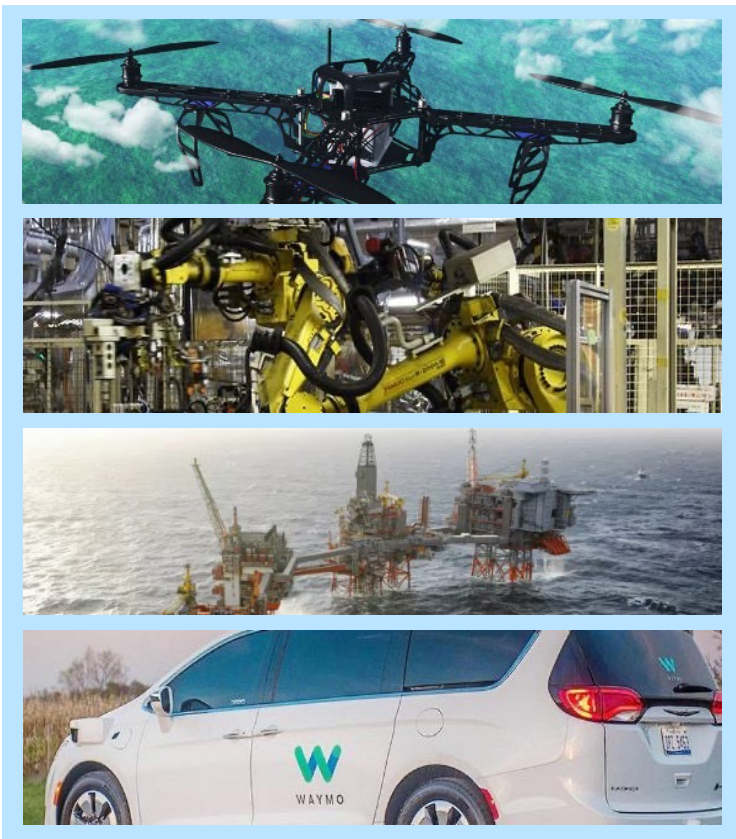
- IDC Customer Insights & Analysis

# AI/ML project pipeline: High level

Major phases and focus on storage

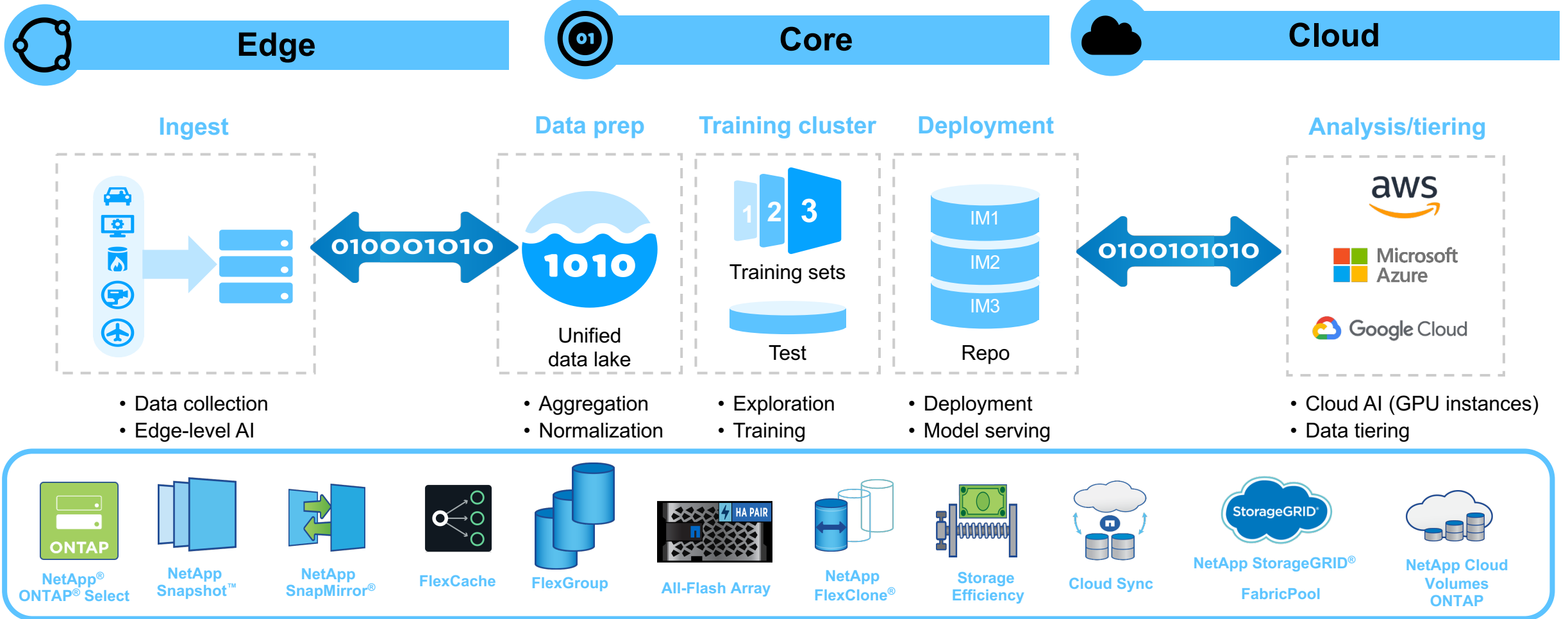


# 核心、雲、邊緣 連接？



# How and where does NetApp ONTAP help?

Edge-to-core-to-cloud data pipeline



# How NetApp helps customers in their AI journey



**Integrations and solutions for your AI needs**



**Certified reference architecture systems that combine storage and compute together**



**Guidance and collateral on various tools, integration, and architecture**



**NetApp® AI partner network**



## Matthew Hull, Vice President, Global AI Data Center Solution Sales at NVIDIA -

*“Simplifying and accelerating the adoption of AI remains a fundamental challenge for many enterprises.”*

*NVIDIA and NetApp are helping organizations achieve their AI goals more quickly with DGX systems. Our latest joint offerings provide multiple paths to success, simplifying the deployment of proven AI infrastructure.”*



*“If data is the fuel of AI, NetApp is the fuel of NVIDIA”* **Jensen Huang - CEO NVIDIA**

\*Quote taken from recent call between Jensen & George Kurian – NetApp CEO

# NetApp ONTAP AI

Accelerate your AI data pipeline for deep learning

## Simple

- Eliminates design complexity and guesswork
- Partners deliver complete solution

## Integrated

- Intelligently manage your data across edge, core and cloud
- Deploy AI frameworks with confidence



## Powerful

- Scale without limits
- Start small and grow non-disruptively



# NetApp ONTAP AI 整合式解決方案

NetApp 和 NVIDIA 提供的規範式產品可簡化AI 工作負載的相關採購、設定、安裝和支援工作



## 規模

四種組態：入門、小、中、大型



## 容量

每項產品都有兩個容量選項，  
並可隨需擴充



## 支援

一通電話即可取得支援

預先定義的組態，帶來簡易的操作體驗，多種組態更可滿足您的業務需求

# Portfolio of AI infrastructure bundles

Meet today's needs and scale in-place to meet tomorrow's



	Starter	Small	Medium	Large
GPU Compute	1x DGX A100	2x DGX A100	4x DGX A100	8x DGX A100
File Storage	1x NetApp A400	1x NetApp A400	1x NetApp A700	1x NetApp A800
Networking	1x SN3700, 1x AS4610	2x Mellanox SN3700, 2x Mellanox QM8700, 1x Mellanox AS4610		
Infrastructure	Mgt Svrs, 1x Rack, PDUs	Management servers, racks, PDUs		
Integration	Performed by SI			
Installation	Performed by SI			
Coordinated Support	Delivered by NVIDIA enterprise support			

All bundles are available through single SKUs at Arrow Electronics

# 效能驗證

Figure 11) MLPerf v0.7 average images per second.

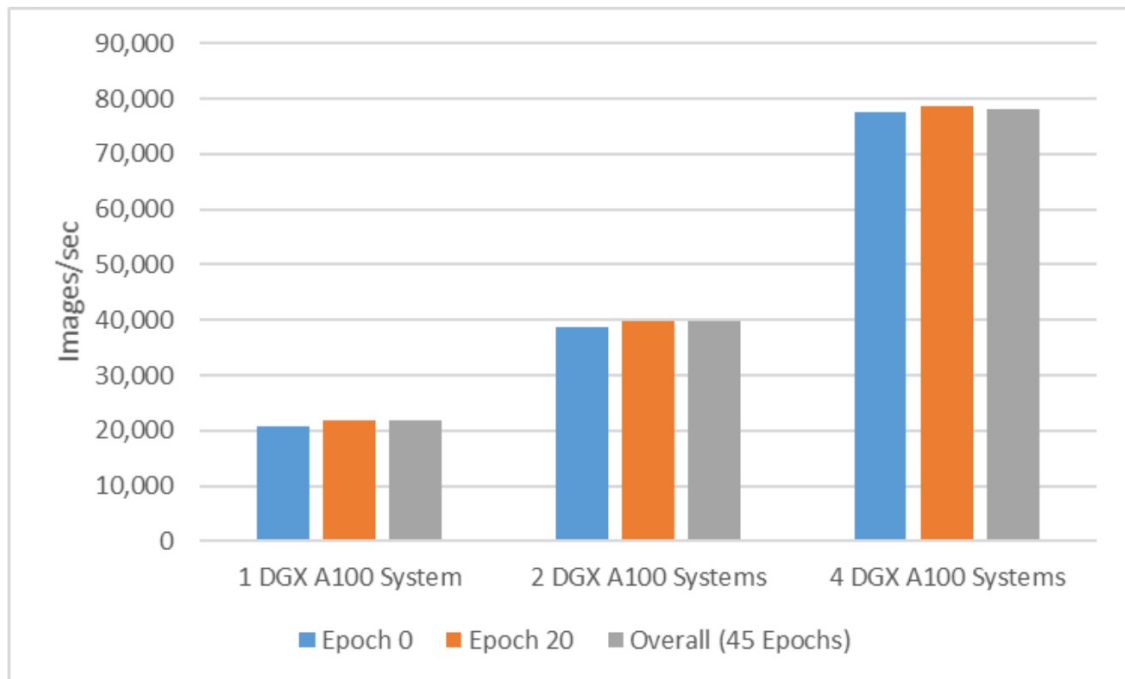
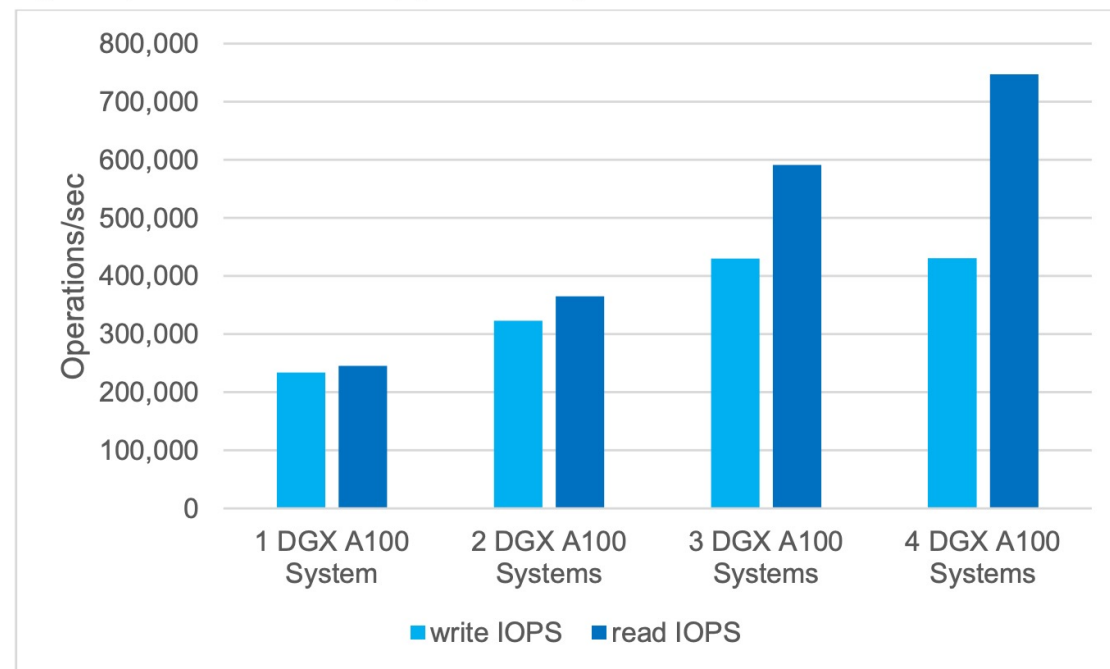
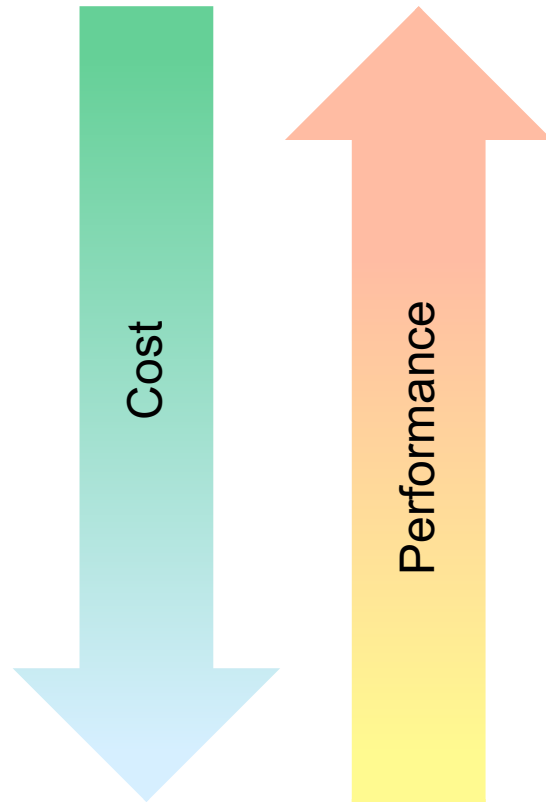


Figure 10) FIO IOPS test results (operations/sec).



# ONTAP Systems for AI/ML training

Start small, scale big



		Throughput**	Raw capacity (typical / max)	Connectivity	Number of NVIDIA A100 GPUs supported***
AFF A800	1 HA pair *	25GBps	368TB/3.6PB	100GbE	08 - 64
	4 HA pairs	100GBps	1.47PB/14.4PB		256
	12 HA pairs	300GBps	4.4PB /43.2PB		768
AFF A400	1 HA pair	11GB/s	182TB/14.7PB	40/100GbE	06 - 32
	4 HA pairs	44GBps	728TB /58.8PB		128
	12 HA pairs	132GBps	2.1PB /176.4PB		384
AFF A250	1 HA pair	7.4GBps	91.2TB /4.4PB	25GbE	04 - 08
	4 HA pairs	29.6GBps	364.8TB /17.6PB		16 - 32
AFF A220	1 HA pair	5.1GBps	91.2TB/4.4PB	10GbE	02 - 04
	4 HA pairs	20.4GBps	364.8TB/17.6PB		08 - 16

\* 1 NetApp® AFF = 1 HA pair = 2 nodes

\*\* 100% sequential read

\*\*\* Estimates based on workload testing in NVA-1153

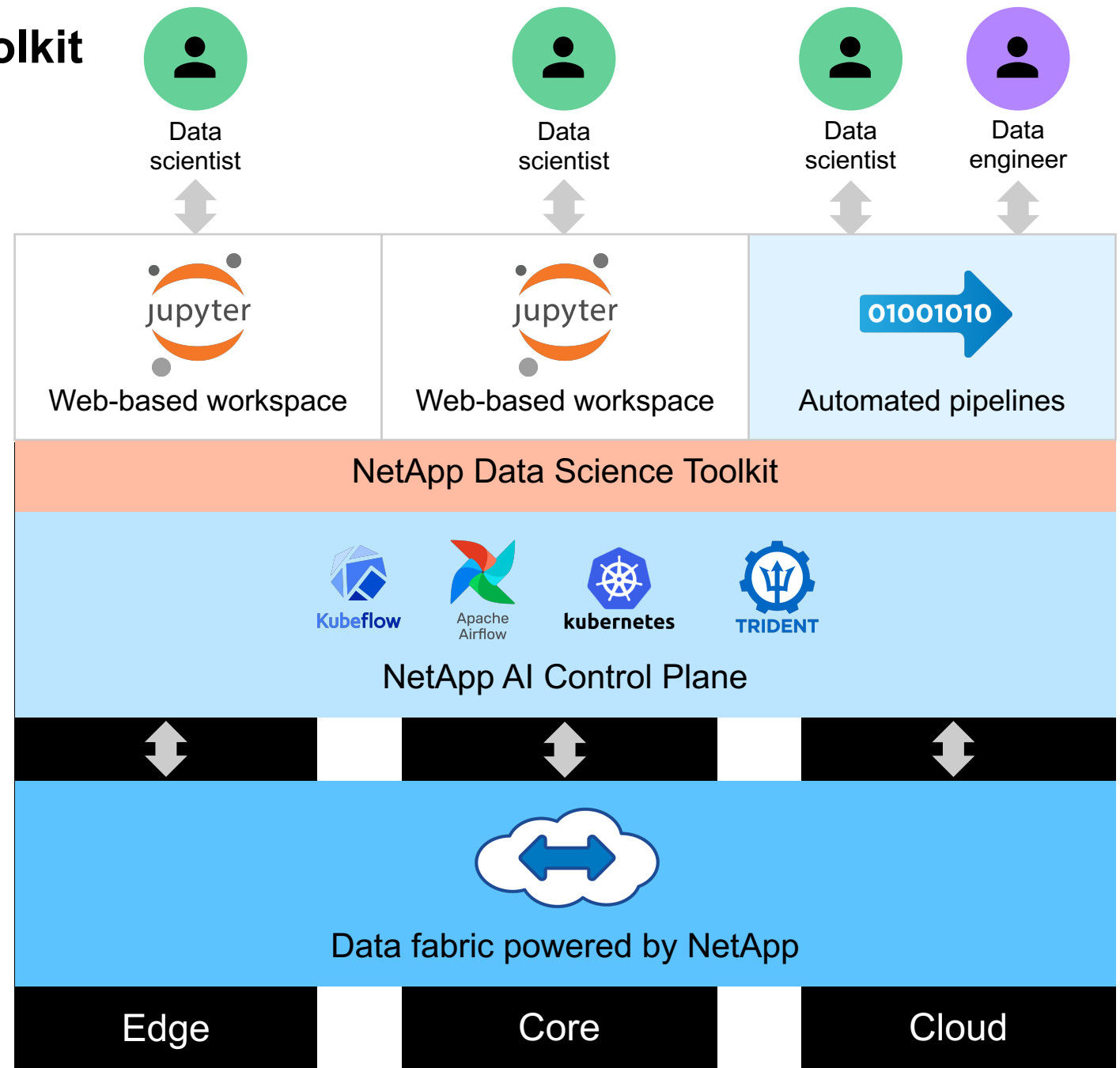
# AI Control Plane / NetApp DataOps Toolkit

Multiple deployment options

Easy to Use

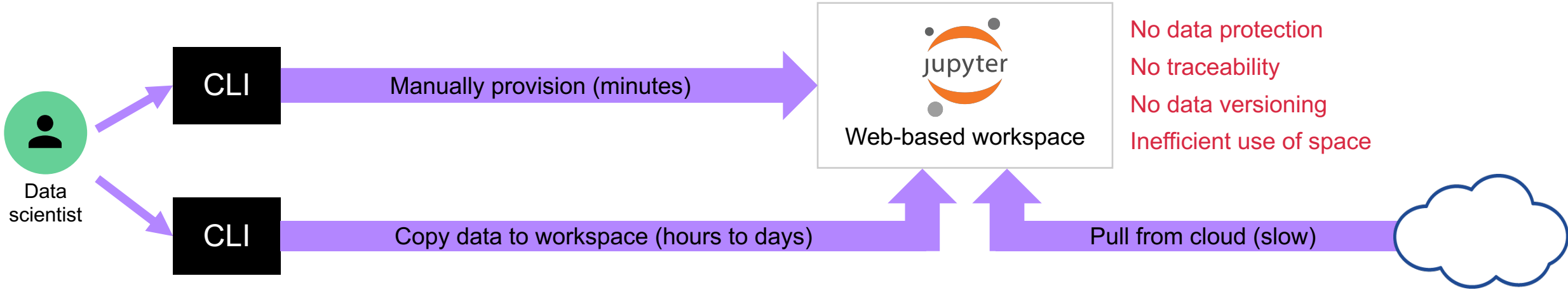
Simple to Management

Automate your data workflows

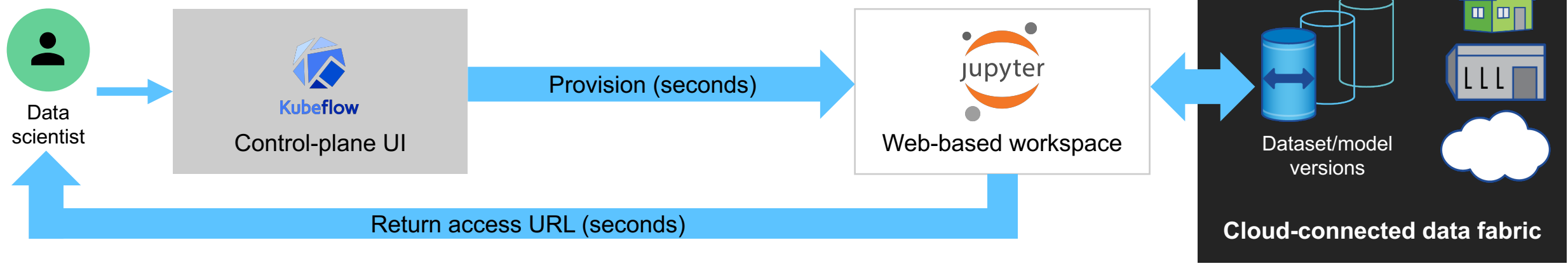


# Data science workspace creation

Common DIY method



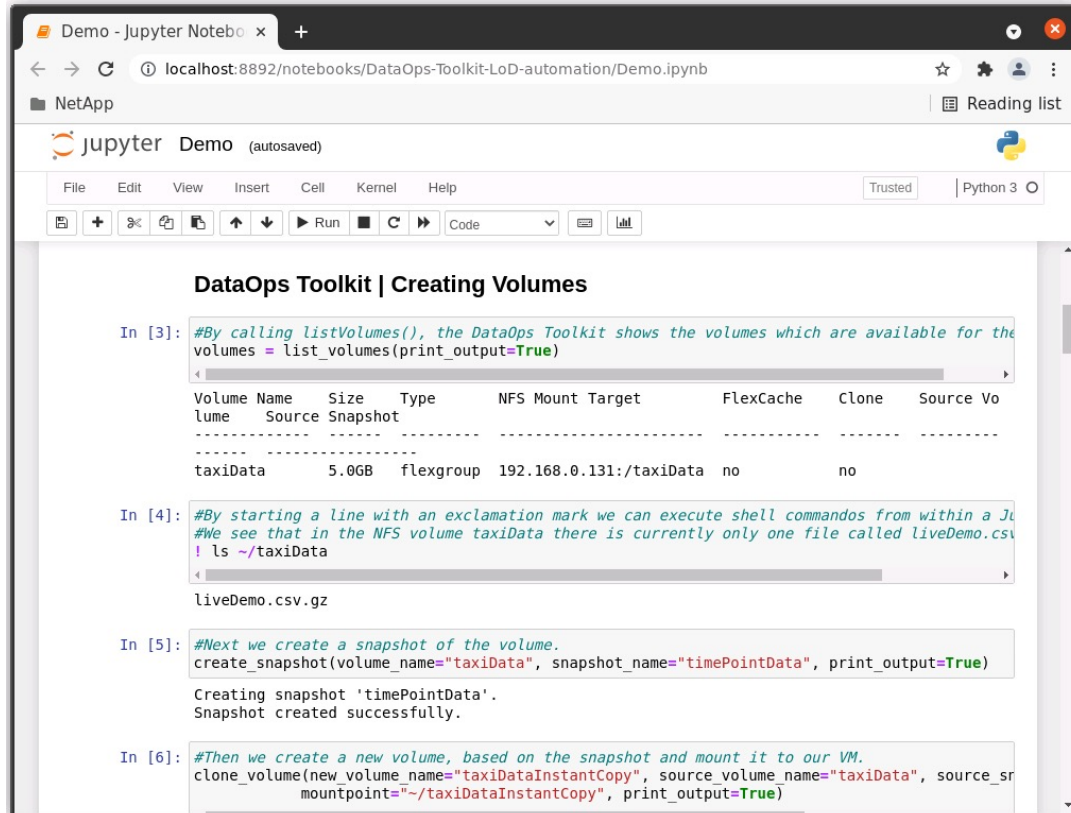
# NetApp® AI Control Plane



# NetApp DataOps Toolkit (Traditional)

Simplify access to NetApp solutions from Data Science environments

## Jupyter Notebook:



```

In [3]: #By calling listVolumes(), the DataOps Toolkit shows the volumes which are available for the
volumes = list_volumes(print_output=True)

Volume Name  Size  Type      NFS Mount Target      FlexCache  Clone  Source Vo
lume        Source Snapshot
-----
taxiData     5.0GB flexgroup 192.168.0.131:/taxiData no          no

In [4]: #By starting a line with an exclamation mark we can execute shell commands from within a Jupyter cell
#We see that in the NFS volume taxiData there is currently only one file called liveDemo.csv
! ls ~/taxiData

liveDemo.csv.gz

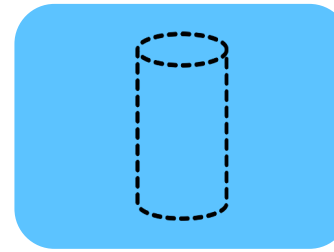
In [5]: #Next we create a snapshot of the volume.
create_snapshot(volume_name="taxiData", snapshot_name="timePointData", print_output=True)

Creating snapshot 'timePointData'.
Snapshot created successfully.

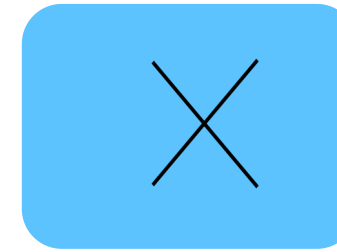
In [6]: #Then we create a new volume, based on the snapshot and mount it to our VM.
clone_volume(new_volume_name="taxiDataInstantCopy", source_volume_name="taxiData", source_sr
mountpoint=~ /taxiDataInstantCopy", print_output=True)

```

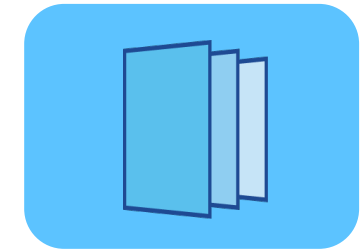
## Key Functions:



Cloning of volumes



Deleting volumes



Creating Snapshots



Pull Content from S3



Push files to S3

# 2 Seconds

vs 17 hours

## 高效加速 AI 訓練工作流程

- Copying a dataset takes seconds instead of hours/days
  - Example: 10TB dataset – 2 seconds vs 17 hours
- Copies of datasets are stored efficiently
  - Example: 10 copies of each dataset – up to 90% reduction in storage space

# Near-instantaneous dataset clones

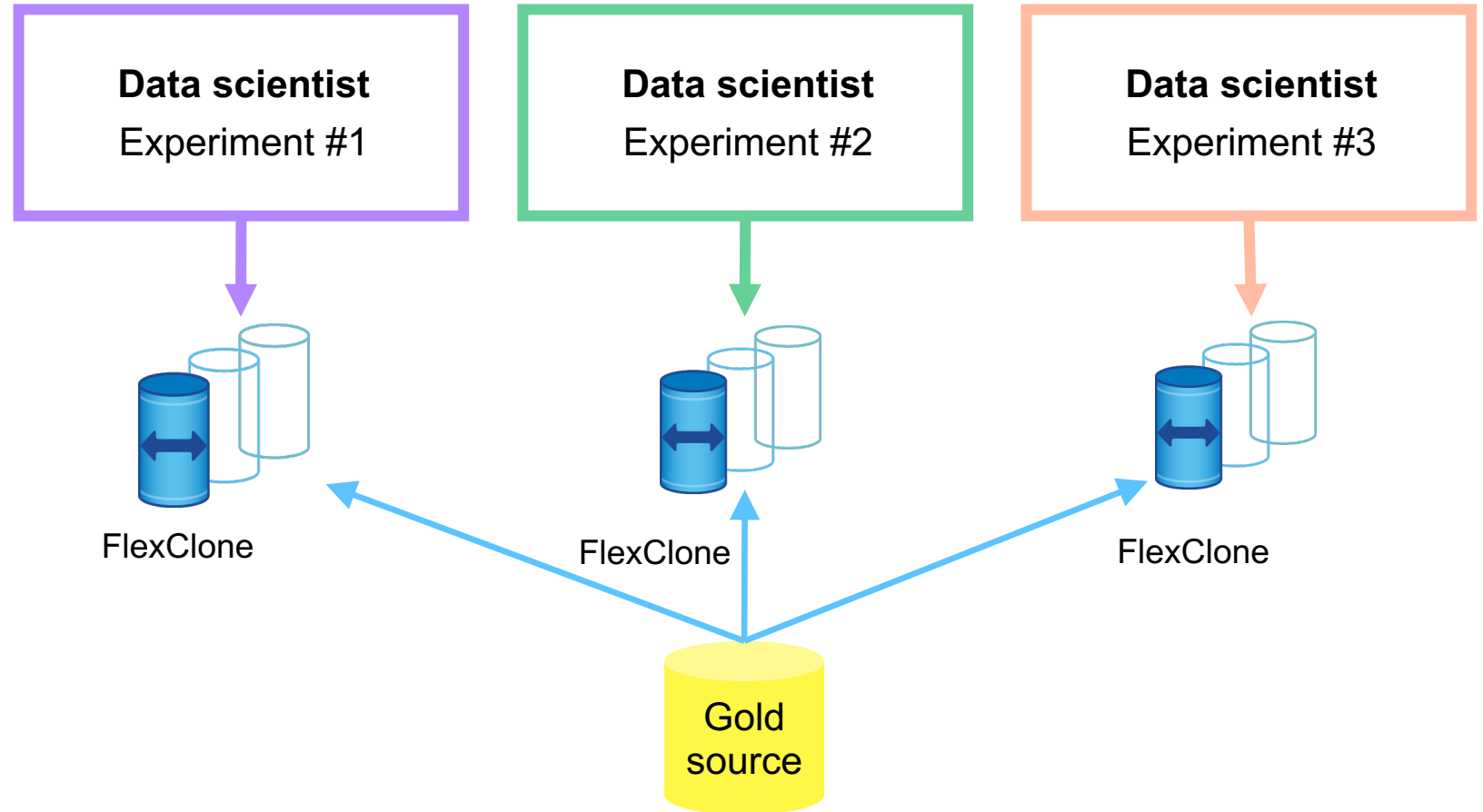
Accelerate the AI training workflow

## Challenges

- Copying data takes hours and days to finish
- Huge storage footprint

## Solution

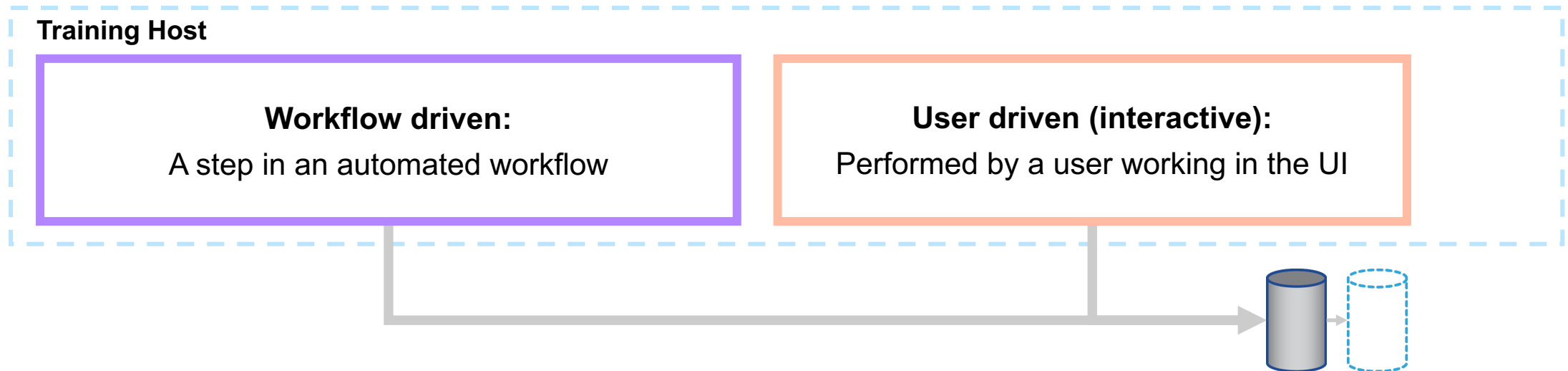
- With NetApp® FlexClone® it take seconds instead of hours or days
- Copies are stored efficiently



# Dataset and model versioning

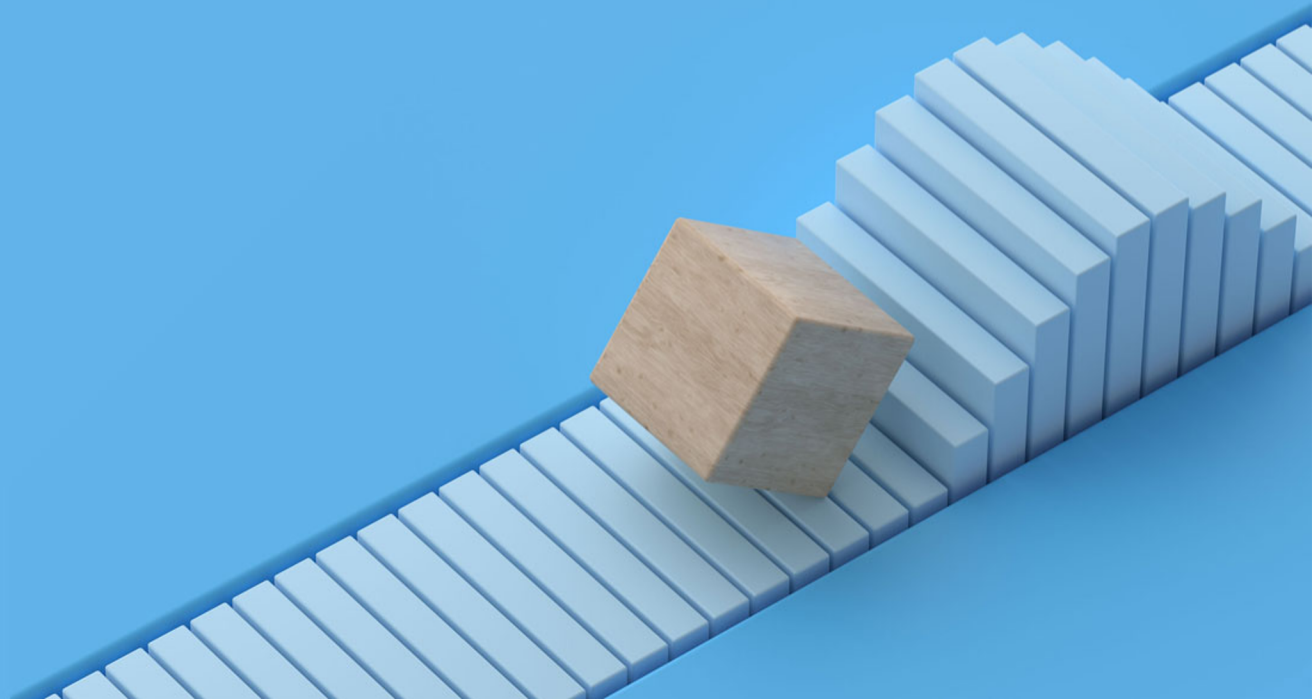
A new approach is needed

- The challenge:
  - As AI moves into highly regulated/sensitive industries – dataset-to-model traceability and reproducibility is a **mandatory requirement**
  - Traditional versioning tools (e.g. Git) are designed for code, not data
- The solution:
  - Utilize existing, proven tech: **volume snapshots**—long and proven history of being used to **quickly and efficiently version datasets**
  - Snapshot name can be saved in code or model repository to implement traceability



# NetApp AFF 系統

協助資料中心順利轉型到 Flash 架構



# NetApp is No.1 in Taiwan

CY2021 Q2 and Rolling 4 Quarters

- **102.4%** year-on-year (Y/Y) growth In the Open Networked segment, outpacing market growth of 25.5%. Top position in the segment with market share of **38.4%**
- **158.7%** Y/Y growth in AFA, outpacing market growth of 31.9%. Top position in AFA with market share of **45.6%**
- NetApp is the **#1 vendor in Taiwan** based on rolling 4 quarters average, which account for **29.1%** market share in Open Networked (#1) and **33.7%** in AFA (#1)

Taiwan Market Share CY2021 Q2

OPEN NETWORKED CY2021 Q2		
Ranking	Vendor	Shares
<b>1</b>	<b>NetApp</b>	<b>38.4%</b>
2	Dell	22.2%
3	HPE	16.9%
4	IBM	5.1%
5	Hitachi	3.8%

AFA CY2021 Q2		
Ranking	Vendor	Shares
<b>1</b>	<b>NetApp</b>	<b>45.6%</b>
2	HPE	19.0%
3	Dell	17.9%
4	IBM	10.6%
5	Pure	4.9%

OPEN NETWORKED Rolling 4 Quarters		
Ranking	Vendor	Share
<b>1</b>	<b>NetApp</b>	<b>29.1%</b>
2	Dell	26.3%
3	HPE	20.6%
4	IBM	6.5%
5	Hitachi	4.8%

AFA Rolling 4 Quarters		
Ranking	Vendor	Share
<b>1</b>	<b>NetApp</b>	<b>33.7%</b>
2	HPE	23.0%
3	Dell	21.7%
4	IBM	12.2%
5	Pure	6.3%

Source: IDC Worldwide Quarterly Enterprise Storage Systems Tracker

# NetApp's Solutions for Customer Multicloud/Hybrid Cloud Solutions

**YOUR APPS**

**ENTERPRISE APPS**  
SAP, Databases, HPC

**FILE SERVICES**  
Sharing Linux & Windows clients

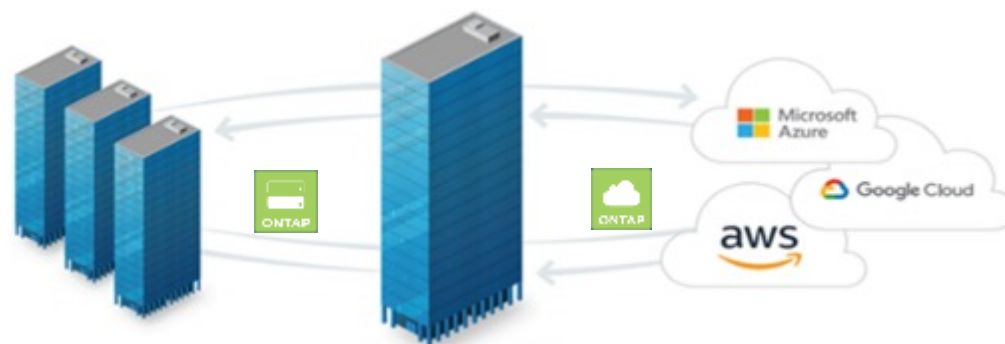
**CLOUD NATIVE APPS**  
Containers, Web apps, persistent storage

**AI & ML**  
AI and deep learning workloads

## YOUR ENVIRONMENT

### Your Data On-premises

- All Flash AFF and FAS
- Software-Defined Converged Infrastructure
- FlexPod
- E-Series
- StorageGRID



### Your Data in the Cloud

- Azure NetApp Files
- Amazon FSx for NetApp ONTAP
- Cloud Volumes Service
- Cloud Volumes ONTAP
- Cloud Manager

### Your Global Workforce

- Virtual Desktop
- Global File Cache

## YOUR DATA - MANAGED

### Backup Service

- Cloud Backup

### Compliance & Security

- Cloud Data Sense
- Cloud Secure

### Data Migration

- Cloud Sync
- Cloud Tiering

### Containerization

- Astra
- Trident

### Resource Management

- Cloud Insights
- spot Compute & Storage

## AFF A 系列：功能強大且值得信賴的智慧型端點對端點 NVMe 系統



**AFF A800**

適合需要高效能及超低延遲的最嚴苛工作負載



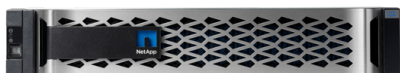
**AFF A700**

適合需要效能、容易維修及恢復能力的  
關鍵任務型工作負載及資料中心



**AFF A400**

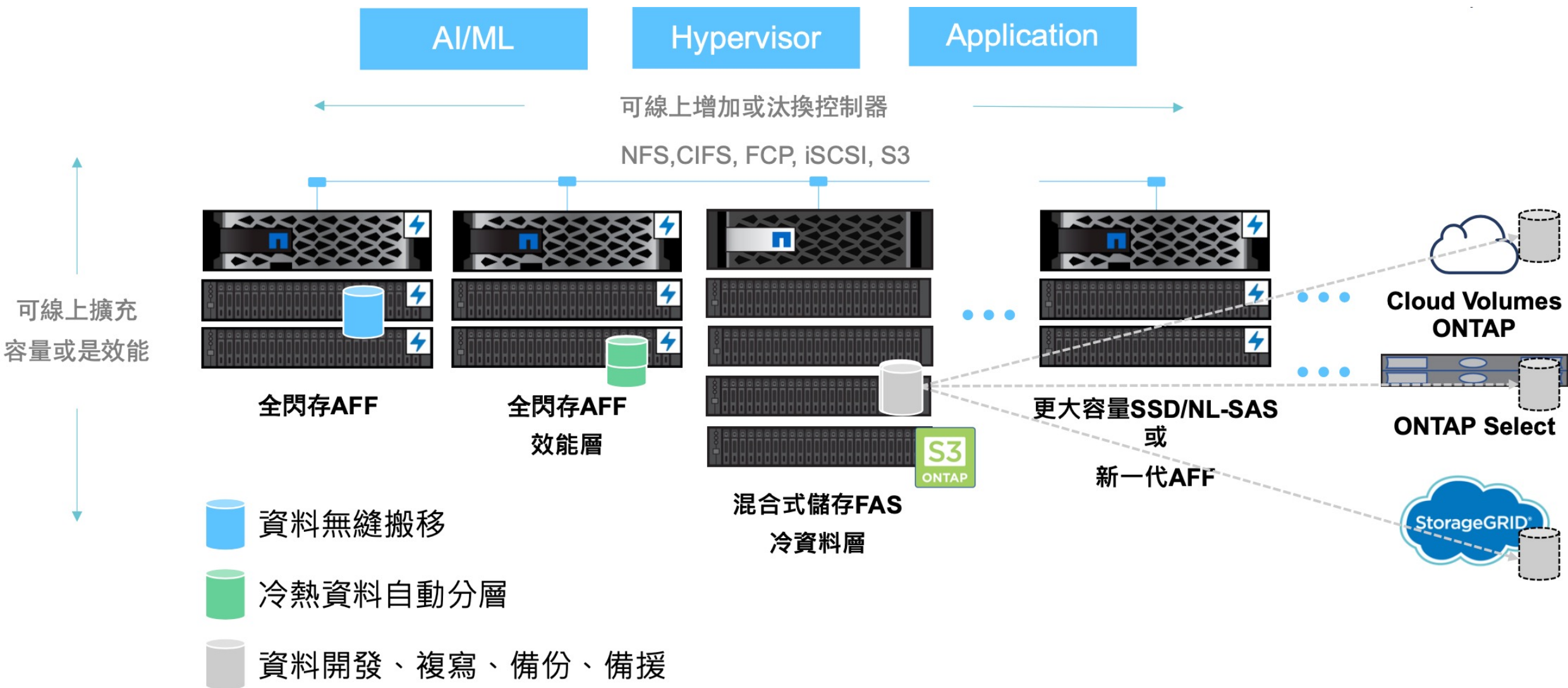
適合需要在效能與豐富功能之間取得平衡的  
大部分企業應用程式



**AFF A250 新推出**

適合需要物超所值效能的 MSE 及小型企業

# 企業級資源池分享



# 客戶案例分享



# Cambridge Consultants

Digital Greenhouse

## Business challenge

- Get ahead of the trends in AI
- Require extreme performance for cutting edge innovation and modeling

## Solution

- ONTAP® AI, 1 x A300, 2 x NVIDIA DGX-1, ONTAP FlexClone® and data protection features

## Use case

- Advanced AI research

## Why NetApp

- Cutting edge performance and ONTAP features
- Proven data management for demanding environments

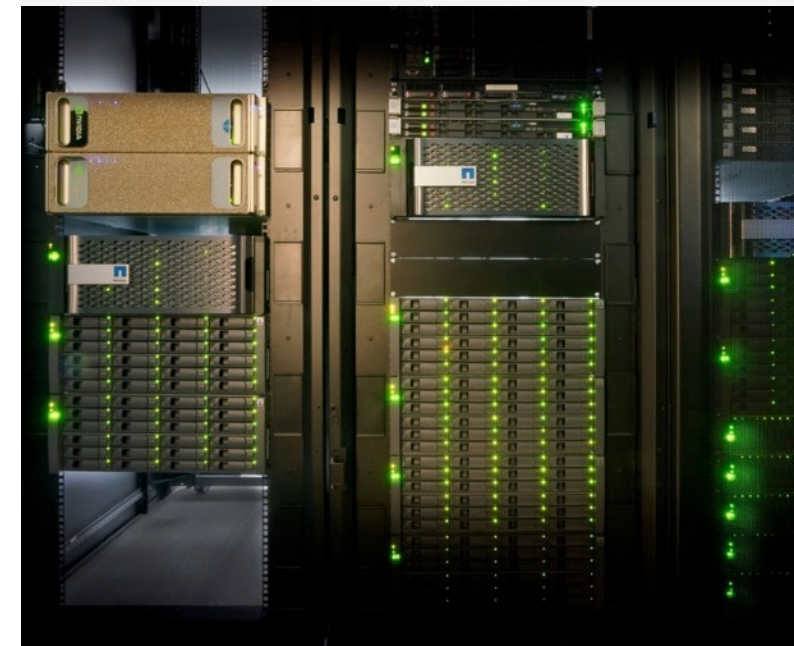
“One of the key reasons our deep learning storage is based on NetApp is the positive experience we have had with the technology in the past.”

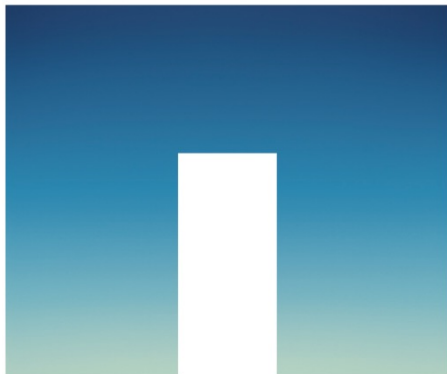
- Tim Ensor, Director of Artificial Intelligence, Cambridge Consultants

- [Success story](#)



[Cambridge Consultants](#) is a global product development and technology consultancy firm providing outsourced research and development services

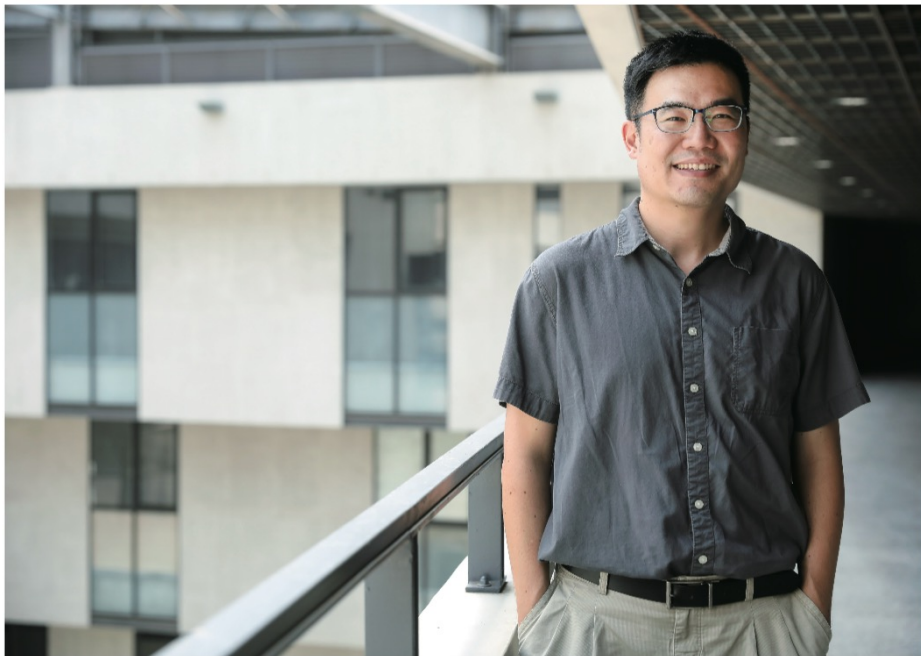




穩定性高、擴充便利 NetApp FAS2750 保護資料最佳選擇

## 元智大學人工智慧跨域創新應用中心發展有成，產學合作效益陸續浮現

人工智慧成為顯學，遠東集團旗下多家子公司均有意運用該技術強化競爭力，元智大學被賦予協助集團人工智慧發展的重責大任，在 2018 年設立人工智慧跨域創新應用中心。NetApp FAS2750 儲存設備具備極佳穩定性，可讓寶貴資料能獲得妥善保護，同時加速研究專案執行速度。



在全球吹起數位轉型的風潮下，帶動企業積極運用 AI、大數據、物聯網等創新技術，優化企業營運流程或發展創新服務。被英國泰晤士報評為全球前 250 名最佳年輕大學的元智大學，培育科技人才不遺餘力及產學合作成果豐碩，在 2018 年即成立人工智慧跨域創新應用中心，擴大與遠東集團旗下公司的合作範疇，運用高可靠度的

化，創造更大效益。元智在設立人工智慧跨域創新應用中心時，即認為基礎資訊架構穩定性非常重要，因此需要高可靠度、穩定性佳的儲存設備，以支撐 AI 專案的穩定運作。在泰瑩科技的協助下，最終選擇深受好評的 NetApp FAS2750 儲存設備。

### 滿足未來資料成長需求

#### NetApp FAS2750 雀屏中選

NetApp FAS2750 儲存設備是針對中小企業或分據點設計的產品，具備與中高階機種相同的穩定性、可靠度，與多種強大的資料儲存與管理技術，可降低資料複雜度與提高管理人員的工作效率。如此一來，讓用戶能以符合成本效益的方

**NetApp unlocks  
the best of cloud**

