Referência

NetApp Data Classification

NetApp
January 14, 2026

Índice

Referência

Tipos de instância de NetApp Data Classification com suporte

O software NetApp Data Classification deve ser executado em um host que atenda a requisitos específicos do sistema operacional, requisitos de RAM, requisitos de software e assim por diante. Ao implantar a Classificação de Dados na nuvem, recomendamos que você use um sistema com características "grandes" para obter funcionalidade completa.

Você pode implantar a Classificação de Dados em um sistema com menos CPUs e menos RAM, mas há algumas limitações ao usar esses sistemas menos potentes. "Saiba mais sobre essas limitações" .

Nas tabelas a seguir, se o sistema marcado como "padrão" não estiver disponível na região onde você está instalando o Data Classification, o próximo sistema na tabela será implantado.

Tipos de instância da AWS

Tamanho do sistema	Especificações	Tipo de instância
Extra grande	32 CPUs, 128 GB de RAM, 1 TiB gp3 SSD	"m6i.8xlarge"(padrão)
Grande	16 CPUs, 64 GB de RAM, SSD de 500 GiB	"m6i.4xlarge"(padrão) m6a.4xlarge m5a.4xlarge m5.4xlarge m4.4xlarge
Médio	8 CPUs, 32 GB de RAM, SSD de 200 GiB	"m6i.2xlarge"(padrão) m6a.2xlarge m5a.2xlarge m5.2xlarge m4.2xlarge
Pequeno	8 CPUs, 16 GB de RAM, SSD de 100 GiB	"c6a.2xlarge"(padrão) c5a.2xlarge c5.2xlarge c4.2xlarge

Tipos de instância do Azure

Tamanho do sistema	Especificações	Tipo de instância
Extra grande	32 CPUs, 128 GB de RAM, disco do sistema operacional (2.048 GiB, taxa de transferência mínima de 250 MB/s) e disco de dados (SSD de 1 TiB, taxa de transferência mínima de 750 MB/s)	"Standard_D32_v3"(padrão)
Grande	16 CPUs, 64 GB de RAM, SSD de 500 GiB	"Standard_D16s_v3"(padrão)

Tipos de instância do GCP

Tamanho do sistema	Especificações	Tipo de instância
Grande	16 CPUs, 64 GB de RAM, SSD de 500 GiB	"n2-padrão-16"(padrão) n2d-standard-16 n1-standard-16

Metadados coletados de fontes de dados na NetApp Data Classification

A NetApp Data Classification coleta determinados metadados ao executar verificações de classificação nos dados de suas fontes de dados e sistemas. A Classificação de Dados pode acessar a maioria dos metadados necessários para classificar seus dados, mas há algumas fontes nas quais não conseguimos acessar os dados necessários.

	Metadados	CIFS	NFS
Carimbos de tempo	*Tempo de criação*	Disponível	Não disponível (sem suporte no Linux)
	Último horário de acesso	Disponível	Disponível
	Hora da última modificação	Disponível	Disponível
Permissões	*Permissões abertas*	Se o grupo "TODOS" tiver acesso ao arquivo, ele será considerado "Aberto à organização".	Se "Outros" tiver acesso ao arquivo, ele será considerado "Aberto à organização".
	Acesso de usuários/grupos	As informações de usuários e grupos são obtidas do LDAP	Não disponível (os usuários do NFS geralmente são gerenciados localmente no servidor, portanto, o mesmo indivíduo pode ter um UID diferente em cada servidor)

- A Classificação de Dados não extrai o "último horário de acesso" das fontes de dados do banco de dados.
- Versões mais antigas do sistema operacional Windows (por exemplo, Windows 7 e Windows 8) desabilitam a coleta do atributo "hora do último acesso" por padrão, pois isso pode afetar o desempenho do sistema. Quando esse atributo não for coletado, as análises de Classificação de Dados baseadas no "último horário de acesso" serão afetadas. Você pode habilitar a coleta do último horário de acesso nesses sistemas Windows mais antigos, se necessário.

Carimbo de data e hora do último acesso

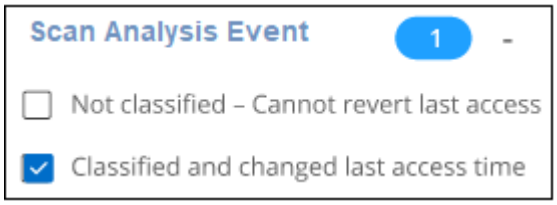
Quando a Classificação de Dados extrai dados de compartilhamentos de arquivos, o sistema operacional considera que está acessando os dados e altera o "último horário de acesso" de acordo. Após a digitalização, a Classificação de Dados tenta reverter o último horário de acesso para o registro de data e hora original. Se a Classificação de Dados não tiver permissões de gravação de atributos no CIFS ou permissões de gravação no NFS, o sistema não poderá reverter o último horário de acesso para o registro de data e hora original. Os volumes ONTAP configurados com SnapLock têm permissões somente leitura e também não podem reverter o último horário de acesso para o registro de data e hora original.

Por padrão, se a Classificação de Dados não tiver essas permissões, o sistema não verificará esses arquivos em seus volumes porque a Classificação de Dados não pode reverter o "último horário de acesso" para o registro de data e hora original. No entanto, se você não se importa se o último horário de acesso será redefinido para o horário original em seus arquivos, você pode selecionar a opção **Verificar quando faltarem permissões de "atributos de gravação"** na parte inferior da página Configuração para que a Classificação

de Dados verifique os volumes independentemente das permissões.

Essa funcionalidade é aplicável a sistemas ONTAP locais, Cloud Volumes ONTAP, Azure NetApp Files, Amazon FSx for NetApp ONTAP e compartilhamentos de arquivos de terceiros.

Há um filtro na página Investigação chamado *Evento de Análise de Verificação* que permite exibir os arquivos que não foram classificados porque a Classificação de Dados não conseguiu reverter o último horário de acesso ou os arquivos que foram classificados mesmo que a Classificação de Dados não tenha conseguido reverter o último horário de acesso.



As seleções de filtros são:

- "Não classificado — Não é possível reverter o último horário de acesso" - Isso mostra os arquivos que não foram classificados devido à falta de permissões de gravação.
- "Último horário de acesso classificado e atualizado" - Mostra os arquivos que foram classificados e a Classificação de Dados não conseguiu redefinir o último horário de acesso para a data original. Este filtro é relevante somente para ambientes em que você ativou a opção **Verificar quando faltarem permissões de "gravação de atributos"**.

Se necessário, você pode exportar esses resultados para um relatório para ver quais arquivos estão ou não sendo verificados devido às permissões. "Saiba mais sobre relatórios de investigação de dados".

Efetue login no sistema de NetApp Data Classification

Você precisa fazer login no sistema de NetApp Data Classification para poder acessar arquivos de log ou editar arquivos de configuração.

Quando o Data Classification é instalado em uma máquina Linux em suas instalações ou em uma máquina Linux implantada na nuvem, você pode acessar o arquivo de configuração e o script diretamente.

Quando a Classificação de Dados é implantada na nuvem, você precisa fazer SSH para a instância da Classificação de Dados. Faça login no sistema via SSH inserindo o usuário e a senha ou usando a chave SSH fornecida durante a instalação do agente do Console. O comando SSH é:

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

- `<path_to_the_ssh_key>`= localização das chaves de autenticação ssh
- `<machine_user>`:
  - Para AWS: use o <ec2-user>
  - Para o Azure: use o usuário criado para a instância do Console
  - Para GCP: use o usuário criado para a instância do Console
- `<datasense_ip>`= Endereço IP da instância da máquina virtual

Você precisa modificar as regras de entrada do grupo de segurança para acessar o sistema na nuvem. Para mais detalhes, consulte:

- "Regras de grupo de segurança na AWS"
- "Regras de grupo de segurança no Azure"
- "Regras de firewall no Google Cloud"

APIs de NetApp Data Classification

Os recursos de NetApp Data Classification disponíveis por meio da interface do usuário da Web também estão disponíveis por meio da API REST.

Há quatro categorias definidas na Classificação de Dados que correspondem às guias na IU:

- Investigação
- Conformidade
- Governança
- Configuração

As APIs na documentação do Swagger permitem que você pesquise, agregue dados, rastreie suas verificações e execute ações como copiar, mover e excluir.

Visão geral

A API permite que você execute as seguintes funções:

- Informações de exportação
  - Tudo o que está disponível na IU pode ser exportado por meio da API (com exceção de relatórios)
  - Os dados são exportados em formato JSON (fácil de analisar e enviar para aplicativos de terceiros, como o Splunk)
- Crie consultas usando instruções "AND" e "OR", inclua e exclua informações e muito mais.

  Por exemplo, você pode localizar arquivos *sem* Informações Pessoais Identificáveis (PII) específicas (funcionalidade não disponível na interface do usuário). Você também pode excluir campos específicos para a operação de exportação.

- Executar ações
  - Atualizar credenciais CIFS
  - Visualizar e cancelar ações

- Verifique novamente os diretórios
- Exportar dados

A API é segura e usa o mesmo método de autenticação da interface do usuário. Você pode encontrar informações sobre a autenticação no"Documentação do REST API" .

Acessando a referência da API do Swagger

Para acessar o Swagger, você precisará do endereço IP da sua instância de Classificação de Dados. No caso de uma implantação na nuvem, você usará o endereço IP público. Então você precisará acessar este endpoint:

https://<ip_de_classificação>/documentação

Exemplo usando as APIs

O exemplo a seguir mostra uma chamada de API para copiar arquivos.

Solicitação de API

Inicialmente, você precisará obter todos os campos e opções relevantes para que um sistema visualize todos os filtros na guia de investigação.

```
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"
```

Resposta

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

Usaremos essa resposta em nossos parâmetros de solicitação para filtrar os arquivos desejados que queremos copiar.

Você pode aplicar uma ação em vários itens. Os tipos de ação suportados incluem: mover, excluir e copiar.

Criaremos a ação de cópia:

Solicitação de API

A próxima API é a API de ação e permite que você crie múltiplas ações.

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

Resposta

A resposta retornará o objeto de ação, então você pode usar as APIs get e delete para obter o status da ação ou cancelá-la.

```
{
  "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"
}
```

Informações sobre direitos autorais

Informações sobre marcas comerciais