参照
NetApp Data Classification

NetApp
January 14, 2026

目次

参照

サポートされている**NetApp Data Classification**インスタンスタイプ

NetApp Data Classificationソフトウェアは、特定のオペレーティング システム要件、RAM 要件、ソフトウェア要件などを満たすホスト上で実行する必要があります。クラウドでデータ分類を展開する場合は、完全な機能を実現するために「大規模」特性を持つシステムを使用することをお勧めします。

CPU 数と RAM 数が少ないシステムでもデータ分類を展開できますが、これらの非力なシステムを使用する場合はいくつかの制限があります。"これらの制限について学ぶ" 。

次の表で、「デフォルト」とマークされているシステムが、Data Classification をインストールするリージョンで使用できない場合は、表の次のシステムが展開されます。

AWSインスタンスタイプ

システムサイズ	仕様	インスタンスタイプ
特大	32 個の CPU、128 GB の RAM、1 TiB の gp3 SSD	"m6i.8xlarge"（デフォルト）
大規模	16 CPU、64 GB RAM、500 GiB SSD	"m6i.4xlarge"(デフォルト) m6a.4xlarge m5a.4xlarge m5.4xlarge m4.4xlarge
中	8 CPU、32 GB RAM、200 GiB SSD	"m6i.2xlarge"(デフォルト) m6a.2xlarge m5a.2xlarge m5.2xlarge m4.2xlarge
小規模	8 CPU、16 GB RAM、100 GiB SSD	"c6a.2xlarge"(デフォルト) c5a.2xlarge c5.2xlarge c4.2xlarge

Azureインスタンスの種類

システムサイズ	仕様	インスタンスタイプ
特大	32 個の CPU、128 GB の RAM、OS ディスク (2,048 GiB、最小 250 MB/秒のスループット)、およびデータ ディスク (1 TiB SSD、最小 750 MB/秒のスループット)	"標準_D32_v3"（デフォルト）
大規模	16 CPU、64 GB RAM、500 GiB SSD	"標準_D16s_v3"（デフォルト）

GCPインスタンスタイプ

システムサイズ	仕様	インスタンスタイプ
大規模	16 CPU、64 GB RAM、500 GiB SSD	"n2-標準-16"(デフォルト) n2d-standard-16 n1-standard-16

NetApp Data Classificationのデータソースから収集されたメタデータ

NetApp Data Classification は、データ ソースおよびシステムからのデータの分類スキャンを実行するときに、特定のメタデータを収集します。データ分類では、データを分類するために必要なメタデータのほとんどにアクセスできますが、必要なデータにアクセスできないソースもいくつかあります。

	メタデータ	CIFS	NFS
タイムスタンプ	作成時間	利用可能	利用できません（Linuxではサポートされていません）
	最終アクセス時間	利用可能	利用可能
	最終更新日時	利用可能	利用可能
権限	オープン権限	「EVERYONE」グループがファイルにアクセスできる場合、そのファイルは「組織に公開」されているとみなされます。	「その他」がファイルにアクセスできる場合、そのファイルは「組織に公開」されているとみなされます。
	ユーザー/グループアクセス	ユーザーとグループの情報はLDAPから取得されます	利用できません (NFS ユーザーは通常、サーバー上でローカルに管理されるため、同じユーザーが各サーバーで異なる UID を持つことができます)

- データ分類では、データベース データ ソースから「最終アクセス時刻」を抽出しません。
- 古いバージョンの Windows OS (Windows 7 や Windows 8 など) では、システム パフォーマンスに影響を及ぼす可能性があるため、「最終アクセス時刻」属性の収集がデフォルトで無効になっています。この属性が収集されない場合、「最終アクセス時刻」に基づくデータ分類分析に影響が出ます。必要に応じて、これらの古い Windows システムでの最終アクセス時刻の収集を有効にすることができます。

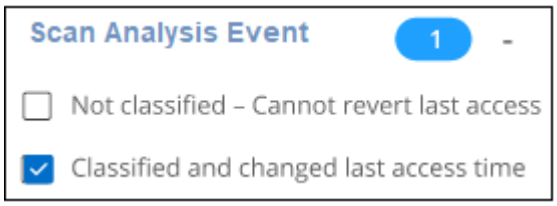
最終アクセス時間のタイムスタンプ

データ分類がファイル共有からデータを抽出すると、オペレーティング システムはそれをデータへのアクセスと見なし、それに応じて「最終アクセス時刻」を変更します。スキャン後、データ分類は最終アクセス時刻を元のタイムスタンプに戻そうとします。データ分類に CIFS での属性書き込み権限または NFS での書き込み権限がない場合、システムは最終アクセス時刻を元のタイムスタンプに戻すことができません。 SnapLockが設定されたONTAPボリュームには読み取り専用権限があり、最終アクセス時刻を元のタイムスタンプに戻すこともできません。

デフォルトでは、データ分類にこれらの権限がない場合、データ分類は「最終アクセス時刻」を元のタイムスタンプに戻すことができないため、システムはボリューム内のこれらのファイルをスキャンしません。ただし、ファイルの最終アクセス時刻が元の時刻にリセットされても構わない場合は、構成ページの下部にある「属性の書き込み」権限がない場合にスキャンする スイッチを選択して、権限に関係なくデータ分類がボリュームをスキャンするようにすることができます。

この機能は、オンプレミスのONTAPシステム、 Cloud Volumes ONTAP、 Azure NetApp Files、 Amazon FSx for NetApp ONTAP管理、およびサードパーティのファイル共有に適用できます。

調査ページには、「スキャン分析イベント」というフィルターがあり、これを使用すると、データ分類で最終アクセス時刻を戻せなかったために分類されなかったファイル、またはデータ分類で最終アクセス時刻を戻せなかったにもかかわらず分類されたファイルを表示できます。



フィルターの選択肢は次のとおりです。

- 「未分類 - 最終アクセス時間を戻すことができません」 - 書き込み権限がないため分類されなかったファイルが表示されます。
- 「分類され、更新された最終アクセス時刻」 - 分類されたファイルが表示され、データ分類では最終アクセス時刻を元の日付にリセットできませんでした。このフィルターは、*「属性の書き込み」権限がない場合にスキャン*をオンにした環境にのみ関連します。

必要に応じて、これらの結果をレポートにエクスポートして、権限によりスキャンされているファイルとスキャンされていないファイルを確認できます。"データ調査レポートの詳細"。

NetApp Data Classificationシステムにログインする

ログ ファイルにアクセスしたり、構成ファイルを編集したりするには、 NetApp Data Classificationシステムにログインする必要があります。

Data Classification がオンプレミスの Linux マシンまたはクラウドに展開した Linux マシンにインストールされている場合は、構成ファイルとスクリプトに直接アクセスできます。

Data Classification をクラウドにデプロイする場合は、Data Classification インスタンスに SSH で接続する必要があります。ユーザー名とパスワードを入力するか、コンソール エージェントのインストール時に指定し

た SSH キーを使用して、システムに SSH 接続します。 SSH コマンドは次のとおりです。

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

- `<path_to_the_ssh_key>`= SSH認証キーの場所
- `<machine_user>`:
  - AWSの場合: <ec2-user> を使用します
  - Azureの場合: コンソールインスタンス用に作成されたユーザーを使用します
  - GCPの場合: コンソールインスタンス用に作成されたユーザーを使用します
- `<datasense_ip>`= 仮想マシンインスタンスのIPアドレス

クラウド内のシステムにアクセスするには、セキュリティ グループの受信ルールを変更する必要があります。詳細については、以下を参照してください。

- "AWSのセキュリティグループルール"
- "Azure のセキュリティ グループ ルール"
- "Google Cloud のファイアウォール ルール"

NetApp Data ClassificationAPI

Web UI を通じて利用できるNetApp Data Classification機能は、REST API を通じても利用できます。

データ分類内には、UI のタブに対応する 4 つのカテゴリが定義されています。

- 調査
- コンプライアンス
- ガバナンス
- 構成

Swagger ドキュメントの API を使用すると、検索、データの集約、スキャンの追跡、コピー、移動、削除などのアクションを実行できます。

概要

API を使用すると、次の機能を実行できます。

- 輸出情報
  - UI で利用できるものはすべて API 経由でエクスポートできます (レポートを除く)
  - データは JSON 形式でエクスポートされます (簡単に解析でき、Splunk などのサードパーティ アプリケーションにプッシュできます)
- 「AND」および「OR」ステートメントを使用してクエリを作成し、情報を含めたり除外したりします。

  たとえば、特定の個人識別情報 (PII) を含まないファイルを見つけることができます (UI では機能は使用で

きません)。エクスポート操作から特定のフィールドを除外することもできます。

- アクションを実行する
  - CIFSクレデンシャルの更新
  - アクションの表示とキャンセル
  - ディレクトリを再スキャン
  - データをエクスポートする

API は安全で、UI と同じ認証方法を使用します。認証に関する情報は、"REST API ドキュメント" 。

Swagger APIリファレンスへのアクセス

Swagger にアクセスするには、データ分類インスタンスの IP アドレスが必要です。クラウド展開の場合は、パブリック IP アドレスを使用します。次に、このエンドポイントにアクセスする必要があります。

https://<分類IP>/documentation

APIの使用例

次の例は、ファイルをコピーするための API 呼び出しを示しています。

API 要求

調査タブですべてのフィルターを表示するには、最初にシステムの関連するフィールドとオプションをすべて取得する必要があります。

```
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"
```

応答

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

リクエストパラメータでその応答を使用して、コピーする目的のファイルをフィルタリングします。

複数のアイテムにアクションを適用できます。サポートされているアクション タイプには、移動、削除、コピーなどがあります。

コピーアクションを作成します。

API 要求

次の API はアクション API であり、複数のアクションを作成できます。

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

応答

応答ではアクション オブジェクトが返されるため、get および delete API を使用してアクションのステータスを取得したり、アクションをキャンセルしたりできます。

```
{
  "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"
}
```