Fsas Technologies

P3AG-6562-02Z0

ETERNUS AX series オールフラッシュアレイ，ETERNUS HX series ハイブリッドアレイ

ONTAP File System Analytics ベストプラクティスガイド ソリューション展開

目次

図目次

表目次

はじめに

本書では、ONTAP File System Analytics のアーキテクチャ、ONTAP System Manager との統合、外部アプリケーションとの REST API 通信、分析における課題、およびお客様が直面する可能性のあるさまざまな問題を ONTAP File System Analytics がどのように解決するかについて説明します。

第 2 版
2025 年 3 月

登録商標

本製品に関連する他社商標については、以下のサイトを参照してください。
https://www.fujitsu.com/jp/products/computing/storage/trademark/
本書では、本文中の™、® などの記号は省略しています。

本書の読み方

対象読者

本書は、ETERNUS AX/HX の設定、運用管理を行うシステム管理者、または保守を行うフィールドエンジニアを対象としています。必要に応じてお読みください。

関連マニュアル

ETERNUS AX/HX に関連する最新の情報は、以下のサイトで公開されています。
https://www.fujitsu.com/jp/products/computing/storage/manual/

本書の表記について

本文中の記号

本文中では、以下の記号を使用しています。

注意	お使いになるときに注意していただきたいことを記述しています。必ずお読みください。
備考	本文を補足する内容や、参考情報を記述しています。

1. ONTAP File System Analytics

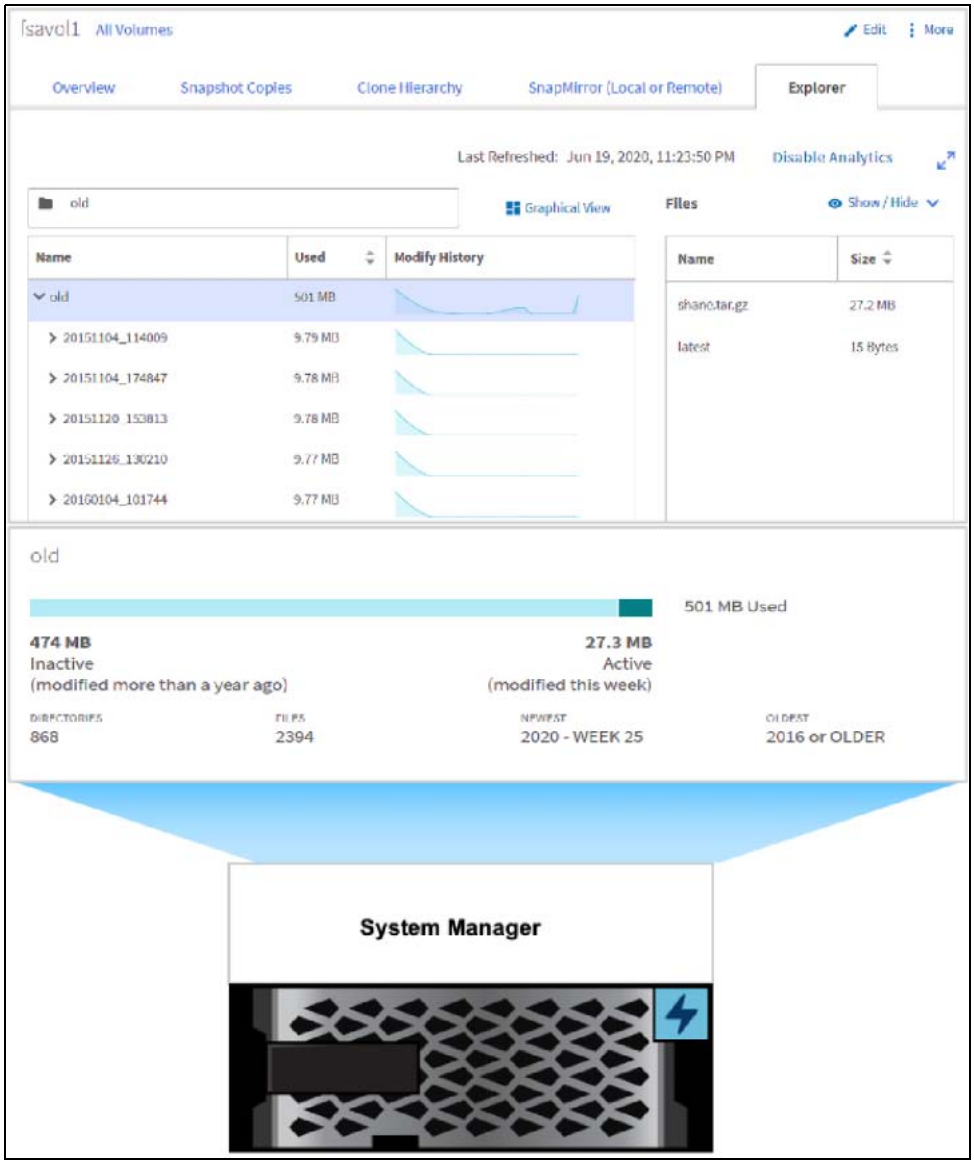
ONTAP File System Analytics は、ONTAP 9.8 で導入された、FlexGroup および FlexVol ボリュームの内容に関するデータを収集して表示するためのフレームワークです。これにより、外部ツールを使用しなくても、長期間にわたるファイルおよびディレクトリの容量アクセスと使用状況を可視化します。このリアルタイムの可視性により、スループットに対する QoS の変更や、プライマリストレージおよびセカンダリストレージへのボリュームの移動など、効果的なデータ管理と運用が可能になります。

1.1 技術概要

ONTAP File System Analytics は、データをリアルタイムで統合・収集し、ファイル数、ディレクトリ数、ファイル作成後の経過時間を示すヒストグラム、容量使用率などのパラメータに関する詳細情報を提供します。

ホット、ウォーム、コールドなどの業界標準の基準を使用してデータを視覚化できます。File System Analytics は、アクティブ、通常、非アクティブの各ラベルを使用して同様のデータを提供します。

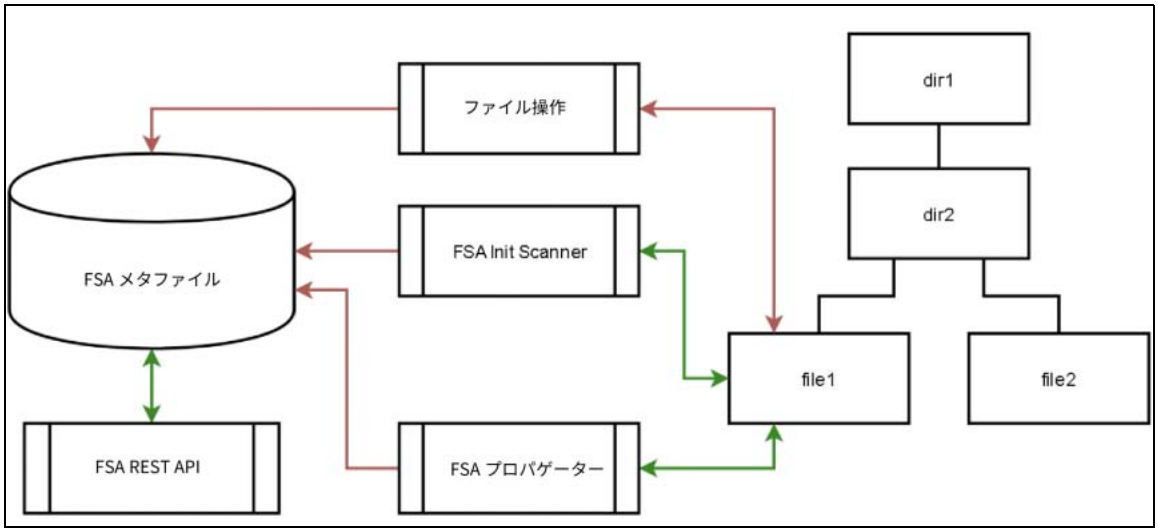
図 1.1 ONTAP File System Analytics: 高レベルデータ



1.2 アーキテクチャ

ONTAP File System Analytics(FSA) アーキテクチャには、Init Scanner、Propagator、メタファイル、ファイル操作、REST API の 5 つのコンポーネントがあります。(図 1.2 の赤の矢印は読み取り / 書き込み操作、緑の矢印は読み取り専用操作を示します。)

図 1.2 ONTAP File System Analytics アーキテクチャ



admin などの特権ユーザーが System Manager、CLI、または REST API を使用して分析を有効にすると、ONTAP ファイルシステムの Init Scanner は FlexVol ボリューム内の複数のディレクトリで並行して実行されます。Init Scanner は FlexGroup ボリューム内の複数の構成要素に対しても同時に実行され、収集された詳細情報をメタファイルに渡します。スキャナの実行時間は、inode の数に比例します。各 FlexVol ボリュームにつき 1 つのメタファイルがあり、FlexGroup では各構成要素につき 1 つのメタファイルがあります。メタファイルには、データベースのようなキー値形式で詳細が格納されます。キーはディレクトリの ID、値は File System Analytics データです。

File System Analytics Propagator は、累積された File System Analytics の変更をディレクトリ階層の各レベルにプロパゲートします。FlexGroup ボリュームと FlexVol ボリュームの両方をほぼリアルタイムに集約するために、バックグラウンドでプロセスとして実行されます。ファイル操作では、ファイルシステムの変更が記録され、メタファイルにレポートされます。この動作はインラインかつ軽量のため、パフォーマンスへの影響が最小限に抑えられています。

REST API を使用すると、ファイルシステムをスキャンすることなく、メタファイルから File System Analytics レコードを収集するディレクトリをクエリで検索できます。デフォルトでは、結果は JSON 形式で表示され、並べ替えとフィルタを実行したり、返されるレコードのタイムアウトと制限を設定したりできます。ONTAP REST API を使用して ONTAP File System Analytics をオンまたはオフにすることもできます。

2. ONTAP File System Analytics と System Manager の統合

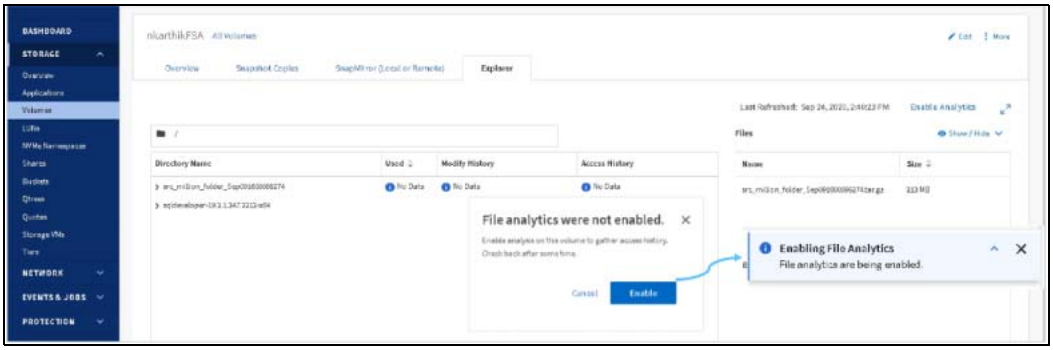
ONTAP File System Analytics は、ONTAP System Manager を使用して表示できます。通常、System Manager は admin などの特権ユーザーによって操作されます。このユーザーには、任意のディレクトリレベルから System Manager を介して File System Analytics から詳細を収集できる ONTAP API 要求への読み取り専用アクセス権が必要です。

2.1 File System Analytics の有効化

File System Analytics を有効にして、System Manager、ONTAP CLI、または REST API を使用して使用状況データを収集および表示できます。新しいボリュームを作成するとき、または ONTAP 9.8 以降のボリュームを含むシステムをアップグレードするときに、File System Analytics を有効にできます。

File System Analytics を有効にするには、次のタスクを実行します。

1. Storage > Volumes を選択し、任意のボリュームを選択します。
2. Explorer を起動して、分析を有効または無効にします。



ボリュームのサイズと内容によっては、ONTAP がボリューム内のデータを処理する間、分析を有効にするのに時間がかかる場合があります。進行状況は、System Manager または ONTAP CLI から `volume analytics show` コマンドを使用して表示できます。

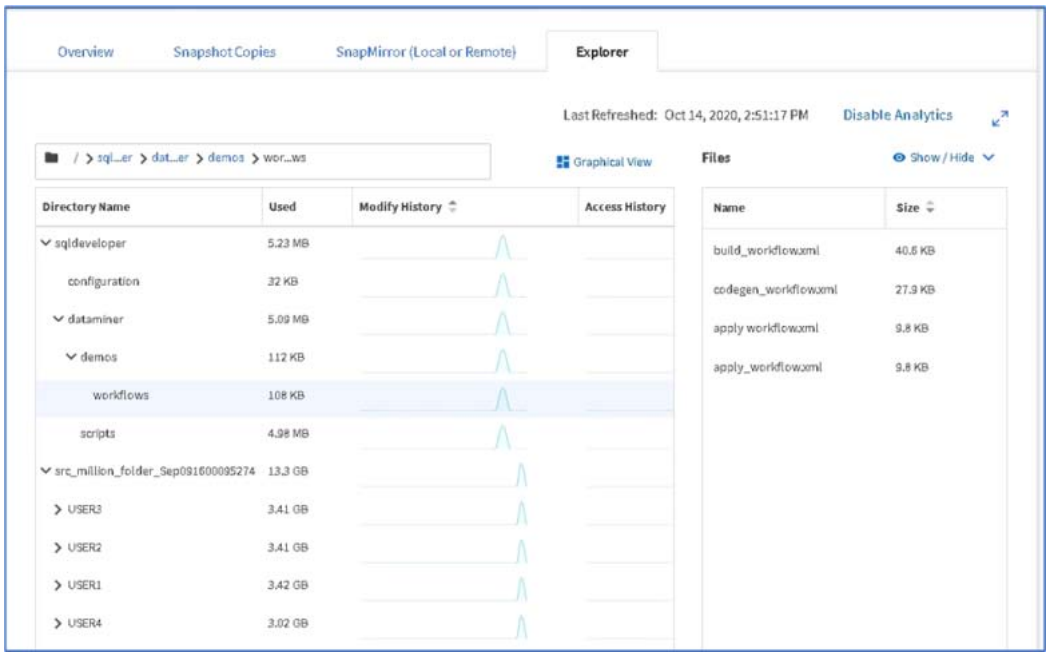
2.2 File System Analytics 用の System Manager ビュー

System Manager には、File System Analytics 用にリストビュー、グラフィックビュー、ファイルビュー、サマリービューという 4 つのビューがあります。

リストビュー

- フォルダとサブフォルダのツリー表示
- 名前、使用サイズ、変更履歴、およびアクセス履歴でソートされた、展開可能なリストです。

図 2.1 ONTAP File System Analytics 用の System Manager リストビューとテーブルビュー



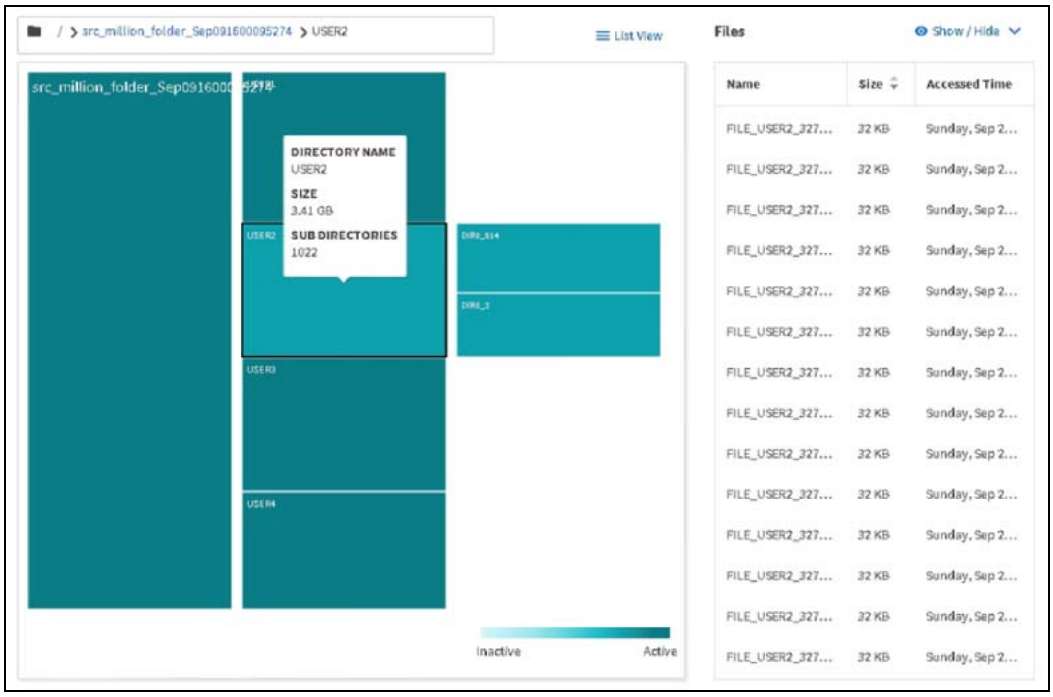
備考

現在のリリースでは、同じレベルまたは単一のフォルダ内に 3000 を超えるフォルダがある場合、デフォルトのソートパラメータは無効になっています。ただし、API 要求は制限なく使用できます。

グラフィックビュー

- ディレクトリのボックスサイズは、ディレクトリのサイズに比例します。
- フォルダの上部にカーソルを移動すると、そのフォルダ内のサブディレクトリの名前、サイズ、数を確認できます。
- 最初のリリースでは、最初の 3 つのフォルダが表示されます。

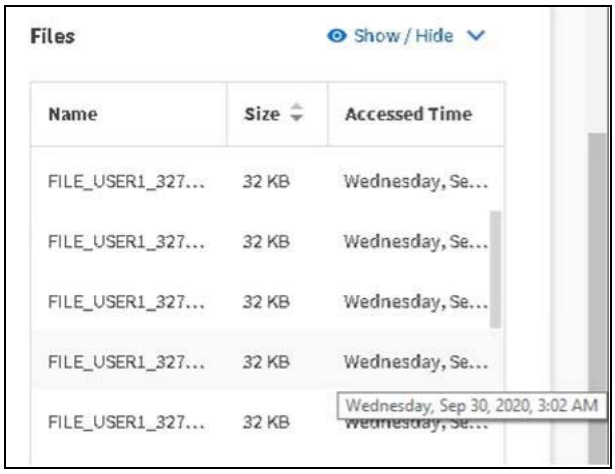
図 2.2 System Manager グラフィックビュー、ONTAP File System Analytics



ファイルビュー

- ファイルビューには、ディレクトリリストで選択したオブジェクトのファイル名、サイズ、アクセス時間が表示されます。
- アクセス日時を表示するファイルを選択します。

図 2.3 ONTAP File System Analytics 用の System Manager テーブルビュー



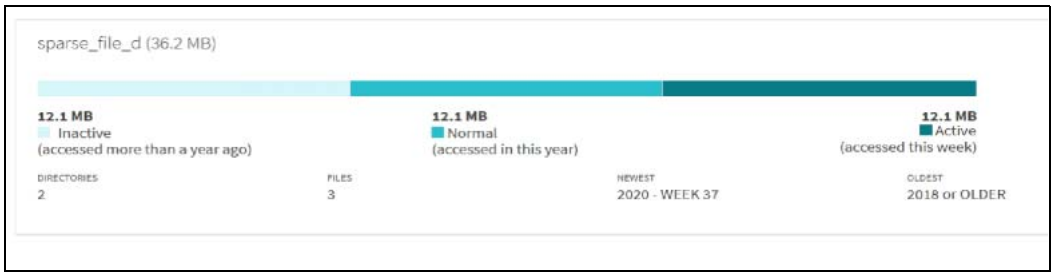
備考

ファイルビューの自動更新はありません。したがって、新しいファイルをフォルダに追加したときは、別のタブを選択してから元のページに戻ることで、更新された情報を表示できます。

サマリパネル

- ディレクトリ数、ファイル数、および作成されてからの経過時間情報も表示します。
- アクティブ (Active)、通常 (Normal)、非アクティブ (Inactive) のラベルでデータを表示します。
  - アクティブデータは 1 週間以内にアクセスされています。
  - 非アクティブデータは 1 年以上アクセスされていません。
  - 通常のデータは 1 週間から 1 年以内にアクセスされています。

図 2.4 ONTAP File System Analytics 用の System Manager サマリパネル



備考

ONTAP File System Analytics は、FlexCache ボリュームおよび SAN ボリュームと互換性がありません。このリリースでは、NAS ボリュームでのみ動作します。

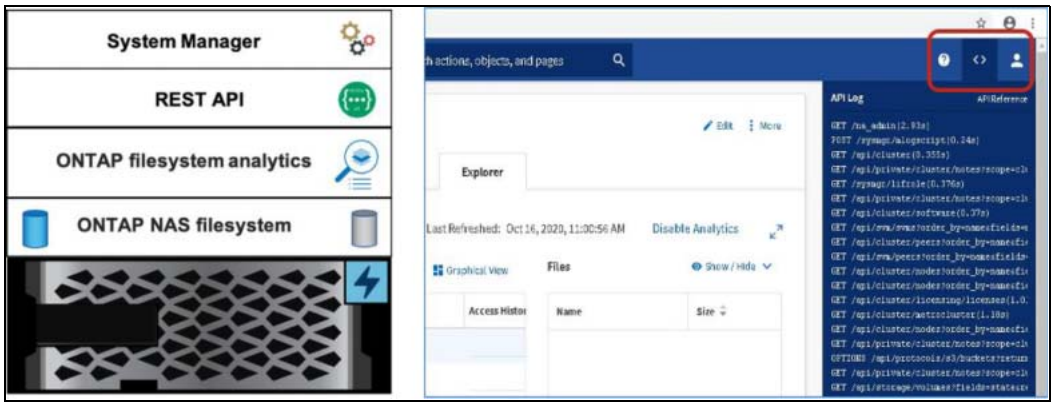
3. ONTAP File System Analytics REST API

サービスアプリケーションの開発には REST API が重要になります。現在、ユーザは独自開発の製品において、REST API のサポートが必要です。

3.1 System Manager と ONTAP File System Analytics の統合

System Manager は、REST API を使用して ONTAP File System Analytics からデータを収集し、リストビュー、グラフィカルビュー、ツリービュー、およびサマリービューを構築します。System Manager は、REST API 通信のワークフロー全体を自動化します。

図 3.1 System Manager と、ONTAP File System Analytics および API リファレンスの統合



System Manager では、Explore タブと画面右上端の＜＞アイコンをクリックすることで API リファレンスの早見表を表示します。この早見表を使用すると、各操作で System Manager が使用する API を確認できます。

備考

ONTAP File System Analytics にはロールベースアクセスコントロール (RBAC) はありませんが、ユーザの分析 API を無効にすることで、そのユーザが CLI を使用できないようにすることができます。

3.2 ONTAP File System Analytics とのアプリケーション統合

一般的に、お客様は ONTAP File System Analytics を Grafana などのアプリケーションと統合したいと考えます。ONTAP File System Analytics は REST API をサポートしており、System Manager にない独自のビューを作成できます。

降順でソートしたフォルダ数とファイル数を取得するには、次の例を参照してください。次の例では、変数は user、pw、および ontap_cluster_management_ip で、値はすでに指定されています。

```
root@scspr1882834014 demo]# curl -siku $user:$pw  --request GET
"https://${ontap_cluster_management_ip}/api/storage/volumes/1f911d57-e279-11ea-90fc-
005056a7adc6/files?fields=analytics.file_count&order_by=analytics.file_count+desc" | egrep
'name|file_count'
      "name": ".",
        "file_count": 735
      "name": "logs",
        "file_count": 628
      "name": "www",
        "file_count": 76
      "name": "misc",
        "file_count": 26
      "name": ".."
      "name": ".snapshot"
      "href": "/api/storage/volumes/1f911d57-e279-11ea-90fc-
005056a7adc6/files?fields=analytics.file_count&order_by=analytics.file_count+desc"
```

この例を昇順で表示するには、キーワード desc を asc に変更します。

ブラウザの ONTAP cluster management IP API ドキュメントを使用すると、ONTAP API、具体的には分析 API と通信できます。ボリュームのルートフォルダからの分析情報を提供する次の例を参照してください。

図 3.2　分析 API に対する ONTAP File System Analytics API ドキュメントのフロー例

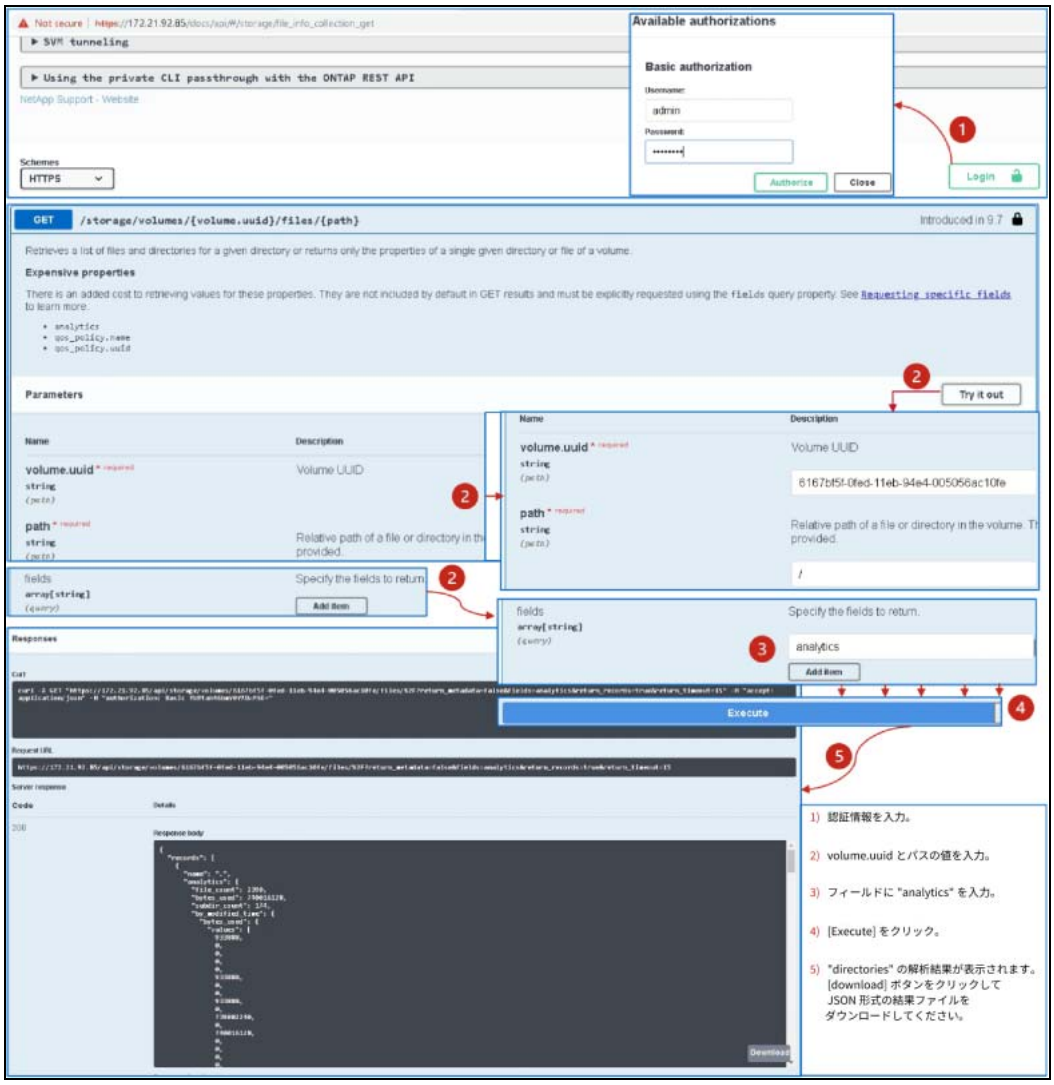


表 3.1 には、Analytics REST API の詳細が記載されています。

表 3.1 REST API の詳細

Analytics REST API	タイプ	意味
analytics.by_modified_time.bytes_used.oldest_label	文字列	analytics.by_modified_time.bytes_used.oldest_label によるフィルタリング
analytics.by_modified_time.bytes_used.percentages	数字	analytics.by_modified_time.bytes_used.percentages によるフィルタリング
analytics.by_modified_time.bytes_used.values	整数	analytics.by_modified_time.bytes_used.values によるフィルタリング
analytics.by_modified_time.bytes_used.newest_label	文字列	analytics.by_modified_time.bytes_used.newest_label によるフィルタリング
analytics.by_modified_time.bytes_used.labels	文字列	analytics.by_modified_time.bytes_used.labels によるフィルタリング
analytics.file_count	整数	analytics.file_count によるフィルタリング
analytics.bytes_used	整数	analytics.bytes_used によるフィルタリング
analytics.subdir_count	整数	analytics.subdir_count によるフィルタリング
analytics.by_accessed_time.bytes_used.oldest_label	文字列	analytics.by_accessed_time.bytes_used.oldest_label によるフィルタリング
analytics.by_accessed_time.bytes_used.percentages	数字	analytics.by_accessed_time.bytes_used.percentages によるフィルタリング
analytics.by_accessed_time.bytes_used.values	整数	analytics.by_accessed_time.bytes_used.values によるフィルタリング
analytics.by_accessed_time.bytes_used.newest_label	文字列	analytics.by_accessed_time.bytes_used.newest_label によるフィルタリング
analytics.by_accessed_time.bytes_used.labels	文字列	analytics.by_accessed_time.bytes_used.labels によるフィルタリング

analytics.by_access_time.bytes_used API と analytics.by_modify_time.bytes_user API には、labels、latest_label、および oldest_label という 3 つの項目があります。Labels は、アクセス / 変更ヒストグラム値において、対応するデータがゼロ以外の値で関連付けられている Newest_label (直近ラベル) と Oldest_label (最古ラベル) で期間を示す文字列です。

次の例を参照してください。analytics.file_count API の結果は、付録で提供されているテンプレートスクリプトに基づいて降順で表示されています。

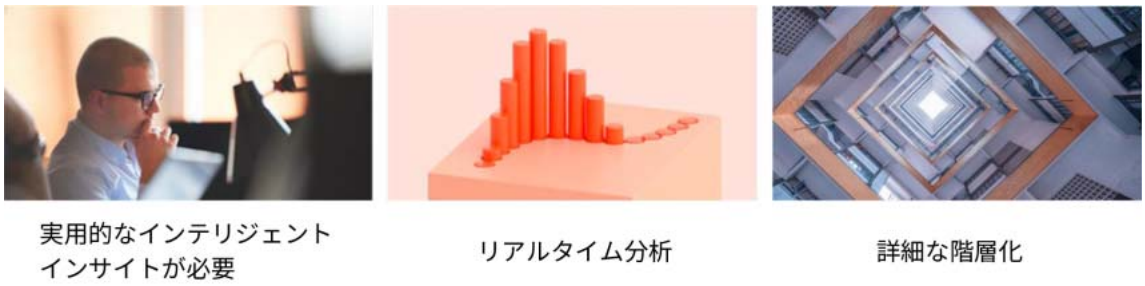
```
[root@scspr1936701025 ~]# bash apitemplate.sh
User : admin
Password :
Server : 10.236.153.165
Volume UUID : 2538a931-13c1-11eb-bfde-005056a772a7
choose the API from the list if not enter the API name:
 1. analytics.file_count
 2. analytics.by_modified_time.bytes_used.oldest_label
 3. analytics.by_modified_time.bytes_used.percentages
 4. analytics.by_modified_time.bytes_used.values
 5. analytics.by_modified_time.bytes_used.oldest_label
 6. analytics.by_modified_time.bytes_used.newest_label
 7. analytics.by_modified_time.bytes_used.labels
 8. analytics.bytes_used
 9. analytics.subdir_count
 10. analytics.by_accessed_time.bytes_used.oldest_label
 11. analytics.by_accessed_time.bytes_used.percentages
 12. analytics.by_accessed_time.bytes_used.values
 13. analytics.by_accessed_time.bytes_used.newest_label
 14. analytics.by_accessed_time.bytes_used.labels
 0. Others
choose the number or '0' for provide API name : 1
API choosen : analytics.file_count
Full path. default(/): /
Do you want to call the API with default options. (y/n): default(y):n
order_by filedname. default is API name(analytics.file_count):
order_by. asc/desc : default (asc): Desc
Number of seconds to allow the call to execute before returning: default (15): 10
user pw server volumeuuid apiname path fieldname order_by return_records max_records
return_timeout
admin xxxx 10.236.153.165 2538a931-13c1-11eb-bfde-005056a772a7 analytics.file_count %2F
analytics.file_count desc 10
curl -siku admin:xxx  --request GET "https://10.236.153.165/api/storage/volumes/2538a931-13c1-
11eb-bfde-
005056a772a7/files/%2F?fields=analytics.file_count&order_by=analytics.file_count+desc&return_time
out=10"
HTTP/1.1 200 OK
Date: Mon, 26 Oct 2020 21:18:04 GMT
Server: libzapid-httpd
X-Content-Type-Options: nosniff
Cache-Control: no-cache,no-store,must-revalidate
Content-Type: application/hal+json
Transfer-Encoding: chunked

{
  "records": [
    {
      "name": ".",
      "analytics": {
        "file_count": 137355
      }
    },
    {
      "name": "src_million_folder_Sep091600095274",
      "analytics": {
        "file_count": 134861
      }
    },
    {
      "name": "sqldeveloper-19.2.1.247.2212-x64",
      "analytics": {
        "file_count": 2390
      }
    },
```

```
    {
      "name": "20190626_162237",
      "analytics": {
        "file_count": 31
      }
    },
    {
      "name": "20190626_161806",
      "analytics": {
        "file_count": 24
      }
    },
    {
      "name": "20190626_162128",
      "analytics": {
        "file_count": 24
      }
    },
    {
      "name": "20170504_083851",
      "analytics": {
        "file_count": 8
      }
    },
    {
      "name": "20170504_090829",
      "analytics": {
        "file_count": 8
      }
    },
    {
      "name": "20170504_090917",
      "analytics": {
        "file_count": 8
      }
    },
    {
      "name": ".."
    },
    {
      "name": "src_million_folder_Sep091600095274.tar.gz"
    },
    {
      "name": ".snapshot"
    }
  ],
  "num_records": 12,
  "_links": {
    "self": {
      "href": "/api/storage/volumes/2538a931-13c1-11eb-bfde-
005056a772a7/files/%2F?fields=analytics.file_count&order_by=analytics.file_count+desc&return_time
out=10"
    }
  }
}[root@scspr1936701025 ~]#
```

4. 課題

NAS データのファイル分析を行う際には、次の 3 つの課題があります。

図 4.1　ONTAP File System Analytics - お客様の課題



実用的なインテリジェントインサイトが必要

ファイルとフォルダからシステムの状態を見通すことが求められています。たとえば、変更時刻やアクセス時刻に基づいて、ほとんどの変更がどこで行われたかを知ることができます。最も大きいフォルダはどれか？ データのアクセス頻度は高いか？　このような情報は、適切な QoS 制御を使用して、アクセス頻度の高いデータを高速なプライマリストレージ、アクセス頻度の低いデータをセカンダリストレージまたはクラウドストレージに移動するなどの必要なアクションを実行するために必要です。

リアルタイム分析

重要な情報をステークホルダーに提供するために、リアルタイムのオンラインファイル分析レポートが必要とされています。システム履歴を含むレポートは、以前のデータの動作を示し、データ管理用の追加ツールを提供します。これらの機能は、機密データに特に役立ちます。

詳細な階層化

ユーザーのホームフォルダを表示し、データのサイズ、ファイル数、変更時刻、アクセス時刻に基づいてさらに細分化することが求められます。大規模な組織では、ホームフォルダの動作を評価することが困難な場合があります。

5. 使用例

File System Analytics の主な使用例は 5 つあります。

- データ検出
- ストレージの最適化
- 稼働状態とパフォーマンス
- ビジネスインテリジェンス
- データガバナンス

ONTAP 9.8 では、データの検出とストレージの最適化に重点を置いています。

図 5.1　ONTAP File System Analytics の主な使用例

データ検出

データ検出とは、ディレクトリレベルでのデータ分散に関するものです。お客様は、プロジェクトで使用される容量、ユーザー ID、サブフォルダ情報、ファイル数の多いディレクトリ、平均ノードサイズなど、ワークロードに関する情報を必要としています。サードパーティのツールを使用してこの情報を取得することもできますが、これを行うと、システムのパフォーマンスに影響を与える時間のかかるプロセスとなり、ファイルシステムが時間の経過とともにファイルを変更するときに一貫性のない情報が生成される可能性があります。この場合、オンボックス分析ツールが最適なソリューションです。

ストレージの最適化

最も効果的な方法でストレージリソースを使用し、データの経過時間に基づいて適切なストレージ階層にデータを移動することが求められます。管理者がセカンダリストレージ階層またはクラウドにデータを移動またはアーカイブできるように、変更またはアクセスされていない古いデータを特定したいと考えています。

稼働状態とパフォーマンス

ワークロードの多い環境では、よりアクティブなディレクトリとツリー、アクセス頻度の高いファイルと箇所、オーナーシップに基づくトップユーザーとクライアントの特定など、詳細なトラブルシューティングが求められます。これらの詳細情報は、お客様がパフォーマンスの問題を特定して解決し、ワークロードストレージコントローラの負荷を分散してリソースの使用率を向上させるのに役立ちます。

ビジネスインテリジェンス

お客様は、時系列データ、履歴データ、トレンドデータに基づいて適切な意思決定を行う必要があります。これらのデータは、データ増加率に関するレポートの生成、将来のデータ増加の予測、データポイント、平均、チャージバックレポートの比較に役立ちます。このリリースでは ONTAP File System Analytics によるビジネスインテリジェンスのサポートを提供していませんが、ONTAP API との外部サードパーティアプリケーションの統合を利用して、この使用例に関するサポートをお客様に提供できます。

データガバナンス

データガバナンスでは、データの使用、アクセス、および動作パターンを利用して、組織内のデータ資産を規則的に管理します。また、データガバナンスはこれらのパターンを視覚化し、ランサムウェアなどの脅威から組織を保護するためにユーザーの行動異常を検出します。このリリースでは ONTAP File System Analytics によるデータガバナンスをサポートしていませんが、ONTAP API との外部サードパーティアプリケーションの統合を利用して、この使用例に関するサポートをお客様に提供できます。

6. ONTAP File System Analytics の性能

この章では、ONTAP File System Analytics でお客様の課題に対処する方法について説明します。

6.1 変更の少ないワークロードと変更の多いワークロード

ほとんどの場合、トランザクションワークロードには他のワークロードタイプよりも多くの変更が含まれています。このようなワークロードを管理するためには、最も変更が多い / 少ないワークロードとそのフォルダの情報が求められます。

ONTAP File System Analytics を使用して、GUI と REST API という 2 つの方法でこのニーズに対応しています。

GUI ベースのソリューション

以下のリストビューでは、変更量とアクセス履歴に基づいた各フォルダの挙動をしめします。名前、使用量、変更量、アクセス履歴 (Volume > Explorer > Modify History) に基づいてフォルダを並べ替えることができます。「Modify History」オプションでは、変更が最も多いフォルダと最も少ないフォルダが昇順または降順で表示されます。フォルダ名を基準にすると、そのフォルダで実行されているワークロードを確認できます。

図 6.1 最も変更が多い / 少ないワークロードの System Manager ビュー

REST API ベースのソリューション

REST API を使用してデータにアクセスすることが望まれる場合もあります。以下の例では、最も変更の少ないフォルダを示しています。最も変更が多いフォルダを参照する場合は、`order_by` を `analytics.by_modified_time.bytes_used+desc` に変更します。

```
[root@scspr1936701025 FSAFG]# curl -siku ${user}:${pw} --request GET
"https://${ontap_cluster_management_ip}/api/storage/volumes/${volume_uuid}/files/%2F?fields=analy
tics.by_modified_time.bytes_used&order_by=analytics.by_modified_time.bytes_used+asc"
HTTP/1.1 200 OK
Date: Wed, 21 Oct 2020 20:15:04 GMT
Server: libzapid-httpd
X-Content-Type-Options: nosniff
Cache-Control: no-cache,no-store,must-revalidate
Content-Type: application/hal+json
Transfer-Encoding: chunked

{
  "records": [
    {
      "name": ".."
    },
    {
      "name": "src_million_folder_Sep091600095274.tar.gz"
    },
    {
      "name": ".snapshot"
    },
    {
      "name": "20170504_090917",
      "analytics": {
        "by_modified_time": {
          "bytes_used": {
            "values": [
              12288,
              0,
              0,
              0,
              0,
              12288,
              0,
              0,
              12288,
              0,
              0,
              0,
              12288,
              0,
              0,
              110592,
              0,
              0
            ],
            "percentages": [
              10.00,
              0.00,
              0.00,
              0.00,
              0.00,
              10.00,
              0.00,
              0.00,
              10.00,
              0.00,
              0.00,
              0.00,
              10.00,
              0.00,
              0.00,
              90.00,
              0.00,
              0.00
            ],
            "newest_label": "2020-W43",
            "oldest_label": 2017
          }
        }
      }
    },
```

```
    {
      "name": "20170504_090829",
      "analytics": {
        "by_modified_time": {
          "bytes_used": {
            "values": [
              12288,
              0,
              0,
              0,
              0,
              12288,
              0,
              0,
              12288,
              0,
              0,
              0,
              12288,
              0,
              0,
              122880,
              0,
              0
            ],
            "percentages": [
              9.09,
              0.00,
              0.00,
              0.00,
              0.00,
              9.09,
              0.00,
              0.00,
              9.09,
              0.00,
              0.00,
              0.00,
              9.09,
              0.00,
              0.00,
              90.91,
              0.00,
              0.00
            ],
            "newest_label": "2020-W43",
            "oldest_label": 2017
          }
        }
      }
    },
   …
…
<removed some results to save page spaces>
},
    {
      "name": ".",
      "analytics": {
        "by_modified_time": {
          "bytes_used": {
            "values": [
              1228800,
              0,
              0,
              0,
              224616448,
              1228800,
              5210320896,
              0,
              1228800,
              5210320896,
              739082240,
              0,
              5950631936,
```

```
            1130496,
            0,
            393216,
            0,
            0
          ],
          "percentages": [
            0.02,
            0.00,
            0.00,
            0.00,
            3.77,
            0.02,
            87.54,
            0.00,
            0.02,
            87.54,
            12.42,
            0.00,
            99.97,
            0.02,
            0.00,
            0.01,
            0.00,
            0.00
          ],
          "newest_label": "2020-W43",
          "oldest_label": 2017
        }
      }
    }
  }
 ],
 "num_records": 12,
 "analytics": {
   "by_modified_time": {
     "bytes_used": {
       "labels": [
          "2020-W43",
          "2020-W42",
          "2020-W41",
          "2020-W40",
          "2020-W39",
          "2020-10",
          "2020-09",
          "2020-08",
          "2020-Q4",
          "2020-Q3",
          "2020-Q2",
          "2020-Q1",
          "2020",
          "2019",
          "2018",
          "2017",
          "--2016",
          "unknown"
        ]
      }
    }
  },
  "_links": {
    "self": {
      "href": "/api/storage/volumes/2538a931-13c1-11eb-bfde-
005056a772a7/files/%2F?fields=analytics.by_modified_time.bytes_used&order_by=analytics.by_modifie
d_time.bytes_used+Asc"
    }
  }
}[root@scspr1936701025 FSAFG]# curl -siku ${user}:${pw} --request GET
"https://${ontap_cluster_management_ip}/api/storage/volumes/${volume_uuid}/files/%2F?fields=analy
tics.by_modified_time.bytes_used&order_by=analytics.by_modified_time.bytes_used+Asc"  | less
[root@scspr1936701025 FSAFG]#
```


上記の例では、analytics.by_modified_time.bytes_used.values、analytics.by_modified_time.bytes_used.labels、および analytics.by_modified_time.bytes_used.percentageなどのサブフィールドもフィールド値として使用できます。

6.2 最大フォルダ

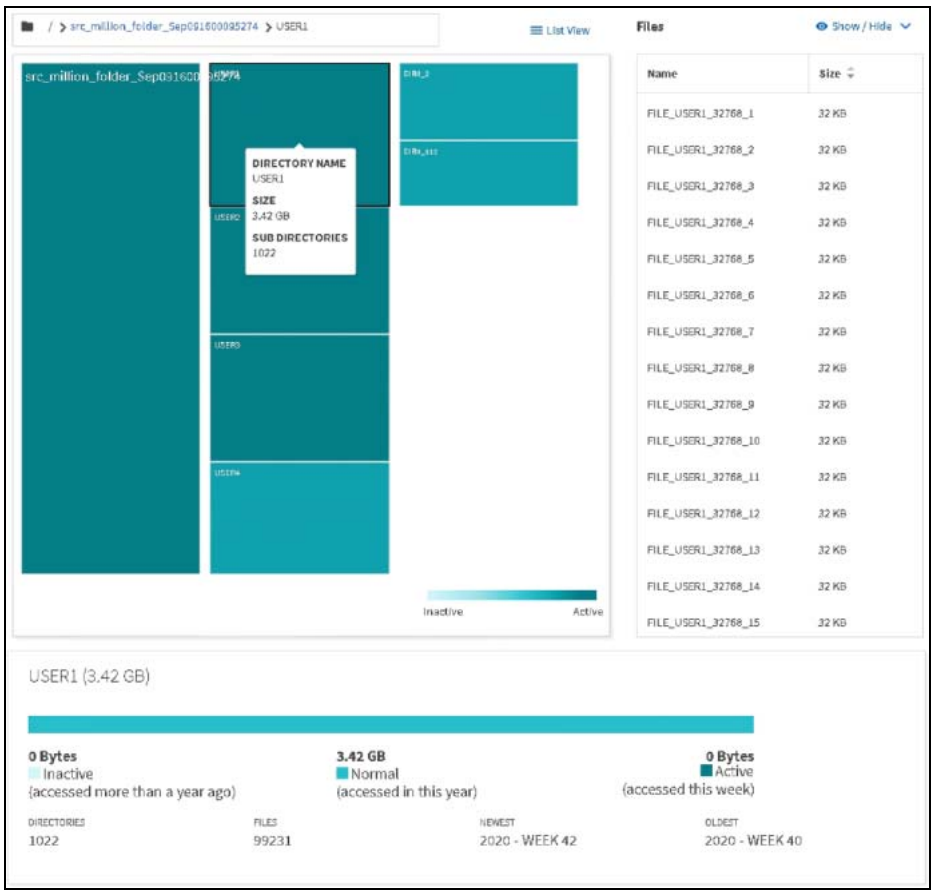
NAS ボリューム内で、最も多くの容量を消費するフォルダはどれかを把握したくなることは少なくありません。たとえば、どのユーザーのホームフォルダが最も大きいかを知りたいという場合です。

GUI または REST API を使用した ONTAP File System Analytics を通じてソリューションを提供できます。

GUI ソリューション

ONTAP File System Analytics は、外部ツールを使用しないリアルタイム分析を提供します。サイズに基づいて特定のフォルダの詳細を掘り下げることができます (Volume > Explorer でグラフィックビューを表示します)。ツールチップを使用して、サブフォルダの数と各フォルダのサイズを確認することもできます。たとえば、特定のユーザーのホームフォルダを選択して、ファイル数、フォルダ数、最新および最も古いファイル、アクティブ、通常、非アクティブのデータのサイズなどの詳細情報を取得できます。ファイルビューには、ユーザーのホームフォルダにあるファイルのリストも表示されます。

図 6.2 最大サイズのフォルダの System Manager ビュー



REST API ソリューション

次の例は、REST API を使用して最大のフォルダを収集するオプションを示しています。

```
[root@scspr1936701025 FSAFG]# curl -siku ${user}:${pw} --request GET
"https://${ontap_cluster_management_ip}/api/storage/volumes/${volume_uuid}/files/%2F?fields=analy
tics.bytes_used&order_by=analytics.bytes_used+desc" | egrep 'name|bytes_used'
      "name": ".",
        "bytes_used": 5952155648
      "name": "src_million_folder_Sep091600095274",
        "bytes_used": 4985806848
      "name": "sqldeveloper-19.2.1.247.2212-x64",
        "bytes_used": 740012032
      "name": "20190626_162237",
        "bytes_used": 765952
      "name": "20190626_162128",
        "bytes_used": 262144
      "name": "20190626_161806",
        "bytes_used": 253952
      "name": "20170504_083851",
        "bytes_used": 172032
      "name": "20170504_090829",
        "bytes_used": 135168
      "name": "20170504_090917",
        "bytes_used": 122880
      "name": ".."
      "name": "src_million_folder_Sep091600095274.tar.gz"
      "name": ".snapshot"
      "href": "/api/storage/volumes/2538a931-13c1-11eb-bfde-
005056a772a7/files/%2F?fields=analytics.bytes_used&order_by=analytics.bytes_used+desc"
[root@scspr1936701025 FSAFG]#
```

備考

REST API の例では、%2F は / を意味します。

6.3 多数のファイルを検索する

大きなファイルやフォルダを特定することで、それらを適切なストレージに移動することができます。サイズが大きいファイルの中には、非アクティブデータでプライマリストレージ内の容量を大量に消費している場合があります。こういったフォルダはセカンダリもしくはクラウドストレージに移動することができます。NAS ボリューム内の多数のファイルを検索すると、適切なデータの場所を選択するのに役立ちます。

ONTAP System Manager には UI インターフェースはありませんが、大量のファイルを取得するための REST API サポートを提供しています。フォルダ内の多数のファイルを昇順、降順で取得する場合は、次の例を参照してください。

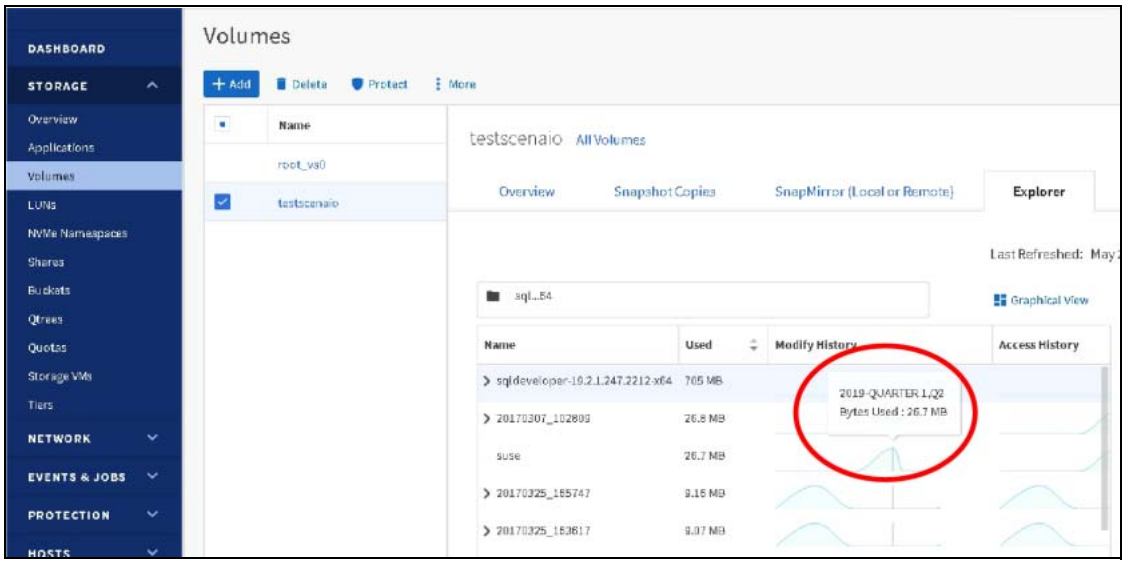
```
[root@scspr1936701025 FSAFG]# curl -siku ${user}:${pw} --request GET
"https://${ontap_cluster_management_ip}/api/storage/volumes/${volume_uuid}/files/%2F?fields=analy
tics.file_count&order_by=analytics.file_count+asc" | egrep 'name|file_count'
      "name": ".."
      "name": "src_million_folder_Sep091600095274.tar.gz"
      "name": ".snapshot"
      "name": "20170504_083851",
        "file_count": 8
      "name": "20170504_090829",
        "file_count": 8
      "name": "20170504_090917",
        "file_count": 8
      "name": "20190626_161806",
        "file_count": 24
      "name": "20190626_162128",
        "file_count": 24
      "name": "20190626_162237",
        "file_count": 31
      "name": "sqldeveloper-19.2.1.247.2212-x64",
        "file_count": 2390
      "name": "src_million_folder_Sep091600095274",
        "file_count": 133679
      "name": ".",
        "file_count": 136173
      "href": "/api/storage/volumes/2538a931-13c1-11eb-bfde-
005056a772a7/files/%2F?fields=analytics.file_count&order_by=analytics.file_count+asc"
[root@scspr1936701025 FSAFG]# curl -siku ${user}:${pw} --request GET
"https://${ontap_cluster_management_ip}/api/storage/volumes/${volume_uuid}/files/%2F?fields=analy
tics.file_count&order_by=analytics.file_count+desc" | egrep 'name|file_count'
      "name": ".",
        "file_count": 136173
      "name": "src_million_folder_Sep091600095274",
        "file_count": 133679
      "name": "sqldeveloper-19.2.1.247.2212-x64",
        "file_count": 2390
      "name": "20190626_162237",
        "file_count": 31
      "name": "20190626_161806",
        "file_count": 24
      "name": "20190626_162128",
        "file_count": 24
      "name": "20170504_083851",
        "file_count": 8
      "name": "20170504_090829",
        "file_count": 8
      "name": "20170504_090917",
        "file_count": 8
      "name": ".."
      "name": "src_million_folder_Sep091600095274.tar.gz"
      "name": ".snapshot"
      "href": "/api/storage/volumes/2538a931-13c1-11eb-bfde-
005056a772a7/files/%2F?fields=analytics.file_count&order_by=analytics.file_count+desc"
[root@scspr1936701025 FSAFG]#
```

6.4 アクティビティの動作

機密データの異常な動作を特定するために、アクセスまたはその他のパラメータの予期しないスパイクを検出できます。たとえば、ランサムウェア攻撃によって一部のフォルダが影響を受ける可能性がある場合、Volume > Foloder > Folder List > 異常スパイクのあるフォルダを選択することで、管理者はこれらのフォルダとそのデータを別のセカンダリストレージに移動して、さらに調査することができます。

以下のように GUI を使用してアクティビティを監視できます。

図 6.3 System Manager の異常アクティビティ検出



6.5 Analytics REST API の JSON 形式

表 6.1 に、Analytics REST API の JSON 形式とその説明を示します。この情報は開発者が Analytics API に基づいてアプリケーションを設計するのに役立ちます。

表 6.1 Analytics REST API JSON

フィールド	サブフィールド	説明
Name		ディレクトリエントリ名
File_count		ディレクトリルートのファイル数
Bytes_used		ディレクトリのサイズ
by_modified_time		変更履歴
	bytes used	4 年以上前、3 年以内、2 年以内、4ヶ月以内、3ヶ月以内に変更されたデータのサイズ
	percentages	総ディレクトリサイズに対するデータ変更量の割合 (%)
	newest folder	直近で変更したフォルダ
	oldest folder	最古のディレクトリ
Access history		アクセス履歴
	bytes used	4 年以上前、3 年以内、2 年以内、4ヶ月以内、3ヶ月以内にアクセスしたデータのサイズ
	percentages	総ディレクトリサイズに対するデータ変更量の割合 (%)
	newest folder	直近にアクセスしたフォルダの名前
	oldest folder	1 年以上アクセスされていない最古のディレクトリ

7. まとめ

ONTAP File System Analytics には、次の利点があります。

- 最も変更されたデータまたは最も変更されていないデータの特定、最大のフォルダの特定、ファイルの最大数の場所の特定など、ファイル構造のマッピングと特徴づけ
- 異常なデータ動作の識別
- リアルタイム分析
- REST API を介した ONTAP File System Analytics とのアプリケーション統合
- データ検出とストレージ最適化のサポート
- サマリパネルには、ホット (アクティブ)、コールド (非アクティブ)、およびウォーム (正常) データが示されます。
- 変更およびアクセス時間に基づくデータのソート
- プライマリストレージとセカンダリストレージに関するコストパフォーマンスに優れた意思決定
- 詳細なディレクトリアクセス

A. 付録

```
[root@scspr1936701025 ~]# cat apitemplate.sh
#! /usr/software/bin/bash

path_to_chars()
{
        result="$1"
        count=`echo ${result}|wc -m`
        result_temp=''
        for cursor in `seq 1 $count`
        do
                c1=`echo $result|cut -c${cursor}`
                if test "$c1" = "/"
                then
                        c1="%2F"
                fi
                result_temp+=${c1}
        done
        result=${result_temp}
        eval $2='$result'
}

echo -n "User : "
read user
[[ -z $user ]] && { echo "User name is empty"; exit 1; }

echo -n "Password : "
read -s pw
echo
[[ -z $pw ]] && { echo "Password is empty"; exit 1; }

echo -n "ONTAP Cluster Management IP/Name : "
read server
[[ -z $ontap_cluster_management_ip ]] && { echo "ONTAP Cluster Management ip or name is empty";
exit 1; }

echo -n "Volume UUID : "
read volumeuuid;
[[ -z $volumeuuid ]] && { echo "volume uuid is empty"; exit 1; }

echo "choose the API from the list if not enter the API name:"

echo -e " 1. analytics.file_count"
echo -e " 2. analytics.by_modified_time.bytes_used.oldest_label"
echo -e " 3. analytics.by_modified_time.bytes_used.percentages"
echo -e " 4. analytics.by_modified_time.bytes_used.values"
echo -e " 5. analytics.by_modified_time.bytes_used.oldest_label"
echo -e " 6. analytics.by_modified_time.bytes_used.newest_label"
echo -e " 7. analytics.by_modified_time.bytes_used.labels"
echo -e " 8. analytics.bytes_used"
echo -e " 9. analytics.subdir_count"
echo -e " 10. analytics.by_accessed_time.bytes_used.oldest_label"
echo -e " 11. analytics.by_accessed_time.bytes_used.percentages"
echo -e " 12. analytics.by_accessed_time.bytes_used.values"
echo -e " 13. analytics.by_accessed_time.bytes_used.newest_label"
echo -e " 14. analytics.by_accessed_time.bytes_used.labels"
echo -e " 0. Others"

echo -n "choose the number or '0' for provide API name : "
read apiname
[[ -z $apiname ]] && { echo "apiname is empty"; exit 1; }

#echo $apiname
case $apiname in
        "1")
                apiname="analytics.file_count"; echo "API choosen : $apiname"
                ;;
        "2")
                apiname="analytics.by_modified_time.bytes_used.oldest_label"; echo "API choosen :
$apiname"
                ;;
```

```
        "3")
                apiname="analytics.by_modified_time.bytes_used.percentages"; echo "API choosen :
$apiname"
                ;;
        "4")
                apiname="analytics.by_modified_time.bytes_used.values"; echo "API choosen :
$apiname"
                ;;
        "5")
                apiname="analytics.by_modified_time.bytes_used.oldest_label"; echo "API choosen :
$apiname"
                ;;
        "6")
                apiname="analytics.by_modified_time.bytes_used.newest_label"; echo "API choosen :
$apiname"
                ;;
        "7")
                apiname="analytics.by_modified_time.bytes_used.labels"; echo "API choosen :
$apiname"
                ;;
        "8")
                apiname="analytics.bytes_used"; echo "API choosen : $apiname"
                ;;
        "9")
                apiname="analytics.subdir_count"; echo "API choosen : $apiname"
                ;;
        "10")
                apiname="analytics.by_accessed_time.bytes_used.oldest_label"; echo "API choosen :
$apiname"
                ;;
        "11")
                apiname="analytics.by_accessed_time.bytes_used.percentages"; echo "API choosen :
$apiname"
                ;;
        "12")
                apiname="analytics.by_accessed_time.bytes_used.values"; echo "API choosen :
$apiname"
                ;;
        "13")
                apiname="analytics.by_accessed_time.bytes_used.newest_label"; echo "API choosen :
$apiname"
                ;;
        "14")
                apiname="analytics.by_accessed_time.bytes_used.labels"; echo "API choosen :
$apiname"
                ;;
        "0")
                echo -n " Enter the API Name : "
                read apiname
                echo "API name : $apiname"
                ;;
        *)
                echo "Entries not exist";
                exit 1;
                ;;
esac

echo -n "Full path. default(/): "
read path_input;
[[ -z $path_input ]] && { path_input="/"; }
path_to_chars $path_input path

echo -n "Do you want to call the API with default options. (y/n): default(y):"
read default_yes_or_no;
[[ -z $default_yes_or_no ]] && { default_yes_or_no="y"; }
default_yes_or_no="${default_yes_or_no,,}"
if test $default_yes_or_no = "y" -o $default_yes_or_no = "yes"
then
        #API call to ontap filesystem analytics
        echo "curl -siku $user:xxx  --request GET
\"https://${ontap_cluster_management_ip}/api/storage/volumes/${volumeuuid}/files/${path}?fields=$
{apiname}\" "
        curl -siku $user:$pw  --request GET
```

```
"https://${ontap_cluster_management_ip}/api/storage/volumes/${volumeuuid}/files/${path}?fields=${
apiname}"
else
        #order_by : field  - filedname. default is API name
        echo -n "order_by filedname. default is API name($apiname): "
        read order_by_fieldname;
        [[ -z $order_by_fieldname ]] && { order_by_fieldname=$apiname; }

        #order_by : order_by. default assending order
        echo -n "order_by. asc/desc : default (asc): "
        read order_by;
        order_by="${order_by,,}"
        [[ -z $order_by ]] && { order_by="asc"; }

        #initialize extra options
        extra_options='';
        #return_timeout - default 15 sec
        echo -n "Number of seconds to allow the call to execute before returning: default (15): "
        read return_timeout;
        [[ -z $return_timeout ]] && { return_timeout=15; }
        if test $return_timeout -ne 15
        then
                extra_options+="&return_timeout=${return_timeout}"
        fi
        #API call to ontap filesystem analytics
        echo "user pw ontap_cluster_management_ip volumeuuid apiname path fieldname order_by
return_records max_records return_timeout"
        echo $user xxxx $ontap_cluster_management_ip $volumeuuid $apiname $path
$order_by_fieldname $order_by $return_records $max_records $return_timeout
        echo "curl -siku ${user}:xxx  --request GET
\"https://${ontap_cluster_management_ip}/api/storage/volumes/${volumeuuid}/files/${path}?fields=$
{apiname}&order_by=${order_by_fieldname}+${order_by}${extra_options}\" "
        curl -siku $user:$pw  --request GET
"https://${ontap_cluster_management_ip}/api/storage/volumes/${volumeuuid}/files/${path}?fields=${
apiname}&order_by=${order_by_fieldname}+${order_by}${extra_options}"
fi
[root@scspr1936701025 ~]#
```

ETERNUS AX series オールフラッシュアレイ , ETERNUS HX series ハイブリッドアレイ
ONTAP File System Analytics ベストプラクティスガイド ソリューション展開

P3AG-6562-02Z0

発行年月　2025 年 3 月
発行責任　エフサステクノロジーズ株式会社

Fsas Technologies