



Documentazione di AI Data Engine

AI Data Engine

NetApp

March 13, 2026

This PDF was generated from <https://docs.netapp.com/it-it/ai-data-engine/index.html> on March 13, 2026.
Always check docs.netapp.com for the latest.

Sommario

Documentazione di AI Data Engine	1
Note di rilascio	2
Novità di AI Data Engine	2
Novità nella versione iniziale di AIDE 9.18.1	2
Limitazioni note di AI Data Engine	3
Limitazioni note per la versione iniziale di AIDE 9.18.1	3
Inizia	6
Scopri il tuo sistema AI Data Engine	6
Scopri AI Data Engine	6
Architettura di AI Data Engine	7
Componenti e responsabilità di AIDE per ruolo	11
Avvio rapido per AI Data Engine	16
Installa AIDE	17
Requisiti per l'installazione di AI Data Engine	17
Installa il sistema storage AFX per AI Data Engine	18
Installa i tuoi nodi di elaborazione dati	18
Configura il tuo sistema AIDE	28
Imposta AI Data Engine	28
Installa una licenza AIDE nel tuo sistema AFX	30
Configura OpenID Connect per AIDE in ONTAP	31
Configura gli spazi di lavoro	34
Preparare i dati per AI Data Engine	34
Crea uno spazio di lavoro in AI Data Engine	35
Crea uno spazio di lavoro	35
Rivedi i dettagli dell'area di lavoro	36
Aggiornamenti e versioning dell'area di lavoro	36
Assegnare l'accesso utente agli spazi di lavoro di AI Data Engine	37
Amministrare e monitorare	39
Monitorare i processi del cluster	39
Visualizza lo stato del sistema AIDE e del cluster	39
Visualizza approfondimenti per ottimizzare il tuo sistema AIDE	41
Visualizza eventi di sistema AIDE, lavori e registro di controllo	42
Gestire gli spazi di lavoro di AI Data Engine	44
Esaminare lo stato dell'area di lavoro	44
Modifica le proprietà dell'area di lavoro e la pianificazione dell'intervallo di refresh	45
Aggiungi container di dati a un workspace esistente	45
Rimuovi i container di dati da un workspace	46
Gestisci gli utenti dell'area di lavoro	46
Eliminare un workspace	46
Aggiorna e mantieni il tuo sistema AIDE	47
Aggiornamenti e compatibilità del sistema AI Data Engine	47
Aggiorna AI Data Engine software	48
Aggiungi nodi di elaborazione dati al tuo AIDE cluster	49

Sostituisci un nodo nel tuo cluster AIDE	51
Gestisci la vettorializzazione e le raccolte di dati	54
Avvio rapido Data-to-RAG per AI Data Engine	54
Esplora i metadati del workspace nella AI Data Engine Console	55
Sign in to AIDE Console come data engineer o data scientist	55
Visualizza i tuoi spazi di lavoro accessibili	55
E ora?	56
Crea raccolte di dati in AI Data Engine Console	56
Creare una raccolta di dati dai metadati del workspace	56
Pubblica una raccolta di dati	57
Aggiornare o eliminare una raccolta di dati	57
E ora?	58
Visualizza le raccolte di dati in AI Data Engine	58
Visualizza raccolte di dati a livello di cluster	59
Monitorare i lavori e gli eventi correlati alla raccolta	59
Visualizza le raccolte di dati dalla AIDE Console	60
E ora?	60
Implementare guardrail	61
Definisci le tue policy Data Guardrails in AI Data Engine per il tuo data estate	61
Comprendere i tipi di policy	61
Abilita classificatori	61
Gestisci le categorie del classificatore	62
Crea e gestisci le policy Data Guardrails	63
Come le policy interagiscono con gli spazi di lavoro	64
Informazioni correlate	64
FAQ per NetApp AI Data Engine	65
Nozioni di base su AIDE	65
Utenti e ruoli	65
Requisiti e deployment	65
Gestione e interfacce	66
Funzionalità e capacità	66
Integrazione e interoperabilità	67
Distribuzione e licenze	67
Note legali	69
Copyright	69
Marchi	69
Brevetti	69
Informativa sulla privacy	69
Open source	69
AI Data Engine	69

Documentazione di AI Data Engine

Note di rilascio

Novità di AI Data Engine

AI Data Engine (AIDE) 9.18.1 è la versione iniziale della piattaforma di NetApp per la gestione dei dati AI. Questa versione introduce un motore di metadati e flussi di lavoro di gestione che consentono alle organizzazioni di catalogare e organizzare dati non strutturati per carichi di lavoro AI, fornendo le basi per funzionalità avanzate di governance e vettorializzazione. La governance avanzata (guardrails) e la vettorializzazione sono disponibili per i clienti che dispongono delle licenze AI Data Engine appropriate.

Novità nella versione iniziale di AIDE 9.18.1

AIDE 9.18.1 introduce le seguenti funzionalità fondamentali:

"Metadata Engine per la catalogazione dei dati AI"

La versione iniziale include un Metadata Engine che cataloga file e oggetti nei cluster ONTAP.

Le caratteristiche principali includono:

- Estrazione automatizzata di metadati (attributi principali ed estesi, tag di oggetti) da volumi ONTAP locali e remoti su cluster peered.
- API REST centralizzate per query e filtraggio per applicazioni che richiedono una vista globale dei dati aziendali.
- Archiviazione scalabile dei metadati.
- Estrazione automatica dei metadati attivata durante la creazione dell'area di lavoro.

"Gestione dello spazio di lavoro"

Gli spazi di lavoro forniscono un raggruppamento logico delle fonti di dati (volumi) per i progetti AI.

La versione iniziale supporta:

- Creazione di spazi di lavoro che si estendono su volumi ONTAP locali e remoti (utilizzando il peering dei cluster).
- Assegnazione dei controlli di accesso agli spazi di lavoro, supportando ambienti multiutente e multi-tenant.
- Estrazione automatica dei metadati e popolamento del catalogo al momento della creazione dell'area di lavoro.

"Data Sync per la valuta automatizzata dei dati"

Data Sync mantiene aggiornati i cataloghi di metadati e le raccolte di dati man mano che i dati di origine cambiano, senza alcun intervento manuale.

Le caratteristiche principali includono:

- Sincronizzazione automatica dei dati da cluster ONTAP remoti o locali utilizzando la replica SnapMirror basata su policy.
- Aggiornamenti incrementali che propagano solo i dati modificati, riducendo l'overhead.

- Intervalli di aggiornamento configurabili per workspace.
- Monitoraggio a livello di workspace dello stato e dell'attività di sincronizzazione.

"Configurazione e gestione del cluster"

La versione iniziale include i seguenti flussi di lavoro:

- Rilevamento e aggiunta di nodi di elaborazione dati (DCN) durante la configurazione del cluster.
- Creazione di VM di storage dedicate ai metadati per il Metadata Engine.
- Configurazione delle interfacce del servizio Data Engine per l'accesso ai metadati a livello di cluster.
- Peering con altri cluster ONTAP per estendere la catalogazione dei metadati nell'intero patrimonio di dati.

"Autenticazione OpenID Connect (OIDC)"

- Autenticazione basata su OIDC/OAuth per l'accesso sicuro a ONTAP System Manager e Data Engine Console con Microsoft Entra ID e Active Directory Federation Services (ADFS).
- Controlli di accesso basati sui ruoli per la gestione dell'area di lavoro e dei metadati.

"Funzionalità avanzate di cura e governance dei dati"

Le seguenti funzionalità sono disponibili per i clienti che dispongono delle licenze AI Data Engine appropriate:

- **Vettorizzazione e RAG:** creazione di raccolte di dati, embedding ed endpoint di recupero nella AI Data Engine Console, utilizzando metadati dagli spazi di lavoro AIDE.
- **Governance basata su guardrail:** definizione di guardrail policy nella AI Data Engine Console e associazione di tali policy con i workspace in ONTAP System Manager.

Hardware e piattaforme supportati

AI Data Engine 9.18.1 viene eseguito su cluster ONTAP AI data platform che combinano:

- Nodi di storage AFX 1K
- Nodi di calcolo dati NetApp

Informazioni correlate

- ["Limitazioni note di AI Data Engine"](#)
- ["Scopri l'architettura e i componenti di AI Data Engine"](#)

Limitazioni note di AI Data Engine

Le limitazioni note identificano piattaforme, dispositivi o funzioni non supportati da questa versione del prodotto o che non interagiscono correttamente con esso. Esaminare attentamente queste limitazioni.

Limitazioni note per la versione iniziale di AIDE 9.18.1

Queste limitazioni si applicano al Metadata Engine, ai data compute nodes e ai flussi di lavoro di gestione in AIDE 9.18.1.

Requisiti e gestione dei nodi di elaborazione

- **Requisito minimo del nodo di calcolo dati**

I cluster AIDE richiedono un minimo e un massimo di 3 nodi di elaborazione dati (DCN) per la funzionalità Metadata Engine. I cluster con meno di 3 DCN non possono abilitare le funzionalità Metadata Engine.

- **Nessun supporto per l'aggiunta di nodi DCN tramite NetApp Console**

Gli aggiornamenti e le aggiunte dei nodi DCN devono essere eseguiti utilizzando ONTAP System Manager, non tramite NetApp Console.

Fonti dati supportate

- **Nessun supporto per bucket ONTAP S3 o StorageGRID come origini dati**

Solo i volumi ONTAP (locali o remoti) sono supportati come origini dati per gli spazi di lavoro e la catalogazione dei metadati. I bucket ONTAP S3 e gli oggetti StorageGRID non possono essere aggiunti agli spazi di lavoro e non sono indicizzati dal Metadata Engine in questa versione.

- **Nessun supporto per la creazione di spazi di lavoro con volumi FlexCache**

FlexCache volumes non possono essere aggiunti come origini dati agli spazi di lavoro.

Limitazioni di aggiornamento e ripristino del software

- **Aggiornamenti software manuali solo per nodi DCN**

Gli aggiornamenti software automatici per i cluster DCN non sono supportati in AIDE 9.18.1. Il software del nodo DCN può essere aggiornato solo caricando l'immagine da un client locale. Il download di immagini da server esterni (HTTP/FTP) non è supportato.

- **Nessun ripristino del DCN cluster software**

Il software del cluster DCN non può essere ripristinato a una versione precedente. Sono consentiti solo gli aggiornamenti a versioni successive.

- **Nessun ripristino ONTAP dei cluster di storage AFX**

I cluster di storage AFX non possono essere ripristinati a versioni precedenti di ONTAP. Sono consentiti solo gli aggiornamenti a versioni successive.

Ciclo di vita del workspace e configurazione dell'accesso

- **Nessuna eliminazione temporanea o ripristino per i workspace**

L'eliminazione di un'area di lavoro è definitiva. Non è possibile ripristinare le aree di lavoro eliminate.

- **Nessun supporto per la configurazione OIDC durante la configurazione iniziale del cluster**

La configurazione OIDC/OAuth deve essere eseguita dopo la creazione del cluster utilizzando ONTAP System Manager.

Informazioni correlate

- "Novità in AI Data Engine"

Inizia

Scopri il tuo sistema AI Data Engine

Scopri AI Data Engine

La piattaforma NetApp AI Data Engine (AIDE) è una piattaforma enterprise progettata per accelerare e semplificare l'elaborazione, la gestione e la governance dei dati guidate dall'intelligenza artificiale. AIDE può aiutare a trasformare grandi quantità di dati non strutturati in dataset strutturati e pronti per l'AI. È progettata per soddisfare le esigenze dei moderni carichi di lavoro di machine learning (ML) e generative AI (GenAI), supportando sia le operazioni IT tradizionali che i nuovi ruoli incentrati sull'AI.

AIDE affronta le sfide dell'AI

AIDE è progettato per aiutare le organizzazioni a gestire i dati per i carichi di lavoro AI e offre le seguenti funzionalità chiave:

- **Gestione centralizzata dei metadati:** AIDE raccoglie e cataloga metadati dai volumi ONTAP, rendendo possibile la ricerca, la classificazione e l'applicazione di policy di governance ai dataset.
- **Elaborazione automatizzata dei dati:** AIDE supporta la creazione di pipeline di dati per carichi di lavoro AI e ML, inclusa la possibilità di generare vector embeddings per la ricerca semantica (con licenze appropriate).
- **Isolamento dei dati e controllo degli accessi:** AIDE applica controlli di accesso e isolamento di base dei dati per più team o progetti.
- **Integrazione con gli strumenti NetApp:** AIDE funziona con ONTAP System Manager per l'amministrazione dello storage e fornisce un'interfaccia dedicata (AI Data Engine Console) per data engineer e data scientist per gestire raccolte di dati e flussi di lavoro.

Caratteristiche di progettazione di alto livello

Le seguenti caratteristiche di progettazione definiscono come AI Data Engine è costruito per soddisfare le esigenze dei carichi di lavoro AI:

- **Servizi basati su microservizi:** utilizza Kubernetes per orchestrare servizi modulari e resilienti per la catalogazione dei metadati, la ricerca vettoriale e la gestione dell'infrastruttura.
- **Sicurezza enterprise:** implementa crittografia, controllo degli accessi in base al ruolo (RBAC) e auditing su tutti i dati e metadati.
- **Accesso ai dati multiprotocollo:** supporta NFS e SMB per una raccolta e un recupero flessibili dei dati.
- **Pipeline di dati automatizzate:** tiene traccia delle modifiche dei dati, crea incorporamenti e gestisce database vettoriali per applicazioni AI.

Come i dati fluiscono attraverso AIDE

Comprendere come i dati fluiscono attraverso AIDE aiuta a illustrare il valore della piattaforma per i team di AI/ML:

1. **Ingestione dati:** i file vengono archiviati in volumi ONTAP utilizzando protocolli standard (NFS e SMB). I dati possono risiedere su storage AIDE locale (il cluster AFX all'interno della tua distribuzione AIDE) o su

cluster ONTAP remoti. I dati provenienti dai cluster remoti vengono sincronizzati con il cluster AFX locale tramite ONTAP SnapMirror, quindi tutti i dati elaborati da AIDE sono infine archiviati e accessibili localmente.



I bucket S3 non sono supportati come origini dati per spazi di lavoro o raccolte di dati.

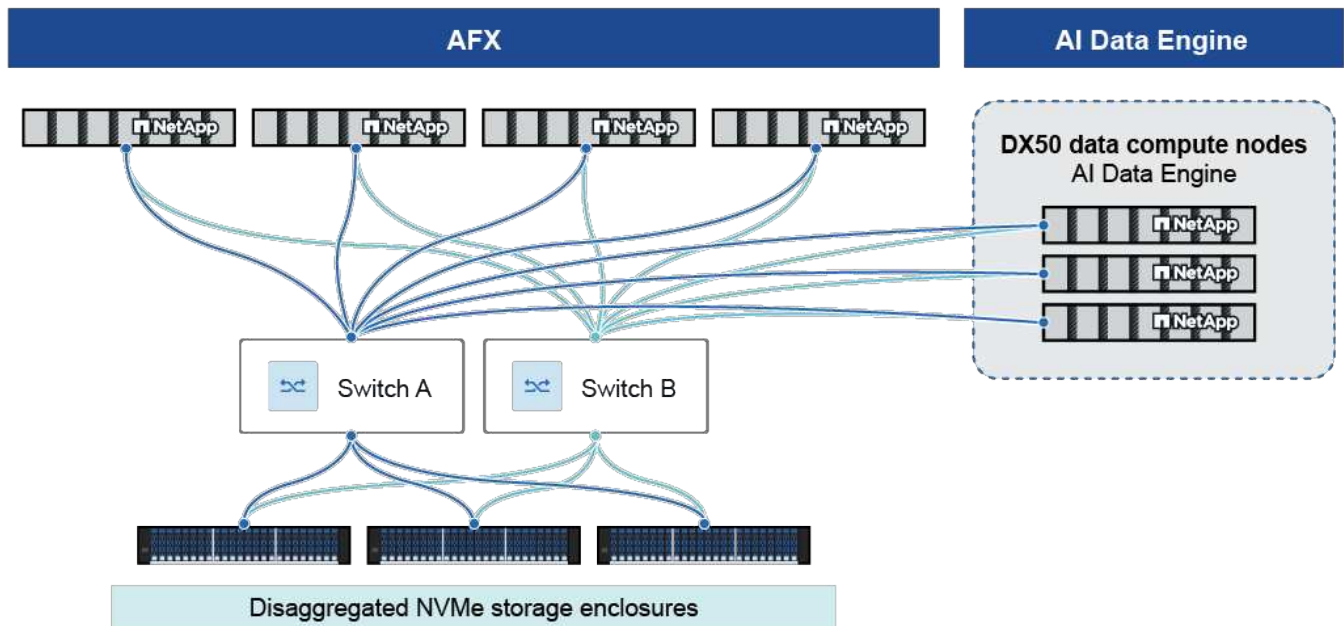
1. **Creazione di aree di lavoro:** gli amministratori di storage definiscono le aree di lavoro in ONTAP System Manager, raggruppando i volumi ONTAP correlati per progetti, team o flussi di lavoro specifici. Le autorizzazioni di accesso e le policy di governance vengono assegnate a livello di area di lavoro.
2. **Estrazione dei metadati:** AIDE analizza automaticamente file e oggetti negli spazi di lavoro, estraendo metadati (tipo di file, dimensione, timestamp, attributi personalizzati) e memorizzandoli in un catalogo centralizzato. Questo avviene continuamente man mano che i dati cambiano.
3. **Classificazione e governance:** i classificatori analizzano i dati alla ricerca di informazioni sensibili (PII, dati finanziari) o tipologie di documenti (legali, HR). Le policy Guardrail applicano automaticamente la redazione o le restrizioni di accesso.
4. **Creazione di raccolte di dati:** gli ingegneri dei dati e gli scienziati dei dati utilizzano l'AI Data Engine Console per interrogare il catalogo dei metadati, filtrare i risultati e assemblare raccolte di dati curate per specifici compiti di AI.
5. **Vettorizzazione:** per le raccolte che richiedono una ricerca semantica, AIDE genera incorporamenti utilizzando i modelli AI selezionati. I vettori vengono memorizzati nel database di vettori per un recupero dalle performance elevate.
6. **Consumo di AI/ML:** le applicazioni accedono ai dati attraverso più percorsi:
 - Accesso diretto a file/oggetti tramite NFS o SMB
 - Query di ricerca semantica sul database vettoriale
 - Endpoint RAG che combinano il recupero dei dati con l'integrazione del modello GenAI
 - Accesso REST API per flussi di lavoro programmatici

Questo flusso di lavoro automatizzato e basato su policy riduce il tempo e lo sforzo manuale necessari per preparare i dati per l'AI, consentendo ai team di concentrarsi sullo sviluppo del modello e sulle informazioni anziché sulla gestione dei dati.

Architettura di AI Data Engine

AIDE è basato su un'architettura scalabile e fault-tolerant che separa storage ed elaborazione, consentendo elevate prestazioni e flessibilità per i carichi di lavoro AI.

Componenti fisici



Nodi controller AFX

I nodi controller AFX eseguono una personalità specializzata del software ONTAP progettata per supportare i requisiti dell'ambiente AFX. I client accedono ai nodi tramite diversi protocolli, tra cui NFS e SMB. Ogni nodo ha una vista completa dello storage, a cui può accedere in base alle richieste del client. I nodi sono stateful con memoria non volatile per mantenere le informazioni di stato critiche e includono ulteriori miglioramenti specifici per i carichi di lavoro di destinazione.

Per garantire elevata disponibilità e prestazioni, sono necessari almeno quattro nodi controller AFX per le distribuzioni AIDE.

Nodi di elaborazione dati

I nodi di elaborazione dati (DCN) sono server basati su Linux con elevate risorse di CPU, RAM e GPU, dedicati alle attività di elaborazione dati di intelligenza artificiale. Ospitano servizi specifici per l'intelligenza artificiale, come la catalogazione dei metadati, la ricerca vettoriale e le pipeline di embedding.

Sono necessari esattamente *tre* DCN per le distribuzioni AIDE.

Switch di cluster/storage

Switch ridondanti ad alta velocità (100GbE o superiore) collegano ONTAP e DCN per il trasferimento dei dati a bassa latenza e per un'elevata disponibilità.

Scaffali di storage

Gli scaffali NVMe-oF con SSD ad alta densità garantiscono una latenza ultra-bassa e ridondanza, supportando lo storage su scala PB.

Networking

Tutti i DCN e i nodi di storage ONTAP sono collegati tramite switch di cluster ridondanti ad alta velocità (minimo 100GbE). Questa architettura separa le risorse di calcolo e di storage, consentendo a ciascuna di scalare in modo indipendente e ottimizzando sia le prestazioni che l'utilizzo delle risorse.

Il networking tra DCN e nodi ONTAP è isolato utilizzando VLAN dedicate e IPspaces sugli switch del cluster. Ciò garantisce che tutte le comunicazioni, come l'accesso ai dati, le API di gestione e il traffico dei servizi interni, rimangano sicure, efficienti e non interferiscano con altre operazioni di rete.

Caratteristiche principali di AI Data Engine

Le funzionalità principali di AI Data Engine (AIDE) lavorano insieme per automatizzare, proteggere e accelerare il ciclo di vita dei dati AI. Ogni funzionalità è implementata come un set di microservizi in esecuzione su DCN, integrati con lo storage ONTAP ed esposti tramite API REST e interfacce di gestione.

Metadata Engine

Il Metadata Engine genera automaticamente una vista strutturata, aggiornata e interattiva del tuo patrimonio di dati NetApp.

Licenza e accesso

Il Metadata Engine è incluso con la licenza base ONTAP One ed è disponibile all'installazione di AIDE.

Puoi accedervi tramite ONTAP System Manager.

Capacità

- Cataloga i metadati per tutte le fonti di dati, inclusi i volumi archiviati localmente sul cluster AFX e quelli sincronizzati dai cluster ONTAP remoti.
- Estrae automaticamente i metadati e popola il catalogo man mano che i dati vengono acquisiti o modificati.
- Fornisce l'accesso all'API REST per l'interrogazione dei metadati, consentendo ai professionisti dei dati e agli amministratori di storage di scoprire, classificare e comprendere i dati.
- Scarica le query sui metadati dal percorso dei dati, riducendo il carico del traffico NFS sui sistemi di storage.
- Supporta grandi record di metadati con funzionalità di indicizzazione e ricerca.
- Si integra con le astrazioni dell'area di lavoro e della raccolta di dati per applicare il controllo degli accessi e la governance.

Sincronizzazione dati

Data Sync è un servizio automatizzato in background che garantisce che il catalogo dei metadati e le raccolte di dati rimangano aggiornati e coerenti con le fonti di dati sottostanti, anche quando i dati di origine cambiano.

Licenza e accesso

La funzionalità Data Sync è inclusa nella licenza base ONTAP One ed è disponibile al momento dell'installazione di AIDE.

Capacità

- Sincronizza i dati da cluster ONTAP remoti o locali utilizzando la replica SnapMirror basata su policy. I dati dai cluster remoti vengono copiati sul cluster AFX locale per l'elaborazione AIDE.
- Aggiorna in modo incrementale in base alle modifiche rilevate, propagando solo i dati modificati.
- Fornisce mobilità dei dati sicura e incrementale e sincronizzazione nell'intero patrimonio di dati.
- Pianifica e monitora gli intervalli di sincronizzazione con frequenze di aggiornamento configurabili per workspace.
- Si integra con i flussi di lavoro di creazione dell'area di lavoro per estrarre e aggiornare i metadati man mano che vengono aggiunte nuove fonti di dati.

Data Guardrails

Il servizio Data Guardrails fornisce governance e protezione continue e automatizzate per i dati sensibili durante l'intero ciclo di vita dell'AI.

Licenza e accesso

La funzionalità Data Guardrails non è inclusa nella licenza base ONTAP One e richiede una licenza AIDE separata.

È possibile accedere alla funzionalità Data Guardrails tramite la AI Data Engine Console.

Capacità

- Esegue continuamente la scansione, la classificazione e la categorizzazione dei dati.
- Identifica i dati sensibili e i rischi utilizzando classificatori integrati e personalizzabili per attività quali il rilevamento di PII.
- Automatizza la gestione dei dati sensibili tramite redazione basata su policy, mascheramento e restrizioni di accesso.
- Applica gli standard aziendali e normativi tramite Data Guardrails associati agli spazi di lavoro.
- Limita l'accesso ai file o ai volumi sensibili come configurato, con audit logging e reporting di conformità.
- Si integra con la gestione dell'area di lavoro e della raccolta di dati per applicare Data Guardrails in modo coerente nei flussi di lavoro dei dati AI.

Data Curator

Il servizio Data Curator consente una rapida scoperta, ricerca, vettorializzazione e recupero dei dati per applicazioni AI e GenAI.

Licenza e accesso

La funzionalità Data Curator non è inclusa nella licenza base ONTAP One e richiede una licenza AIDE separata.

È possibile accedere a Data Curator tramite la AI Data Engine Console.

Capacità

- Cerca dati rilevanti nello storage utilizzando il catalogo centralizzato dei metadati.
- Fornisce strumenti per gli scienziati dei dati per creare raccolte di dati curate.
- Genera automaticamente incorporamenti vettoriali a livello di storage.
- Fornisce un endpoint di recupero sicuro per le applicazioni AI, supportando la ricerca semantica vettoriale e il re-ranking.
- Si integra con strumenti e tecnologie di AI, inclusi pipeline di Retrieval-Augmented Generation (RAG) e framework di AI agentica.
- Fornisce API REST per l'accesso programmatico alle raccolte di dati, alla ricerca vettoriale e agli endpoint di recupero.

Sicurezza e multi-tenancy

La piattaforma applica sia il controllo degli accessi in base al ruolo (RBAC) sia le liste di controllo degli accessi a livello di risorsa (ACL). Tutte le azioni API e utente vengono sottoposte ad audit e tutti i dati sono crittografati a riposo e in transito. I singoli tenant sono isolati per dati e metadati.

Informazioni correlate

- ["Installa una licenza AIDE"](#)
- ["Avvio rapido Data-to-RAG"](#)

Componenti e responsabilità di AIDE per ruolo

Componenti di AI Data Engine e interazioni in base al ruolo

AI Data Engine (AIDE) è costituito da numerosi componenti principali che lavorano insieme per fornire una piattaforma completa di gestione ed elaborazione dei dati per i carichi di lavoro AI. Questi componenti includono workspaces, raccolte di dati, vector databases, guardrails, cataloghi di metadati, retrieval endpoints e classificatori. Ogni componente svolge un ruolo specifico nel consentire un’efficiente scoperta, curation, governance e integrazione dei dati con applicazioni AI/ML.

Ogni utente AIDE interagisce con i componenti AIDE in modo diverso a seconda del proprio ruolo.

Ruoli utente focalizzati su storage e dati

AIDE introduce nuovi ruoli utente, continuando a supportare i tradizionali ruoli di amministrazione del sistema ONTAP:

Utenti dello storage

- **Amministratore di storage:** gestisce la configurazione dei cluster AFX e AIDE, la rete, il provisioning dello storage e l’accesso degli utenti.

Utenti dei dati

- **Data engineer:** crea e ottimizza pipeline AI/ML, gestisce data collection e integra modelli AI.
- **Data scientist:** scopre, seleziona e analizza dataset, crea data collection e sfrutta gli endpoint di retrieval per le applicazioni GenAI.

Ruolo (nome RBAC)	Descrizione
Amministratore dello storage (admin)	Gestisce la configurazione dei cluster AFX e AIDE, il networking, il provisioning dello storage e l’accesso degli utenti. Assegna ruoli RBAC agli utenti che determinano il livello di accesso alle interfacce e alle funzionalità di AIDE. Questo ruolo di amministratore ha pieno accesso alla gestione tramite ONTAP System Manager e AI Data Engine Console.
Ingegnere dei dati (data-engineer)	Crea e ottimizza pipeline AI/ML, gestisce raccolte di dati e integra modelli AI. Questo ruolo ha accesso alla AI Data Engine Console per i workflow di data engineering.
Scienziato dei dati (data-scientist)	Individua, seleziona e analizza set di dati, crea raccolte di dati e sfrutta gli endpoint di recupero per le applicazioni GenAI. Questo ruolo ha accesso all’AI Data Engine Console per i flussi di lavoro di data science.

Componenti del sistema AIDE

Ogni utente AIDE (amministratori di storage, data engineer e data scientist) interagisce con i componenti AIDE in base al proprio ruolo.

Spazi di lavoro

Un'area di lavoro è un segmento logico di dati all'interno del cluster, che raggruppa volumi per uno specifico progetto, team o workflow. Le aree di lavoro definiscono l'ambito di visibilità, accesso e governance dei dati in AIDE.

Catalogo dei metadati

Un database centralizzato e scalabile che memorizza i record dei metadati per tutti i file e gli oggetti nel cluster locale, inclusi i dati sincronizzati dai cluster ONTAP remoti tramite ONTAP SnapMirror o cluster peering.

Consente una ricerca e un filtraggio avanzati e interattivi.

Classificatori

I classificatori sono strumenti (integrati o personalizzati) che analizzano e contrassegnano i file per tipi specifici di dati sensibili (ad esempio, PII, finanziari, sanitari) o categorizzano i documenti per tipo (ad esempio, legali, HR, vendite).

Raccolta di dati

Una raccolta di dati è un gruppo selezionato di file o oggetti correlati provenienti da un'area di lavoro, definito da una query specificata dall'utente per l'utilizzo nei workflow GenAI. Il contenuto dei file nella raccolta di dati, dopo la pubblicazione, è disponibile per la ricerca semantica tramite API per applicazioni GenAI.

database vettoriale

Il vector database memorizza gli embeddings generati dalle raccolte di dati, consentendo la ricerca e il recupero semantico dalle performance elevate per le applicazioni AI e GenAI.

Guardrail

I guardrail sono meccanismi basati su policy che applicano la governance dei dati, la classificazione e la protezione (come la redazione o le restrizioni di accesso) durante l'intero ciclo di vita dei dati AI.

Endpoint di recupero (RAG endpoint)

Un endpoint di recupero (talvolta denominato Retrieval-Augmented Generation o "RAG" endpoint) è un'API sicura che consente alle applicazioni AI e GenAI di accedere a dati, contesto o embedding rilevanti da raccolte curate e dal vector database.

Gli endpoint RAG sono progettati per supportare flussi di lavoro di AI avanzati, come la ricerca semantica e le risposte contestuali nei modelli di AI generativa. Collegando le applicazioni di AI a un endpoint di retrieval, puoi migliorare l'accuratezza e la pertinenza del modello fornendo accesso real-time a dataset curati e AI-ready gestiti da AIDE.

Informazioni correlate

- ["Come gli amministratori di storage AIDE lavorano con i componenti AIDE"](#)
- ["Come gli ingegneri dei dati AIDE lavorano con i componenti AIDE"](#)

- ["Come gli data scientist di AIDE lavorano con i componenti di AIDE"](#)

Interfacce di AI Data Engine

AI Data Engine (AIDE) fornisce tre interfacce principali per l'interazione e l'automazione degli utenti. Ogni ruolo, come amministratori di storage, data engineer e data scientist, utilizza queste interfacce in base ai propri compiti e responsabilità specifici.

ONTAP System Manager

ONTAP System Manager è un'interfaccia basata su web progettata per gli amministratori di storage. Fornisce flussi di lavoro per la configurazione del cluster, la gestione degli spazi di lavoro, il monitoraggio DCN e l'associazione di policy di guardrail.

AI Data Engine Console

La AI Data Engine Console è un'interfaccia dedicata per data engineer e data scientist. Consente agli utenti di esplorare fonti di dati, creare e gestire raccolte di dati, configurare pipeline di dati, applicare classificatori e interagire con guardrail e funzionalità di ricerca vettoriale. La console fornisce strumenti avanzati per la scoperta, la cura e l'integrazione dei dati con workflow AI/ML.

REST API

AIDE espone l'API REST ONTAP per l'automazione, l'integrazione e l'accesso programmatico. L'API supporta la configurazione del cluster, la gestione di workspace e raccolte, le query sui metadati, la ricerca vettoriale e gli endpoint di recupero.

Scopri come gli amministratori di storage di AI Data Engine lavorano con i componenti AIDE

In qualità di amministratore dello storage, gestisci l'infrastruttura AIDE tramite ONTAP e la AIDE Console, effettuando il provisioning degli spazi di lavoro, applicando guardrail policy e monitorando lo stato del sistema. Il tuo ruolo si concentra sull'assicurare uno storage dei dati affidabile, sicuro e conforme per i carichi di lavoro AI.

Accesso al componente amministratore dello storage

Componente	Livello di accesso	Flusso di lavoro dell'amministratore dello storage
ONTAP System Manager	Gestisci (crea, modifica, elimina)	Utilizzi ONTAP System Manager come interfaccia principale per l'amministrazione del cluster, il provisioning dell'area di lavoro, la gestione delle policy guardrail e il monitoraggio dello stato del sistema.
AI Data Engine Console	Gestisci (crea, modifica, elimina)	Utilizzi la AI Data Engine Console per monitorare gli spazi di lavoro, visualizzare lo stato della raccolta e supervisionare l'attività del sistema nell'ambiente AIDE.

Componente	Livello di accesso	Flusso di lavoro dell'amministratore dello storage
API REST ONTAP	Gestisci (crea, modifica, elimina)	Utilizzi l'API REST per automatizzare le attività infrastrutturali, gestire gli spazi di lavoro e le policy di guardrail a livello di programmazione e integrare l'amministrazione di AIDE con strumenti e flussi di lavoro esterni.
Spazi di lavoro	Gestisci (crea, modifica, elimina)	Crei e gestisci gli spazi di lavoro utilizzando ONTAP System Manager. Selezioni quali fonti dati sono incluse, assegni autorizzazioni a data engineer e data scientist e alleggi policy di guardrail per applicare governance e conformità. Monitori anche lo stato di salute e l'accesso agli spazi di lavoro.
Raccolta di dati	Visualizza (sola lettura)	Visualizzi lo stato e l'integrità delle raccolte di dati all'interno di ogni workspace utilizzando System Manager. Garantisci che le fonti di dati sottostanti siano disponibili e protette, ma non crei né modifichi le raccolte.
Guardrail	Gestisci (crea, modifica, elimina)	Definisci e associ le guardrail policy agli spazi di lavoro utilizzando System Manager. Monitori lo stato delle guardrail e i report di conformità. Garantisci che le policy siano applicate e aggiornate secondo necessità.
Catalogo dei metadati	Monitor (visualizza stato di salute, stato, attività)	Ti assicuri che il catalogo dei metadati sia popolato e aggiornato. Monitori lo stato del catalogo e supporti il controllo degli accessi.
database vettoriale	Provision/Monitor (distribuisci, configura, visualizza stato)	Ti occupi del provisioning e del monitoraggio dell'infrastruttura del database vettoriale, assicurandoti che i nodi di elaborazione dati siano dotati di risorse GPU e di licenze adeguate. Supporti l'ambiente, ma non gestisci direttamente gli embeddings o le query.
Classificatori	Gestisci (crea, modifica, elimina)	Crei, configuri e gestisci classificatori e le relative categorie. Applichi i classificatori agli spazi di lavoro e ne monitori l'efficacia.

Scopri come gli ingegneri dei dati e gli scienziati dei dati di AI Data Engine lavorano con i componenti AIDE

In qualità di data engineer o data scientist, utilizzi l'AI Data Engine Console per esplorare i workspace a cui hai accesso, creare e gestire raccolte di dati, eseguire ricerche semantiche e integrare retrieval endpoint nei workflow AI/ML.

Gli ingegneri dei dati si concentrano sulla trasformazione dei dati raw in set di dati pronti per l'AI creando raccolte, configurando pipeline di embedding e controllando quali utenti possono accedere alle raccolte pubblicate. I data scientist si concentrano sullo sfruttamento di set di dati curati per l'analisi, l'addestramento

dei modelli e le applicazioni GenAI, senza gestire il controllo degli accessi o l'infrastruttura.

Accesso al componente utente dati

Componente	Livello di accesso	Flusso di lavoro dell'ingegnere dei dati	Flusso di lavoro del data scientist
AI Data Engine Console	Gestisci (crea, modifica, elimina)	La AI Data Engine Console è la tua interfaccia principale per le attività quotidiane, tra cui la scoperta dei dati, la gestione della raccolta, la configurazione della pipeline e la pubblicazione di endpoint RAG o di recupero, per gli spazi di lavoro a cui sei autorizzato ad accedere.	La AI Data Engine Console è la tua interfaccia principale per l'esplorazione dei dati, il perfezionamento e il versioning delle raccolte all'interno degli spazi di lavoro a cui puoi accedere, e per la connessione di dataset curati ed endpoint di retrieval ad analisi, modellazione e workflow GenAI.
API REST ONTAP	Gestisci (crea, modifica, elimina)	Utilizzi l'API REST per automatizzare le operazioni del ciclo di vita della raccolta, attivare e monitorare le pipeline di embedding e integrare programmaticamente i flussi di lavoro dei dati con strumenti esterni.	Utilizzi l'API REST per accedere programmaticamente alle raccolte di dati, eseguire query di ricerca vettoriale e integrare endpoint di recupero in applicazioni AI/ML e framework agentici.
Spazi di lavoro	Visualizza/utilizza (sola lettura)	Esplori gli spazi di lavoro assegnati per identificare e comprendere le fonti di dati disponibili prima di creare raccolte.	Effettui ricerche negli spazi di lavoro assegnati per individuare file e oggetti rilevanti per specifiche attività di ricerca o modellazione.
Raccolta di dati	Gestisci (crea, modifica, elimina)	È possibile creare raccolte di dati selezionando e filtrando i dati sorgente tramite tag, classificazione e altri attributi, e gestire l'intero ciclo di vita della raccolta dalla creazione e dal versioning fino alla pubblicazione come endpoint RAG per l'uso da parte dell'AI. È inoltre possibile gestire quali data scientist e altri utenti possono accedere a ciascuna raccolta.	Crei, selezioni, annoti, versioni e perfezioni raccolte di dati all'interno degli spazi di lavoro a cui ti è stato dato accesso. Utilizzi queste raccolte come base per la ricerca semantica e i workflow GenAI.

Componente	Livello di accesso	Flusso di lavoro dell'ingegnere dei dati	Flusso di lavoro del data scientist
Catalogo dei metadati	Query/utilizzo (utilizzo per pipeline)	Utilizzi il catalogo dei metadati per valutare e selezionare le origini dati per l'ingestione, eseguendo query per individuare i file rilevanti e confermare che soddisfano i requisiti delle raccolte che stai creando all'interno degli spazi di lavoro assegnati.	Puoi cercare e filtrare i metadati negli spazi di lavoro a cui puoi accedere per individuare file e oggetti necessari per l'analisi o l'addestramento del modello, facendo affidamento sulla struttura del catalogo che è stata creata e mantenuta dai data engineer.
database vettoriale	- Gestisci incorporamenti/ricerca (data engineer) - Utilizzo/ricerca (data scientist)	È possibile attivare pipeline di embedding, monitorare lo stato di vettorializzazione, configurare i parametri di chunking ed embedding ed esporre endpoint di retrieval supportati dalla ricerca vettoriale. Applicazioni e agenti interrogano quindi questi endpoint tramite l'API per la ricerca semantica e i workflow RAG.	Si eseguono query di ricerca semantica sugli embeddings generati dalle pipeline gestite dagli ingegneri dei dati e si integrano i risultati del recupero nei flussi di lavoro GenAI o RAG per risposte di modelli contestuali. Non si configurano chunking, embeddings o parametri della pipeline.
Classificatori	Utilizza (consuma dati classificati)	Utilizzi i risultati della classificazione per annotare e contrassegnare i dati sorgente durante la preparazione della raccolta, assicurando che i contenuti in ingresso nelle pipeline siano etichettati correttamente per i flussi di lavoro AI a valle.	Si utilizzano dati pre-classificati per garantire che nelle analisi e nella modellazione vengano utilizzati solo contenuti conformi e pertinenti.

Avvio rapido per AI Data Engine

Per iniziare a utilizzare il tuo sistema AI Data Engine, devi installare i componenti hardware, configurare il cluster, impostare l'accesso ai dati dagli host al sistema storage e predisporre lo storage.

1

AIDE verrà installato con un nuovo cluster AFX o con uno esistente?

È necessario decidere se AIDE e AFX verranno installati contemporaneamente oppure se AIDE verrà integrato con un cluster AFX esistente.

2

Installa e configura il tuo hardware

"[Installare e configurare](#)" i nodi di calcolo del cluster AIDE. A seconda dell'ambiente di installazione, assicurati anche di "[installare](#)" l'hardware AFX.

3

Configura il tuo cluster

Utilizza ONTAP System Manager per guidarti attraverso un processo rapido e semplice per ["Configurare AIDE con un cluster AFX"](#).

4

Configurare gli spazi di lavoro e l'accesso ai dati

["Imposta un'area di lavoro e gli utenti per tale area di lavoro che possono accedere ai dati di AI Data Engine"](#).

E ora?

Ora puoi utilizzare ONTAP System Manager per gestire il tuo AI Data Engine e consentire ai tuoi data engineer e data scientist di iniziare con i loro workspace e configurazioni.

Installa AIDE

Requisiti per l'installazione di AI Data Engine

Esaminare i requisiti per l'installazione di AI Data Engine. AIDE richiede un sistema storage AFX, un minimo di tre Data Compute Node, switch di rete e cavi.

Requisiti hardware

L'AI Data Engine richiede un sistema storage AFX e un minimo di tre Data Compute Node. Il sistema AFX fornisce l'infrastruttura di storage, mentre i Data Compute Node ospitano i componenti software AIDE che consentono la gestione, la cura dei dati e le funzionalità AI.

- **AFX storage system:** include un controller AFX, uno shelf di dischi e uno switch di rete. L'AFX storage system è necessario per l'implementazione di AIDE.
- **Nodi di elaborazione dati:** sono richiesti almeno tre nodi di elaborazione dati. I nodi di elaborazione dati sono nodi hardware forniti da NetApp che ospitano il software AIDE, inclusi Metadata Engine, Data Sync, Data Curator e Data Guardrails.
 - Ogni Data Compute Node ha gli slot I/O 4 e 5 disponibili per la connettività. Lo slot 3 è riservato per la GPU. Gli slot 1 e 2 non sono popolati né accessibili.
 - Le porte e4a ed e5a vengono utilizzate per le connessioni cluster.
 - Le porte e4b ed e5b sono utilizzate per le connessioni di rete host.

Requisiti dello switch di rete

AI Data Engine richiede switch di rete per abilitare la connettività di rete host e la comunicazione tra nodi per i Data Compute Node.

- Switch client per la connettività di rete host (Cisco Nexus 9332D-GX2B o Cisco Nexus 9364D-GX2A)
- Switch di cluster per la comunicazione tra nodi (Cisco Nexus 9332D-GX2B o Cisco Nexus 9364D-GX2A)
- Switch di gestione opzionali per la gestione della rete

Requisiti di cablaggio

I seguenti cavi sono necessari per collegare i Data Compute Node agli switch di rete e alla rete di gestione.

- Cavi breakout da 400-GbE a 100-GbE (4x100GbE) per collegare i nodi agli switch client e cluster

Supporto multi-cluster

Dopo che AIDE è stato distribuito con AFX, può connettersi e gestire i dati di altri cluster ONTAP 9.18.1 e versioni successive utilizzando SnapMirror e il peering dei cluster.

Installa il sistema storage AFX per AI Data Engine

Installare il sistema storage AFX come primo passo per l'implementazione di AI Data Engine. Il sistema storage AFX fornisce la base dell'infrastruttura di storage ed è necessario prima di installare i data compute nodes.

Seguire le ["Documentazione di installazione di AFX 1K"](#) per installare il sistema storage AFX.

Cosa c'è dopo

Dopo aver completato l'installazione del sistema storage AFX, ["installare i Data Compute Node"](#).

Installa i tuoi nodi di elaborazione dati

Flusso di lavoro di installazione e configurazione per i data compute nodes per AI Data Engine

Per installare e configurare i nodi di elaborazione dati (DCN), è necessario esaminare i requisiti hardware, preparare il sito, installare e cablare i componenti hardware, accendere il sistema e configurare il cluster ONTAP.

1

["Esaminare i requisiti di installazione hardware"](#)

Assicurarsi di avere un sistema storage AFX 1K esistente installato, quindi verificare i requisiti hardware per l'installazione dei nodi di elaborazione dati per AIDE. Per informazioni sull'installazione del sistema storage AFX 1K, consulta ["Documentazione di installazione del sistema storage AFX 1K"](#).

2

["Preparati a installare i tuoi data compute nodes"](#)

Per prepararsi all'installazione dei nodi di elaborazione dati, è necessario predisporre il sito, verificare i requisiti ambientali ed elettrici e assicurarsi che ci sia spazio sufficiente nell'armadio. Quindi, disimballare l'attrezzatura, confrontarne il contenuto con la bolla di accompagnamento e registrare l'hardware per accedere ai vantaggi dell'assistenza. Sono necessari almeno tre nodi di elaborazione dati per implementare l'AI Data Engine.

3

["Installa l'hardware per i tuoi nodi di elaborazione dati"](#)

Installa i kit di guide per i tuoi data compute nodes. Fissa i tuoi data compute nodes all'interno dell'armadio. Infine, collega i dispositivi di gestione dei cavi sul retro del sistema per un instradamento organizzato dei cavi.

4

["Cabla i tuoi nodi di elaborazione dati"](#)

Per cablare l'hardware, collegare innanzitutto i nodi di elaborazione dati alla rete dati, quindi collegare i nodi di elaborazione dati agli switch del cluster.

"Accendi i nodi di elaborazione dati"

Dopo aver installato l'hardware del rack e cablato i nodi di elaborazione dati, è necessario accendere i DCN e i nodi controller per il sistema storage AFX, se non sono già accesi.

Requisiti di installazione per i nodi di elaborazione dati per AI Data Engine

Esaminare l'attrezzatura necessaria e le precauzioni di sollevamento per i nodi di calcolo dati per AI Data Engine.

Prerequisiti

Prima di installare i nodi di elaborazione dati per AIDE, assicurati di avere:

- Un sistema storage AFX 1K.



Per informazioni sull'installazione del sistema storage AFX 1K, consulta ["Documentazione di installazione del sistema storage AFX 1K"](#).

Attrezzatura necessaria per l'installazione

Per installare i nodi di elaborazione dati per AIDE, sono necessari le seguenti apparecchiature e strumenti.

- Accesso a un browser web per configurare i nodi di data compute
- Cinturino di scarica elettrostatica (ESD)
- Torcia elettrica
- Laptop o console con connessione USB/seriale
- Cacciavite Phillips #2

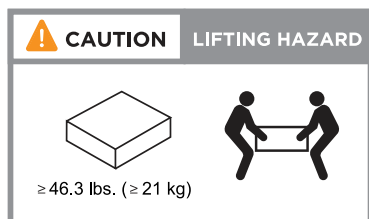
Precauzioni di sollevamento

I nodi di elaborazione dati sono pesanti. Prestare attenzione quando si sollevano e si spostano questi elementi.

Pesi dei nodi di calcolo dei dati

Prendere le necessarie precauzioni quando si sposta o si solleva il data compute node.

Un nodo di elaborazione dati può pesare fino a 46,3 lbs (21 kg). Per sollevare il nodo di elaborazione dati, utilizzare due persone o un sollevatore idraulico.



Informazioni correlate

- ["Informazioni sulla sicurezza e avvisi normativi"](#)

E ora?

Dopo aver esaminato i requisiti hardware, ["preparati a installare i tuoi data compute nodes"](#).

Preparati a installare i nodi di elaborazione dati per AI Data Engine

Preparati a installare i nodi di elaborazione dati per AI Data Engine preparando il sito, disimballando le scatole e confrontando il contenuto delle scatole con la bolla di accompagnamento, e registrando il sistema per accedere ai vantaggi dell'assistenza.

Fase 1: prepara il sito

Per installare i nodi di data compute, assicurati che il sito e l'armadio o il rack che intendi utilizzare soddisfino le specifiche per la tua configurazione.

Passaggi

1. Usa ["NetApp Hardware Universe"](#) per confermare che il tuo sito soddisfa i requisiti ambientali ed elettrici per i nodi di elaborazione dati.
2. Assicuratevi di disporre di spazio adeguato nell'armadio per i nodi di elaborazione dati all'interno dell'installazione esistente del sistema storage AFX 1K.
 - 1U per ogni nodo di elaborazione dati
 - 2U per ogni nodo controller AFX 1K
 - 2U per ogni shelf NX224
 - 1U o 2U per switch, a seconda del modello di switch.

Fase 2: disimballare le scatole

Dopo esserti assicurato che il sito e l'armadio o il rack che intendi utilizzare per i nodi di elaborazione dati soddisfano le specifiche richieste, disimballa tutte le scatole e confronta il contenuto con gli articoli indicati sulla bolla di accompagnamento.

Passaggi

1. Aprire con attenzione tutte le scatole e disporre i contenuti in modo ordinato.
2. Confronta gli articoli non imballati con la bolla di accompagnamento e annota eventuali discrepanze.

Puoi ottenere la lista dei colli scansionando il codice QR sul lato della scatola di spedizione.

I seguenti elementi sono alcuni dei contenuti che potresti vedere nelle scatole.

Hardware	Cavi	
• Lunetta • Kit rail con istruzioni • Nodo/i di elaborazione dati	• cavi di alimentazione	

Passaggio 3: registra i tuoi nodi di elaborazione dati

Dopo che ti sei assicurato che il tuo sito soddisfi i requisiti per le specifiche del nodo di elaborazione dati e hai verificato di avere tutte le parti che hai ordinato, dovresti registrare il tuo sistema.

Passaggi

1. Individua i numeri di serie dei tuoi nodi di elaborazione dati.

Puoi trovare i numeri di serie nelle seguenti posizioni:

- Sulla bolla di accompagnamento
- Nella tua email di conferma
- Su ciascun nodo di elaborazione dati o, in alcuni sistemi, sul modulo di gestione del sistema di ciascun nodo di elaborazione dati.

SSN: XXXXXXXXXXXXX



2. Naviga verso il ["Sito di supporto NetApp"](#).
3. Determina se è necessario registrare il tuo sistema storage:

Se sei un...	Segui questi passaggi...
Cliente NetApp esistente	a. Sign in con il tuo username e password. b. Seleziona Systems > My Systems. c. Confermare che il nuovo numero di serie sia elencato. d. Se il numero di serie non è elencato, seguire le istruzioni per i nuovi clienti NetApp.
Nuovo NetApp customer	a. Fai clic su Register Now per creare un account. b. Selezionare Systems > Register Systems. c. Inserire il numero di serie del sistema storage e i dettagli richiesti. Una volta che NetApp approva la tua registrazione, puoi scaricare il software necessario. L'approvazione richiede fino a 24 ore.

E ora?

Dopo aver preparato l'installazione dei nodi di elaborazione dati, ["installare i Data Compute Node"](#).

Installa i tuoi nodi di calcolo dati per AI Data Engine

Installa e proteggi i nodi di elaborazione dati nell'armadio.

Prima di iniziare

- Assicurati di avere le istruzioni incluse nel rail kit.
- Comprendere i problemi di sicurezza correlati al peso del nodo di elaborazione dati, del sistema storage e dello storage shelf.
- Comprendere che il flusso d'aria attraverso il sistema storage entra dalla parte anteriore, dove sono installati la cornice o i cappucci terminali, e fuoriesce dalla parte posteriore, dove si trovano le porte.



In generale, gli switch dovrebbero essere installati al centro dell'armadio. I ripiani storage dovrebbero essere installati sotto lo switch e sopra un secondo switch installato. I nodi controller possono essere installati sopra o sotto gli switch all'interno dell'armadio. I nodi di calcolo dati possono essere installati sopra o sotto i nodi controller all'interno dell'armadio.

Passaggi

- 1. Installa i kit di binari per i data compute nodes, secondo necessità, utilizzando le istruzioni incluse nei kit.
- 2. Installa e proteggi i nodi di elaborazione dati nell'armadio:
 - a. Posizionare il nodo di elaborazione dati sulle guide al centro dell'armadio, quindi sostenere l'appliance dal basso e farlo scorrere in posizione.
 - b. Fissa il nodo di elaborazione dati all'armadio utilizzando le viti di montaggio incluse.
- 3. Fissare la cornice alla parte anteriore del data compute node.

E ora?

Dopo aver installato i nodi di elaborazione dati, ["cablare i nodi di elaborazione dati per AIDE"](#).

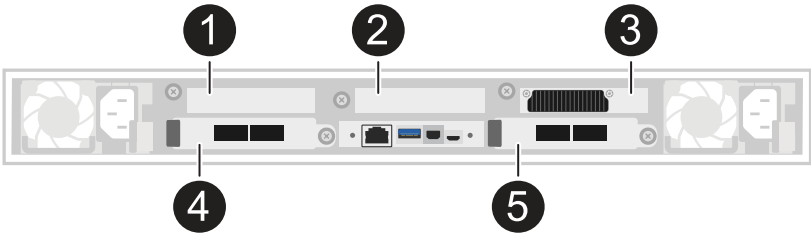
Requisiti per il cablaggio dei Data Compute Node per AI Data Engine

I Data Compute Node si integrano con il tuo sistema storage AFX 1K tramite connessioni di rete host e cluster. Esamina la configurazione degli slot I/O, i tipi di cavi e i requisiti di connessione per la tua implementazione.

Configurazione del cablaggio

I nodi Data Compute si collegano agli stessi cluster switch dei nodi controller AFX 1K, estendendo il sistema storage con risorse di calcolo ottimizzate per carichi di lavoro AI e machine learning.

La configurazione iniziale di AI Data Engine (AIDE) supporta un minimo di tre Data Compute Node. Per dettagli completi sulla configurazione e sulle priorità degli slot, vedere ["NetApp Hardware Universe"](#).



1	Slot inutilizzato sul Data Compute Node.
2	Slot inutilizzato sul Data Compute Node.
3	GPU slot sul Data Compute Node.
4	slot sul Data Compute Node.

Configurazione dello slot I/O

Il Data Compute Node utilizza uno schema di numerazione degli slot specifico che differisce dalle configurazioni server standard. Comprendere la disposizione degli slot è essenziale per un cablaggio corretto.

- **Slot 3:** riservato per GPU (non accessibile per il cablaggio I/O)
- **Slot 4 e 5:** slot I/O utilizzati per le connessioni di rete
 - Porta a: connessioni di rete del cluster
 - Porta b: connessioni di rete host
- **Slot 1 e 2:** non popolati e inaccessibili per l'uso

Connessioni di rete

I nodi di elaborazione dati richiedono due tipi di connessioni di rete per integrarsi con il sistema storage AFX 1K.

• Connessioni di rete host

Le connessioni di rete host forniscono l'accesso ai dati client e consentono ai Data Compute Node di elaborare i carichi di lavoro. Ogni Data Compute Node utilizza le porte e4b ed e5b per connessioni ridondanti a switch di rete host separati.

Assegnazioni delle porte:

- e4b: si connette allo switch di rete host A
- e5b: si connette allo switch di rete B

• Connessioni di rete cluster

Le connessioni di rete del cluster consentono la comunicazione tra i Data Compute Node e i nodi controller AFX 1K all'interno del cluster storage. Ogni Data Compute Node utilizza le porte e4a ed e5a per connessioni ridondanti a switch di rete separati del cluster.

Assegnazioni delle porte:

- e4a: si collega allo switch di rete del cluster A
- e5a: si collega allo switch di rete del cluster B

Componenti hardware supportati

I Data Compute Node richiedono cavi e switch specifici per garantire la corretta connettività e le prestazioni con il sistema storage AFX 1K.

Data Compute Node	Switch supportati	Cavi supportati
Data compute nodes (ne sono richiesti almeno tre)	• Cisco Nexus 9332D-GX2B (400GbE) • Cisco Nexus 9364D-GX2A (400GbE)	• Cavi breakout QSFP-DD 400GbE a 4x100GbE QSFP56 per connessioni ai Data Compute Node: ◦ Porte di rete da 100GbE verso il cluster Data Compute Node (e4a, e5a) ◦ 100GbE alle porte di rete host del Data Compute Node (e4b, e5b) • Cavi RJ-45 per connessioni di gestione

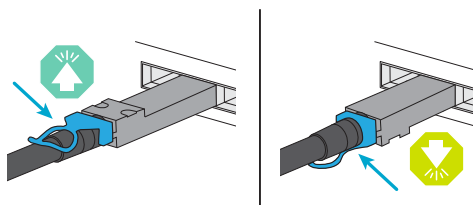


I cavi breakout forniscono quattro connessioni 100GbE da ciascuna porta dello switch 400GbE. Collegare l'estremità 400GbE agli switch e l'estremità 100GbE alle porte I/O del Data Compute Node.

Orientamento del cavo

Quando si collegano i cavi ai Data Compute Node, il corretto orientamento dei cavi garantisce connessioni affidabili.

La grafica di cablaggio nelle procedure di installazione mostra delle icone a forma di freccia che indicano il corretto orientamento (verso l'alto o verso il basso) della linguetta di estrazione del connettore del cavo quando si inserisce un connettore in una porta. Quando si inserisce il connettore, si dovrebbe sentire uno scatto che lo fissa in posizione. Se non si sente lo scatto, rimuoverlo, capovolgerlo e riprovare.



Maneggiare con cura i delicati componenti del connettore quando li si agganciano in posizione.

E ora?

Dopo aver esaminato la configurazione del cablaggio, ["cablare l'hardware"](#) per i Data Compute Node.

Collega i tuoi Data Compute Node per AI Data Engine

Collega i tuoi Data Compute Node alla rete host e agli switch di rete del cluster per abilitare l'elaborazione dei carichi di lavoro AI e l'integrazione con il tuo sistema storage AFX 1K. Questa procedura utilizza connessioni 100GbE sia per l'accesso alla rete host che per la comunicazione del cluster, consentendo ai nodi di sfruttare l'infrastruttura del cluster esistente senza spegnere il sistema AFX.

Informazioni su questa attività

Queste procedure mostrano configurazioni comuni. Il cablaggio specifico dipende dai componenti ordinati per il tuo sistema storage. Per dettagli completi sulla configurazione e sulle priorità degli slot, vedere ["NetApp Hardware Universe"](#).



Non è necessario spegnere il sistema storage AFX 1K durante il cablaggio dei data compute node. È possibile aggiungere i data compute node a un sistema storage AFX 1K già acceso e configurato.

Prima di iniziare

- Hai già installato un sistema storage AFX 1K. Per informazioni sull'installazione del sistema storage AFX 1K, consulta ["Documentazione di installazione del sistema storage AFX 1K"](#).
- Gli switch di rete necessari sono installati e configurati. Contattare l'amministratore di rete per informazioni sulla connessione del sistema agli switch di rete.
- Hai esaminato il ["requisiti di cablaggio per i Data Compute Node"](#).



È necessario un minimo di tre nodi di calcolo dati per implementare l'AI Data Engine.

Passaggio 1: collegare i Data Compute Node alla rete host

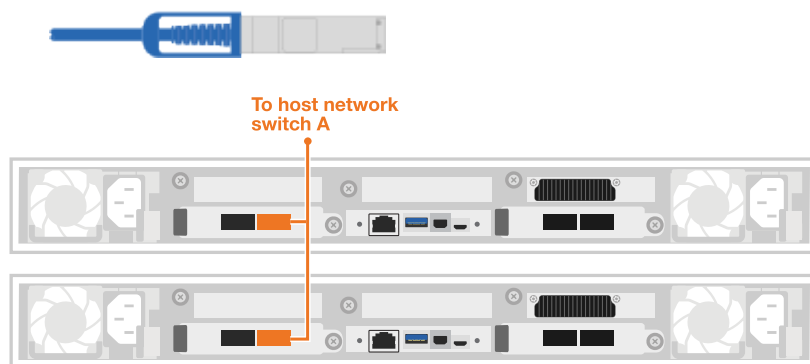
È possibile collegare le porte del Data Compute Node alla rete host.

Passaggi

1. Collegare la porta e4b dai seguenti Data Compute Node allo switch di rete Ethernet A:

- Data Compute Node 1, porta e4b
- Data Compute Node 2, porta e4b

Cavi 100GbE

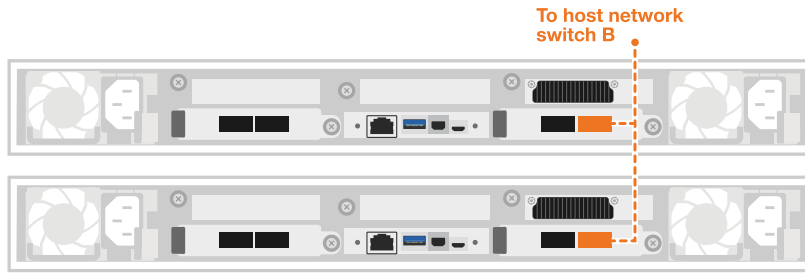


2. Collegare la porta e5b dai seguenti Data Compute Node allo switch di rete Ethernet B:

- Data Compute Node 1, porta e5b
- Data Compute Node 2, porta e5b

Cavi 100GbE





Passaggio 2: collegare le connessioni del cluster del Data Compute Node

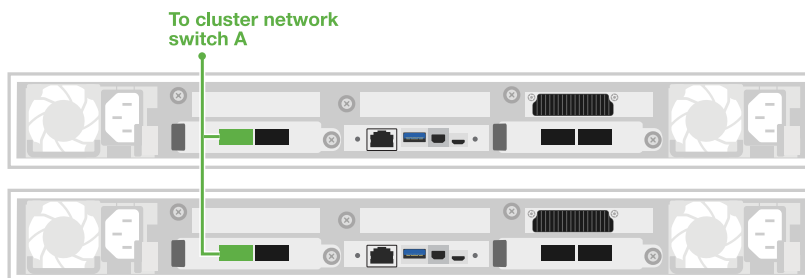
Per i nodi di elaborazione dati, utilizzare cavi breakout 4x100GbE per collegare le porte e4a/e5a per le connessioni del cluster.

Passaggi

1. Collegare la porta e4a dai seguenti Data Compute Node a una porta non-ISL sullo switch di rete del cluster A:

- Data Compute Node 1, porta e4a
- Data Compute Node 2, porta e4a

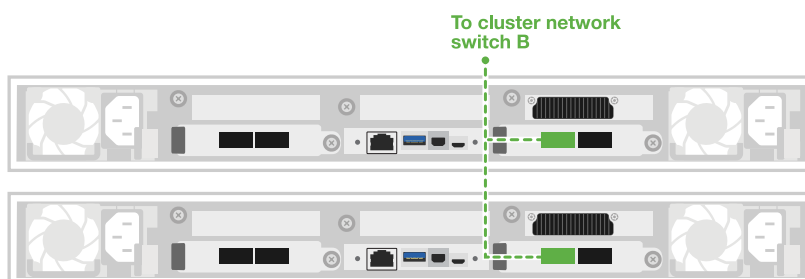
Cavi breakout 4x100GbE



2. Collegare la porta e5a dai seguenti Data Compute Node a una porta non ISL sullo switch di rete del cluster B:

- Data Compute Node 1, porta e5a
- Data Compute Node 2, porta e5a

Cavi breakout 4x100GbE



E ora?

Dopo aver cablato l'hardware, ["accendi i tuoi Data Compute Node"](#).

Accendi i nodi di elaborazione dati per AI Data Engine

Dopo aver installato l'hardware del rack e cablato i nodi di elaborazione dati, è necessario accendere i DCN e i nodi controller per il sistema storage AFX, se non sono già accesi.

Prima di iniziare

- Assicuratevi che i vostri scaffali siano accesi e che a ciascuno sia assegnato un ID scaffale univoco. Per informazioni sull'assegnazione degli ID scaffale per il sistema storage AFX, consultate il ["documentazione sull'assegnazione di ID shelf univoci"](#).

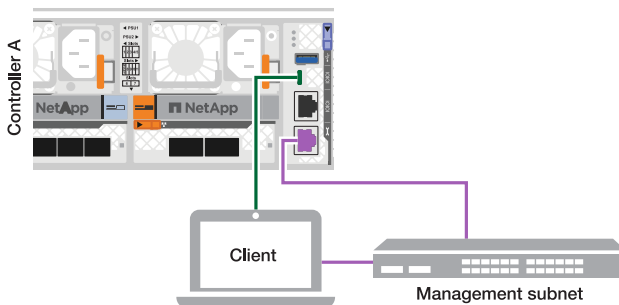
Passaggi

Dopo aver acceso gli scaffali di archiviazione e assegnato gli ID univoci, accendi i DCN e accendi i nodi dello storage controller se non sono già accesi.

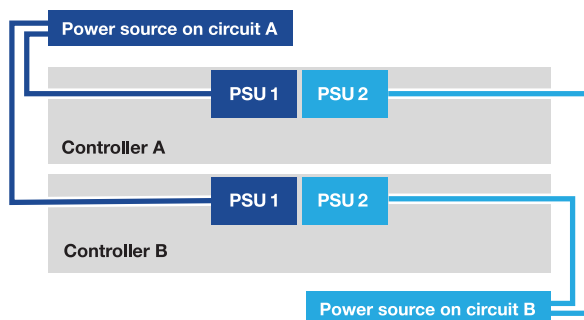
- Collega il tuo laptop alla porta seriale della console. Questo ti permette di monitorare la sequenza di avvio quando i controller sono accesi.
 - Impostare la porta della console seriale sul laptop a 115,200 baud con N-8-1.

Consultare la guida in linea del laptop per istruzioni su come configurare la porta della console seriale.

- Collega il cavo della console al laptop e collega la porta seriale della console sul controller utilizzando il cavo della console fornito con il tuo sistema storage.
- Collegare il laptop allo switch sulla subnet di gestione.



- Assegna un indirizzo TCP/IP al laptop, utilizzandone uno che si trova sulla subnet di gestione.
- Collega i cavi di alimentazione agli alimentatori del controller, quindi collegali alle fonti di alimentazione su circuiti diversi.



- Il sistema inizia ad avviarsi. L'avvio iniziale potrebbe richiedere fino a otto minuti.
 - I LED lampeggiano e le ventole si avviano, il che indica che i controller si stanno accendendo.
 - Le ventole potrebbero essere rumorose all'avvio, il che è normale.
4. Collega i cavi di alimentazione agli alimentatori del nodo di elaborazione dati, quindi collegali alle fonti di alimentazione su circuiti diversi.
 5. Fissa i cavi di alimentazione utilizzando il dispositivo di fissaggio su ciascun alimentatore.
 6. Accendere i nodi di data compute.

Potrebbe essere necessario rimuovere la cornice per accedere all'interruttore di alimentazione; in tal caso, ricordarsi di reinstallarla in seguito.

E ora?

Dopo aver attivato i nodi di data compute, ["configurare un ONTAP AIDE cluster"](#).

Configura il tuo sistema AIDE

Imposta AI Data Engine

In qualità di amministratore dello storage ONTAP, puoi integrare AI Data Engine (AIDE) con la distribuzione del tuo sistema AFX. Puoi configurare AIDE come parte di una distribuzione iniziale del sistema storage AFX o aggiungere AIDE a un cluster AFX già distribuito e che sta già servendo dati. In entrambi i casi, dovrai anche completare la configurazione di AIDE dopo aver terminato la configurazione di base.

1. Configurare AIDE e integrarlo con un sistema storage AFX

Esistono due modi per configurare AIDE e integrarlo con un sistema storage AFX. Seleziona l'opzione più adatta al tuo ambiente.



Prima di configurare AIDE, assicurati che i Data Compute Node siano installati e collegati agli switch del cluster. Vedi ["Installa AIDE"](#) per ulteriori informazioni.

Configurare AIDE con un nuovo cluster AFX

Se stai configurando AIDE come parte di una distribuzione iniziale del sistema AFX, il processo di configurazione di AIDE fa parte del flusso di lavoro standard di configurazione del cluster AFX. Durante la configurazione del cluster ONTAP, System Manager rileva automaticamente i DCN connessi e li include come parte della configurazione del cluster. Dopo di ciò, sei pronto per completare la configurazione di AIDE, inclusa la configurazione della licenza richiesta.

Configurare AIDE con un cluster AFX esistente

Se si aggiunge un cluster di Data Compute Node a un sistema storage AFX esistente, è necessario configurare i DCN e completare l'impostazione del cluster AIDE.

Aggiungi i Data Compute Node al cluster

ONTAP rileva dinamicamente i nuovi Data Compute Node connessi alla rete del cluster e li visualizza in System Manager. È possibile ["aggiungi nodi di elaborazione dati al tuo cluster AIDE"](#). Si noti inoltre quanto

segue:

- I Data Compute Node devono essere fisicamente connessi agli switch del cluster. Sono richiesti esattamente tre DCN quando si crea un AIDE cluster.
- È necessario disporre di indirizzi IP sufficienti per la rete DCN e delle altre informazioni di configurazione di rete richieste.

Completa la configurazione del tuo cluster AIDE

Dopo aver integrato i Data Compute Node e completato la configurazione iniziale della rete, puoi terminare la configurazione del tuo nuovo AIDE cluster.

Informazioni su questa attività

Sul lato destro della pagina viene visualizzata una finestra di benvenuto con le istruzioni di configurazione necessarie. Un segno di spunta viene visualizzato a sinistra di ogni passaggio o elemento di azione già completato. Una freccia (→) viene visualizzata per gli elementi non ancora completati. Per ciascuna delle azioni incomplete:

- Seleziona il collegamento al titolo abilitato e completa l'azione di configurazione per il tuo ambiente.
- Se il collegamento è disabilitato, passa il mouse sul titolo per visualizzare l'azione richiesta per abilitarlo e continuare con l'azione di setup.

Ogni passaggio è corredato, ove possibile, da un collegamento a documentazione aggiuntiva. Alcune delle configurazioni potrebbero essere incluse come parte della configurazione del cluster AFX. Vedere il ["Documentazione AFX"](#) per ulteriori informazioni.

Passaggi

1. Selezionare **Configura link aggregation groups e VLAN**.

Se necessario per il tuo ambiente, puoi configurare LAG e VLAN a livello di cluster. Selezionando questa opzione si apre la pagina delle porte Ethernet.

2. Seleziona **Configura protocolli di rete**.

È possibile configurare i protocolli di rete della storage VM (SVM).

3. Seleziona **Aggiorna software di elaborazione dati**.

Se un'immagine del software data compute è disponibile per l'aggiornamento, questa opzione sarà abilitata. Selezionarla per avviare la pagina di aggiornamento del software data compute. Per ulteriori informazioni, consulta ["Aggiorna AI Data Engine software"](#).

4. Selezionare **Configure data engine networking**.

Questa opzione visualizza la pagina per configurare l'indirizzo di rete del data engine.

5. Seleziona **Crea interfacce di rete intercluster**

Questa opzione avvia la pagina per configurare le interfacce di rete intercluster come parte di una procedura in due fasi se si intende catalogare dati remoti.

6. Selezionare **Peer con altri sistemi storage ONTAP**.

Dopo aver creato le interfacce di rete intercluster, è possibile stabilire relazioni di peering con altri sistemi

storage ONTAP.

7. Seleziona **Add a data container**.

È possibile aggiungere un volume da utilizzare da AIDE e associarlo a una specifica SVM.

8. Seleziona **Add Workspaces**.

Verrà avviata la pagina degli spazi di lavoro di Data Engine, in cui è possibile ["creare uno spazio di lavoro"](#).

9. Seleziona **Configura OpenID Connect**.

Sarà necessario ["configurare l'autenticazione OpenID"](#) per abilitare l'accesso alla AI Data Engine Console.

Dopo aver finito

Dopo che tutte le action item sono state completate, la finestra di guida viene nascosta. Puoi chiuderla e riapirla manualmente secondo necessità.

2. Completare la configurazione di AIDE

Dopo che il cluster AIDE è stato configurato e integrato con il sistema storage AFX, è necessario completare la configurazione di AIDE.

Passaggi

1. ["Aggiungi le licenze AIDE richieste"](#).

È necessario installare una licenza AI Data Engine per usufruire di tutte le funzionalità di AIDE, comprese le capacità di vettorizzazione e guardrails.

2. ["Configura l'autenticazione OIDC"](#).

Assicurati di configurare OIDC per tutte le distribuzioni. Devi configurare OIDC per accedere all'AI Data Engine Console.

Installa una licenza AIDE nel tuo sistema AFX

È necessario installare una licenza NetApp per accedere a tutte le funzionalità di AI Data Engine (AIDE). Come amministratore del cluster ONTAP, è possibile gestire le licenze utilizzando ONTAP System Manager.

Preparati a ottenere la licenza per l'AI Data Engine

Una licenza è un record di uno o più diritti software. Tutte le licenze AFX vengono fornite come file di licenza NetApp (NLF), ovvero un singolo file che abilita più funzionalità. Ci sono diversi aspetti da considerare prima di installare una licenza sul sistema storage AFX a supporto di AIDE.

Tipi di licenze

Sono necessari due tipi principali di licenze per iniziare a utilizzare AIDE tramite il tuo sistema storage AFX.

Licenza ONTAP One

La licenza Metadata Engine Basic è in genere installata in fabbrica come parte della licenza ONTAP One. Consente l'accesso alla funzionalità Metadata Engine che fornisce la base essenziale per le operazioni AIDE.

Sono incluse anche tutte le funzionalità principali necessarie per amministrare il sistema ONTAP.

AI Data Engine

È necessario acquistare e installare la licenza AI Data Engine per accedere ai servizi premium necessari per attivare tutte le funzionalità di AIDE. La licenza sblocca i nodi di elaborazione dati, abilitando funzionalità di AI come la vettorizzazione, i guardrail di governance, l'inferenza e un'esperienza UI integrata. La licenza include il conteggio delle GPU e una data di scadenza.

Requisiti di installazione della licenza

Dovrai acquistare la licenza AIDE e scaricare il file NLF associato sul tuo sistema locale. Potrai quindi caricare il file sul tuo sistema storage AFX tramite System Manager. Assicurati inoltre di disporre di quanto segue:

- Credenziali di amministratore per accedere a ONTAP System Manager
- Un cluster AFX che esegue ONTAP 9.18.1 e versioni successive

Installa una licenza sul tuo sistema AFX

È possibile installare una licenza AIDE per attivare le funzionalità AIDE aggiuntive necessarie per il tuo sistema storage AFX.

Passaggi

1. In System Manager, seleziona **Cluster** e poi **Settings**.
2. Accanto a **Licenses**, seleziona ➔.
3. Selezionare la scheda **Funzionalità** per visualizzare le funzionalità ONTAP disponibili.
4. Per installare una licenza, selezionare la scheda **Installed licenses**.
5. Seleziona **+ Add**.
6. Seleziona un file di licenza locale e seleziona **Aggiungi**.

Informazioni correlate

- ["Panoramica delle licenze ONTAP"](#)
- ["Come scaricare le licenze NLF dal NetApp Support Site"](#)
- ["CLI ONTAP: system license add"](#)

Configura OpenID Connect per AIDE in ONTAP

In qualità di amministratore del cluster ONTAP, puoi utilizzare ONTAP System Manager per configurare l'autenticazione OpenID Connect (OIDC) per un cluster AI Data Engine (AIDE). Ciò fornisce un accesso sicuro e centralizzato tramite un provider di identità (IdP) esterno.



È necessario configurare OIDC per accedere all'AI Data Engine Console. Quando è configurato, tutta l'autenticazione avviene tramite OIDC. Se OIDC non è configurato, la console non sarà disponibile sia agli amministratori che ai data engineer e ai data scientist. In questo caso, l'accesso a System Manager torna all'autenticazione locale.

Si noti inoltre quanto segue sulla configurazione OIDC per l'accesso AIDE:

- Non è possibile modificare una configurazione OIDC esistente. Se è necessario apportare una modifica, eliminare prima la configurazione e crearne una nuova con le impostazioni desiderate.
- Se si disabilita o si rimuove OIDC, System Manager tornerà all'autenticazione utente locale ONTAP.

Panoramica di OIDC

OpenID Connect (OIDC) è un protocollo di autenticazione basato sul framework OAuth 2.0. Estende OAuth 2.0, utilizzato principalmente per l'autorizzazione, aggiungendo un livello di identità. OIDC introduce il concetto di ID token, che è un JSON Web Token (JWT) contenente claim sull'evento di autenticazione e sull'identità dell'utente.

È necessario selezionare e configurare un provider di identità esterno (IdP) supportato da AFX con AIDE. L'IdP autentica gli utenti e rilascia token che AFX, tramite System Manager, può utilizzare per concedere l'accesso alla AI Data Engine Console.

Configura i provider di identità di terze parti

Per autenticarsi tramite OIDC, è necessario innanzitutto configurare un IdP esterno. L'implementazione ONTAP di OIDC utilizza le attestazioni di ruolo nei token per applicare il RBAC. Quando si configura un IdP, assicurarsi che sia configurato per restituire le attestazioni di ruolo nel token id e nel token di accesso. ONTAP supporta due IdPs per l'autenticazione OIDC: Entra ID e Active Directory Federation Services (AD FS).

Entra ID

È possibile configurare Entra ID utilizzando i seguenti passaggi high-level:

1. Crea una nuova App Registration nella pagina di configurazione di Entra ID.
2. Impostare il valore Redirect URI (Web) su `https://$CLUSTER_MGMT_IP/oidc/callback`, sostituendo l'indirizzo IP di gestione del cluster appropriato o il FQDN.
3. Crea i ruoli richiesti in App Roles e assegnali ai tuoi utenti.
4. Aggiorna le richieste del token in Token Configuration per restituire i ruoli in id-token e access-token.

Vedi ["Configura un provider OpenID Connect con Entra ID"](#) per maggiori informazioni.

Active Directory Federation Services

È possibile configurare AD FS utilizzando i seguenti high-level passaggi:

1. Crea un nuovo Application Group e seleziona **Server application accessing a web API**.
2. Impostare il valore Redirect URI (Web) su `https://$CLUSTER_MGMT_IP/oidc/callback`, sostituendo l'indirizzo IP di gestione del cluster appropriato o il FQDN.
3. Configura le attestazioni per restituire i ruoli nei token.

Vedi ["Aggiungi AD FS come provider di identità OpenID Connect"](#) per maggiori informazioni.

Configura OIDC in System Manager

Dopo aver configurato l'IdP, è possibile impostare l'autenticazione OIDC in System Manager per abilitare l'accesso sicuro alla AI Data Engine Console.

Prima di iniziare

- È necessario disporre dell'accesso come amministratore a System Manager.
- Il tuo provider di identità OIDC deve essere configurato e accessibile.

Passaggi

1. In System Manager, seleziona **Cluster** e poi **Settings**; individua la scheda OpenID Connect.
2. Se OIDC è già configurato, è possibile modificare o disattivare la configurazione. Se OIDC non è configurato selezionare  per avviare il processo di configurazione.
3. In Configura OpenID Connect, fornisci valori per i seguenti campi:
 - Fornitore
 - Emittente
 - URI del set di chiavi web JSON
 - Endpoint di autorizzazione
 - Endpoint del token
 - Endpoint di fine sessione
 - Emittente del token di accesso (opzionale)
4. In Configurazione client, fornire valori per i seguenti campi:
 - ID client
 - Richiesta utente remoto
 - Intervallo di aggiornamento
5. In Dettagli connessione, fornire valori per i seguenti campi:
 - Indirizzo IP del cluster o FQDN
 - proxy in uscita (facoltativo)
6. In Mapping ruoli esterni, selezionare un mapping ruoli esistente o definire un nuovo ruolo per l'utente `admin`ONTAP.
7. Selezionare **Abilita ora** e poi **Salva**. System Manager verrà aggiornato per applicare le nuove impostazioni di autenticazione.
8. Accedi con le tue credenziali IdP; dopo l'autenticazione riuscita verrai restituito a System Manager.

Informazioni correlate

- ["OpenID Foundation"](#)
- ["Implementazione ONTAP OAuth 2.0"](#)

Configura gli spazi di lavoro

Preparare i dati per AI Data Engine

Dopo aver creato un cluster, crea un data container che contenga i dati che intendi utilizzare con AI Data Engine (AIDE). Questo data container deve essere un volume ONTAP, locale o di un cluster ONTAP peered che esegue ONTAP 9.

Anziché caricare manualmente i dati del cluster ONTAP su AIDE, sarà necessario eseguire il peering dei cluster e delle SVM di destinazione con il cluster AIDE e quindi decidere quali volumi NFS utilizzare con il catalogo dei metadati AIDE. Dopo aver creato il data container, è possibile creare un workspace e associare il data container a quel workspace. Gli utenti di quel workspace potranno quindi accedere e interagire con le raccolte di dati e le risorse affiliate al workspace per i propri carichi di lavoro AI.

Informazioni su questa attività

È necessario eseguire il peering di ogni SVM contenente i dati che si desidera utilizzare con AI Data Engine. Il peering del solo cluster non è sufficiente. Questo garantisce che AI Data Engine possa accedere, integrare e indicizzare i dati come previsto.

Dovresti collegare l'SVM nel cluster AIDE che fungerà da destinazione di SnapMirror.

Non è necessario creare relazioni SnapMirror tra i cluster ONTAP e il cluster AIDE. Queste relazioni verranno create automaticamente durante la creazione dell'area di lavoro.

Prima di iniziare

- Sono necessari privilegi di *storage administrator* per eseguire il peering di cluster e SVM e selezionare data container.
- Hai identificato i cluster ONTAP e gli SVM che contengono i dati che desideri utilizzare con AI Data Engine.
- Hai confermato che i volumi delle origini dati hanno soddisfatto i seguenti requisiti:
 - I volumi sono online e accessibili.
 - Il protocollo NFS è abilitato. Solo i volumi abilitati per NFS sono supportati come container di dati per AI Data Engine. I volumi SMB e CIFS non sono supportati.
 - I volumi non sono volumi FlexCache.
 - Le origini dati sono volumi di lettura-scrittura. I volumi di protezione dati non sono supportati.

Passaggi

1. ["Esegui il peering di ogni cluster ONTAP e SVM"](#) che contiene i dati che vuoi utilizzare con AI Data Engine.
2. ["Seleziona i volumi che desideri utilizzare con AI Data Engine"](#).

Per ogni volume, annotare le seguenti informazioni:

- Nome volume
- UUID
- Nome SVM e UUID
- Nome del cluster e UUID

E ora?

"[Crea uno spazio di lavoro](#)" e associare i data container che hai creato a un workspace.

Informazioni correlate

- "[Opzioni di migrazione dei dati](#)"

Crea uno spazio di lavoro in AI Data Engine

Dopo aver configurato un cluster, puoi creare un workspace. I workspace consentono di segmentare i dati sul cluster, controllare l'accesso ai dati per i singoli utenti ed escludere i dati a cui AI Data Engine (AIDE) non deve accedere.

Se gestisci l'archiviazione, utilizzerai ONTAP System Manager per creare e gestire workspaces.

Le organizzazioni creano spazi di lavoro in base a team, progetti, livelli di sensibilità dei dati o altri criteri rilevanti. Ad esempio, se lavori nel settore sanitario, potresti segmentare i dati clinici in uno spazio di lavoro ma escludere i dati relativi a IT, legale o altri reparti.

Informazioni su questa attività

I limiti di elaborazione del sistema influiscono sulla creazione di workspace (in genere fino a 15 GB al giorno per cluster). Se crei più workspace in parallelo o in rapida successione, ciascun workspace potrebbe richiedere più tempo per essere elaborato e potresti riscontrare ritardi significativi.

Monitora lo stato di creazione dell'area di lavoro dalla pagina [Inventario delle aree di lavoro](#). Per risultati ottimali, evita di creare molte aree di lavoro contemporaneamente se hai bisogno di accedere immediatamente a queste funzionalità.

Prima di iniziare

- Sono necessari i privilegi di *storage administrator* per creare spazi di lavoro e associare raccolte di dati.
- Hai determinato le origini dati remote (peered) e locali che intendi utilizzare con l'area di lavoro e con AI Data Engine.
- Hai "[creato almeno un container di dati](#)" a disposizione che l'area di lavoro può utilizzare, come un volume locale o un volume da un cluster peered.



Aggiungi un volume a un'area di lavoro che non eliminerai durante il periodo di validità previsto di tale area di lavoro. Se elimini un volume dopo averlo aggiunto a un'area di lavoro, l'area di lavoro entrerà in stato di errore. Conferma la fattibilità a lungo termine del volume prima di creare un'area di lavoro.

- Assicurarsi che NFS sia abilitato sul volume ma che CIFS non sia abilitato. Le aree di lavoro supportano solo volumi con NFS. I volumi con CIFS (SMB) non sono supportati.

Crea uno spazio di lavoro

Crea un workspace e associa i container di dati che contengono i dati che desideri utilizzare con AI Data Engine.

Passaggi

1. In ONTAP System Manager, vai su **Data Engine > Workspaces**.
2. Seleziona **Aggiungi**.

3. Nella finestra di dialogo **Aggiungi area di lavoro**, seleziona almeno un container di dati disponibile da associare all'area di lavoro.
4. Configura **"cluster peered"** in modo che i dati di tali cluster possano essere accessibili all'interno dell'area di lavoro
5. Se desideri configurare l'accesso utente all'area di lavoro, puoi farlo ora o **"attendere fino a quando l'area di lavoro è stata creata"**.
6. Configura un intervallo di refresh per la frequenza con cui l'area di lavoro si sincronizza con i container di dati associati per acquisire dati nuovi o aggiornati (ad esempio, sei ore).



Scegli un intervallo che bilanci la freschezza dei dati con le prestazioni del sistema. Se aggiungi un container di dati a più aree di lavoro, il sistema utilizza automaticamente l'intervallo più aggressivo (il più breve). Per ulteriori informazioni, consulta la documentazione su [aggiornamenti e versioning dell'area di lavoro](#).

7. Seleziona **Continua**.
8. Nella finestra di dialogo **Finalize workspace**, inserire un nome e una descrizione per l'area di lavoro.
9. Selezionare **Aggiungi** per creare l'area di lavoro.

Risultato

Il processo di creazione dell'area di lavoro richiede da diversi minuti a diverse ore per essere completato, a seconda del set di dati associato, del numero di file, delle dimensioni dei file e di altri fattori.

Il sistema estrae automaticamente i metadati da tutte le fonti di dati e li memorizza in un catalogo di metadati che gli utenti possono utilizzare per individuare i file necessari per i loro progetti. Dopo aver assegnato gli utenti all'area di lavoro, gli utenti data engineer possono configurare e interagire con i componenti affiliati all'area di lavoro dalla AI Data Engine Console.

Il nuovo spazio di lavoro viene visualizzato nella pagina Spazi di lavoro in `Creating` stato finché il processo non viene completato e lo stato cambia in `ready`.

Rivedi i dettagli dell'area di lavoro

Dopo la creazione dell'area di lavoro, rivedi i dettagli dell'area di lavoro.

Passaggi

1. Esamina i dettagli dell'area di lavoro, incluse le dimensioni totali, la percentuale della capacità del cluster utilizzata e la data dell'aggiornamento più recente dell'area di lavoro.
2. Selezionare il nome dell'area di lavoro per aprire la pagina dei dettagli.
3. Nella scheda Panoramica, visualizza i dettagli dell'area di lavoro che includono contenitori di dati associati, utenti e attività.

Aggiornamenti e versioning dell'area di lavoro

Ogni aggiornamento dell'area di lavoro crea una versione immutabile che cattura lo stato corrente di tutti i file e gli oggetti nell'area di lavoro. Le versioni includono metadati completi, riferimenti agli snapshot utilizzati durante l'estrazione e un ID di processo per la tracciabilità. Ciò supporta la data lineage, la riproducibilità e l'audit.

Gli aggiornamenti vengono eseguiti in base alla pianificazione configurata (ad esempio ogni sei ore) oppure quando vengono attivati manualmente. L'intervallo di refresh minimo supportato è di un'ora; il massimo è di un anno. Se un container di dati è incluso in più aree di lavoro, il sistema utilizza l'intervallo di refresh più

frequente e di durata più breve per pianificare l'estrazione dei metadati.

Per impostazione predefinita, il sistema conserva le versioni precedenti, attuali e successive (in corso). Il sistema conserva le versioni meno recenti in base alla policy della tua organizzazione e può eliminarle se necessario.

È possibile elencare tutte le versioni di un'area di lavoro e visualizzare le differenze tra le versioni per identificare quali file o oggetti sono stati aggiunti, modificati o eliminati. Ciò consente di monitorare le modifiche nel tempo e comprendere l'evoluzione dei dati dell'area di lavoro.

Assegnare l'accesso utente agli spazi di lavoro di AI Data Engine

In qualità di amministratore dello storage, assegna gli utenti a un'area di lavoro in base ai loro ruoli: data engineer, data scientist o altri ruoli a seconda della struttura e delle esigenze della tua organizzazione. Gli utenti accedono alla AI Data Engine (AIDE) Console utilizzando le proprie credenziali e accedono alle risorse del container di dati all'interno dell'area di lavoro assegnata.

ONTAP System Manager consente di gestire quali utenti hanno accesso alle aree di lavoro di AI Data Engine. È possibile aggiungere o rimuovere utenti per controllare chi può visualizzare, modificare e interagire con i dati e le attività dell'area di lavoro.

Prima di iniziare

- Sono necessari i privilegi di *amministratore di storage* per gestire l'accesso degli utenti alle aree di lavoro.
- Conferma di aver creato il workspace e che sia attivo nell'inventario dei Workspaces.
- Confermare che tutti i container di dati pertinenti siano stati aggiunti all'area di lavoro e siano accessibili.
- ["Conferma che OIDC è abilitato e configurato per il cluster"](#). La mappatura dei ruoli da IdP a ONTAP deve essere completata per ogni utente o gruppo IdP di data engineer e data scientist pertinente.

Passaggi

- Aggiungi un utente a un'area di lavoro:
 - a. In ONTAP System Manager, vai su **Data Engine > Workspaces**.
 - b. Selezionare il nome dell'area di lavoro per aprire la pagina dei dettagli.
 - c. Vai alla scheda Utenti.
 - d. Seleziona il pulsante **Aggiungi** per aprire la finestra di dialogo per aggiungere utenti.
 - e. Inserisci i dettagli di uno o più utenti. Inserisci i dettagli come un elenco di utenti OIDC separati da virgole.
 - f. Seleziona **Aggiungi** per concedere all'utente l'accesso all'area di lavoro.
- Rimuovere un utente da un workspace:
 - a. Nella scheda Utenti della pagina dei dettagli del workspace, individua e seleziona l'utente che desideri rimuovere.
 - b. Seleziona il pulsante **Rimuovi**.
 - c. Conferma la rimozione nella finestra di dialogo.
 - d. Il sistema rimuove immediatamente l'utente e revoca il loro accesso all'area di lavoro.

Risultato

Solo gli utenti elencati nella scheda Utenti dell'area di lavoro possono accedere e interagire con i dati e le attività dell'area di lavoro.

Amministrare e monitorare

Monitorare i processi del cluster

Visualizza lo stato del sistema AIDE e del cluster

In qualità di amministratore dello storage, puoi utilizzare ONTAP System Manager per accedere alla dashboard e visualizzare lo stato del cluster. Questo è un buon primo passo prima di iniziare le attività amministrative di AIDE o se sospetti un problema operativo.

Prima di iniziare

- Per eseguire attività amministrative correlate ad AIDE ONTAP sono necessari i privilegi di *amministratore di storage*.

Monitora lo stato e la capacità di AIDE dalla dashboard

1. Connettersi a ONTAP System Manager utilizzando l'indirizzo di gestione del cluster:

```
https://$FQDN_OR_IP/
```

2. Sign in con un account amministratore.
3. Seleziona **Dashboard** nel riquadro di navigazione a sinistra.
4. Esamina il riquadro **Health**:
 - Confermare lo stato di salute generale del cluster.
 - Verificare il conteggio e lo stato dei **Data compute nodes**.
 - Controlla gli avvisi:
 - Problemi del nodo DCN o problemi di connettività
 - Aree di lavoro o raccolte di dati in errore (ad esempio, errori di pubblicazione della raccolta)
5. Esamina il riquadro **Capacity**:
 - Annotare la capacità totale del cluster e la capacità utilizzata.
 - Per i cluster AIDE, verificare:
 - Capacità utilizzata dai metadati AIDE e dai volumi delle applicazioni (metadata Storage VM)
 - Capacità utilizzata dagli spazi di lavoro e dalle raccolte di dati (se disponibili)
6. Facoltativamente, esaminare i riquadri **Rete** e **Prestazioni** per comprendere il comportamento a livello di cluster che potrebbe influire sui carichi di lavoro AIDE (ad esempio, congestione della rete o ritardo di protezione).

Visualizza i dati DCN sullo stato e l'utilizzo

1. Nel riquadro di navigazione, seleziona **Cluster** e poi **Overview**.
2. Seleziona la scheda **Data compute**.

Questa scheda mostra tutti i nodi DCN nel cluster con:

- Nome nodo, modello, seriale e versione software
 - Stato generale del nodo
 - Utilizzo della CPU e della memoria
 - Utilizzo della GPU (se sono presenti GPU)
 - Eventuali indicatori di errore a livello di nodo
3. Espandi un nodo DCN per aprire la vista dettagliata e controllare:
 - Utilizzo della CPU e della memoria di sistema
 - Utilizzo della memoria GPU
 - Problemi hardware o di servizio segnalati
 4. Selezionare **Cabling** nella pagina **Cluster > Overview** per verificare che i nodi DCN siano cablati correttamente agli switch del cluster e per identificare eventuali problemi di porta o di collegamento.

Monitora gli spazi di lavoro e l'impronta dei metadati

1. Nel riquadro di navigazione, seleziona **Data engine** e poi **Workspaces**.
2. Rivedi il riepilogo dell'area di lavoro nella parte superiore della pagina:
 - Conteggio degli spazi di lavoro e dei relativi stati (ad esempio, Processing, Healthy, Error).
 - Dimensione totale dell'area di lavoro.
 - Percentuale della capacità del cluster consumata da tutti gli spazi di lavoro.
3. Esaminare la griglia dell'area di lavoro:
 - Confermare che gli spazi di lavoro critici mostrino uno stato **Healthy**.
 - Verifica le dimensioni dello workspace e il consumo di capacità.
 - Cerca eventuali spazi di lavoro in **Error** o in stati di **Processing** esecuzione prolungata.
4. Per rivedere i dettagli di uno specifico workspace, seleziona il suo nome:
 - Nella scheda **Panoramica**, conferma:
 - Stato e dimensione dell'area di lavoro
 - Contenitori di dati (volumi) inclusi e conteggio dei relativi elementi
 - Ora dell'ultimo aggiornamento per ciascuna fonte dati
 - Nella scheda **Raccolta di dati**, confermare:
 - Quali raccolte di dati esistono per quell'workspace (le raccolte di dati sono di sola lettura in System Manager)
 - Il loro stato, dimensione e ora dell'ultimo aggiornamento
 - Nella scheda **Utenti**, controlla quali utenti di AI Data Engine Console hanno accesso.

Monitora metadati Storage VM e protezione gestita da AIDE

1. Nel riquadro di navigazione, seleziona **Cluster** e poi **Storage VMs**.
2. Individuare la Storage VM con sottotipo `data-engine` (la SVM dei metadati):
 - Confermare che la SVM dei metadati sia online.
 - Facoltativamente, apri i dettagli per vedere i conteggi per:

- Volumi
- LIFs con tipo `Data compute network` (utilizzati per la comunicazione DCN-ONTAP)

3. Seleziona **Protezione** e poi **Relazioni** per visualizzare la protezione per le origini dati remote utilizzate negli spazi di lavoro:

- Identificare le relazioni SnapMirror create da AIDE tramite il modello di denominazione:
 - Volume di destinazione: `<source_volume_name>_dest_<source_volume_UUID>`
 - Politica: `<source_volume_name>_dest_aide_policy_<source_volume_UUID>`
- Utilizzare questa vista per verificare che le relazioni siano sane e che il tempo di ritardo sia in linea con le aspettative di aggiornamento del workspace.



Non modificare la Storage VM dei metadati, le relazioni SnapMirror create da AIDE o gli snapshot gestiti da AIDE (o le relative pianificazioni) direttamente in ONTAP. Le modifiche possono compromettere la cronologia delle versioni di AIDE. ["Regola le impostazioni di aggiornamento dell'area di lavoro"](#) se è necessario modificare il comportamento di aggiornamento.

Esaminare gli avvisi e le notifiche relativi ad AIDE

1. Nel riquadro di navigazione, seleziona **Eventi & processi** e poi **Avvisi di sistema**.
2. Esaminare eventuali avvisi attivi relativi a:
 - Stato di salute o connettività del nodo DCN
 - Problemi di rete del data engine
 - Errori nell'area di lavoro o nella raccolta di dati
 - Mancata corrispondenza delle versioni software tra ONTAP e DCN cluster
3. Se necessario, configura le destinazioni di notifica (ad esempio, email, syslog) in **Cluster > Settings > Notification management** per garantire che gli avvisi relativi ad AIDE vengano inoltrati ai tuoi strumenti operativi.

Informazioni correlate

["Preparati ad amministrare il tuo sistema storage AFX"](#)

Visualizza approfondimenti per ottimizzare il tuo sistema AIDE

In qualità di amministratore dello storage, puoi utilizzare la funzionalità *Insights* di ONTAP System Manager per visualizzare gli aggiornamenti di configurazione suggeriti in linea con le best practice di NetApp. Queste modifiche possono ottimizzare la sicurezza e le prestazioni del tuo AIDE cluster.

Informazioni su questa attività

Ogni approfondimento viene presentato come un riquadro o una scheda separata sulla pagina che puoi scegliere di implementare o ignorare. Puoi anche selezionare il link alla documentazione associata per saperne di più su una tecnologia specifica.

Passaggi

1. In System Manager, seleziona **Analysis** e poi **Insights**.

2. Esaminare le raccomandazioni disponibili.

Cosa c'è dopo

Eseguire una delle azioni consigliate per implementare le best practice di configurazione.

Visualizza eventi di sistema AIDE, lavori e registro di controllo

In qualità di amministratore dello storage, puoi esaminare gli eventi, i processi e i messaggi del registro di controllo generati da AIDE per monitorare l'elaborazione interna e diagnosticare potenziali problemi. Il sistema AIDE può essere configurato per inoltrare queste informazioni, insieme ad altri dati correlati, per ulteriore elaborazione e archiviazione.

Prima di iniziare

- Per eseguire attività amministrative correlate ad AIDE ONTAP sono necessari i privilegi di *amministratore di storage*.

Monitora le attività, gli eventi e i lavori di AIDE

È possibile utilizzare la vista centralizzata **Attività** per monitorare eventi e lavori specifici di AIDE in tutte le aree di lavoro o rivedere l'attività limitata alle singole aree di lavoro.

Visualizza l'attività AIDE a livello di cluster

Monitora le operazioni dell'area di lavoro, risolvi i problemi di estrazione dei metadati e monitora la pubblicazione della raccolta di dati nell'intera distribuzione AIDE.

1. In ONTAP System Manager, nel riquadro di navigazione seleziona **Data engine** e poi **Activity**.
2. Seleziona la scheda **Eventi**:
 - Esaminare i recenti eventi specifici di AIDE, come i seguenti:
 - Creazione, aggiornamento o eliminazione dell'area di lavoro
 - Operazioni di aggiunta/rimozione del data container
 - Pubblicazione della raccolta di dati (se presente)
 - Utilizza filtri (per gravità, tipo di oggetto, workspace o intervallo di tempo) per concentrarti sugli eventi attivi o critici.
3. Seleziona un singolo evento da aprire e rivedere:
 - Descrizione e timestamp
 - Area di lavoro, raccolta di dati o origine dati interessati
 - Azione consigliata, se fornita
4. Selezionare la scheda **Jobs**:
 - Monitora i lavori di lunga durata come:
 - Estrazione iniziale dei metadati per uno spazio di lavoro
 - Lavori di refresh dell'area di lavoro/aggiornamento del catalogo
 - Lavori di pubblicazione o refresh della raccolta di dati
 - Controlla lo stato e l'avanzamento del lavoro.

5. Seleziona un job per aprire la peek view e rivedere:
 - Ora di inizio e fine
 - Percentuale di avanzamento e fase (per esempio, **Scanning**, **Publishing**)
 - Area di lavoro, raccolta di dati o origine dati interessati
 - Messaggi di errore per lavori non riusciti

Visualizza l'attività specifica dell'area di lavoro

Per risolvere i problemi di uno specifico workspace, apri i dettagli del workspace (**Data engine > Workspaces**, seleziona un workspace) e utilizza la scheda **Attività**:

- Esamina gli eventi e i lavori limitati solo a quell'area di lavoro.
- Utilizzare questa vista per isolare problemi quali un singolo spazio di lavoro bloccato in `Processing` stato.

Visualizza gli eventi a livello di cluster

Esamina i messaggi di evento per avere una preziosa traccia dell'attività del sistema. Ogni evento include una descrizione e un identificatore univoco insieme a un'azione consigliata.

1. In ONTAP System Manager, selezionare **Events & jobs** e quindi **Events**.
2. Rivedi e rispondi alle azioni consigliate nella parte superiore della pagina, come ad esempio abilitare l'aggiornamento automatico.
3. Selezionare la scheda **Events log** per visualizzare un elenco dei messaggi.
4. Seleziona un messaggio di evento per esaminarlo in modo più dettagliato, incluso il numero di sequenza, la descrizione, l'evento e l'azione consigliata.
5. Facoltativamente, seleziona la scheda **Suggerimenti Active IQ** e registrati ad Active IQ per ottenere informazioni dettagliate sui rischi per il cluster.

Visualizza i lavori a livello di cluster

Visualizza tutti i lavori in esecuzione sul cluster AIDE, inclusi i lavori specifici di AIDE e i lavori generali ONTAP.

1. In ONTAP System Manager, selezionare **Events & Jobs** e quindi **Jobs**.
2. Personalizza la visualizzazione e cerca e scarica le informazioni sul lavoro secondo necessità.

Visualizza il registro di controllo

Utilizza il registro di controllo per esaminare un record dell'attività del sistema in base all'uso di protocolli di accesso come HTTP.

1. In ONTAP System Manager, selezionare **Events & jobs** e quindi **Audit logs**.
2. Seleziona **Impostazioni** per abilitare o disabilitare le operazioni che vengono monitorate.

Gestisci notifiche

Configura le destinazioni di notifica per inoltrare automaticamente gli eventi AIDE e i registri di controllo.

Passaggi

1. In ONTAP System Manager, selezionare **Cluster** e quindi **Settings**.

2. Vai su **Gestione notifiche** e seleziona .
3. Selezionare l'azione appropriata per visualizzare o configurare le destinazioni utilizzate da AIDE:
 - a. Destinazioni evento: seleziona **Visualizza destinazioni evento**
 - b. Destinazioni del registro di controllo: seleziona **View audit destinations**
4. Selezionare **Aggiungi** come appropriato e fornire le informazioni sulla destinazione.
5. Seleziona **Save**.

Informazioni correlate

- ["Monitoraggio di eventi, prestazioni e salute ONTAP"](#)

Gestire gli spazi di lavoro di AI Data Engine

Un'area di lavoro è un insieme di origini dati (volumi) che l'AI Data Engine (AIDE) utilizza per creare e aggiornare un catalogo di metadati per un progetto o un caso d'uso specifico. In qualità di amministratore dello storage, puoi utilizzare ONTAP System Manager per monitorare lo stato dell'area di lavoro, modificarne la configurazione, controllare le origini dati, gestire gli utenti ed eliminare le aree di lavoro quando non sono più necessarie.

Prima di iniziare

- Per gestire gli spazi di lavoro sono necessari i privilegi di *storage administrator*.

Esaminare lo stato dell'area di lavoro

Esaminare lo stato dell'area di lavoro, l'utilizzo della capacità e lo stato dei metadati per garantire che il Metadata Engine funzioni come previsto e non consumi risorse impreviste.

Passaggi

1. Da ONTAP System Manager, nel riquadro di navigazione selezionare **Data engine > Workspaces**.
2. Consultare il riepilogo in cima alla pagina per il numero totale di spazi di lavoro, lo stato generale degli spazi di lavoro e l'utilizzo della capacità.
3. Per informazioni specifiche sull'area di lavoro, seleziona un nome area di lavoro. Nella scheda **Panoramica** conferma:

- Stato e dimensioni dell'area di lavoro.
- Contenitori di dati (volumi) inclusi nel workspace.
- Conteggio degli elementi e ora dell'ultimo aggiornamento per data source.
- Eventuali avvisi a livello di workspace.



Se un'area di lavoro o una raccolta di dati mostra uno stato di errore, verificare che tutti i volumi di origine siano online e accessibili.

4. Selezionare la scheda **Data collections** per visualizzare:
 - Tutte le raccolte di dati associate a questa workspace.
 - Stato (ad esempio `Published` o `Error`), dimensione e ora dell'ultimo aggiornamento.



System Manager è di sola lettura per le raccolte dati. I data engineer possono creare e gestire le raccolte dati nell'AI Data Engine Console.

5. Selezionare la scheda **Utenti** per visualizzare:
 - L'elenco degli utenti con accesso a questo workspace.
6. Selezionare la scheda **Attività** per visualizzare eventi e lavori relativi solo a questa area di lavoro.

Modifica le proprietà dell'area di lavoro e la pianificazione dell'intervallo di refresh

È possibile modificare il nome, la descrizione, l'intervallo di refresh e (se concesso in licenza) la policy di guardrail di un'area di lavoro.

Passaggi

1. Da **Data engine > Workspaces**, seleziona ⓘ accanto a un workspace e seleziona **Modifica**.
2. Modifica le proprietà dell'area di lavoro:
 - Aggiorna **Name** e **Description** secondo necessità.
 - Regola l'**intervallo di refresh** (frequenza di aggiornamento dei metadati) entro l'intervallo consentito (ore e giorni).
 - Se è installata una licenza AIDE, è possibile selezionare **Guardrail policy**.
3. Seleziona **Save**.



Le modifiche all'intervallo di refresh o all'elaborazione dei metadati potrebbero influire sulla frequenza con cui le relazioni remote di SnapMirror vengono aggiornate per questa area di lavoro.

Aggiungi container di dati a un workspace esistente

È possibile aggiungere volumi montati aggiuntivi (locali o da un cluster remoto peered) in modo che i relativi metadati vengano inclusi nel catalogo dell'area di lavoro.

1. Da **Data engine > Workspaces**, eseguire una delle seguenti operazioni:
 - Seleziona ⓘ accanto all'area di lavoro e seleziona **Aggiungi container di dati**.
 - Aprire l'area di lavoro, selezionare la scheda **Overview** e quindi selezionare **Add** nella sezione data containers.
2. Nella finestra di dialogo **Aggiungi container di dati all'area di lavoro**:
 - Individua i volumi locali sul cluster AIDE.
 - Espandi i cluster peered per selezionare volumi remoti (i volumi remoti richiedono il peering di cluster e SVM).



Possono essere selezionati solo i volumi idonei, online, che non sono esclusi a livello globale e non fanno già parte dell'area di lavoro.

3. Se viene richiesto il mapping del volume remoto:
 - Selezionare la VM di archiviazione di destinazione sul cluster AIDE per ricevere le destinazioni SnapMirror per i volumi remoti scelti.

4. Seleziona **Aggiungi**.
5. Utilizzare la scheda **Attività** dell'area di lavoro o **Motore dati > Attività** per monitorare l'estrazione dei metadati e qualsiasi inizializzazione di SnapMirror per nuove origini dati.

Rimuovi i container di dati da un workspace

È possibile rimuovere un container di dati quando non è più rilevante per lo scopo dell'area di lavoro o se si desidera ridurre l'ambito della gestione dei metadati per tale area di lavoro. La rimozione di un container di dati interrompe l'aggiornamento dei metadati per quel volume e rimuove i relativi metadati dal catalogo dei metadati.



Non eliminare un volume di origine da ONTAP che è stato aggiunto a un'area di lavoro. Se elimini il volume, l'area di lavoro entrerà in stato di errore. Rimuovere sempre prima il container di dati dall'area di lavoro prima di eliminare qualsiasi volume ONTAP sottostante.

Passaggi

1. Passare a **Data engine > Workspaces** e selezionare il workspace che contiene il data container.
2. Nella scheda **Overview**, individua il container da rimuovere.
3. Selezionare **Rimuovi** accanto al data container.
4. Rivedi la finestra di dialogo di conferma e seleziona **Remove**.



La rimozione di un container di dati da un'area di lavoro non elimina il volume ONTAP sottostante né la sua relazione di SnapMirror. Influisce solo sull'utilizzo dei metadati all'interno dell'AI Data Engine.

Gestisci gli utenti dell'area di lavoro

È possibile concedere o revocare l'accesso per gli utenti data engineer e data scientist a un'area di lavoro. Questi utenti sono definiti nel provider di identità (OIDC) e mappati ai ruoli ONTAP. Consultare la ["assegna gli utenti agli spazi di lavoro"](#) documentazione per informazioni su come gestire l'accesso degli utenti.

Eliminare un workspace

È possibile eliminare un'area di lavoro per rimuovere la definizione dell'area di lavoro e i metadati AIDE associati. Anche tutte le raccolte di dati e gli incorporamenti vettoriali associati all'area di lavoro vengono rimossi.



I dati ONTAP sottostanti (volumi, SnapMirror relationship) non vengono eliminati.

Passaggi

1. In **Data engine > Workspaces**, esegui una delle seguenti operazioni:
 - Elimina un singolo workspace, seleziona e seleziona **Elimina**.
 - Elimina più aree di lavoro, seleziona le caselle di controllo per le aree di lavoro, quindi seleziona **Elimina**.
2. Nella finestra di dialogo di conferma, esaminare gli impatti dell'azione prima di procedere:
 - I metadati del workspace vengono eliminati definitivamente.
 - Le raccolte di dati e gli embedding associati all'area di lavoro vengono eliminati definitivamente.



Non esiste alcuna opzione di soft-delete o di ripristino.

3. Seleziona la casella di controllo per confermare la tua comprensione e seleziona **Delete**.

Informazioni correlate

- ["Assegna utenti agli spazi di lavoro"](#)

Aggiorna e mantieni il tuo sistema AIDE

Aggiornamenti e compatibilità del sistema AI Data Engine

Mantieni i componenti del sistema AI Data Engine (AIDE) aggiornati per garantire prestazioni ottimali e accesso a nuove funzionalità. Aggiorna i componenti dopo l'implementazione, quando diventano disponibili nuovi software o firmware, quando aggiungi o sostituisci nodi, o periodicamente per gli aggiornamenti delle funzionalità.

Componenti del sistema AIDE

Ci sono due componenti principali che richiedono aggiornamenti in un sistema AIDE: ONTAP software e AIDE software, che include il firmware DCN.

Software ONTAP

ONTAP è il sistema operativo che viene eseguito sui sistemi di storage NetApp, inclusi quelli utilizzati nelle distribuzioni AIDE. Mantieni ONTAP aggiornato per preservare la stabilità del sistema, la sicurezza e la compatibilità con i componenti AIDE. I componenti AIDE vengono aggiornati separatamente.

Aggiornamenti software AIDE

Gli aggiornamenti del software AIDE Console e del firmware DCN vengono distribuiti insieme come un unico pacchetto (.tgz) e non sono incorporati nelle immagini ONTAP. Gli aggiornamenti garantiscono il corretto funzionamento dei componenti hardware all'interno del sistema AIDE e forniscono nuove funzionalità, miglioramenti delle prestazioni e correzioni di bug.

Comprendere il processo di aggiornamento

Gli aggiornamenti del software AIDE possono essere gestiti tramite ONTAP System Manager.

AIDE non supporta la funzionalità di aggiornamenti automatici del software ONTAP. È possibile registrarsi per ricevere notifiche da ["Download dal sito di supporto NetApp"](#), ma tutti gli aggiornamenti del software AIDE vengono eseguiti manualmente dall'amministratore.

Tipi di release e ambito:

- Le versioni principali di ONTAP (9.x.x) e le versioni principali di AIDE (9.x.x U0) introducono nuove funzionalità, API o modifiche che influiscono sull'integrazione di ONTAP.
- Le patch release ONTAP (9.x.x Px) e le update release AIDE (9.x.x Ux) contengono correzioni e aggiornamenti che non influiscono sull'integrazione ONTAP.

Matrice di compatibilità

Assicurare la compatibilità tra ONTAP e il software AIDE durante la pianificazione degli aggiornamenti.

Il software AIDE viene rilasciato come versione "U". Una release principale di AIDE è una release "U0" e le successive release minori saranno "U1" e successive.

Compatibilità tra ONTAP e AIDE

Versioni di AIDE	Versioni ONTAP supportate
9.18.1 U0	9.18.1 GA e tutti 9.18.1 Px
9.18.1 U1 e versioni successive	9.18.1 GA e tutti 9.18.1 Px



"Px" indica tutte le patch release di ONTAP nella versione principale (ad esempio, 9.18.1 P1, 9.18.2 e altre).

Percorsi di aggiornamento AIDE

Gli esempi seguenti utilizzano ipotetiche versioni future per mostrare i percorsi di upgrade e aggiornamento consentiti da AIDE 9.18.1 U0 e 9.18.1 U1.

Se la tua attuale versione di AIDE è...	E la tua versione di destinazione di AIDE è...	Il tuo percorso di upgrade o aggiornamento è...
9.18.1 U0	9.18.1 U1	Diretto
9.18.1 U0	9.18.1 U3	Diretto (puoi aggiornare da qualsiasi Ux a qualsiasi Ux successiva in 9.18.1)
9.18.1 U1	9.18.1 U3	Diretto (puoi aggiornare da qualsiasi Ux a qualsiasi Ux successiva in 9.18.1)

Ripristina limitazioni

I sistemi AIDE non supportano le operazioni di ripristino per DCN firmware, gli aggiornamenti software AIDE o ONTAP sui sistemi storage AFX. Dopo aver installato un aggiornamento o un upgrade, non è possibile tornare a una versione precedente. Esaminare [requisiti di compatibilità](#) prima di aggiornare o eseguire l'upgrade.

Informazioni correlate

- ["Aggiorna il software ONTAP per i sistemi AFX"](#)
- ["Aggiorna il software AIDE"](#)

Aggiorna AI Data Engine software

In qualità di amministratore dello storage, puoi aggiornare AI Data Engine (AIDE) software, data compute node (DCN) firmware e altri file di sistema sul tuo sistema AIDE utilizzando ONTAP System Manager.

AIDE non supporta la funzionalità di aggiornamenti automatici del software ONTAP. È possibile registrarsi per ricevere notifiche da ["Download dal sito di supporto NetApp"](#), ma tutti gli aggiornamenti del software AIDE vengono eseguiti manualmente dall'amministratore.

Informazioni su questa attività

Il pacchetto software combinato AI Data Engine è significativamente più grande di un tipico pacchetto di aggiornamento ONTAP (il pacchetto AIDE è di circa 40GB). Prevedere tempi di caricamento e installazione più lunghi durante l'aggiornamento del software AIDE.

Prima di iniziare

- È necessario disporre dei privilegi di *storage administrator* per aggiornare DCN firmware e AI Data Engine software.
- Per il tuo account attivo sono necessarie le credenziali del NetApp Support Site.
- ["Garantire la compatibilità tra ONTAP, DCN firmware e AI Data Engine software durante la pianificazione degli aggiornamenti"](#).



La funzionalità di ripristino non è supportata per gli aggiornamenti del firmware DCN o del software AI Data Engine. Dopo l'installazione di un aggiornamento, non è possibile ripristinare una versione precedente.

Passaggi

1. ["Scarica il file di aggiornamento combinato del firmware DCN e del software AIDE su un client locale"](#).
2. In System Manager, seleziona **Cluster > Settings > Software updates**.
3. Accanto a **Software updates**, seleziona ➔.
4. In Aggiornamenti AI Data Engine, seleziona **Aggiungi file software AI Data Engine** e seleziona il pacchetto di aggiornamento.
5. Una volta completato il caricamento del pacchetto, selezionare **Aggiorna** per iniziare a installare gli aggiornamenti sui nodi DCN.



La console AIDE non è disponibile o accessibile quando l'aggiornamento del nodo DCN è in corso.

Risultato

I DCN vengono aggiornati con AI Data Engine software e la versione aggiornata viene visualizzata per ciascun nodo.

Informazioni correlate

- ["Aggiorna il software ONTAP per i sistemi AFX"](#)

Aggiungi nodi di elaborazione dati al tuo AIDE cluster

È possibile aggiungere nodi di elaborazione dati (DCN) durante la creazione di un nuovo cluster AI Data Engine (AIDE) o l'espansione di un cluster esistente. Il flusso di lavoro consiste nell'individuare e configurare i nodi utilizzando ONTAP System Manager.

Preparati ad aggiungere nodi

Ci sono diverse considerazioni quando si aggiungono nodi DCN.

Quando si crea un nuovo AIDE cluster

Sono necessari esattamente tre nodi DCN per un nuovo cluster AIDE.

Installazione hardware e indirizzabilità

Assicurati che siano soddisfatti i seguenti prerequisiti:

- Il nuovo hardware DCN è montato su rack, acceso e cablato agli switch del cluster.
- È disponibile un intervallo di spazio di indirizzi IP per la subnet back-end DCN-ONTAP.
- Il cluster ONTAP è stato inizializzato ed è raggiungibile dalla LIF di gestione del cluster.

Credenziali di System Manager

Sono necessari i privilegi di *storage administrator* per eseguire attività di creazione o espansione del cluster AIDE.

Compatibilità software

Consultare la seguente documentazione per verificare che le versioni hardware e software DCN siano compatibili con il cluster ONTAP:

- ["Il software DCN è compatibile con la versione di ONTAP"](#).

Durante l'operazione di aggiunta del nodo, System Manager confermerà che i nuovi nodi eseguono una versione software compatibile con:

- La versione effettiva del cluster ONTAP (ECV) se si tratta della prima unione DCN.
- La versione del cluster DCN esistente se i DCN sono già presenti.

Se un nodo è incompatibile:

- Viene visualizzato un errore accanto al DCN interessato nella finestra di dialogo **Aggiungi**.
- È necessario innanzitutto aggiornare il DCN software (o ONTAP, a seconda dei casi) a una versione compatibile.

Aggiungi i nodi di elaborazione dei dati

Si aggiungono i nodi DCN durante la creazione di un nuovo cluster AIDE o l'espansione di un cluster esistente.

Passaggi

1. In System Manager, seleziona **Dashboard** nel riquadro di navigazione e poi la scheda **Health**.
2. Confermare che ci sono nodi da aggiungere e selezionare **Visualizza dettagli** per visualizzare l'elenco.

L'elenco contiene i nodi scoperti che non fanno ancora parte dell'AIDE cluster

3. In alternativa, è possibile selezionare **Cluster e Overview** e la scheda **Data compute** per visualizzare l'elenco.
4. Nella parte inferiore della pagina di data compute, seleziona **Aggiungi** sopra l'elenco dei nodi.
5. Nella finestra di dialogo **Aggiungi nodi di elaborazione dati**, seleziona i nodi DCN che desideri aggiungere.

Puoi facoltativamente rinominare i singoli nodi prima di aggiungerli.

6. Se è la prima volta che aggiungi nodi e non esiste alcuna subnet backend, seleziona **Add subnet e fornisci**:
 - Nome subnet (per uso interno)

- Indirizzo e maschera di sottorete
- Intervallo di indirizzi IP per i nodi DCN e ONTAP su questa rete backend

System Manager verifica che l'intervallo includa un numero sufficiente di indirizzi IP liberi per tutti i DCN aggiunti, per tutti i nodi ONTAP nel cluster e per gli IP mobili aggiuntivi a livello di cluster utilizzati per la comunicazione da DCN a ONTAP.

- Che tu abbia aggiunto la subnet backend o che essa esista già:
 - Esaminare gli indirizzi IP disponibili.
 - Se necessario, selezionare **Edit subnet** ed estendere l'intervallo IP.
 - È possibile solo aumentare l'intervallo. La riduzione o la modifica della subnet non sono supportate.
 - La modifica della subnet o dell'intervallo IP potrebbe richiedere la ricreazione del cluster Kubernetes sottostante sui DCN e può richiedere diversi minuti.
- Facoltativamente, è possibile configurare l'interfaccia del servizio Data Engine fornendo:
 - Indirizzo IP del servizio
 - Maschera di rete
 - Gateway (se richiesto per il tuo ambiente)

Il singolo IP verrà bilanciato in base al carico tra i DCN e utilizzato come indirizzo frontend per l'AI Data Engine Console e le relative API.

- Esaminare i nodi selezionati, la subnet backend e le impostazioni dell'interfaccia del servizio Data Engine.
- Seleziona **Aggiungi** e attendi il completamento dell'operazione. System Manager eseguirà le seguenti azioni:
 - Aggiungi i nodi selezionati al cluster DCN
 - Effettua il provisioning della rete backend e unisci i nodi al cluster DCN basato su Kubernetes
 - Aggiorna i metadati interni per la scoperta DCN
- Dopo il completamento, seleziona **Cluster** e **Overview** e conferma:
 - In **Calcolo dati** i nuovi DCN vengono visualizzati come parte del cluster
 - Tutti i nodi sono **Healthy**
 - Verificare che la scheda **Health** della dashboard mostri il conteggio dei nodi aggiornato

Sostituisci un nodo nel tuo cluster AIDE

È necessario sostituire un nodo di elaborazione dati (DCN) nel cluster AI Data Engine (AIDE) se smette di funzionare o deve essere sostituito a causa di un guasto hardware, di un aggiornamento o di manutenzione. Ciò garantisce che il cluster AIDE rimanga integro e operativo. La procedura può essere eseguita senza interrompere i servizi in corso.

Prepararsi a sostituire un nodo

Ci sono diverse cose da considerare prima di sostituire un nodo nel tuo AIDE cluster.

Credenziali di System Manager

Per eseguire le attività di sostituzione dei nodi del cluster AIDE sono necessari i privilegi di *amministratore di storage*.

Restrizioni

È necessario essere a conoscenza delle seguenti restrizioni quando si sostituisce un nodo nel cluster AIDE:

- La sostituzione dei nodi è supportata solo tramite la CLI e facoltativamente l'API REST.
- Non è possibile eseguire la sostituzione del nodo utilizzando System Manager.
- Il nuovo nodo deve corrispondere alla versione software del cluster; ONTAP lo aggiornerà se necessario.
- Per evitare conflitti di indirizzi IP, il nodo non funzionante non deve essere acceso mentre è connesso alla rete del cluster.

Requisiti

Ti serviranno i seguenti elementi:

- Numero di serie per il nuovo nodo sostitutivo

Sostituisci un nodo DCN nel tuo cluster AIDE

È possibile sostituire un nodo DCN nel cluster AIDE utilizzando la seguente procedura.

Passaggi

1. Rimuovere fisicamente il nodo guasto

Spegnere e scollegare il nodo dalla rete del cluster. Assicurarsi che il nodo non sia avviato sulla rete durante il processo di sostituzione.

2. Elimina il nodo non riuscito dal cluster utilizzando il seguente comando:

```
dcn cluster node delete -name <node_name> -force true
```

Fornire il nome effettivo per il valore <node_name>.

3. Collegare fisicamente il nuovo nodo al cluster

Assicurati che il nodo sia cablato, acceso e rilevabile.

4. Visualizza i nodi rilevabili e non configurati per verificare che il nuovo nodo sia online:

```
dcn cluster node show -membership available
```

5. Aggiungi il nodo al cluster utilizzando il seguente comando:

```
dcn cluster node create -serial-number <new_node_serial>
```

ONTAP assegnerà un indirizzo IP al nuovo nodo. Se la versione software del nodo non corrisponde al cluster, ONTAP aggiornerà automaticamente il nodo.

6. Verificare lo stato del cluster e l'integrazione dei nodi con uno dei seguenti comandi:

```
dcn cluster node show
```

```
dcn cluster node show -instance
```

Informazioni correlate

- ["Espandi il tuo cluster compute"](#)

Gestisci la vettorializzazione e le raccolte di dati

Avvio rapido Data-to-RAG per AI Data Engine

Passa da un sistema AI Data Engine (AIDE) appena implementato a un endpoint retrieval-augmented generation (RAG) funzionante utilizzando questo flusso di lavoro. Comprendi come gli amministratori di storage, i data engineer e i data scientist collaborano utilizzando ONTAP System Manager e AIDE Console.

Prima di iniziare

- Hai installato e aggiunto i Data compute nodes (DCN) al cluster ONTAP.
- Hai installato e ottenuto la licenza AI Data Engine software per la vettorializzazione e guardrails.
- Hai configurato ["OpenID Connect \(OIDC\)"](#) e mappato i ruoli di amministratore, data engineer e data scientist.

1

Definire l'ambito e la governance dei dati

In qualità di amministratore di storage o amministratore di sicurezza, desideri preparare l'ambiente in AIDE Console e ONTAP System Manager:

- ["Crea uno o più spazi di lavoro"](#) da fonti di dati locali e remote.
- ["Configura classificatori e policy di protezione"](#) in AIDE Console.
- ["Assegna l'accesso agli spazi di lavoro ai data engineer e ai data scientist"](#).

2

Esplora metadati dello workspace

In qualità di data engineer o data scientist, desideri esplorare i metadati del workspace utilizzando AIDE Console:

- ["Esplora i metadati del workspace"](#) per comprendere i contenuti disponibili.
- Definisci uno o più sottoinsiemi logici di dati che dovrebbero alimentare RAG (ad esempio, articoli di supporto, manuali di prodotto o note cliniche anonimizzate).

3

Creare e pubblicare una raccolta di dati

In qualità di data engineer o data scientist, vuoi trasformare il sottoinsieme scelto in una raccolta RAG-ready:

- ["Crea una raccolta di dati"](#) dall'area di lavoro utilizzando i filtri selezionati.
- ["Pubblica la raccolta di dati"](#) e monitorare l'indicizzazione finché non raggiunge Ready stato.
- Copia l'URI dell'endpoint di recupero per la raccolta scelta e fornisci agli data scientist o agli sviluppatori di applicazioni.
- ["Visualizza lo stato della raccolta di dati e l'impronta vettoriale"](#) secondo necessità.

E ora?

- ["Definisci il tuo patrimonio di dati e le policy di guardrail in AI Data Engine"](#)

- ["Esplora i metadati del workspace nella AI Data Engine Console"](#)
- ["Crea raccolte di dati in AI Data Engine Console"](#)

Esplora i metadati del workspace nella AI Data Engine Console

Come data engineer o data scientist, il tuo primo compito in AI Data Engine (AIDE) è capire quali dati sono disponibili nei tuoi spazi di lavoro. Utilizzi AI Data Engine Console per interrogare il catalogo dei metadati, cercare file rilevanti e identificare sottoinsiemi di dati da trasformare in raccolte di dati.

Prima di iniziare

- Sono necessari i privilegi di *data engineer* o *data scientist* in AI Data Engine Console e l'accesso ad almeno un workspace.
- Un amministratore di storage ha:
 - Creato uno o più spazi di lavoro in ONTAP System Manager.
 - Assegnato al tuo utente o gruppo l'accesso alle aree di lavoro pertinenti.
- L'estrazione dei metadati per l'area di lavoro è stata completata e l'area di lavoro è in Ready stato.
- I classificatori sono abilitati, in modo che i metadati includano tag di classificazione (ad esempio, indicatori PII).

Sign in to AIDE Console come data engineer o data scientist

Passaggi

1. In un browser, vai all'URL della AIDE Console:

```
https://<cluster_management_ip>/console
```

2. Autenticati tramite il provider OIDC della tua organizzazione.
3. Verifica che il tuo ruolo sia riconosciuto come data engineer o data scientist (ad esempio, in base all'area di lavoro disponibile e alle azioni di raccolta di dati). Per ulteriori informazioni, consulta ["Documentazione del ruolo AIDE"](#) per scoprire come data engineer e data scientist lavorano con i componenti AIDE.

Risultato

Hai effettuato l'accesso all'AI Data Engine Console e vedi solo gli spazi di lavoro a cui ti è stato concesso l'accesso.

Visualizza i tuoi spazi di lavoro accessibili

Passaggi

1. In AI Data Engine Console, vai su **Data Curator > Workspaces**.
2. Rivedi l'elenco degli spazi di lavoro a cui puoi accedere.
3. Seleziona un'area di lavoro per aprirne i dettagli.

Risultato

Ora hai una visualizzazione con ambito area di lavoro del patrimonio di dati che gli amministratori di storage

hanno reso disponibile per i tuoi progetti.

E ora?

- "Crea raccolte di dati per RAG da uno spazio di lavoro"

Crea raccolte di dati in AI Data Engine Console

Le raccolte di dati sono i componenti fondamentali del RAG in AI Data Engine (AIDE). In qualità di data engineer o data scientist, definisci quali file appartengono a una raccolta, configuri le opzioni di incorporamento e indicizzazione e pubblichi la raccolta in modo che le applicazioni possano interrogarla tramite un endpoint di recupero.

Eseguirai tutte le attività di raccolta di dati nella AI Data Engine Console.

Prima di iniziare

- Sono necessari i privilegi di *data engineer* o *data scientist* in AI Data Engine Console (https://<cluster_management_ip>/console).
- Hai accesso ad almeno uno spazio di lavoro con metadati estratti e in Ready stato.
- Hai esplorato i metadati del workspace e identificato query o filtri che definiscono sottoinsiemi significativi di dati.
- La licenza del software AI Data Engine è installata e le funzionalità di inferenza sono abilitate.

Creare una raccolta di dati dai metadati del workspace

Passaggi

1. Vai su **Data Curator > Workspaces** e seleziona il workspace che contiene i dati di destinazione.
2. Seleziona **Add data collection**.
3. Nella pagina **Crea nuova raccolta di dati**, procedi come segue:
 - a. Inserisci un nome e una descrizione per la raccolta (ad esempio, `Support_KB_RAG_EN`).
 - b. Scegli se la raccolta deve essere:
 - **Dinamico**: i nuovi file vengono automaticamente identificati e aggiunti alla raccolta di dati in base ai criteri di filtraggio che definisci. Questo avviene durante gli aggiornamenti dell'area di lavoro.
 - **Statico**: scegli quali file includere nella raccolta. Puoi modificare i file se la raccolta di dati è in `draft` stato. Dopo che la raccolta di dati passa in `Published` stato, non può essere modificata.
4. Specificare il sottoinsieme di origine:
 - a. Utilizza parole chiave e filtri (tipo di file, timestamp e altri attributi) per trovare i file rilevanti da includere.



È possibile selezionare un nome file per aprire una finestra di anteprima del contenuto.

5. Aggiungi questi file alla raccolta di dati.
6. Seleziona **Salva** per finalizzare la raccolta.

Risultato

Hai definito l'ambito della raccolta di dati e aggiunto i file necessari ad essa. AIDE genera gli embedding e crea

l'indice vettoriale quando pubblichi la raccolta.



Crea raccolte di piccole dimensioni e mirate (ad esempio, per caso d'uso o dominio) anziché un'unica raccolta "tutto". Questo migliora la pertinenza del recupero e la gestibilità.

Pubblica una raccolta di dati

Pubblica la raccolta di dati per renderla interrogabile dalle applicazioni AI tramite un endpoint di recupero RAG. La pubblicazione genera incorporamenti vettoriali dai file selezionati e li indicizza per la ricerca semantica. Dopo che la raccolta raggiunge `Ready` stato, il suo endpoint diventa disponibile per i data scientist per integrarlo in notebook, pipeline e applicazioni AI per retrieval-augmented generation (RAG) e ricerca.



Per le raccolte di grandi dimensioni, considera di pianificare la pubblicazione iniziale e le principali ripubblicazioni durante i periodi di minore affluenza per ridurre al minimo la contesa delle risorse.

Passaggi

1. Vai su **Data Curator > Data collections** e seleziona il menu delle opzioni (`...`) per la tua raccolta di dati.
2. Seleziona **Publish**.
3. Seleziona una configurazione di ottimizzazione predefinita o personalizzata.
4. Seleziona **Pubblica** per avviare la trasformazione dei dati.
5. In AIDE Console, apri la vista dei dettagli della raccolta (**Data Curator > Data collections**) per gli aggiornamenti di stato.

Risultato

La raccolta raggiunge lo `Ready` stato ed è disponibile per l'uso da parte di applicazioni downstream e data scientist.

Da **Data Curator > Data collections**, puoi selezionare **Copia URI** per ottenere le informazioni necessarie per accedere alla raccolta di dati tramite un'API.

Aggiornare o eliminare una raccolta di dati

Nel tempo potrebbe essere necessario perfezionare o ritirare raccolte di dati. Perfezionare una raccolta potrebbe comportare la modifica dei filtri per aggiungere o rimuovere file, la modifica delle impostazioni di embedding o l'aggiornamento della descrizione della raccolta. L'eliminazione di una raccolta la rimuove definitivamente e rende non disponibile il relativo endpoint di recupero.

Aggiornare una raccolta di dati

È possibile aggiornare una raccolta di dati quando è in `draft` stato.

Passaggi

1. Vai a **Data Curator > Data collections**.
2. Seleziona la raccolta che desideri modificare.
3. Scegli **Modifica**.
4. Modifica una delle seguenti:
 - Nome e descrizione

- Filtri (percorsi, tipi di file, classification tag).
- Impostazioni di embedding e chunking.

5. Salva le modifiche.

6. Pubblica nuovamente la raccolta in modo che la nuova definizione e gli embeddings abbiano effetto.

Risultato

Viene eseguito un nuovo processo di indicizzazione con la configurazione aggiornata e la raccolta torna a uno stato Ready quando è completo.

Elimina una raccolta

L'eliminazione di una raccolta è definitiva. Assicurarsi che nessuna applicazione di produzione dipenda ancora dall'endpoint di recupero della raccolta prima di eliminarla.

Passaggi

1. Vai su **Data Curator > Data collections** e seleziona il menu delle opzioni (**...**) per la raccolta.
2. Scegli **Elimina**.
3. Conferma l'eliminazione.

Risultato

La definizione della raccolta e i suoi embeddings vengono rimossi da AI Data Engine. Le applicazioni che tentano di interrogare il precedente endpoint di retrieval non riusciranno dopo la rimozione della raccolta.

E ora?

- ["Visualizza raccolte di dati"](#)

Visualizza le raccolte di dati in AI Data Engine

Dopo che gli ingegneri dei dati o gli scienziati dei dati hanno creato e pubblicato raccolte di dati dagli spazi di lavoro, è necessario avere visibilità sul loro stato, dimensioni e impatto sull'AI Data Engine cluster.

Se sei un amministratore di storage, un data engineer o un data scientist, puoi visualizzare le raccolte di dati su ONTAP System Manager e AIDE Console.

Prima di iniziare

- Per visualizzare le raccolte di dati, sono necessari i privilegi di *amministratore dello storage* in ONTAP System Manager oppure i privilegi di *data engineer* o *data scientist* in AI Data Engine Console (https://<cluster_management_ip>/console).
- Esiste almeno un'area di lavoro con metadati estratti correttamente.
- Gli ingegneri dei dati o i data scientist hanno creato e pubblicato almeno una raccolta di dati da AI Data Engine Console.
- La licenza del software AI Data Engine è installata e le funzionalità di inferenza sono abilitate, in modo che gli endpoint di vettorializzazione e recupero siano attivi.

Visualizza raccolte di dati a livello di cluster

Per gli amministratori di storage, ONTAP System Manager fornisce una visualizzazione a livello di cluster delle raccolte di dati e del loro footprint, ma non consente agli amministratori di crearle o modificarle.

Passaggi

1. In System Manager, vai su **Data Engine > Data collections**.
2. Esaminare il riepilogo dell'inventario in cima alla pagina:
 - Numero totale di raccolte di dati per stato
 - Spazio totale occupato dal database vettoriale in tutte le raccolte
 - Spazio vettoriale come percentuale della capacità complessiva del cluster
3. Seleziona una raccolta di dati individuale e rivedila:
 - Nome e descrizione della raccolta
 - UUID
 - Area di lavoro associata
 - Stato
 - Dimensione della raccolta
 - Creatore
 - Ultimo aggiornamento

Risultato

Ora hai una high-level panoramica di tutte le raccolte di dati nel cluster e del loro impatto sullo storage. Utilizza questa vista per identificare le raccolte che sono grandi, obsolete o bloccate in uno stato non pronto.

È anche possibile vedere se una singola raccolta di dati viene aggiornata attivamente e se eventuali errori stanno bloccando l'utilizzo di RAG.

Monitorare i lavori e gli eventi correlati alla raccolta

In qualità di amministratore dello storage, puoi monitorare i job che creano e aggiornano le raccolte dalla pagina **Activity** a livello di cluster e dai dettagli dell'area di lavoro.

Passaggi

1. In System Manager, vai su **Data Engine > Activity**.
2. Nella scheda **Events**:
 - a. Filtra per tipo (ad esempio, workspace, raccolta di dati) o gravità.
 - b. Espandi qualsiasi evento correlato alle raccolte di dati (ad esempio, "Pubblicazione raccolta di dati non riuscita") per visualizzare maggiori dettagli.
3. Nella scheda **Jobs**:
 - a. Filtra per concentrarti sui lavori di indicizzazione e pubblicazione della raccolta di dati.
 - b. Per ogni lavoro, apri la vista peek per vedere:
 - Percentuale di avanzamento.
 - Orari di inizio e fine.
 - Eventuali messaggi di errore o avvisi segnalati.

4. Facoltativamente, torna al workspace interessato (**Data Engine > Workspaces**) e apri la relativa scheda **Attività** per visualizzare gli eventi e i processi limitati solo a quel workspace.

Risultato

È possibile monitorare il ciclo di vita delle raccolte di dati, identificare i processi bloccati o non riusciti e raccogliere informazioni contestuali da trasmettere a data engineers, data scientists o supporto.



Quando una raccolta di dati rimane in `Publishing` stato per un periodo prolungato, verificare la presenza di un processo di lunga durata corrispondente nella pagina Activity prima di presumere un errore.

Visualizza le raccolte di dati dalla AIDE Console

In genere, gli ingegneri dei dati e i data scientist monitorano le raccolte di dati direttamente da AIDE Console, dove vengono create e pubblicate.

Passaggi

1. Accedi ad AIDE Console come data engineer o data scientist.
2. Vai a **Raccolte dati** e seleziona la raccolta di dati desiderata.
3. Per ogni raccolta:
 - a. Controllare lo stato (`Draft`, `Publishing`, `Ready`, o `Failed`).
 - b. Selezionare il nome della raccolta di dati per rivedere i dettagli della definizione (filtri, tipi di file inclusi, opzioni del classificatore, impostazioni di embedding).
 - c. Ispeziona i timestamp dell'ultima pubblicazione o aggiornamento.
4. Se necessario, aprire i dettagli del job o i log (ove disponibili) per comprendere i fallimenti o le esecuzioni incomplete.

Risultato

Gli ingegneri dei dati e i data scientist possono iterare sulle definizioni di raccolta e pubblicarle nuovamente mentre monitorano stato e integrità, senza coinvolgere gli amministratori dello storage.

E ora?

- ["Crea raccolte di dati per RAG in AI Data Engine Console"](#)

Implementare guardrail

Definisci le tue policy Data Guardrails in AI Data Engine per il tuo data estate

In qualità di proprietario di dati o di una piattaforma, utilizzi AI Data Engine (AIDE) Console per definire quali dati rientrano nell'ambito dell'AI, quali dati sono sempre off-limits e quali regole di sicurezza si applicano quando tali dati vengono utilizzati per la classificazione e la retrieval-augmented generation (RAG).

Utilizza queste procedure per definire tali policy in AIDE Console in modo che ONTAP System Manager possa applicarle a tutti i dati nei workspace.

Prima di iniziare

- Per creare e gestire policy globali, sono necessari i privilegi di *amministratore di storage* in AI Data Engine Console (https://<cluster_management_ip>/console).
- Si dispone di un AIDE cluster con nodi di data compute distribuiti e funzionanti.
- "OpenID Connect (OIDC)" è configurato e il tuo ruolo IdP è mappato a un ruolo di amministratore AIDE che consente la gestione delle policy sui dati.
- La licenza del software AI Data Engine è installata in modo che le funzionalità guardrail e inferencing siano abilitate.
- Esiste almeno un'area di lavoro oppure hai concordato con l'amministratore per capire quali origini dati (volumi) verranno utilizzate nelle aree di lavoro.

Comprendere i tipi di policy

AIDE Console espone questi tipi di policy che modellano il patrimonio dei tuoi dati:

- **Classificatori:** abilita i classificatori per rilevare PII, problemi di sicurezza o altri modelli in tutte le aree di lavoro.
- **Categorie di classificazione:** Raggruppa i classificatori in categorie di conformità per l'organizzazione e la gestione.
- **Guardrail policies:** regole di sicurezza e redazione applicate al momento del recupero o dell'inferenza.

Non è possibile utilizzare ONTAP System Manager per creare o gestire queste Data Guardrails policy. Le legge e le applica solo quando un amministratore di storage le applica agli spazi di lavoro. Tutta la definizione e la manutenzione delle policy avviene nella AI Data Engine Console.

Abilita classificatori

I classificatori analizzano sia i metadati che il contenuto per annotare file e oggetti (ad esempio, rilevando PII o categorie sensibili). Prima che i classificatori possano essere eseguiti sui dati dell'area di lavoro, è necessario abilitarli in AI Data Engine Console.

Informazioni su questa attività

Il comportamento del classificatore è controllato a livello globale in AIDE Console. Tutti i classificatori abilitati vengono eseguiti su ogni workspace. Poiché vengono applicati a livello globale, non possono essere abilitati o disabilitati per un singolo workspace. Possono essere abilitati o disabilitati solo a livello globale.

Passaggi

1. Nella AIDE Console, vai su **Data Guardrails > Classifiers**.
2. Seleziona una categoria di classificatori per visualizzare i classificatori che contiene.
3. Seleziona le caselle di controllo per i classificatori che desideri abilitare o seleziona tutte le righe per abilitare i classificatori in blocco.
4. Seleziona **Abilita**.



Utilizza l'opzione di selezione in blocco per abilitare più classificatori contemporaneamente. Ogni volta che abiliti un classificatore, viene attivato un aggiornamento in tutte le aree di lavoro. Per ridurre al minimo gli aggiornamenti non necessari, abilita più classificatori contemporaneamente anziché uno alla volta.

Risultato

Tutti gli spazi di lavoro appena creati ed esistenti eseguono i classificatori abilitati durante l'elaborazione dei metadati.

I tag di classificazione vengono scritti nel catalogo dei metadati e diventano disponibili agli ingegneri dei dati per il filtraggio durante la creazione di raccolte di dati.

Gestisci le categorie del classificatore

I classificatori sono organizzati in categorie (ad esempio "PII" o "Financial data"). Le categorie aiutano a raggruppare i classificatori correlati per una gestione più semplice e una maggiore visibilità sulla conformità. È possibile utilizzare le categorie predefinite che AIDE fornisce o creare categorie personalizzate per soddisfare i requisiti di conformità.

Passaggi

1. Nella AIDE Console, vai su **Data Guardrails > Classifiers**.
2. Visualizza le categorie di classificazione esistenti. Esistono due principali categorie di classificazione:
 - Contenuto o dati: rileva particolari tipi di dati all'interno dei file.
 - Documento: classifica il tipo di documento in base al contenuto.
3. Determina se le sottocategorie del classificatore predefinito sono sufficienti o se desideri creare una tua sottocategoria.
 - Se si utilizza una sottocategoria di classificazione predefinita (ad esempio, **General Privacy**):
 - i. Seleziona il nome della categoria in Classifier categories per visualizzare i classificatori associati.
 - ii. Esamina l'elenco dei classificatori.
 - iii. Seleziona **Aggiungi** per trovare e aggiungere classificatori non elencati dall'elenco completo dei classificatori disponibili.
 - Se vuoi creare una categoria personalizzata, seleziona **+ Add**.
 - i. Aggiungi un nome univoco, una descrizione e assegna i classificatori disponibili alla categoria.
 - ii. Seleziona **Add**
4. Per disabilitare un classificatore all'interno di una categoria, seleziona **...** per il classificatore e scegli **Disabilita**. Puoi anche selezionare tutte le righe per apportare modifiche di stato in blocco.

Risultato

Le categorie organizzano i classificatori per la visibilità della conformità. Gli ingegneri dei dati possono utilizzare i tag di classificazione durante il filtraggio e la creazione di raccolte di dati.

Crea e gestisci le policy Data Guardrails

Le policy Guardrail determinano come AIDE risponde quando i classificatori rilevano contenuti sensibili o quando i prompt e i risultati del recupero violano le regole sui contenuti.

I comportamenti tipici dei Data Guardrails includono:

- Mascheramento o oscuramento delle PII dai frammenti recuperati.
- Blocco delle risposte che violano le regole di conformità.
- Registrazione o etichettatura delle violazioni per audit.

Informazioni su questa attività

Puoi creare e gestire le policy Data Guardrails solo in AIDE Console.

È possibile associare gli spazi di lavoro in ONTAP System Manager a una sola policy di guardrail alla volta.

Passaggi

1. In AI Data Engine Console, vai a **Data Guardrails > Guardrail policies**.
2. Seleziona **Aggiungi**.
3. Inserisci un nome e una descrizione che descrivano chiaramente l'ambito (ad esempio, `Customer PII redaction for support KB`).
4. Configura le condizioni guidate dal classificatore di dati richieste per l'attivazione del guardrail:
 - a. Definisci le condizioni per l'attivazione del guardrail:
 - i. Scegli la categoria del classificatore o il tipo di classificatore per ogni condizione.
 - ii. Aggiungi e definisci condizioni aggiuntive secondo necessità.
 - iii. Definisci criteri di ricerca specifici in **Search**, quindi seleziona **Accept**.
 - b. Definire le azioni per la policy Data Guardrails, come l'anonimizzazione dei contenuti o il blocco e la rimozione di un file da una raccolta di dati.
5. Seleziona l'area di lavoro a cui verrà applicato il guardrail.
6. Imposta lo stato della policy:
 - **Abilitato**: attiva immediatamente la policy.
 - **Modalità test**: consente di convalidare l'impatto della policy prima di attivarla.
 - **Disabilitato**: salva il guardrail senza applicarlo.
7. Selezionare **Add** per salvare il criterio e applicarlo all'area di lavoro.



Utilizzare la **Modalità di test** con un'area di lavoro pilota e una raccolta di dati non di produzione per capire quante risposte sarebbero interessate prima di abilitare l'applicazione di un controllo rigoroso.

Risultato

La nuova policy Data Guardrails è attiva e applicata all'area di lavoro selezionata.

Come le policy interagiscono con gli spazi di lavoro

Dopo che le policy sono state definite:

- L'amministratore dello storage utilizza ONTAP System Manager per creare spazi di lavoro, selezionare data container e associare una guardrail policy.
- I classificatori vengono eseguiti automaticamente sui contenuti dell'area di lavoro in base a ciò che hai abilitato.
- I Data Guardrails associati all'area di lavoro influenzano il comportamento degli endpoint di recupero.

Per data engineer e data scientist:

- Il patrimonio di dati visibile (aree di lavoro e raccolte dati) è già filtrato in base all'assegnazione dei ruoli.
- I metadati che interroghi (ad esempio, i tag PII) sono determinati dai classificatori abilitati.
- Le risposte che le tue pipeline RAG ricevono sono limitate dai guardrail configurati a livello di area di lavoro.

Informazioni correlate

- ["Visualizza le raccolte di dati in AI Data Engine"](#)
- ["Avvio rapido Data-to-RAG per AI Data Engine"](#)

FAQ per NetApp AI Data Engine

Questa FAQ copre le domande più comuni su NetApp AI Data Engine (AIDE), inclusa la sua architettura, distribuzione, tipologie di utenti, caratteristiche tecniche, integrazione e licenze.

Nozioni di base su AIDE

Che cos'è NetApp AI Data Engine (AIDE)?

NetApp AI Data Engine (AIDE) è un servizio dati integrato nello storage che copre l'intero ciclo di vita dell'AI, dalla scoperta e preparazione dei dati raw alla fornitura di endpoint di recupero per alimentare l'AI generativa (GenAI), la Retrieval-Augmented Generation (RAG), l'AI agentic e le AI factory. AIDE automatizza la sincronizzazione e il rilevamento delle modifiche, fornendo una vista unificata e aggiornata dei dati selezionati per la scoperta e la cura dei dati.

Come funziona AIDE?

AIDE si integra direttamente con i sistemi di storage NetApp ONTAP per creare una vista globale e strutturata dell'intero patrimonio di dati NetApp con rilevamento automatico delle modifiche e sincronizzazione. AIDE offre vettorizzazione real-time con compressione e deduplicazione, Data Guardrails basati su policy e integrazione con strumenti di AI.

Utenti e ruoli

Chi utilizza l'AI Data Engine?

Gli utenti principali di AIDE includono:

1. **Amministratori di storage ONTAP:** gestiscono l'infrastruttura, le esigenze di storage specifiche per l'AI, la sicurezza e la conformità.
2. **Data engineers:** gestiscono spostamento dei dati, preparazione e integrazione tra gli ambienti.
3. **Data scientists:** preparano e trasformano i dati rilevanti per il consumo da parte dell'AI.

Requisiti e deployment

Quale hardware è necessario?

AIDE richiede sistemi AFX per l'implementazione (inclusi un controller AFX, uno shelf di dischi e uno switch di rete), ma può utilizzare i dati dei cluster che eseguono ONTAP 9 tramite SnapMirror e peering di cluster. Per le implementazioni di AIDE sono necessari almeno quattro nodi controller AFX per garantire elevata disponibilità e prestazioni.

AIDE viene eseguito su un nodo di calcolo dati NetApp (DCN). Sono necessari tre DCN. Il DCN ospita il software AIDE, che include Metadata Engine, Data Sync, Data Curator e Data Guardrails.

Posso usare il mio DCN?

No. Il DCN è un nodo hardware di elaborazione dati fornito da NetApp ed è l'unico meccanismo di distribuzione per l'AI Data Engine.

Qual è il numero minimo di DCN richiesto?

Sono necessari esattamente tre DCN.

Quale sistema operativo è in esecuzione sui DCN?

I DCN eseguono uno stack software fornito da NetApp con AIDE.

È possibile implementare AIDE senza AFX?

No. AIDE richiede AFX per il deployment. AIDE utilizza Trident per consumare i volumi AFX per lo storage interno (volumi persistenti). Il cluster AFX che fornisce storage per AIDE può essere collegato in peering con un sistema o cluster ONTAP 9. Utilizza il peering tra cluster e SnapMirror per sincronizzare i dati dal cluster ONTAP remoto al sistema AFX.

Gestione e interfacce

AIDE Console fa parte di NetApp Console o è un'interfaccia separata?

La AIDE Console è un'interfaccia di gestione separata che viene eseguita sui DCN. Utilizzi la AIDE Console per gestire i servizi AIDE, come Data Guardrails e Data Curator. Puoi anche utilizzare ONTAP System Manager per monitorare il cluster AIDE.

Funzionalità e capacità

Quali sono le caratteristiche principali di AIDE?

Ci sono quattro caratteristiche principali di AIDE:

Metadata Engine

- Genera automaticamente una visualizzazione strutturata, aggiornata e interattiva dei tuoi dati.
- Funziona con i dati memorizzati su ONTAP.
- Consente ai professionisti dei dati di collaborare con gli amministratori dello storage per trovare e comprendere i dati.
- Le API interrogano i metadati per fornire funzionalità riducendo il carico del traffico NFS sui sistemi di storage.
- La funzionalità di estrazione e catalogazione dei metadati è stata sviluppata appositamente per AIDE e funziona in modo continuo, sfruttando le funzionalità ONTAP come gli snapshot.

Sincronizzazione dati

- Mantiene automaticamente l'attualità dei dati man mano che i dati di origine cambiano senza intervento manuale.
- Gli amministratori definiscono l'intervallo di refresh in giorni o ore.
- Fornisce mobilità incrementale dei dati e sincronizzazione tra i dati per eliminare copie ridondanti dei dati di AI.

Data Guardrails

- Identifica e protegge automaticamente i dati sensibili durante l'intero ciclo di vita dell'AI. È accessibile tramite AI Data Engine Console.
- Esegue continuamente la scansione, la classificazione e la categorizzazione dei dati.
- Identifica dati sensibili (come PII) e rischi.
- Facilita la creazione di policy per la gestione automatica dei dati sensibili in linea con gli standard aziendali e normativi.
- Fornisce la redazione automatica delle informazioni sensibili per la protezione dei dati.

- Limita l'accesso ai file sensibili secondo necessità.

Data Curator

- Consente agli scienziati dei dati di cercare dati rilevanti attraverso lo storage.
- Crea raccolte di dati curate con i dati esistenti sui volumi AFX.
- Genera incorporamenti vettoriali nel layer di storage per ridurre la crescita eccessiva dei dati e aumentare le prestazioni.
- Fornisce un endpoint di recupero per applicazioni AI con ricerca semantica vettoriale e re-ranking.

Integrazione e interoperabilità

AIDE supporta metadati federati su più cluster ONTAP?

AIDE può connettersi a più cluster ONTAP tramite SnapMirror e il cluster peering, consentendo la visibilità centralizzata dei metadati.

Dove vengono archiviati i metadati?

AIDE memorizza i metadati sul cluster AFX connesso utilizzando un volume persistente fornito da AFX. I DCN utilizzano l'archiviazione locale per le operazioni interne.

AIDE Metadata Engine classifica i dati?

No. Il Metadata Engine cataloga i metadati del filesystem e fornisce API per interrogare questi metadati catalogati.

Quali fonti di dati sono supportate?

AIDE supporta volumi ONTAP (locali o remoti) come sorgenti di dati. I cluster ONTAP remoti devono eseguire ONTAP 9 ed essere connessi tramite cluster peering e SnapMirror.

I bucket ONTAP S3 e gli oggetti StorageGRID non sono supportati come origini dati in AIDE 9.18.1.

Quali tipi di file può AIDE elaborare per la classificazione, la vettorializzazione e la ricerca semantica?

AIDE supporta un'ampia gamma di tipi di file, tra cui PDF, DOCX, PPTX, TXT e file immagine con funzionalità OCR.

AIDE supporta la classificazione di dati non inglesi?

AIDE supporta solo dati in lingua inglese.

Quali integrazioni supporta AIDE?

AIDE fornisce un endpoint API RAG accessibile tramite chiamate API dirette o tramite un server Model Context Protocol (MCP). Questo supporta l'integrazione con framework e strumenti agentic AI.

Distribuzione e licenze

Quali sono le opzioni di distribuzione?

AIDE è distribuito on-premises sull'infrastruttura AFX con DCN. Si integra direttamente con NetApp ONTAP AFX.

Come viene concessa in licenza AIDE?

AIDE richiede una licenza software per eseguire Data Guardrails e Data Curator.

Se hai bisogno solo del Metadata Engine, la licenza ONTAP One, inclusa in tutti i sistemi AFX, fornisce il diritto

alle funzionalità solo del Metadata Engine.

Informazioni correlate

- ["Installare le licenze AIDE in ONTAP System Manager"](#)
- ["Scopri l'architettura e i componenti di AIDE"](#)

Note legali

Le note legali forniscono accesso a dichiarazioni di copyright, marchi, brevetti e altro.

Copyright

["https://www.netapp.com/company/legal/copyright/"](https://www.netapp.com/company/legal/copyright/)

Marchi

NETAPP, il logo NETAPP e i marchi elencati nella pagina dei marchi di NetApp sono marchi di NetApp, Inc. Altri nomi di aziende e prodotti possono essere marchi dei rispettivi proprietari.

["https://www.netapp.com/company/legal/trademarks/"](https://www.netapp.com/company/legal/trademarks/)

Brevetti

Un elenco aggiornato dei brevetti di proprietà di NetApp è disponibile all'indirizzo:

<https://www.netapp.com/pdf.html?item=/media/11887-patentspage.pdf>

Informativa sulla privacy

["https://www.netapp.com/company/legal/privacy-policy/"](https://www.netapp.com/company/legal/privacy-policy/)

Open source

I file di avviso forniscono informazioni sui diritti d'autore e sulle licenze di terze parti utilizzati nel software NetApp.

AI Data Engine

["Avviso per AIDE 9.18.1"](#)

Informazioni sul copyright

Copyright © 2026 NetApp, Inc. Tutti i diritti riservati. Stampato negli Stati Uniti d'America. Nessuna porzione di questo documento soggetta a copyright può essere riprodotta in qualsiasi formato o mezzo (grafico, elettronico o meccanico, inclusi fotocopie, registrazione, nastri o storage in un sistema elettronico) senza previo consenso scritto da parte del detentore del copyright.

Il software derivato dal materiale sottoposto a copyright di NetApp è soggetto alla seguente licenza e dichiarazione di non responsabilità:

IL PRESENTE SOFTWARE VIENE FORNITO DA NETAPP "COSÌ COM'È" E SENZA QUALSIVOGLIA TIPO DI GARANZIA IMPLICITA O ESPRESSA FRA CUI, A TITOLO ESEMPLIFICATIVO E NON ESAUSTIVO, GARANZIE IMPLICITE DI COMMERCIALIZZABILITÀ E IDONEITÀ PER UNO SCOPO SPECIFICO, CHE VENGONO DECLINATE DAL PRESENTE DOCUMENTO. NETAPP NON VERRÀ CONSIDERATA RESPONSABILE IN ALCUN CASO PER QUALSIVOGLIA DANNO DIRETTO, INDIRETTO, ACCIDENTALE, SPECIALE, ESEMPLARE E CONSEGUENZIALE (COMPRESI, A TITOLO ESEMPLIFICATIVO E NON ESAUSTIVO, PROCUREMENT O SOSTITUZIONE DI MERCI O SERVIZI, IMPOSSIBILITÀ DI UTILIZZO O PERDITA DI DATI O PROFITTI OPPURE INTERRUZIONE DELL'ATTIVITÀ AZIENDALE) CAUSATO IN QUALSIVOGLIA MODO O IN RELAZIONE A QUALUNQUE TEORIA DI RESPONSABILITÀ, SIA ESSA CONTRATTUALE, RIGOROSA O DOVUTA A INSOLVENZA (COMPRESA LA NEGLIGENZA O ALTRO) INSORTA IN QUALSIASI MODO ATTRAVERSO L'UTILIZZO DEL PRESENTE SOFTWARE ANCHE IN PRESENZA DI UN PREAVVISO CIRCA L'EVENTUALITÀ DI QUESTO TIPO DI DANNI.

NetApp si riserva il diritto di modificare in qualsiasi momento qualunque prodotto descritto nel presente documento senza fornire alcun preavviso. NetApp non si assume alcuna responsabilità circa l'utilizzo dei prodotti o materiali descritti nel presente documento, con l'eccezione di quanto concordato espressamente e per iscritto da NetApp. L'utilizzo o l'acquisto del presente prodotto non comporta il rilascio di una licenza nell'ambito di un qualche diritto di brevetto, marchio commerciale o altro diritto di proprietà intellettuale di NetApp.

Il prodotto descritto in questa guida può essere protetto da uno o più brevetti degli Stati Uniti, esteri o in attesa di approvazione.

LEGENDA PER I DIRITTI SOTTOPOSTI A LIMITAZIONE: l'utilizzo, la duplicazione o la divulgazione da parte degli enti governativi sono soggetti alle limitazioni indicate nel sottoparagrafo (b)(3) della clausola Rights in Technical Data and Computer Software del DFARS 252.227-7013 (FEB 2014) e FAR 52.227-19 (DIC 2007).

I dati contenuti nel presente documento riguardano un articolo commerciale (secondo la definizione data in FAR 2.101) e sono di proprietà di NetApp, Inc. Tutti i dati tecnici e il software NetApp forniti secondo i termini del presente Contratto sono articoli aventi natura commerciale, sviluppati con finanziamenti esclusivamente privati. Il governo statunitense ha una licenza irrevocabile limitata, non esclusiva, non trasferibile, non cedibile, mondiale, per l'utilizzo dei Dati esclusivamente in connessione con e a supporto di un contratto governativo statunitense in base al quale i Dati sono distribuiti. Con la sola esclusione di quanto indicato nel presente documento, i Dati non possono essere utilizzati, divulgati, riprodotti, modificati, visualizzati o mostrati senza la previa approvazione scritta di NetApp, Inc. I diritti di licenza del governo degli Stati Uniti per il Dipartimento della Difesa sono limitati ai diritti identificati nella clausola DFARS 252.227-7015(b) (FEB 2014).

Informazioni sul marchio commerciale

NETAPP, il logo NETAPP e i marchi elencati alla pagina <http://www.netapp.com/TM> sono marchi di NetApp, Inc. Gli altri nomi di aziende e prodotti potrebbero essere marchi dei rispettivi proprietari.