Referenz

NetApp Data Classification

NetApp
January 14, 2026

Inhalt

Referenz

Unterstützte NetApp Data Classification Instanztypen

Die NetApp Data Classification -Software muss auf einem Host ausgeführt werden, der bestimmte Betriebssystemanforderungen, RAM-Anforderungen, Softwareanforderungen usw. erfüllt. Wenn Sie die Datenklassifizierung in der Cloud bereitstellen, empfehlen wir Ihnen, für die volle Funktionalität ein System mit der Eigenschaft „groß“ zu verwenden.

Sie können die Datenklassifizierung auf einem System mit weniger CPUs und weniger RAM bereitstellen, bei der Verwendung dieser weniger leistungsstarken Systeme gibt es jedoch einige Einschränkungen. "Erfahren Sie mehr über diese Einschränkungen" .

Wenn in den folgenden Tabellen das als „Standard“ gekennzeichnete System in der Region, in der Sie Data Classification installieren, nicht verfügbar ist, wird das nächste System in der Tabelle bereitgestellt.

AWS-Instanztypen

Systemgröße	Technische Daten	Instanztyp
Extragroß	32 CPUs, 128 GB RAM, 1 TiB gp3 SSD	"m6i.8xlarge"(Standard)
Groß	16 CPUs, 64 GB RAM, 500 GiB SSD	"m6i.4xlarge"(Standard) m6a.4xlarge m5a.4xlarge m5.4xlarge m4.4xlarge
Medium	8 CPUs, 32 GB RAM, 200 GiB SSD	"m6i.2xlarge"(Standard) m6a.2xlarge m5a.2xlarge m5.2xlarge m4.2xlarge
Klein	8 CPUs, 16 GB RAM, 100 GiB SSD	"c6a.2xlarge"(Standard) c5a.2xlarge c5.2xlarge c4.2xlarge

Azure-Instanztypen

Systemgröße	Technische Daten	Instanztyp
Extragroß	32 CPUs, 128 GB RAM, Betriebssystemfestplatte (2.048 GiB, min. 250 MB/s Durchsatz) und Datenfestplatte (1 TiB SSD, min. 750 MB/s Durchsatz)	"Standard_D32_v3"(Standard)
Groß	16 CPUs, 64 GB RAM, 500 GiB SSD	"Standard_D16s_v3"(Standard)

GCP-Instanztypen

Systemgröße	Technische Daten	Instanztyp
Groß	16 CPUs, 64 GB RAM, 500 GiB SSD	"n2-Standard-16"(Standard) n2d-standard-16 n1-standard-16

Aus Datenquellen in der NetApp Data Classification erfasste Metadaten

NetApp Data Classification sammelt bestimmte Metadaten, wenn Klassifizierungsscans für die Daten aus Ihren Datenquellen und Systemen durchgeführt werden. Die Datenklassifizierung kann auf die meisten Metadaten zugreifen, die wir zur Klassifizierung Ihrer Daten benötigen. Es gibt jedoch einige Quellen, bei denen wir nicht auf die benötigten Daten zugreifen können.

	Metadaten	CIFS	NFS
Zeitstempel	*Erstellungszeit*	Verfügbar	Nicht verfügbar (wird unter Linux nicht unterstützt)
	Letzter Zugriffszeitpunkt	Verfügbar	Verfügbar
	Letzte Änderungszeit	Verfügbar	Verfügbar
Berechtigungen	*Öffnen Sie Berechtigungen*	Wenn die Gruppe „JEDER“ Zugriff auf die Datei hat, gilt sie als „Für die Organisation offen“.	Wenn „Andere“ Zugriff auf die Datei haben, gilt sie als „Für die Organisation offen“.
	Benutzer-/Gruppenzugriff	Benutzer- und Gruppeninformationen werden aus LDAP übernommen	Nicht verfügbar (NFS-Benutzer werden normalerweise lokal auf dem Server verwaltet, daher kann dieselbe Person auf jedem Server eine andere UID haben)

- Die Datenklassifizierung extrahiert nicht die „Zeit des letzten Zugriffs“ aus den Datenbankdatenquellen.
- Ältere Versionen des Windows-Betriebssystems (z. B. Windows 7 und Windows 8) deaktivieren die Erfassung des Attributs „Zeit des letzten Zugriffs“ standardmäßig, da dies die Systemleistung beeinträchtigen kann. Wenn dieses Attribut nicht erfasst wird, sind Datenklassifizierungsanalysen, die auf der „Zeit des letzten Zugriffs“ basieren, davon betroffen. Sie können die Erfassung der letzten Zugriffszeit auf diesen älteren Windows-Systemen bei Bedarf aktivieren.

Zeitstempel des letzten Zugriffs

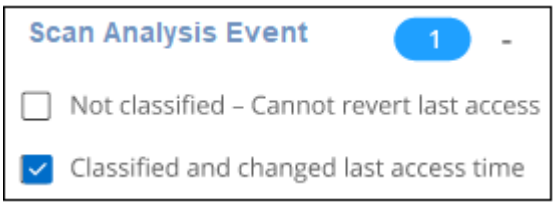
Wenn die Datenklassifizierung Daten aus Dateifreigaben extrahiert, betrachtet das Betriebssystem dies als Zugriff auf die Daten und ändert die „letzte Zugriffszeit“ entsprechend. Nach dem Scannen versucht die Datenklassifizierung, die letzte Zugriffszeit auf den ursprünglichen Zeitstempel zurückzusetzen. Wenn die Datenklassifizierung keine Schreibberechtigungen für Attribute in CIFS oder Schreibberechtigungen in NFS hat, kann das System die letzte Zugriffszeit nicht auf den ursprünglichen Zeitstempel zurücksetzen. Mit SnapLock konfigurierte ONTAP Volumes verfügen über schreibgeschützte Berechtigungen und können den letzten Zugriffszeitpunkt auch nicht auf den ursprünglichen Zeitstempel zurücksetzen.

Wenn Data Classification nicht über diese Berechtigungen verfügt, scannt das System diese Dateien in Ihren Volumes standardmäßig nicht, da Data Classification die „letzte Zugriffszeit“ nicht auf den ursprünglichen Zeitstempel zurücksetzen kann. Wenn es Ihnen jedoch egal ist, ob die letzte Zugriffszeit in Ihren Dateien auf

die ursprüngliche Zeit zurückgesetzt wird, können Sie unten auf der Konfigurationsseite den Schalter **Scannen, wenn Berechtigungen zum Schreiben von Attributen fehlen** auswählen, damit die Datenklassifizierung die Volumes unabhängig von den Berechtigungen scannt.

Diese Funktionalität ist auf lokale ONTAP Systeme, Cloud Volumes ONTAP, Azure NetApp Files, Amazon FSx for NetApp ONTAP Management und Dateifreigaben von Drittanbietern anwendbar.

Auf der Untersuchungsseite gibt es einen Filter namens „Scan-Analyse-Ereignis“, mit dem Sie entweder die Dateien anzeigen können, die nicht klassifiziert wurden, weil die Datenklassifizierung den letzten Zugriffszeitpunkt nicht zurücksetzen konnte, oder die Dateien, die klassifiziert wurden, obwohl die Datenklassifizierung den letzten Zugriffszeitpunkt nicht zurücksetzen konnte.



Die Filterauswahl ist:

- „Nicht klassifiziert – Letzter Zugriffszeitpunkt kann nicht zurückgesetzt werden“ – Hier werden die Dateien angezeigt, die aufgrund fehlender Schreibberechtigungen nicht klassifiziert wurden.
- „Klassifiziert und letzte Zugriffszeit aktualisiert“ – Hier werden die Dateien angezeigt, die klassifiziert wurden und bei denen die Datenklassifizierung die letzte Zugriffszeit nicht auf das ursprüngliche Datum zurücksetzen konnte. Dieser Filter ist nur für Umgebungen relevant, in denen Sie **Scannen, wenn Berechtigungen zum Schreiben von Attributen fehlen** aktiviert haben.

Bei Bedarf können Sie diese Ergebnisse in einen Bericht exportieren, sodass Sie sehen können, welche Dateien aufgrund von Berechtigungen gescannt werden und welche nicht. "Erfahren Sie mehr über Data Investigation-Berichte".

Melden Sie sich beim NetApp Data Classification System an

Sie müssen sich beim NetApp Data Classification System anmelden, damit Sie auf Protokolldateien zugreifen oder Konfigurationsdateien bearbeiten können.

Wenn Data Classification auf einem Linux-Computer bei Ihnen vor Ort oder auf einem Linux-Computer installiert ist, den Sie in der Cloud bereitgestellt haben, können Sie direkt auf die Konfigurationsdatei und das Skript zugreifen.

Wenn die Datenklassifizierung in der Cloud bereitgestellt wird, müssen Sie per SSH auf die Datenklassifizierungsinstanz zugreifen. Sie können per SSH auf das System zugreifen, indem Sie den

Benutzer und das Kennwort eingeben oder den SSH-Schlüssel verwenden, den Sie während der Installation des Konsolenagenten angegeben haben. Der SSH-Befehl lautet:

```
ssh -i <path_to_the_ssh_key> <machine_user>@<datasense_ip>
```

- `<path_to_the_ssh_key>`= Speicherort der SSH-Authentifizierungsschlüssel
- `<machine_user>`:
  - Für AWS: Verwenden Sie den <ec2-user>
  - Für Azure: Verwenden Sie den für die Konsoleninstanz erstellten Benutzer
  - Für GCP: Verwenden Sie den für die Konsoleninstanz erstellten Benutzer
- `<datasense_ip>`= IP-Adresse der virtuellen Maschineninstanz

Sie müssen die eingehenden Regeln der Sicherheitsgruppe ändern, um auf das System in der Cloud zuzugreifen. Weitere Informationen finden Sie unter:

- "Sicherheitsgruppenregeln in AWS"
- "Sicherheitsgruppenregeln in Azure"
- "Firewall-Regeln in Google Cloud"

NetApp Data Classification APIs

Die über die Web-Benutzeroberfläche verfügbaren NetApp Data Classification sind auch über die REST-API verfügbar.

Innerhalb der Datenklassifizierung sind vier Kategorien definiert, die den Registerkarten in der Benutzeroberfläche entsprechen:

- Untersuchung
- Einhaltung
- Führung
- Konfiguration

Mit den APIs in der Swagger-Dokumentation können Sie suchen, Daten aggregieren, Ihre Scans verfolgen und Aktionen wie Kopieren, Verschieben und Löschen ausführen.

Überblick

Mit der API können Sie die folgenden Funktionen ausführen:

- Exportinformationen
  - Alles, was in der Benutzeroberfläche verfügbar ist, kann über die API exportiert werden (mit Ausnahme von Berichten).
  - Daten werden im JSON-Format exportiert (einfach zu analysieren und an Anwendungen von Drittanbietern wie Splunk weiterzuleiten).
- Erstellen Sie Abfragen mit „UND“- und „ODER“-Anweisungen, schließen Sie Informationen ein und aus und mehr.

Sie können beispielsweise Dateien *ohne* spezifische personenbezogene Daten (PII) suchen (Funktion in der Benutzeroberfläche nicht verfügbar). Sie können auch bestimmte Felder vom Exportvorgang ausschließen.

- Aktionen ausführen
  - CIFS-Anmeldeinformationen aktualisieren
  - Aktionen anzeigen und abbrechen
  - Verzeichnisse erneut scannen
  - Daten exportieren

Die API ist sicher und verwendet dieselbe Authentifizierungsmethode wie die Benutzeroberfläche. Informationen zur Authentifizierung finden Sie im"REST API-Dokumentation" .

Zugriff auf die Swagger-API-Referenz

Um auf Swagger zuzugreifen, benötigen Sie die IP-Adresse Ihrer Datenklassifizierungsinstanz. Bei einer Cloud-Bereitstellung verwenden Sie die öffentliche IP-Adresse. Dann müssen Sie zu diesem Endpunkt gelangen:

https://<Klassifizierungs-IP>/documentation

Beispiel für die Verwendung der APIs

Das folgende Beispiel zeigt einen API-Aufruf zum Kopieren von Dateien.

API-Anforderung

Sie müssen zunächst alle relevanten Felder und Optionen für ein System abrufen, um alle Filter auf der Registerkarte „Untersuchung“ anzuzeigen.

```
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"
```

Antwort

```
{
  "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": "system-type",
    "operators": [
      "IN",
      "NOT_IN"
    ],
    "server_data": true,
    "type": "SELECT"
  },
  {
    "active_directory_affected": false,
    "data_mode": "ALL_EXTRACTABLE",
    "field": "ENVIRONMENT",
    "name": "system",
    "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"
    }
  ]
}
```

Wir werden diese Antwort in unseren Anforderungsparametern verwenden, um die gewünschten Dateien zu filtern, die wir kopieren möchten.

Sie können eine Aktion auf mehrere Elemente anwenden. Zu den unterstützten Aktionstypen gehören: Verschieben, Löschen und Kopieren.

Wir erstellen die Kopieraktion:

API-Anforderung

Diese nächste API ist die Aktions-API und ermöglicht Ihnen die Erstellung mehrerer Aktionen.

```
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"]}]}}"
```

Antwort

Die Antwort gibt das Aktionsobjekt zurück, sodass Sie die APIs zum Abrufen und Löschen verwenden können, um den Status der Aktion abzurufen oder sie abzubrechen.

```
{
  "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"
}
```