



Vendor: Network Appliance

Exam Code: NS0-901

Exam Name: NetApp Certified AI Expert Exam

Version: DEMO

QUESTION 1

An AI architect needs to design a complete, end-to-end data pipeline for a new generative AI application at a financial services firm. The application will allow internal analysts to query a massive, 500 TB archive of historical market data and reports to generate summaries. The firm has the following environment and requirements:

Data_Sources: A mix of on-premises ONTAP filers and StorageGRID S3 buckets.

Requirement_1: All queries must be answered using only the private data archive.

Requirement_2: All generated summaries must provide citations to the source reports.

Requirement_3: All data containing client PII must be identified and excluded from the LLM context.

Requirement_4: The solution must be cost-effective for the large, mostly-read data archive.

Which set of actions and technologies constitutes the most robust and compliant solution? (Select all that apply.)

- A. Use NetApp XCP to perform a one-time migration of all data from the ONTAP filers to the StorageGRID data lake.
- B. Implement a Retrieval-Augmented Generation (RAG) architecture to meet the requirements for private data usage and source citation.
- C. Deploy BlueXP classification to scan the entire StorageGRID data lake to identify and tag all files containing PII.
- D. Fine-tune a foundation model on the entire 500 TB dataset to ensure it understands the financial context.
- E. Use SnapMirror to replicate the StorageGRID data lake to a high-performance NetApp ASA system for faster query performance.
- F. During the RAG retrieval step, filter out any documents tagged as containing PII by BlueXP classification before sending them to the LLM.

Answer: ABCF

QUESTION 2

What is the primary role of NetApp Trident in a Kubernetes environment designed for AI workloads?

- A. To directly accelerate GPU computations using specialized drivers.
- B. To function as a container runtime interface for executing AI models.
- C. To act as a dynamic storage orchestrator, provisioning persistent storage from NetApp backends on-demand for containerized applications.
- D. To provide a web-based IDE, like Jupyter, for data scientists to develop models.

Answer: C

QUESTION 3

A data scientist needs to launch a Jupyter notebook as a pod in a Kubernetes cluster. The pod requires a 50 Gi persistent volume for storing datasets and notebooks. The cluster administrator has configured a default Trident StorageClass for general-purpose use. The data scientist has the following PersistentVolumeClaim (PVC) manifest:

```
apiVersion: v1
kind: PersistentVolumeClaim
metadata:
```

```
name: jupyter-pvc
spec:
accessModes:
- ReadWriteOnce
resources:
requests:
storage: 50Gi
```

When this PVC is applied to the cluster, what will be the result?

- A. The PVC will fail because a 'storageClassName' is not explicitly defined.
- B. The PVC will remain in a "Pending" state until a PersistentVolume is manually created.
- C. Trident will automatically provision a 50 Gi volume on its default backend and bind it to the PVC.
- D. Trident will create a 1 Gi volume, as this is the default size for all PVCs.

Answer: C

QUESTION 4

An MLOps engineer is deploying a training pod that requires a high-performance volume. After applying the pod and PVC manifests, the pod remains in a 'Pending' state. The engineer runs 'kubectl describe pod training-pod-7d8c' and sees the following event:

Events:

Type Reason Age From Message

Warning FailedScheduling 2m15s

default-scheduler 0/4 nodes are available: 1 node(s) had volume node affinity conflict, 3 node(s) didn't find available persistent volume to bind.

The engineer then inspects the associated PVC and sees its status is also 'Pending'. What is the most likely cause of this issue?

- A. The Trident controller pod has crashed and needs to be restarted.
- B. The 'storageClassName' specified in the PVC does not match any existing StorageClass managed by Trident.
- C. The ONTAP backend is out of available capacity to provision a new volume.
- D. The Kubernetes scheduler is malfunctioning and cannot assign pods to nodes.

Answer: B

QUESTION 5

An organization's AI platform team needs to provide two distinct tiers of storage for their data scientists on a single Kubernetes cluster:

1. 'gold-tier': Extremely low-latency storage for active model training, using an all-flash NetApp ASA system.
2. 'bronze-tier': Cost-effective, high-capacity storage for data staging and archiving, using a NetApp StorageGRID system.

How should the platform team configure NetApp Trident to meet these requirements? (Select all that apply.)

- A. Create a single Trident backend that points to both the ASA and StorageGRID systems simultaneously.
- B. Create two separate Trident backend configurations, one for the ASA ('backend-asa.yaml') and one for StorageGRID ('backend-sgrid.yaml').
- C. Create a single StorageClass named 'multi-tier-storage' that references both backends.
- D. Create two distinct StorageClasses, 'gold-tier' and 'bronze-tier'.
- E. Configure the 'gold-tier' StorageClass to reference the ASA backend.
- F. Configure the 'bronze-tier' StorageClass to reference the StorageGRID backend.

Answer: BDEF

QUESTION 6

A data science team reports that their Jupyter notebook pod, which was previously working, is now failing to start. The pod's status is 'CrashLoopBackOff'. An MLOps engineer investigates and finds that the pod's PersistentVolumeClaim (PVC) is bound, but the pod logs show a "Permission denied" error when trying to write to its '/data' mount point.

The engineer checks the Trident backend configuration associated with the pod's StorageClass:

```
apiVersion: trident.netapp.io/v1
kind: TridentBackendConfig
metadata:
  name: ontap-nas-eco
spec:
  version: 1
  storageDriverName: ontap-nas
  managementLIF: 10.10.20.5
  dataLIF: 10.10.20.10
  svm: svm-prod-ds
  exportPolicy: read-only-policy
```

What is the most likely cause of the "Permission denied" error?

- A. The 'dataLIF' is configured incorrectly and is unreachable from the Kubernetes nodes.
- B. The 'storageDriverName' should be 'ontap-san' for all AI workloads.
- C. The Trident backend is configured to use an export policy ('read-only-policy') that does not grant write permissions to the Kubernetes nodes.
- D. The Kubernetes pod has an invalid 'securityContext' that prevents it from writing to any volume.

Answer: C

QUESTION 7

An architect is designing a scalable, automated MLOps platform using Kubeflow on a Kubernetes cluster. The platform must support the entire AI lifecycle for multiple teams, with different storage requirements at each stage.

The key requirements are:

- Data Ingestion: A pipeline step needs a shared, read-write volume accessible by multiple pods to stage raw data.
- Experimentation: Data scientists need individual, isolated volumes for their Jupyter notebooks.
- Training: Distributed training jobs require a high-performance, parallel-access filesystem for reading training data.
- Automation: All storage must be provisioned automatically via Kubeflow pipeline definitions

without manual intervention.

Which combination of technologies and configurations would create the most effective solution?

- A. Use the NetApp DataOps Toolkit for all storage provisioning, bypassing Trident and Kubernetes PVCs.
- B. Create a single, large NFS volume and mount it to all pods using a static PersistentVolume.
- C. Configure multiple Trident backends (e.g., 'ontap-nas' for standard volumes, 'ontap-nas-flexgroup' for parallel access) and corresponding StorageClasses.
- D. Use the NetApp DataOps Toolkit for Python within the Kubeflow pipeline components to dynamically create and manage Trident PVCs for each stage.
- E. Rely on hostPath volumes for all storage to ensure the highest performance.

Answer: CD

QUESTION 8

What is the primary architectural benefit of using technologies like RDMA (Remote Direct Memory Access) and GPUDirect Storage in a high-performance AI training cluster?

- A. They reduce the power consumption of the storage array by offloading checksum calculations.
- B. They allow GPUs to communicate directly with each other and with NVMe storage, bypassing the host CPU and system memory, which significantly reduces I/O latency and CPU overhead.
- C. They enable the use of lower-cost Ethernet switches in place of InfiniBand fabrics without any performance degradation.
- D. They automatically encrypt data in-flight between the GPU servers and the storage system.

Answer: B

QUESTION 9

A company is running its AI training workloads on a NetApp AFF A-Series system. To manage costs, they want to automatically move inactive training datasets and older model checkpoints from the high- performance all-flash tier to a lower-cost object storage tier, such as an on-premises StorageGRID or a public cloud bucket. The process must be transparent to the data scientists and not require changes to their scripts or file paths.

Which two NetApp technologies should be combined to achieve this goal? (Choose 2.)

- A. NetApp SnapMirror
- B. NetApp FabricPool
- C. NetApp StorageGRID
- D. NetApp XCP
- E. NetApp FlexCache

Answer: BC

QUESTION 10

An organization recently suffered a ransomware attack that encrypted several volumes on their primary ONTAP storage system, including a critical volume containing curated training data. The security team needs to implement a solution that can proactively detect and block ransomware-like file I/O patterns and automatically create a secure Snapshot copy before any damage is done.

The current ONTAP configuration is as follows:

ONTAP_Version: 9.12.1

Security_Features: SnapLock (Compliance Mode) on archive volumes

Anti-Virus_Scan: Enabled (Vscan)

Ransomware_Detection: Not configured

Which ONTAP feature should be enabled to provide this proactive, automated protection?

- A. Increase the frequency of scheduled Snapshots.
- B. Enable Multi-Admin Verification (MAV) on all volumes.
- C. Enable Autonomous Ransomware Protection (ARP).
- D. Use BlueXP backup and recovery to back up the volumes to the cloud more frequently.

Answer: C

QUESTION 11

An AI infrastructure architect is tasked with designing a solution to address two critical challenges in a large, multi-petabyte AI environment:

1. Cost: A significant portion of the data on the high-performance all-flash storage is inactive but must remain online. The cost of storing this cold data on the performance tier is prohibitive.
2. Traceability: Data scientists need a simple, space-efficient way to version their datasets at key points in their workflow to ensure reproducibility.

The environment consists of NetApp AFF A-Series and NetApp StorageGRID systems. Which combination of NetApp technologies should the architect implement to solve both challenges simultaneously? (Select all that apply.)

- A. Use NetApp XCP to periodically move cold data from the AFF systems to StorageGRID.
- B. Implement NetApp FabricPool to automatically tier cold data blocks from the AFF systems to StorageGRID.
- C. Use NetApp SnapMirror to replicate volumes from the AFF systems to StorageGRID for archival.
- D. Train data scientists to use NetApp Snapshots to create point-in-time, read-only versions of their data volumes.
- E. Implement NetApp FlexClone to create full, writable copies of datasets for each experiment.
- F. Use BlueXP backup and recovery to create backups on StorageGRID, then delete the original volumes from the AFF systems.

Answer: BD

Thank You for Trying Our Product

Lead2pass Certification Exam Features:

- ★ More than **99,900** Satisfied Customers Worldwide.
- ★ Average **99.9%** Success Rate.
- ★ **Free Update** to match latest and real exam scenarios.
- ★ **Instant Download** Access! No Setup required.
- ★ Questions & Answers are downloadable in **PDF** format and **VCE** test engine format.
- ★ Multi-Platform capabilities - **Windows, Laptop, Mac, Android, iPhone, iPod, iPad**.
- ★ **100%** Guaranteed Success or **100%** Money Back Guarantee.
- ★ **Fast**, helpful support **24x7**.



View list of all certification exams: <http://www.lead2pass.com/all-products.html>



Microsoft



ORACLE



CITRIX



JUNIPER
NETWORKS



EMC²
where information lives

10% Discount Coupon Code: ASTR14