참조하십시오

BlueXP classification

NetApp
April 03, 2024

목차

참조하십시오

지원되는 BlueXP 분류 인스턴스 유형

BlueXP 분류 소프트웨어는 특정 운영 체제 요구 사항, RAM 요구 사항, 소프트웨어 요구 사항 등을 충족하는 호스트에서 실행해야 합니다. 클라우드에 BlueXP 분류를 배포할 때는 모든 기능을 위해 "큰" 특성을 가진 시스템을 사용하는 것이 좋습니다.

CPU가 적고 RAM이 적은 시스템에 BlueXP 분류를 배포할 수 있지만 이러한 덜 강력한 시스템을 사용할 때는 몇 가지 제약이 있습니다. "이러한 제한 사항에 대해 자세히 알아보십시오".

다음 표에서 "기본값"으로 표시된 시스템을 BlueXP 분류를 설치하는 지역에서 사용할 수 없는 경우 표의 다음 시스템이 배포됩니다.

AWS 인스턴스 유형

시스템 크기	사양	인스턴스 유형
매우 크게	32개의 CPU, 128GB RAM, 1TiB GP3 SSD	"m6i.8x 대형" (기본값)
대형	CPU 16개, 64GB RAM, 500GiB SSD	"m6i.4xLarge" (기본값) m6a.4xLarge m5a.4xLarge m5.4xLarge m4.4xLarge
중간	CPU 8개, 32GB RAM, 200GiB SSD	"m6i.2xLarge" (기본값) m6a.2xLarge m5a.2xLarge m5.2xLarge m4.2xLarge
작은 크기	CPU 8개, 16GB RAM, 100GiB SSD	"c6a.2xLarge" (기본값) C5A.2xLarge c5.2xLarge c4.2xLarge

Azure 인스턴스 유형

시스템 크기	사양	인스턴스 유형
매우 크게	32개의 CPU, 128GB RAM, OS 디스크(2,048GiB, 최소 250MB/s 처리량) 및 데이터 디스크(1TiB SSD, 최소 750MB/s 처리량)	"Standard_D32_v3" (기본값)
대형	CPU 16개, 64GB RAM, 500GiB SSD	"standard_d16s_v3" (기본값)

GCP 인스턴스 유형

시스템 크기	사양	인스턴스 유형
대형	CPU 16개, 64GB RAM, 500GiB SSD	"N2-표준-16" (기본값) n2d-standard-16 n1-standard-16

데이터 소스에서 수집된 메타데이터

BlueXP 분류는 데이터 소스 및 작업 환경의 데이터에 대한 분류 검사를 수행할 때 특정 메타데이터를 수집합니다. BlueXP 분류는 데이터를 분류하는 데 필요한 대부분의 메타데이터에 액세스할 수 있지만 필요한 데이터에 액세스할 수 없는 소스가 있습니다.

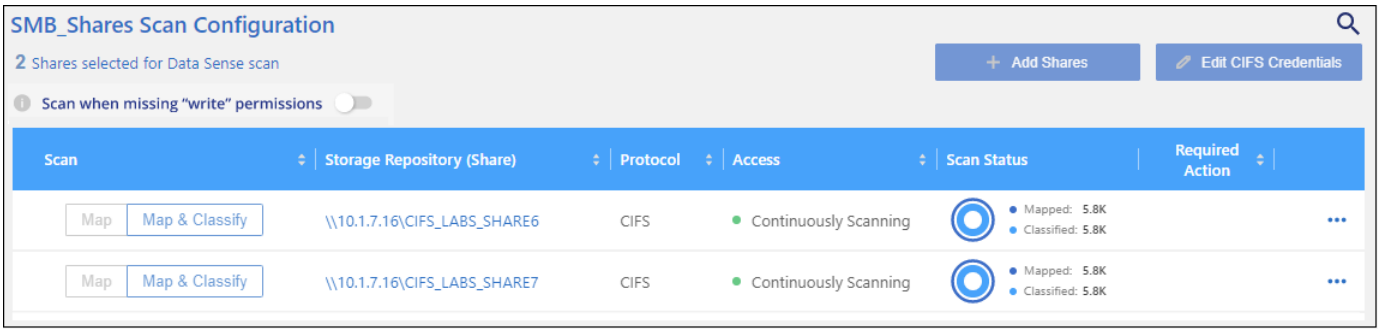
	* 메타데이터 *	**CIFS ***	*** NFS ***
* 타임 스탬프 *	_작성 시간 _	사용 가능	사용할 수 없음(Linux에서는 지원되지 않음)
	_마지막 액세스 시간 _	사용 가능	사용 가능
	_마지막 수정 시간 _	사용 가능	사용 가능
* 권한 *	_권한 열기 _	"Everyone" 그룹이 파일에 액세스할 수 있는 경우 "조직에 열기"로 간주됩니다.	"기타"가 파일에 액세스할 수 있는 경우 "조직에 열기"로 간주됩니다.
	_ 사용자/그룹 액세스 _	사용자 및 그룹 정보는 LDAP에서 가져옵니다	사용할 수 없음(NFS 사용자는 일반적으로 서버에서 로컬로 관리되므로 동일한 개인이 각 서버에서 다른 UID를 가질 수 있음)

- BlueXP 분류는 SharePoint Online, SharePoint 온프레미스(SharePoint Server), OneDrive, Google Drive, Amazon S3, 데이터베이스의 데이터 소스에서는 "마지막 액세스 시간"을 추출하지 않습니다.
- Windows OS의 이전 버전(예: Windows 7 및 Windows 8)은 시스템 성능에 영향을 줄 수 있으므로 기본적으로 "마지막 액세스 시간" 속성의 컬렉션을 사용하지 않습니다. 이 속성을 수집하지 않으면 "마지막 액세스 시간"을 기반으로 하는 BlueXP 분류 분석에 영향을 줍니다. 필요한 경우 이러한 이전 Windows 시스템에서 마지막 액세스 시간 모음을 활성화할 수 있습니다.

마지막 액세스 시간 타임 스탬프입니다

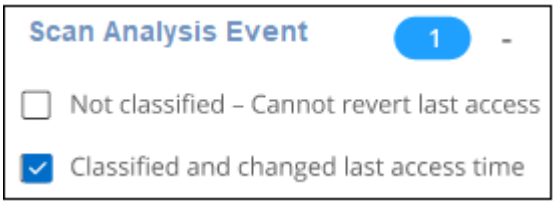
BlueXP 분류는 파일 공유에서 데이터를 추출할 때 운영 체제는 이를 데이터에 액세스하는 것으로 간주하고 그에 따라 "마지막 액세스 시간"을 변경합니다. 검사 후 BlueXP 분류는 마지막 액세스 시간을 원래 타임 스탬프로 되돌려줍니다. BlueXP 분류에 CIFS의 쓰기 속성 권한이나 NFS의 쓰기 권한이 없는 경우 시스템은 마지막 액세스 시간을 원래 타임 스탬프로 되돌릴 수 없습니다. SnapLock로 구성된 ONTAP 볼륨은 읽기 전용 권한을 가지며 마지막 액세스 시간을 원래 타임 스탬프로 되돌릴 수도 없습니다.

기본적으로 BlueXP 분류에 이러한 권한이 없는 경우 BlueXP 분류는 "마지막 액세스 시간"을 원래 타임 스탬프로 되돌릴 수 없기 때문에 시스템에서 볼륨의 해당 파일을 검색하지 않습니다. 그러나 파일의 마지막 액세스 시간이 원래 시간으로 재설정되는 것을 염려하지 않을 경우, BlueXP 분류가 권한에 관계없이 볼륨을 스캔하도록 구성 페이지 하단에 있는 * "쓰기 속성" 권한 * 스위치가 없을 때 * 스캔 을 클릭할 수 있습니다.



이 기능은 온프레미스 ONTAP 시스템, Cloud Volumes ONTAP, Azure NetApp Files, FSx for ONTAP 및 비 NetApp 파일 공유에 적용할 수 있습니다.

조사 페이지에 *Scan Analysis Event* 라는 필터가 있어 BlueXP 분류로 마지막 액세스 시간을 되돌릴 수 없기 때문에 분류되지 않은 파일을 표시할 수 있습니다. 또는 BlueXP 분류로 마지막 액세스 시간을 되돌릴 수 없지만 분류된 파일을 사용할 수 있습니다.



필터 선택 항목은 다음과 같습니다.

- "분류되지 않음 — 마지막 액세스 시간을 되돌릴 수 없음" - 쓰기 권한이 없어 분류되지 않은 파일을 표시합니다.
- "마지막 액세스 시간 분류 및 업데이트" - 분류된 파일과 BlueXP 분류에서 마지막 액세스 시간을 원래 날짜로 다시 설정할 수 없는 파일을 표시합니다. 이 필터는 에서 "쓰기 속성" 권한이 없을 때 * 스캔 을 설정한 환경에만 적용됩니다.

필요한 경우 이러한 결과를 보고서로 내보내면 권한 때문에 어떤 파일이 스캔되고 있는지 또는 스캔되지 않았는지 확인할 수 있습니다. "데이터 조사 보고서에 대해 자세히 알아보십시오".

BlueXP 분류 시스템에 로그인합니다

때때로 BlueXP 분류 시스템에 로그인하여 로그 파일에 액세스하거나 구성 파일을 편집해야 할 수 있습니다.

BlueXP 분류가 온프레미스의 Linux 머신 또는 클라우드에 구축한 Linux 시스템에 설치되면 구성 파일 및 스크립트에 직접 액세스할 수 있습니다.

클라우드에 BlueXP 분류를 배포할 때는 BlueXP 분류 인스턴스에 SSH를 사용해야 합니다. 사용자 및 암호를 입력하거나 BlueXP Connector 설치 중에 제공한 SSH 키를 사용하여 시스템에 SSH를 수행합니다. SSH 명령은 다음과 같습니다.

```
 ssh -i <path_to_the_ssh_key> <machine_user>@<datasense_ip>
* path_to_the_ssh_key> = ssh 인증 키의 위치입니다
* machine_user>:
```

+
AWS의 경우 **<EC2-USER>**를 사용합니다
Azure의 경우: BlueXP 인스턴스에 대해 생성한 사용자를 사용합니다
** GCP의 경우: BlueXP 인스턴스에 대해 생성한 사용자를 사용합니다

- datasense_ip> = 가상 시스템 인스턴스의 IP 주소입니다

클라우드의 시스템에 액세스하려면 보안 그룹 인바운드 규칙을 수정해야 합니다. 자세한 내용은 다음을 참조하십시오.

- "AWS의 보안 그룹 규칙"
- "Azure의 보안 그룹 규칙"
- "Google Cloud의 방화벽 규칙"

BlueXP 분류 API

웹 UI를 통해 사용할 수 있는 BlueXP 분류 기능은 Swagger API를 통해서도 사용할 수 있습니다.

BlueXP 분류 안에는 UI의 탭에 해당하는 4가지 범주가 정의되어 있습니다.

- 조사
- 규정 준수
- 기대치를 설정합니다
- 구성

Swagger 문서의 API를 사용하여 데이터를 검색, 집계, 스캔 추적 및 복사, 이동 등의 작업을 생성할 수 있습니다.

개요

API를 사용하여 다음 기능을 수행할 수 있습니다.

- 정보를 내보냅니다
  - UI에서 사용할 수 있는 모든 것을 API를 통해 내보낼 수 있습니다(보고서 제외).
  - 데이터를 JSON 형식으로 내보내기(분석 및 Splunk와 같은 타사 애플리케이션으로 푸시)
- "AND" AND "OR" 문을 사용하여 쿼리를 만들고, 정보를 포함 및 제외하는 등의 작업을 수행할 수 있습니다.

  예를 들어 FILES_Without_specific 개인 식별 정보(PII)를 찾을 수 있습니다(UI에서는 사용할 수 없는 기능). 내보내기 작업에 대한 특정 필드를 제외할 수도 있습니다.

- 작업을 수행합니다
  - CIFS 자격 증명을 업데이트합니다
  - 작업을 보고 취소합니다
  - 디렉토리를 다시 스캔합니다
  - 데이터를 삭제, 복사, 레이블 지정 및 할당합니다

- 파일 복제 및 복사
- 데이터를 내보냅니다

API는 안전하고 UI와 동일한 인증 방법을 사용합니다. 인증에 대한 자세한 내용은 다음을 참조하십시오. https://docs.netapp.com/us-en/bluexp-automation/platform/get_identifiers.html

Swagger API 참조에 액세스

Swagger로 이동하려면 BlueXP 분류 인스턴스의 IP 주소가 필요합니다. 클라우드 배포의 경우 공용 IP 주소를 사용합니다. 그런 다음 이 끝점에 도달해야 합니다.

https://<classification_ip>/documentation 을 참조하십시오

API 사용 예

다음 예제는 파일을 복사하기 위한 API 호출을 보여 줍니다.

API 요청

처음에는 작업 환경에 대한 모든 관련 필드 및 옵션을 가져와서 조사 탭의 모든 필터를 확인해야 합니다.

```
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"
```

응답


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

요청 매개 변수에서 해당 응답을 사용하여 복사하려는 파일을 필터링합니다.

여러 항목에 작업을 적용할 수 있습니다. 지원되는 작업 유형은 이동, 삭제, 복사, 할당, FlexClone, 데이터 내보내기, 재스캔 및 레이블

다음과 같이 복사 작업을 생성합니다.

API 요청

다음 API는 해당 API이며, 이를 통해 여러 작업을 생성할 수 있습니다.

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

응답

응답은 동작 객체를 반환하므로 API 가져오기 및 삭제 를 사용하여 동작에 대한 상태를 얻거나 취소할 수 있습니다.

```
{
  "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"
}
```