Riferimento

BlueXP classification

NetApp
September 23, 2024

Sommario

Riferimento

Tipi di istanze di classificazione BlueXP supportati

Il software di classificazione BlueXP deve essere eseguito su un host che soddisfi requisiti specifici del sistema operativo, requisiti di RAM, requisiti software e così via. Quando si implementa la classificazione BlueXP nel cloud, si consiglia di utilizzare un sistema con le caratteristiche "grandi" per una funzionalità completa.

È possibile implementare la classificazione BlueXP su un sistema con meno CPU e meno RAM, ma ci sono alcune limitazioni quando si utilizzano questi sistemi meno potenti. "Scopri queste limitazioni".

Nelle tabelle seguenti, se il sistema contrassegnato come "predefinito" non è disponibile nella regione in cui si sta installando la classificazione BlueXP, verrà implementato il sistema successivo nella tabella.

Tipi di istanze AWS

Dimensioni del sistema	Specifiche	Tipo di istanza
Extra large	32 CPU, 128 GB di RAM, 1 TiB SSD GP3	"m6i.8xlarge" (impostazione predefinita)
Grande	16 CPU, 64 GB di RAM, SSD da 500 GiB	"m6i.4xlarge" (impostazione predefinita) m6a.4xlarge m5a.4xlarge m5.4xlarge m4.4xlarge
Medio	8 CPU, 32 GB di RAM, SSD da 200 GiB	"m6i.2xlarge" (impostazione predefinita) m6a.2xlarge m5a.2xlarge m5.2xlarge m4.2xlarge
Piccolo	8 CPU, 16 GB di RAM, SSD da 100 GiB	"c6a.2xlarge" (impostazione predefinita) c5a.2xlarge c5.2xlarge c4.2xlarge

Tipi di istanze di Azure

Dimensioni del sistema	Specifiche	Tipo di istanza
Extra large	32 CPU, 128 GB di RAM, disco OS (2.048 GiB, throughput minimo 250 MB/s) e disco dati (SSD 1 TiB, throughput minimo 750 MB/s)	"Standard_D32_v3" (impostazione predefinita)
Grande	16 CPU, 64 GB di RAM, SSD da 500 GiB	"Standard_D16s_v3" (impostazione predefinita)

Tipi di istanze GCP

Dimensioni del sistema	Specifiche	Tipo di istanza
Grande	16 CPU, 64 GB di RAM, SSD da 500 GiB	"n2-standard-16" (impostazione predefinita) n2d-standard-16 n1-standard-16

Metadati raccolti dalle origini dati

La classificazione BlueXP raccoglie determinati metadati quando si eseguono scansioni di classificazione sui dati provenienti dalle origini dati e dagli ambienti di lavoro. La classificazione BlueXP può accedere alla maggior parte dei metadati necessari per classificare i tuoi dati, ma esistono alcune fonti in cui non siamo in grado di accedere ai dati di cui abbiamo bisogno.

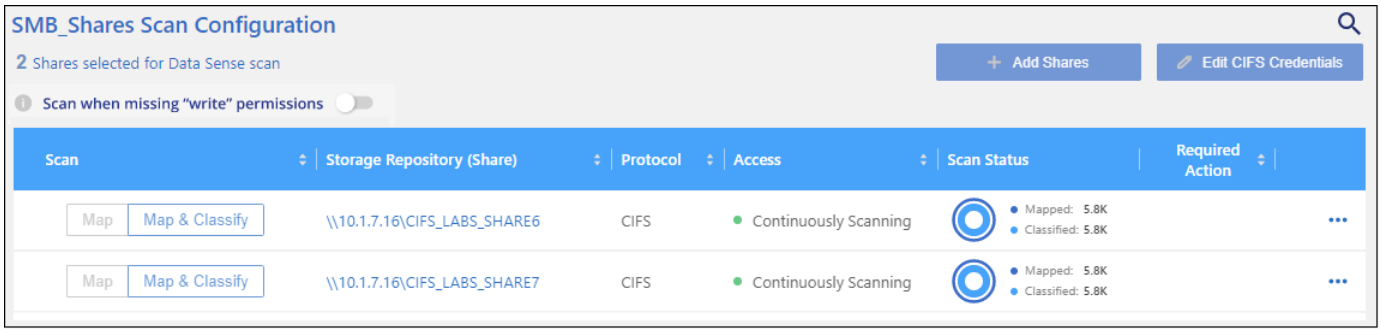
	Metadati	CIFS	NFS
Indicatori di data e ora	*Tempo di creazione*	Disponibile	Non disponibile (non supportato in Linux)
	Ora ultimo accesso	Disponibile	Disponibile
	Ora ultima modifica	Disponibile	Disponibile
Autorizzazioni	*Autorizzazioni aperte*	Se il gruppo "EVERYONE" ha accesso al file, viene considerato "aperto all'organizzazione"	Se "altri" hanno accesso al file, viene considerato "aperto all'organizzazione"
	Accesso utenti/gruppi	Le informazioni relative a utenti e gruppi provengono da LDAP	Non disponibile (gli utenti NFS sono generalmente gestiti localmente sul server, pertanto la stessa persona può avere un UID diverso in ciascun server)

- La classificazione BlueXP non estrae l'ora dell'ultimo accesso dalle origini dei dati del database.
- Le versioni precedenti del sistema operativo Windows (ad esempio, Windows 7 e Windows 8) disattivano per impostazione predefinita la raccolta dell'attributo "ultimo tempo di accesso" perché può influire sulle prestazioni del sistema. Quando questo attributo non viene raccolto, le analisi di classificazione BlueXP basate sull'ultimo tempo di accesso verranno influenzate. È possibile abilitare la raccolta dell'ultimo tempo di accesso su questi sistemi Windows meno recenti, se necessario.

Data e ora dell'ultimo accesso

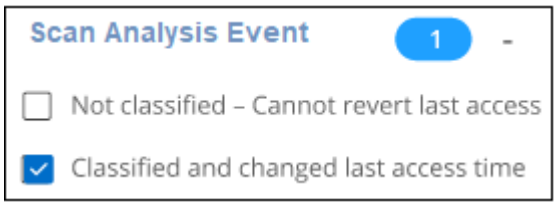
Quando la classificazione BlueXP estrae i dati dalle condivisioni di file, il sistema operativo li considera come utenti che accedono ai dati e modifica di conseguenza il "tempo di accesso ultimo". Dopo la scansione, la classificazione BlueXP tenta di riportare l'ultimo tempo di accesso all'indicatore data e ora originale. Se la classificazione BlueXP non dispone delle autorizzazioni per gli attributi di scrittura in CIFS o di scrittura in NFS, il sistema non può ripristinare l'ultimo orario di accesso all'indicatore data e ora originale. I volumi ONTAP configurati con SnapLock dispongono di autorizzazioni di sola lettura e non possono riportare l'ultimo orario di accesso all'indicatore data e ora originale.

Per impostazione predefinita, se la classificazione BlueXP non dispone di queste autorizzazioni, il sistema non esegue la scansione dei file nei volumi perché la classificazione BlueXP non può riportare l'ultimo tempo di accesso all'indicatore data e ora originale. Tuttavia, se non si ha alcun problema se l'ultimo tempo di accesso viene reimpostato sull'ora originale nei file, è possibile fare clic sull'opzione **scansione quando mancano i permessi di "scrittura attributi"** nella parte inferiore della pagina di configurazione in modo che la classificazione BlueXP scansiona i volumi indipendentemente dalle autorizzazioni.



Questa funzionalità è applicabile ai sistemi ONTAP on-premise, Cloud Volumes ONTAP, Azure NetApp Files, FSX per ONTAP e alle condivisioni di file di terze parti.

Si noti che nella pagina di analisi è presente un filtro denominato *Scan Analysis Event* che consente di visualizzare i file non classificati perché la classificazione BlueXP non ha potuto ripristinare l'ultimo tempo di accesso, Oppure i file classificati anche se la classificazione BlueXP non ha potuto ripristinare l'ultimo tempo di accesso.



Le selezioni dei filtri sono:

- "Non classificato — Impossibile ripristinare l'ultimo tempo di accesso" - Mostra i file che non sono stati classificati a causa di autorizzazioni di scrittura mancanti.
- "Classified and updated last access time" (tempo di accesso ultimo classificato e aggiornato) - Mostra i file classificati e la classificazione BlueXP non è stata in grado di ripristinare l'ultimo tempo di accesso alla data originale. Questo filtro è valido solo per gli ambienti in cui è stata attivata l'OPZIONE **Scan when missing "write attribtributes" permissions**.

Se necessario, è possibile esportare questi risultati in un report in modo da visualizzare i file sottoposti o meno a scansione a causa delle autorizzazioni. "Scopri di più sul Data Investigation Report".

Accedi al sistema di classificazione BlueXP

A volte potrebbe essere necessario accedere al sistema di classificazione BlueXP in modo da poter accedere ai file di log o modificare i file di configurazione.

Quando la classificazione BlueXP è installata su una macchina Linux on-premise o su una macchina Linux implementata nel cloud, puoi accedere direttamente al file di configurazione e allo script.

Quando la classificazione BlueXP viene implementata nel cloud, è necessario eseguire l'SSH nell'istanza di

classificazione BlueXP. Si accede al sistema inserendo l'utente e la password oppure utilizzando la chiave SSH fornita durante l'installazione di BlueXP Connector. Il comando SSH è:

```
 ssh -i <path_to_the_ssh_key> <machine_user>@<datasense_ip>
* <path_to_the_ssh_key> = posizione delle chiavi di autenticazione ssh
* <machine_user>:
```

+
Per AWS: Utilizzare <ec2-user>
Per Azure: Utilizzare l'utente creato per l'istanza di BlueXP
** Per GCP: Utilizzare l'utente creato per l'istanza di BlueXP

- <datasense_ip> = indirizzo IP dell'istanza della macchina virtuale

Nota: Per accedere al sistema nel cloud, è necessario modificare le regole in entrata del gruppo di sicurezza. Per ulteriori informazioni, vedere:

- "Regole del gruppo di sicurezza in AWS"
- "Regole del gruppo di sicurezza in Azure"
- "Regole del firewall in Google Cloud"

API di classificazione BlueXP

Le funzionalità di classificazione BlueXP che sono disponibili tramite l'interfaccia utente web sono anche disponibili tramite l'API Swagger.

Esistono quattro categorie definite nella classificazione BlueXP che corrispondono alle schede dell'interfaccia utente:

- Indagine
- Conformità
- Governance
- Configurazione

Le API nella documentazione di Swagger consentono di cercare, aggregare dati, monitorare le scansioni e creare azioni come copia, spostamento e altro ancora.

Panoramica

L'API consente di eseguire le seguenti funzioni:

- Informazioni sull'esportazione
  - Tutto ciò che è disponibile nell'interfaccia utente può essere esportato tramite l'API (ad eccezione dei report)
  - I dati vengono esportati in formato JSON (semplice da analizzare e inviare ad applicazioni di 3rd parti, come Splunk)
- Creare query utilizzando le istruzioni "AND" e "OR", includere ed escludere informazioni e altro ancora.

Ad esempio, è possibile individuare i file *senza* informazioni personali identificabili (PII) specifiche (funzionalità non disponibile nell'interfaccia utente). È inoltre possibile escludere campi specifici per l'operazione di esportazione.

- Eseguire le azioni
  - Aggiornare le credenziali CIFS
  - Visualizzare e annullare le azioni
  - Eseguire nuovamente la scansione delle directory
  - Esportare i dati

L'API è protetta e utilizza lo stesso metodo di autenticazione dell'interfaccia utente. Le informazioni sull'autenticazione sono disponibili in: https://docs.netapp.com/us-en/bluexp-automation/platform/get_identifiers.html

Accesso al riferimento API Swagger

Per accedere a Swagger, è necessario l'indirizzo IP dell'istanza di classificazione BlueXP. Nel caso di un'implementazione cloud, verrà utilizzato l'indirizzo IP pubblico. Quindi, è necessario accedere a questo endpoint:

https://<classification_ip>/documentazione

Esempio di utilizzo delle API

Nell'esempio seguente viene illustrata una chiamata API per copiare i file.

Richiesta API

Inizialmente, per visualizzare tutti i filtri nella scheda analisi, è necessario ottenere tutti i campi e le opzioni pertinenti per un ambiente di lavoro.

```
curl -X GET "http://{classification_ip}/api/{classification_version}
/search/options?data_mode=ALL_EXTRACTABLE" -H "accept: application/json"
-H "Authorization: Bearer eyJhbGciOiJSUzI1NiIsInR…… " -H "x-agent-id:
hOXsZNvnA5LsthwMILtjL9xZFYBQxAwMclients"
```

Risposta

```
{
  "options": [
    {
      "active_directory_affected": false,
      "data_mode": "ALL_SCANNED",
      "field": "string",
      "is_rulable": true,
      "name": "string",
      "operators": [
        "EQUALS"
      ],
```

```
      "optional_values": [
        {}
      ],
      "secondary": {},
      "server_data": false,
      "type": "TEXT"
    }
  ]
}
{
  "options": [
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "POLICIES",
      "name": "Policies",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "EXTRACTION_STATUS_RANGE",
      "name": "Scan Analysis Status",
      "operators": [
        "IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
      "field": "SCAN_ANALYSIS_ERROR",
      "name": "Scan Analysis Event",
      "operators": [
        "IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
```

```
      "active_directory_affected": false,
      "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
      "field": "PUBLIC_ACCESS",
      "name": "Open Permissions",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": true,
      "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
      "field": "USERS_PERMISSIONS_COUNT_RANGE",
      "name": "Number of Users with Access",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": true,
      "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
      "field": "USER_GROUP_PERMISSIONS",
      "name": "User / Group Permissions",
      "operators": [
        "IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
      "field": "FILE_OWNER",
      "name": "File Owner",
      "operators": [
        "EQUALS",
        "CONTAINS"
      ],
      "server_data": true,
      "type": "TEXT"
    },
```

```
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "ENVIRONMENT_TYPE",
      "name": "Working Environment Type",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "ENVIRONMENT",
      "name": "Working Environment",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_SCANNED",
      "field": "SCAN_TASK",
      "name": "Storage Repository",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
      "field": "FILE_PATH",
      "name": "File / Directory Path",
      "operators": [
        "MULTI_CONTAINS",
        "MULTI_EXCLUDE"
      ],
      "server_data": true,
```

```
    "type": "MULTI_TEXT"
  },
  {
    "active_directory_affected": false,
    "data_mode": "ALL_DASHBOARD_EXTRACTABLE",
    "field": "CATEGORY",
    "name": "Category",
    "operators": [
      "IN",
      "NOT_IN"
    ],
    "server_data": true,
    "type": "SELECT"
  },
  {
    "active_directory_affected": false,
    "data_mode": "ALL_EXTRACTABLE",
    "field": "PATTERN_SENSITIVITY_LEVEL",
    "name": "Sensitivity Level",
    "operators": [
      "IN"
    ],
    "server_data": true,
    "type": "SELECT"
  },
  {
    "active_directory_affected": false,
    "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
    "field": "NUMBER_OF_IDENTIFIERS",
    "name": "Number of identifiers",
    "operators": [
      "IN",
      "NOT_IN"
    ],
    "server_data": true,
    "type": "SELECT"
  },
  {
    "active_directory_affected": false,
    "data_mode": "ALL_EXTRACTABLE",
    "field": "PATTERN_PERSONAL",
    "name": "Personal Data",
    "operators": [
      "IN",
      "NOT_IN"
    ],
```

```
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "PATTERN_SENSITIVE",
      "name": "Sensitive Personal Data",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "DATA_SUBJECT",
      "name": "Data Subject",
      "operators": [
        "EQUALS",
        "CONTAINS"
      ],
      "server_data": true,
      "type": "TEXT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "DIRECTORIES",
      "field": "DIRECTORY_TYPE",
      "name": "Directory Type",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "FILE_TYPE",
      "name": "File Type",
      "operators": [
        "IN",
```

```
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "FILE_SIZE_RANGE",
      "name": "File Size",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
      "field": "FILE_CREATION_RANGE_RETENTION",
      "name": "Created Time",
      "operators": [
        "IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "DISCOVERED_TIME_RANGE",
      "name": "Discovered Time",
      "operators": [
        "IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
      "field": "FILE_LAST_MODIFICATION_RETENTION",
      "name": "Last Modified",
      "operators": [
        "IN"
```

```
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_FILESYSTEM_EXTRACTABLE",
      "field": "FILE_LAST_ACCESS_RANGE_RETENTION",
      "name": "Last Accessed",
      "operators": [
        "IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "FILES",
      "field": "IS_DUPLICATE",
      "name": "Duplicates",
      "operators": [
        "EQUALS",
        "IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "FILES",
      "field": "FILE_HASH",
      "name": "File Hash",
      "operators": [
        "EQUALS",
        "IN"
      ],
      "server_data": true,
      "type": "TEXT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "USER_DEFINED_STATUS",
      "name": "Tags",
      "operators": [
        "IN",
```

```
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    },
    {
      "active_directory_affected": false,
      "data_mode": "ALL_EXTRACTABLE",
      "field": "ASSIGNED_TO",
      "name": "Assigned to",
      "operators": [
        "IN",
        "NOT_IN"
      ],
      "server_data": true,
      "type": "SELECT"
    }
  ]
}
```

Useremo questa risposta nei nostri parametri di richiesta per filtrare i file desiderati che vogliamo copiare.

È possibile applicare un'azione a più elementi. I tipi di azione supportati comprendono: Spostamento, eliminazione, copia, assegnazione a, FlexClone, esportazione di dati, nuova scansione ed etichetta.

Creeremo l'azione di copia:

Richiesta API

Questa API successiva è quella Action API e consente di creare più azioni.

```
curl -X POST "http://
{classification_ip}/api//{classification_version}/actions" -H "accept:
application/json" -H "Authorization: Bearer eyJhbGciOiJSUzI1NiIsInR…… "
-H "x-agent-id: hOXsZNvnA5LsthwMILtjL9xZFYBQxAwMclients " -H "Content-
Type: application/json" -d "{ \"action_type\": \"COPY\", \"data_mode\":
\"FILES\", \"policy_id\": 0, \"request_params\": { destination_nfs_path:
"{ontap_ip}:/{share_name} " },
\"requested_query\":{"condition":"AND","rules":[{"field":"ENVIRONMENT_TYPE
","operator":"IN","value":["ONPREM"]},{"field":"CATEGORY","operator":"IN",
"value":["21"]}]}}"
```

Risposta

La risposta restituirà l'oggetto azione, in modo da poter utilizzare le API Get ed DELETE per ottenere lo stato dell'azione o per annullarla.

```
{
  "action_type": "COPY",
  "creation_time": "2023-08-08T12:37:21.705Z",
  "data_mode": "FILES",
  "end_time": "2023-08-08T12:37:21.705Z",
  "estimated_time_to_complete": 0,
  "id": 0,
  "policy_id": 0,
  "policy_name": "string",
  "priority": 0,
  "request_params": {},
  "requested_query": {},
  "result": {
    "error_message": "string",
    "failed": 0,
    "in_progress": 0,
    "succeeded": 0,
    "total": 0
  },
  "start_time": "2023-08-08T12:37:21.705Z",
  "status": "QUEUED",
  "title": "string",
  "user_id": "string"
}
```

Informazioni sul copyright

Informazioni sul marchio commerciale