



テクニカル レポート

EDAワークロードに対するMicrosoft Azure Cloud Volumes ONTAPのベストプラクティス

Microsoft AzureでのEDAおよび半導体のワークロード 向けにNetApp Cloud Volumes ONTAPを最適化

NetApp
Michael Johnson
2022年11月 | TR-4946

概要

このテクニカルレポートでは、電子設計自動化（EDA）ワークロード向けのNetApp® Cloud Volumes ONTAP® for Microsoft Azureのセットアップとパフォーマンスについて検証します。NetAppのパートナー様、お客様、従業員は、提示された情報を基に、どのワークロードがCloud Volumes ONTAPに適しているかを十分な情報に基づいて判断する必要があります。

<<本レポートは機械翻訳による参考訳です。公式な内容はオリジナルである英語版をご確認ください。>>

目次

はじめに.....	4
ハイブリッドクラウドにおけるEDA向けリファレンスアーキテクチャ	4
EDAストレージのパフォーマンス要件	4
EDAボリュームとディレクトリ	5
EDAデータ型	8
EDAのパフォーマンスとクラウドの拡張性	11
エンタープライズ規模のハイブリッドクラウドデータ管理アーキテクチャ	12
Azure Cloud Volumes ONTAPセットアップのクイックスタート	13
Cloud Volumes ONTAP作成のセットアップ	13
Cloud Volumes ONTAP作成後のセットアップ	22
パフォーマンスの監視	28
容量レポート	29
Cloud Volumes ONTAPインスタンスの起動と停止	29
クイックスタートの実装の詳細.....	29
Cloud Volumes ONTAPインスタンスタイプ	29
追加サービス	30
アベイラビリティゾーン (AZ) を選択	32
NetApp Support Siteアカウントの管理	32
課金方法の選択 (ライセンス)	33
EDAワークロードのCloud Volumes ONTAPサイジングに関する考慮事項	33
Cloud Volumes ONTAPディスクの選択とスループットの要件	34
Cloud Volumes ONTAP作成後のセットアップ	36
追加情報の入手方法.....	39
バージョン履歴.....	40
表一覧	
表1) すべてのジョブの総ターンアラウンド時間で測定されたパフォーマンス	11
表2) シングルノードCloud Volumes ONTAP構成 (VMタイプe48ds_v4に基づく)	19
表3) アクティブ/アクティブHAペアCloud Volumes ONTAP構成 (VMタイプe48ds_v4に基づく)	19
表4) FlexVolサイズ別のinodeのデフォルトと最大値.....	38

図一覧

図1) 一般的なオンプレミス設計インフラ	5
図2) EDAストレージ構成.....	6
図3) ハイブリッドクラウドアーキテクチャとバーストツークラウドアーキテクチャ	8
図4) 一般的なEDAツールとライブラリ UNIXディレクトリ構造	9
図5) EDAプロジェクトのUNIXディレクトリ構造の例.....	10
図6) マルチプロジェクトおよびマルチチーム環境向けのクラウドアーキテクチャ.....	13
図7) 標準NASベンチマーク/	28
図8) 拡張性に応じたEDA読み取りパフォーマンスの最適化	39

はじめに

このドキュメントでは、NetApp Cloud Volumes ONTAP for Electronic Design Automation (EDA) および半導体ワークロードをMicrosoft Azureで使用する際のベストプラクティスを紹介します。このドキュメントでは、一般的なEDAストレージのユースケースと要件について簡単に説明します。

EDAワークロードと半導体ワークロードはハイパフォーマンスコンピューティング (HPC) ワークロードであり、1つのプロジェクトデザインツリーで最大100k以上のコンピュートコアが並列ジョブを実行していることを確認できます。ワークロードの例としては、Synopsys VCS、Cadence Xcellium、Siemens EDA ModelSimシミュレータを実行するフロントエンド検証シミュレーションや、Synopsys PrimetimeやFusionコンパイラなどのチップタイミング解析ジョブなどがあります。多くの場合、デザイン仕様、デザイン・キャプチャー、制約付きランダム・シミュレーション、合成、配置配線 (P&R) 、パワー・パフォーマンス・エリア (PPA) 最適化、フロア・プランニング、チップ仕上げの間で25以上のEDAデザイン・ツールが使用されています。

チップ設計における高度な製造テクノロジにより、EDAワークロード向けのコンピューティングとストレージの増強が常に求められています。容量は事実上無限大で、幅広いコンピューティングとストレージを備えたMicrosoft Azureは、EDAに最適です。Synopsys、Cadence、Siemens EDAなどのEDAベンダーは、EDAワークフローを最適化してクラウドやAIの手法を活用し、大規模な並列コンピューティングリソースを活用してジョブの完了時間を短縮し、結果が得られるまでの時間を短縮し、バグを特定して修正しています。

EDAのお客様は通常、EDAワークロード用に、16~24ノードの大規模なクラスタでハイパフォーマンスなNetApp ONTAPオールフラッシュ (SSD) システムを使用しています。EDAワークロードでは、多くの場合、大量のメタデータとアクセスの並列化が原因で、原因ストレージコントローラのボトルネックが発生します。オンプレミスでは、NetAppのベストプラクティスとして、複数のONTAPノードにまたがるFlexGroupを使用し、コンピューティングサーバの負荷を複数のデータLIFに分散することを推奨します。

このドキュメントでは、利用可能なすべてのオプションを最大化して最適化し、拡張性とスループットを最大化する方法について説明します。NetAppでは、アクティブなEDAワークロードのスループットとIOPSには最大設定を使用することを強く推奨します。

注：このドキュメントでは、階層1（最高のパフォーマンス）構成のみを対象としており、階層2のバックアップ階層構成のベストプラクティスについては説明しません。

ハイブリッドクラウドでのEDA向けリファレンスアーキテクチャ

EDAストレージのパフォーマンス要件

NetAppの最大規模のお客様の中には、100kを超える並列コンピューティングコアを使用しているお客様もいます。大量のファイルを大量に処理する大量のメタデータワークロードを並列処理することで、スケールアウトのパフォーマンスを最適化することが重要になります。設計仕様から設計、タイミング、消費電力、面積最適化、チップ仕上げまで50種類以上のツールが半導体開発プロセスで使用されているため、単一のEDAワークロードはありません。

NetAppは、創業当初から半導体業界のデータ管理とストレージのリーダーであり続けています。オンプレミスのNetAppのお客様は、最新のNetApp ONTAPストレージオペレーティングシステムを実行するNetApp AFF A700sおよびAFF A800クラスのオールフラッシュストレージアプライアンスを使用しています。長年にわたり、A700sやA800システムに見られるような大規模なCPUとメモリ構成を備えたオールフラッシュSSDベースのシステムがEDAワークロードにメリットをもたらすことが実証されました。また、FlexGroupボリュームのスケールアウトパフォーマンスにより、並列ジョブ設計チームが実行できる数が劇的に増加し、EDAジョブの総ターンアラウンド時間が短縮されることも、お客様のワークロードから示されています。

Cloudにより、EDA設計フローでは、利用可能な最新のコンピューティングコアを使用して、ほぼ無制限にコンピューティングを拡張できます。そのため、基盤となるストレージおよびデータ管理システムは、クラウド規模の並列ワークロードをサポートするように設計する必要があります。

本ドキュメントの推奨事項は、1~3ミリ秒以上のレイテンシを維持しながら、最大限の並列処理パフォーマンスを実現するようにCloud Volumes ONTAPを設定することに重点を置いています。コストの最適化は、スケールアウトパフォーマンスに次ぐ2番目の要件です。パフォーマンス要件を決定したら、必要なパフォーマンスとコストのトレードオフに対処できます。

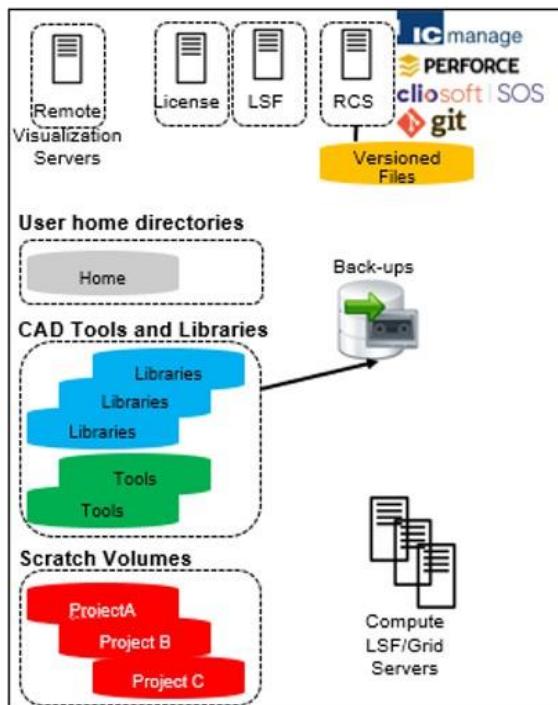
EDAボリュームとディレクトリ

EDAワークフローは、1つのワークフローで多数の異なるボリュームのデータをマウントして使用します。すべてのデータボリュームが同じ特性を持っているわけではなく、各ボリュームがEDAフローでどのように使用されているかを把握しておくと、これらの各データインフラリソースの利用率とランタイムパフォーマンスを最適化するのに役立ちます。

たとえば /mnt/tools、/mnt/libs、は /mnt/libraries、デザインフローで使用されるEDAツールとデザインライブラリの場所を表す一般的なストレージボリューム名です。これらのボリュームには、多くの場合、数十から数百のバージョンのツールやツールがインストールされています。ツールやライブラリは、多くの場合、CADチームによってインストールおよび一元管理され、複数のチームやプロジェクトで共有されます。

ITチームやストレージチームは、通常、ビジネスユニット、チーム、プロジェクト、特定の設計フローのニーズに基づいてストレージをプロビジョニングします。設計フローは企業ごとに固有であるため（ただし、固有ではありません）、データの種類と要件を分類する目的でデータとデータの管理を一般化できます。

図1) 一般的なオンプレミス設計インフラ

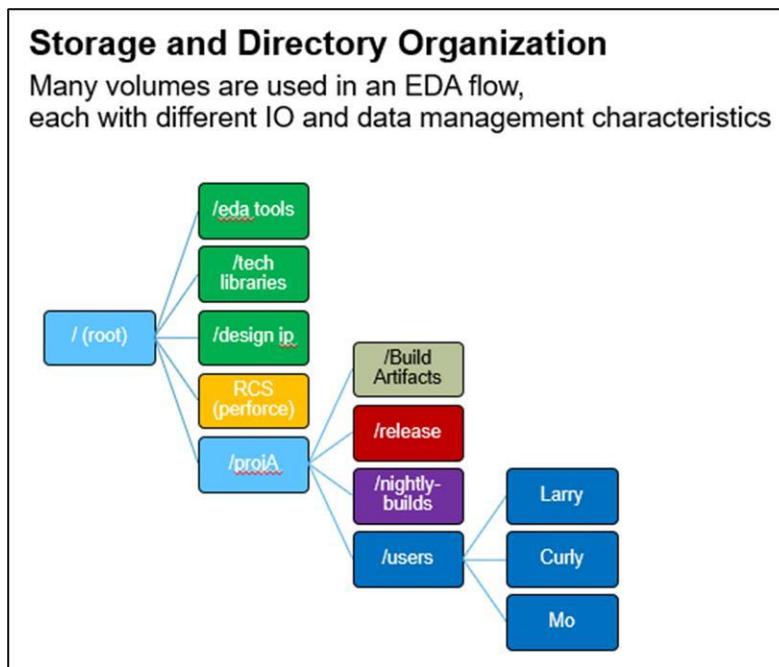


前の図のオンプレミス設計インフラには、通常、次のコンポーネントが含まれています。

- LSF / グリッドサーバのコンピューティング
 - 均一な構成のペアメタルLinuxサーバ。
 - コンピューティングパフォーマンスが最適化されています。
 - さまざまなコア数とメモリサイズ。
- ロードシェアリング (LSF) - ジョブキューイングを管理します。
 - リソースの最適な利用と公平な共有を実現します。

- IBM Spectrum LSF、Univa Grid Engine (Sun Grid Engine)、Slurm。
- FLEXlmライセンスサーバー
 - 最も一般的なEDAライセンスマネージャ。
 - EDAジョブは実行時にライセンスをチェックアウトします。すべてのライセンスが使用されると、追加のジョブは待機するか失敗します。
 - 人件費について、EDAツールライセンスはチップ設計の次の最も高価なコンポーネントです。EDAライセンスは、ツールに応じて、ライセンスあたり5,000ドルから10万ドルまで実行できます。
- リビジョン管理システム (RCS)
 - PerforceとICManageは非常に人気があります。
 - Clio SOS、Synchronicity Design Sync、IBM ClearCaseも一般的に使用されています。
 - Gitまたはアーティファクトクラスのソースとアーティファクトの管理はあまり一般的ではありません。
- データロギング、データベース、データマイニング、AI
 - ログファイルのマイニングやレポート作成には、SplunkやElasticsearchなどのツールが使用されます。
 - データベースはジョブと指標の収集に使用されます。
 - 設計および設計フローの最適化のためのAIの使用が増加しています。
- リモートデスクトップ可視化サーバー
 - ハミングバード、VNC、ターミナルサービス

図2) EDAストレージ構成



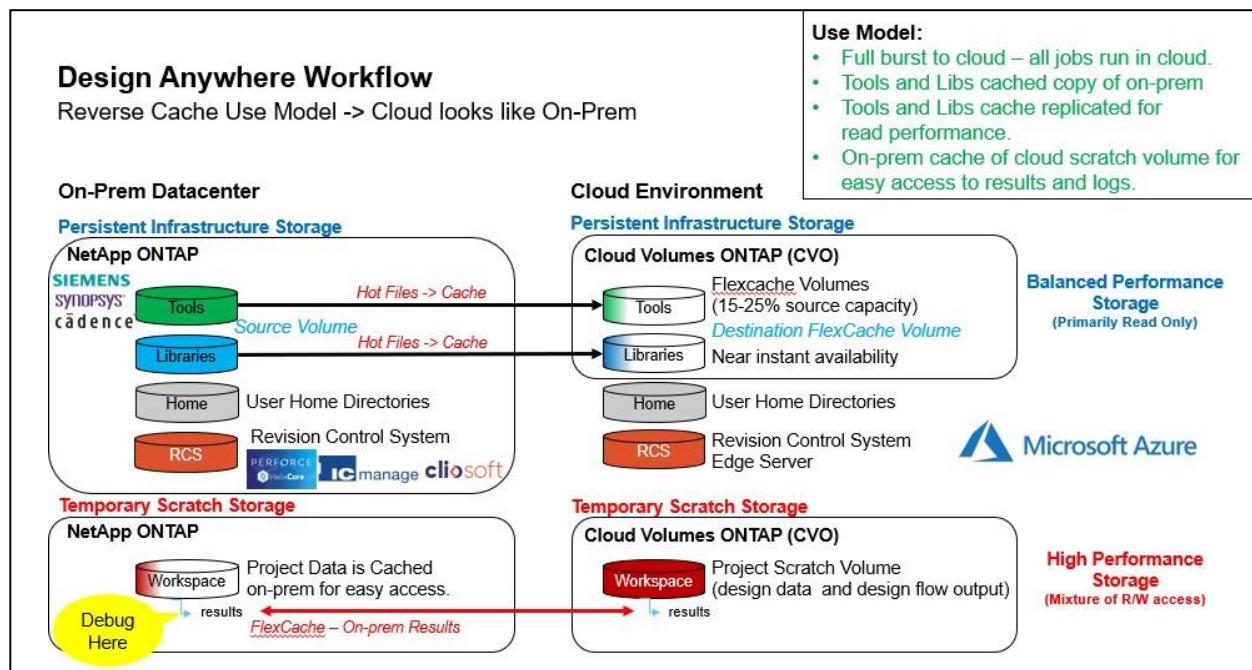
EDAストレージ統合では、通常、次のコンポーネントを使用します。

- 読み取り専用ツールとライブラリボリューム
 - EDAツール (Synopsys、Cadence、Siemens、Ansysなど)
 - テクノロジライブラリ (TSMC、Global Foundryなど)
 - デザインIP (ARM、Synopsys、Cadenceなど)
 - 新しいツールまたはライブラリがインストールされている場合を除き、読み取り専用です。
 - DNSマウントロードバランシングを使用したキャッシングまたはボリュームレプリケーションでは、読み取りパフォーマンスが向上します。

- リビジョン管理システム (RCS)
 - Perforce、ICManage、ClioSoft SOS、ClearCase、またはSynchronicity。
 - Gitと同様ですが、非常に大きなファイルや小さなファイル、テキスト、バイナリに適しています。
 - dbase用のローカル接続SSD (SAN)、ロギング、バージョン管理されたファイルストア用のジャーナリングNFSの組み合わせ。
 - デザインファイルとビルドアーティファクトの両方が非常に大量に存在します。データをPBまで拡張可能
 - マルチサイトの導入やレプリケーションと同様に、コードのチェックアウトに時間がかかることも日々の課題です。ビルドアーティファクトが大きくなるにつれて、データ転送時間も長くなります。
- ビルドアーティファクト
 - ビルドアーティファクトは、EDAとソフトウェアビルドプロセスの再利用可能な出力です。
 - ネットリスト、デザイン・データベース、コンパイル・イメージ、生成されたファイル・リスト、ログ・ファイル、レポートはすべてビルドアーティファクトとみなされる場合があります。
 - ビルドアーティファクトの管理方法は、チームや企業によって異なります。PerforceやICManageなどのツールを使用してバージョンを作成し、アーティファクトを再配布する場合もあります。ファイルシステムに保存されているものもあります。
- リリースビルドボリューム
 - 最終チップ出力のリリース候補がここで生成されます。1つ以上のボリュームを指定できます。通常は設計リードまたは一元的な構築チームによって管理されます。多くの場合、Jenkinsのようなツールによって自動化されます。
 - すべてまたは多くのデザインツールは、ソースファイルの1つのコピーに対して実行されます。
 - データ管理は、DR、バックアップ、HAなどの重要な要素です。
- 夜間またはCIビルド
 - これらは、通常、設計リーダーまたは中央のビルドチームによって管理される、定期的な自動化されたビルドです。これらのジョブは、多くの場合、Jenkins、Bamboo、CircleCIなどのCIツールを使用して自動化されます。
 - これらのビルドでは、スクラッチボリュームまたは通常の保護対象ボリュームが使用されます。
- ユーザワークスペース
 - ユーザワークスペースは、ほとんどの場合、スクラッチボリューム上にプロビジョニングされます。これは、開発者が設計の作成と編集、テストの作成と実行、および設計の品質と機能を向上させるための反復作業のほとんどを行う場所です。
 - 多くの場合、複数のユーザが共有する大容量ボリュームをプロビジョニングします。多くの場合、1人のユーザがボリューム全体を占有しないように、個々のクオータを使用します。

注：通常、新しいジョブを開始する前に、使用可能なストレージ容量を確認する必要があります。ユーザレベルと管理レポートが重要です。

図3) ハイブリッドクラウドアーキテクチャとバーストツークラウドアーキテクチャ



EDAデータ型

読み取りパフォーマンスに最適化されたツールおよびライブラリボリューム

Synopsys、Cadence、Siemens EDAなどのEDAツールは、/mnt/tools/synopsys、などのようにマウントされたツールディレクトリにインストールされます /mnt/tools/cadence。これらのインストールディレクトリには、さまざまなツールとさまざまなバージョンのツールを含めることができます。

同様に、/mnt/libs/tsmc、/mnt/libs/arm/mnt/libs/dolphin次の図に示すように、サードパーティの設計技術ライブラリライブラリは、などのようにマウントされたディレクトリにインストールされます。これらのライブラリには、デザインIP (USBやCPUコアなど) が含まれているか、TSMC、Intel、Global Fundries、UMC、Samsungなどのデバイスマーカーの技術ライブラリが含まれています。これは、プロセスノードのトランジスタ/セル表現に関連する物理的な電力、面積、およびパフォーマンスデータを記述します。技術ライブラリには、通常、増え続けるデータを含む数千のファイルが含まれています。

図4) 一般的なEDAツールとライブラリUNIXディレクトリ構造

EDA Tools (/mnt/tools/...)

```
jmichael@MICHAEL1-PC:~/apple_csg$ find tools/
tools/
tools/synopsys
tools/synopsys/primestime_2020.09.1
tools/synopsys/primestime_2021.06.2
tools/synopsys/primestime_2020.03.2
tools/synopsys/many_other_tools_and_versions
tools/mentor
tools/mentor/many_other_tools_and_versions
tools/cadence
tools/cadence/xcelium_2021.09.1
tools/cadence/xcelium_2022.06.1
tools/cadence/xcelium_2021.09.3
tools/cadence/voltus_2021.03.1
tools/cadence/xcelium_2020.10.4
tools/cadence/many_other_tools_and_versions
tools/cadence/voltus_2022.06.1
tools/cadence/voltus_2021.09.1
```

EDA Libraries (/mnt/libraries/...)

```
jmichael@MICHAEL1-PC:~/apple_csg$ find libraries
libraries
libraries/tsmc
libraries/tsmc/c05
libraries/tsmc/c05/stdcell_lib_c05_and_many_more_versions
libraries/tsmc/c03
libraries/tsmc/c03/stdcell_lib_c03_and_many_more_versions
libraries/tsmc/c10
libraries/tsmc/c10/stdcell_lib_c10_and_many_more_versions
libraries/tsmc/c07
libraries/tsmc/c07/stdcell_lib_c07_20220101
libraries/tsmc/c07/stdcell_lib_c07_and_many_more_versions
libraries/synopsys
libraries/synopsys/usb
libraries/synopsys/usb/usbcore_many_more_versions
libraries/synopsys/mac
libraries/synopsys/mac/mac_many_versions
libraries/arm
libraries/arm/arm10
libraries/arm/arm10/arm10_many_versions
libraries/arm/arm10/arm_many_versions
libraries/arm/arm11
libraries/arm/arm11/arm_many_versions
```

内部ライブラリーまたはデザインIPは、デザインで使用される別のクラスのデータです。これらは、社内のあるチームによって開発され、完全なチップ設計に含めるために別のチームに提供されたサブシステムまたはブロックである可能性があります。たとえば、プロセッサチームによって設計されたARMコアは、コアをモバイルデバイスに統合するチームによって提供され、使用される可能性があります。

ツールとライブラリの主な特徴は、データが通常読み取り専用または読み取り専用大部分であることです。書き込みは、新しいバージョンのツールまたはライブラリがインストールされているときに発生しますが、ツールとライブラリは自動的に上書きされます。新しいツールまたはライブラリのバージョンは、通常、以前のバージョンの隣にインストールされます。そのため、インフラを最適化して読み取り専用のパフォーマンスを最適化することで、このメリットを活用できます。これには、I/Oレイテンシを改善するためにコンピューティングサーバのデータキャッシュを最大化するNFSマウントオプションが含まれる場合があります。

そのため、インフラを最適化して読み取り専用のパフォーマンスを最適化することで、このメリットを活用できます。これには、コンピューティングサーバのデータキャッシュを最大化してI/Oレイテンシを改善するNFSマウントオプションが含まれる場合があります。

ツールとライブラリにはEDAデザインフローがアクセスします。コンピュートファームで同時に実行されている10,000~10,000個の並列ジョブから非常に高い同時アクセスが発生する場合があります。そのため、I/Oパフォーマンスと拡張性が重要です。タイミング解析ツールであるSynopsys PrimeTimeのようなツールは、数千台のサーバから数千のライブラリファイルを同時にロードします。その結果、読み取り負荷が高く、原因のパフォーマンスがボトルネックになる可能性があります。本ドキュメントで紹介するリファレンスアーキテクチャでは、FlexCacheを使用して複数のFlexCachedツールボリュームやライブラリボリューム間でNFSマウントのロードバランシングを行うことで、読み取りパフォーマンスを向上させる方法について説明します。

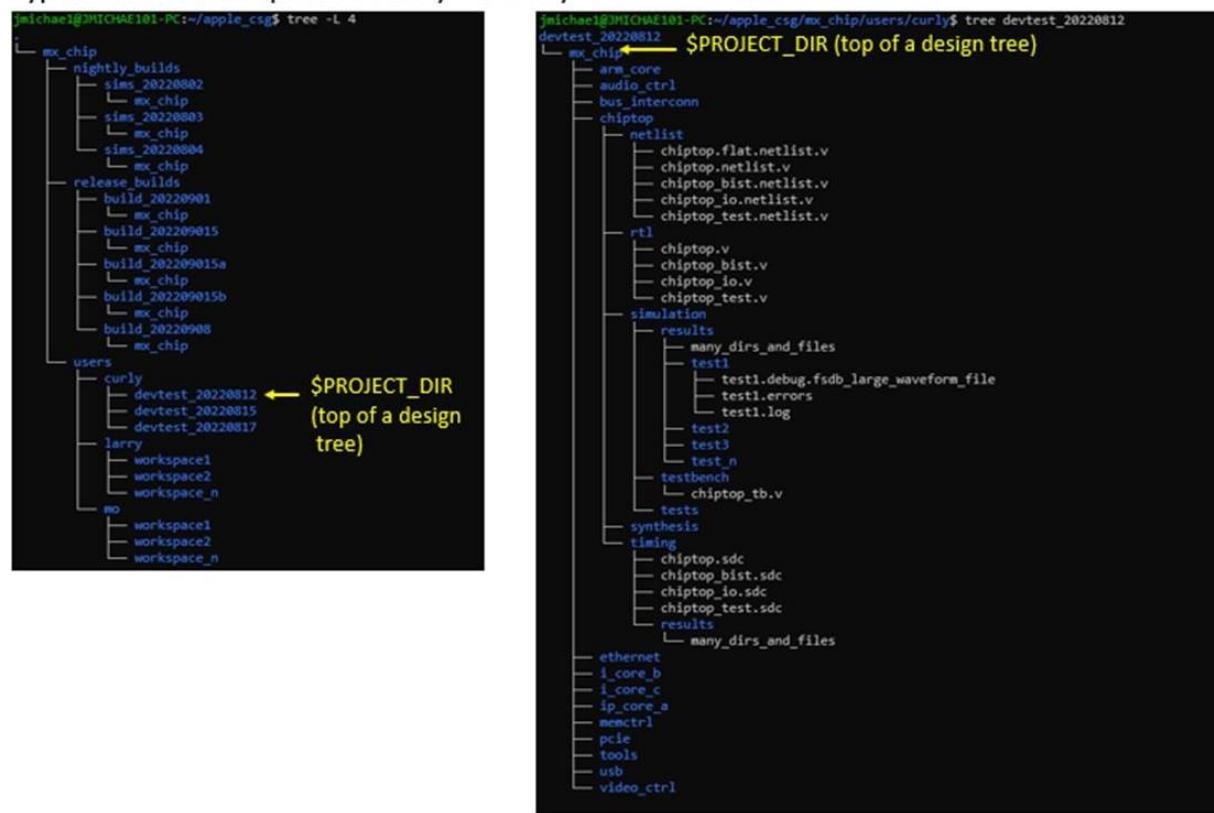
読み取り/書き込みパフォーマンスを最大化するように最適化されたスクラッチボリューム

スクラッチボリュームは、ユーザーワークスペース、夜間の回帰、CI/CDビルドフロー、およびその他多くの開発目的で使用されます。「スクラッチ」という用語は、データが一時的なものであるか、または変更率が非常に高いため、従来のバックアップやアーカイブは不要であるという考えに由来します。

スクラッチは、積極的な開発が行われている場所です。エンジニアはEDAジョブを起動し、完了するまで待機して、設計変更後や新しい機能テスト後にテストが成功したか失敗したかを確認できます。機能の変更や最適化の後、チップの電力性能と面積 (PPA) 特性をチェックするために、何千ものジョブの結果を待っているかもしれません。パフォーマンスは重要であり、容量とコンピューティングの拡張性の両方が重要な要素です。

図5) EDAプロジェクトのUNIXディレクトリ構造の例

Typical EDA scratch space directory hierarchy



Scratchボリュームは、読み取り/書き込みのパフォーマンスとスループットが最大になるように最適化する必要があります。最新のチップ設計フローでは、さまざまなI/Oプロファイル（読み取り/書き込み比率、大容量ファイルと小容量ファイル、シーケンシャルアクセスとランダムアクセス）を備えた100種類ものツールを使用でき、状況は大きく異なる場合があります。これらの環境は複数のフローまたはチームで共有されることが多いため、1つのワークフローに対して最適化できない場合があります。

スクラッチボリュームで実行されるワークフロードは反復性が高いため、多くの場合フォールトトレラントです。サーバが停止した場合や、ストレージのレイテンシが高いためにジョブがハングアップした場合は、多くの場合、ジョブを再起動するだけです。データ損失やジョブの再起動は理想的ではありませんが、通常は致命的ではありません。最悪の問題は、ジョブの再開によって失われた時間です。ジョブによっては数分かかるものもあれば、完了までに数時間から数日かかるものもあります。

インフラの問題に対する耐性はワークフローによって異なります。

以下のEDAベストプラクティスは、ストレージパフォーマンスとインフラ内の障害許容度のバランスを取る方法についてガイダンスを提供します。

最大のデータ保持性を実現するように最適化されたリリースまたはテストリリースのボリューム

ボリュームのもう1つの一般的なクラスは、リリースボリュームまたはプレリリースボリュームです。チップ設計が設計クロージャーとテープアウト（製造工程へのリリースとも呼ばれる）に近づくと、リリース候補と見なされるチップのビルドが実行されます。リリース候補が完全なチップデザインフローを通過する際にすべての品質チェックに合格すると、デザインはリリースの準備が整います。

これらの作業領域は、より重要であり、インフラの障害に対するフォールトトレランスが低いと見なされる可能性があります。

ビルドアーティファクト

ビルドアーティファクトには、EDAまたはソフトウェア開発ワークフローから生成された任意の数のデータを含めることができます。合成ツールを使用してRegister Transfer Level (RTL) コードから生成されるチップまたはIPネットリストは、ビルドアーティファクトです。生成されたゲートレベルのネットリストは、ゲートレベルのシミュレーションのために検証チームによって再利用される可能性があります。また、エミュレーションチームやタイミング解析チーム、フロアプランニングチームが再利用することもあります。現代の半導体は、ソフトIP、ファームIP、ハードIPの両方で構成されています。これらのIPは、あるチームによって開発され、別のチームに提供されて設計に統合されます。IPまたはサブシステム開発チームの成果物は、統合のために別のチームに提供されるIP成果物として再利用される可能性があります。

企業は異なる方法でビルドアーティファクトを管理、共有しています。一部のチームは、Perforce/ICManageなどのツールを使用してビルドアーティファクトをチェックインおよびバージョン管理し、Perforce Helixレプリケーションソフトウェアまたはエッジサーバーを使用して、チームや開発サイトにIPを配布します。データボリュームやディレクトリをレプリケートすることでビルドアーティファクトを配布する企業もあります。また、シンボリックリンクを使用してファイルやディレクトリを開発領域で共有するものもあります。NetAppをご利用の多くのお客様は、SnapMirrorやFlexCacheなどのデータレプリケーションツールを使用して、アーティファクトを複数のサイトに分散しています。

EDAのパフォーマンスとクラウドの拡張性

単一ジョブのパフォーマンスを実測時間で測定

EDAジョブのパフォーマンスは、通常、実測時間で測定されます。1つのチップ設計シミュレーションを実行するのにどれくらいの時間がかかりましたか？タイミング解析または間取り図を完了するのにどれくらいの時間がかかりましたか？

改善ジョブタイムは、設計エンジニアとEDAライセンスの生産性と利用率を向上させるため、非常に重要です。技術者とEDAライセンスは、チップ開発コストの90%以上を占めています。5~10%の壁時計の改善は巨大な勝利と見なされ、お金を費やす価値があります。EDAインフラのコストはチップ開発コストの約10%にすぎず、ストレージはITコストの約4%にすぎません。そのため、エンジニアやライセンスの生産性と利用率が向上すれば、ストレージパフォーマンスにコストをかけることで、本来の効果が得られます。

すべてのジョブの総ターンアラウンド時間によって測定されたパフォーマンス

EDAのパフォーマンスは、総ターンアラウンド時間 (TTT) でも測定されます。これは、ジョブのコレクション全体が完了するまでにかかる時間です。たとえば、1000回のすべてのシミュレーションの実行が開始から終了するまでにどれくらいの時間がかかりますか。この質問に対する回答は、ジョブテストが連続して実行されたか、並行して実行されたか、または何らかの組み合わせで実行されたかによって異なります。LSF、Grid Engine、Slurmなどの負荷分散ツールを使用すると、エンジニアは大量のジョブを送信して並行して実行することができます。並行して実行できるジョブが多いほど、通常はすべてのジョブが完了するまでの時間が短縮されます。

次の例では、20,000個のジョブを実行しています。

表1) すべてのジョブの総ターンアラウンド時間で測定されたパフォーマンス

	合計ジョブ数	最大並列ジョブ数	平均ジョブ実行時間 (分)	TTT (分)	TTT (時間)
例1	20,000	1,000	10	200	3.33
例2	20,000	5,000	10	40	0.67 (5倍高速)

上の表からわかるように、5000個のジョブを並行して実行すると、ジョブのターンアラウンド時間が5倍に短縮されます。これは、バグを見つけるのが5倍速く、設計クロージャーのイテレーションを実行するのが5倍速くなることを意味します。そのため、エンジニアの生産性は5倍になります。

EDA設計アーキテクチャを検討する場合、最大の並列パフォーマンスと拡張性を実現するための最適化は重要なトレードオフです。クラウドは従来のデータセンターの固定リソースと比較して広大で柔軟性に優れているため、クラウドコンピューティングでは並列コンピューティングを無制限に拡張できます。また、クラウドはTTTと生産性を大幅に向上させます。

より多くのジョブを並行して実行する機能は、単一ジョブのウォールクロック時間のパフォーマンスを最適化するのと同じくらい重要です。

エンタープライズ規模のハイブリッドクラウドデータ管理アーキテクチャ

クラウドインフラがスケールアップして複数のEDAワークフロー、プロジェクト、チームに対応できるようになると、次の考慮事項への対応が必要になる場合があります。

スケーリングツールとライブラリ

ツールとライブラリは通常、オンプレミスで共有リソースであるため、クラウド内に配置する必要があります。高度な半導体製造プロセス（7nm、5nm、3nm）は設計の複雑さと規模を増しており、トランジスタの性能、電力、形状を表す複雑さがファイルの大幅な増加を引き起こしているため、設計技術ライブラリは非常に大きくなっています。そのため、ライブラリの読み込み時間が課題となっています。

ツールフローが10k～100kのサーバコアで同時に起動するときにライブラリボリュームが必要になると、ストレージファイラーでブートストームが発生し、読み取りレイテンシが高くなる可能性があります。これは、複数のデザインフロー、プロジェクト、チーム間でツールとライブラリのボリュームを共有することでも複雑になります。FlexCacheボリュームを拡張して、読み取り負荷を並列コンピューティングコア全體に分散させることで、ライブラリとツールのロード時間を大幅に短縮できます。

ツールボリュームとライブラリボリューム（Cloud Volumes ONTAPインスタンス）は複数のプロジェクトやチームで共有されるため、Cloud Volumes ONTAPをHA構成に導入して信頼性、耐久性、アップタイムを向上させることができます。シングルノードのCloud Volumes ONTAPファイルシステムではアップグレード時にダウンタイムが発生しますが、HAインスタンスはアップグレード中に2つのノードのどちらかに一時的にフェイルオーバーできます。HA構成を使用すると、アップグレード時や障害状態時のツールやライブラリボリューム原因のダウンタイムを最小限に抑えることができます。

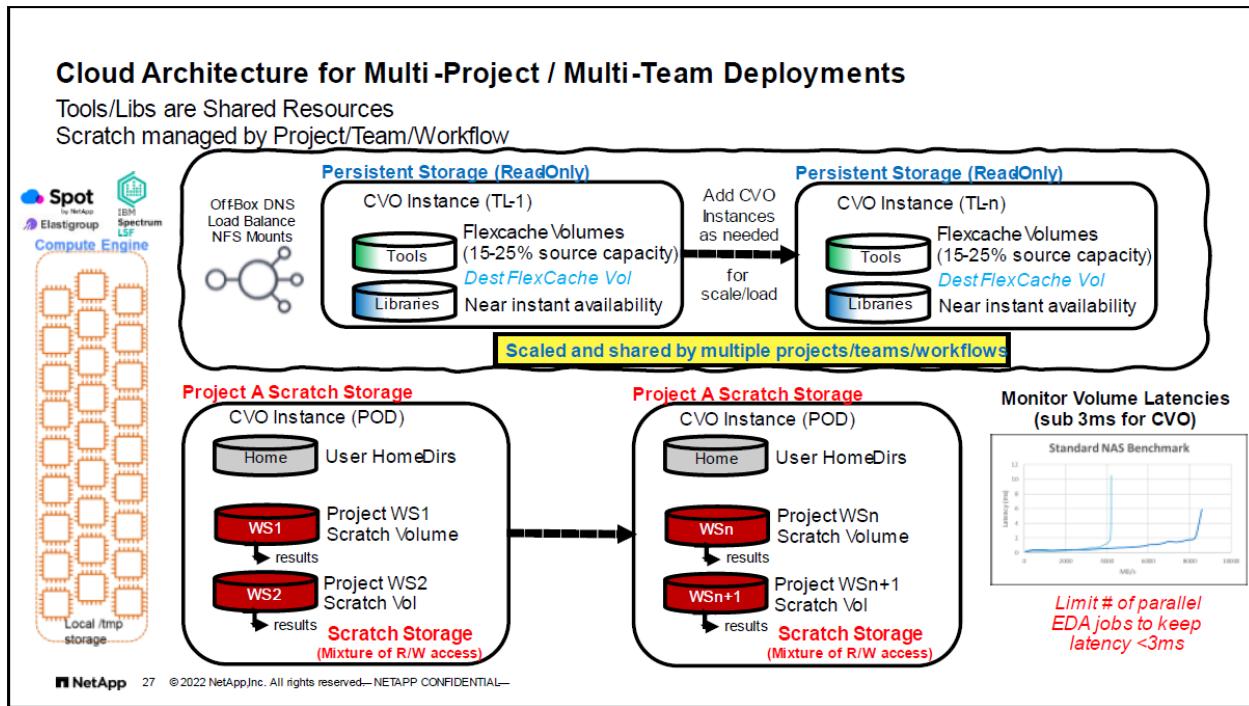
大規模なボリュームをスクラッチ

スクラッチボリュームには、最大の読み取り/書き込みパフォーマンスが必要です。シングルノードCloud Volumes ONTAP構成では、2ノードHA構成と比較して書き込みパフォーマンスが向上する場合があります。ONTAPのアップグレード時や一時的な障害状態時のシステム停止を最小限に抑えるために、フロー、チーム、またはプロジェクト単位でCloud Volumes ONTAPインスタンスを導入できます。Cloud Volumes ONTAPのアップグレードは、フロー、プロジェクト、またはチーム間で調整するよりも、単一のフロー、チーム、またはプロジェクトの所有者と簡単に調整できます。

次のアーキテクチャ図は、増大するツールやライブラリの読み取りワークロードに対応するために、共有Cloud Volumes ONTAPファイルシステムをどのように拡張できるかを示しています。ますます増加する設計ワークロード要件に合わせて柔軟に拡張できるAzure VMが増えるにつれて、ツールとライブラリのFlexCacheボリュームを迅速にスケールアップし、UNIXマウントポイントをDNSロードバランシングしてパフォーマンスを最適化することができます。Azure VMのスケールダウンに伴い、FlexCacheボリュームもスケールバックできます。

次の図は、マルチプロジェクト環境とマルチチーム環境のクラウドアーキテクチャを示しています。ツールとライブラリは共有リソースであり、スクラッチはプロジェクト、チーム、ワークフローによって管理されます。

図6) マルチプロジェクトおよびマルチチーム環境向けのクラウドアーキテクチャ



クラウドは、従来の固定オンプレミスリソースよりもはるかに柔軟なコンピューティングとストレージインフラを提供できます。プロジェクトのスケジュールやワークロードのニーズに基づいてプロジェクトのスケールアップやスケールダウンを行う場合は、Cloud Volumes ONTAPファイルシステムを迅速に作成して、ワークスペース、ワークフロー、またはプロジェクト単位でパフォーマンスの高いストレージを提供できます。

ファイルシステムが過負荷になったタイミングを特定するには、ストレージのレイテンシを監視することを推奨します。負荷の高いファイルシステムでは、レイテンシの増加がレイテンシ曲線の膝のように表されることがよくあります。前の図では、並列EDAジョブの数によって高いI/O負荷が生成されるまで、レイテンシは3ミリ秒未満のままであります。その結果、I/Oレイテンシは3ミリ秒を急速に上回るようになります。

ファイルシステムのレイテンシを監視することは、ファイルのパフォーマンスを公称（3ミリ秒未満のレイテンシ）に保ちながら実行できるEDAジョブの数を調整する良い方法です。プロジェクトのワークスペースやフローを複数のCloud Volumes ONTAPファイルシステムに分散させることは、ストレージワークロードの負荷を分散し、最適な低レイテンシI/Oを実現する優れた方法です。

Azure Cloud Volumes ONTAPセットアップのクイックスタート

このセクションの目的は、ハイパフォーマンスなハイスケールEDAワークロード向けにCloud Volumes ONTAPを最適化するための最新のベストプラクティスを文書化することです。各オプションと構成の詳細と説明は、次の章「[QuickStart Implementation Details](#)」のリンクを参照してください。

Cloud Volumes ONTAP作成のセットアップ

このステップは、ツール、ライブラリ、スクラッチに必要なCloud Volumes ONTAPインスタンスの数に応じて、少なくとも2回実行します。

- スケールアウトツールおよびライブラリのFlexCacheボリューム用のCloud Volumes ONTAPインスタンス。この構成は、非常に高い読み取り専用I/Oパフォーマンスと耐久性を実現するように最適化されています。NetAppでは、HAペアファイルシステム構成を使用することを推奨しています。

- Cloud Volumes ONTAPは、スケールアウトスクラッチボリュームのI/Oを最大化するように最適化されています。この構成は、大規模環境での読み取り/書き込みパフォーマンスに対して絶対最大I/Oスループットを実現するように最適化されていますが、ある程度の耐久性を犠牲にすることはできません。NetAppでは、シングルノードファイルシステム構成の使用を推奨しています。
または
- 高可用性と最大限のI/O、スケールアウトのスクラッチボリュームまたはリリースボリューム向けに最適化されたCloud Volumes ONTAP。この構成は、高可用性を維持しながら、大規模な読み取り/書き込みパフォーマンスで最大のI/Oスループットを実現するように最適化されています。この構成では、シングルノードCloud Volumes ONTAP構成よりも書き込みパフォーマンスが低下します。NetAppでは、HAペアファイルシステム構成の使用を推奨しています。

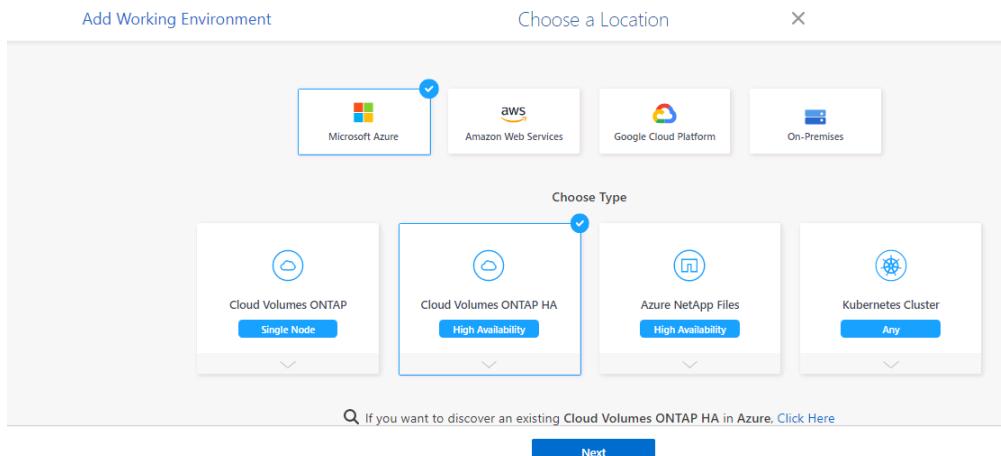
Cloud Managerを開く

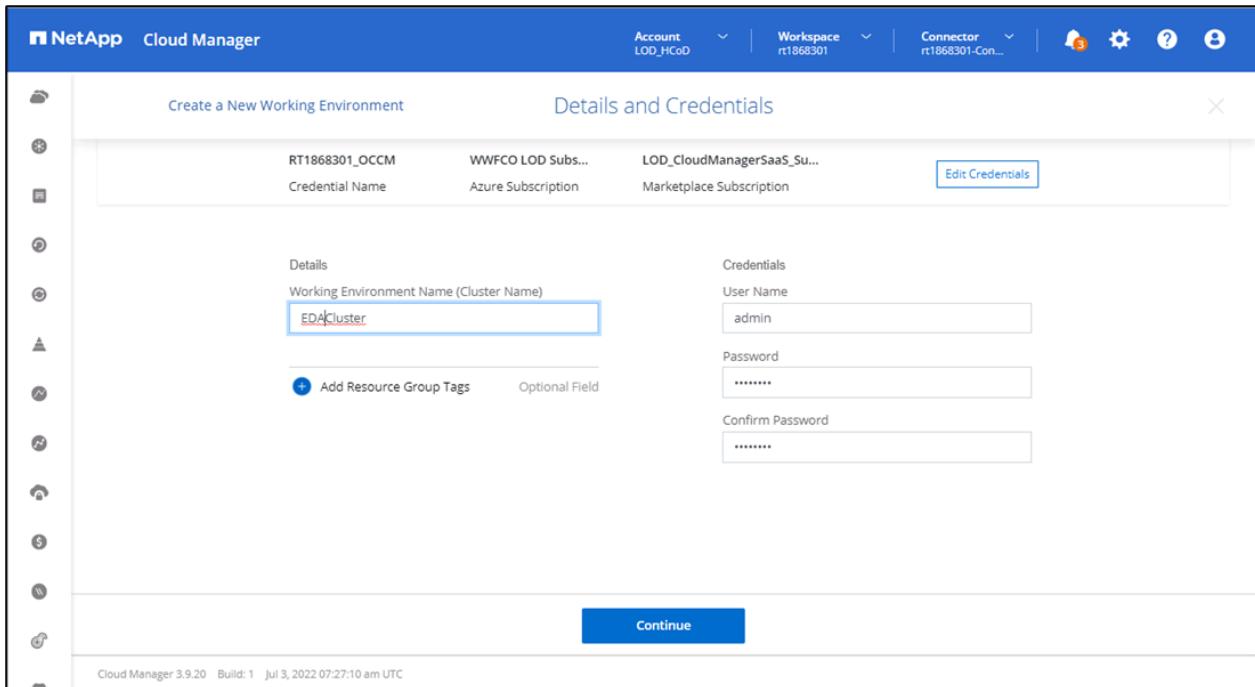
NetApp Support Siteのクレデンシャルを使用して[BlueXPアカウント](#)にログインします。これは、NetApp Active IQに環境を登録するための最初のステップです。

作業環境の追加

[Microsoft Azure]を選択し、必要なCloud Volumes ONTAPファイルシステムのタイプに基づいて[Choose Type]を選択します。

- Cloud Volumes ONTAPは、スケールアウトツールおよびライブラリFlexCacheボリュームのI/Oを最大化するように最適化されています。Cloud Volumes ONTAP HAを選択します。
- Cloud Volumes ONTAPは、スケールアウトスクラッチボリュームのI/Oを最大化するように最適化されています。Cloud Volumes ONTAP Single Nodeを選択します。
- 高可用性と最大限のI/O、スケールアウトスクラッチ/リリースボリューム向けに最適化されたCloud Volumes ONTAP。Cloud Volumes ONTAP HAを選択します。

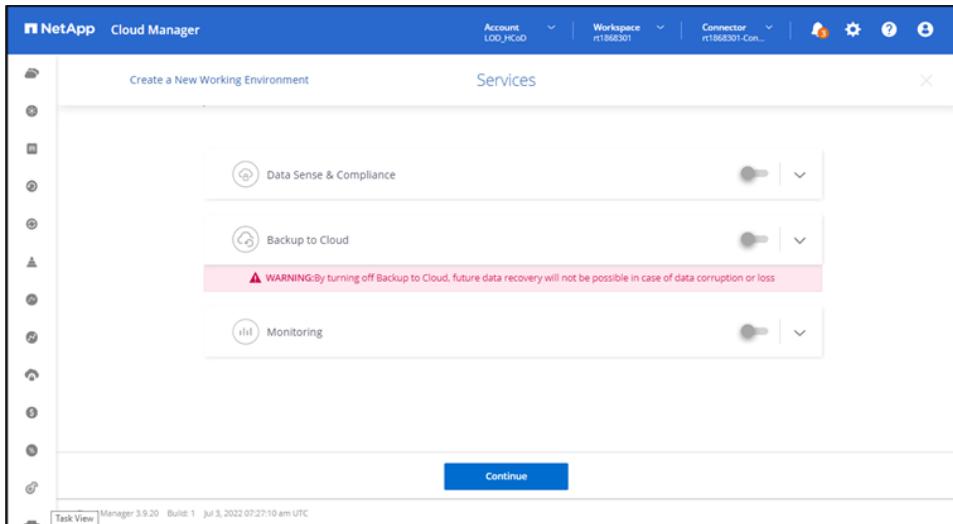




追加されたサービスを選択

この項目の詳細については、「追加サービス」のセクションを参照してください。

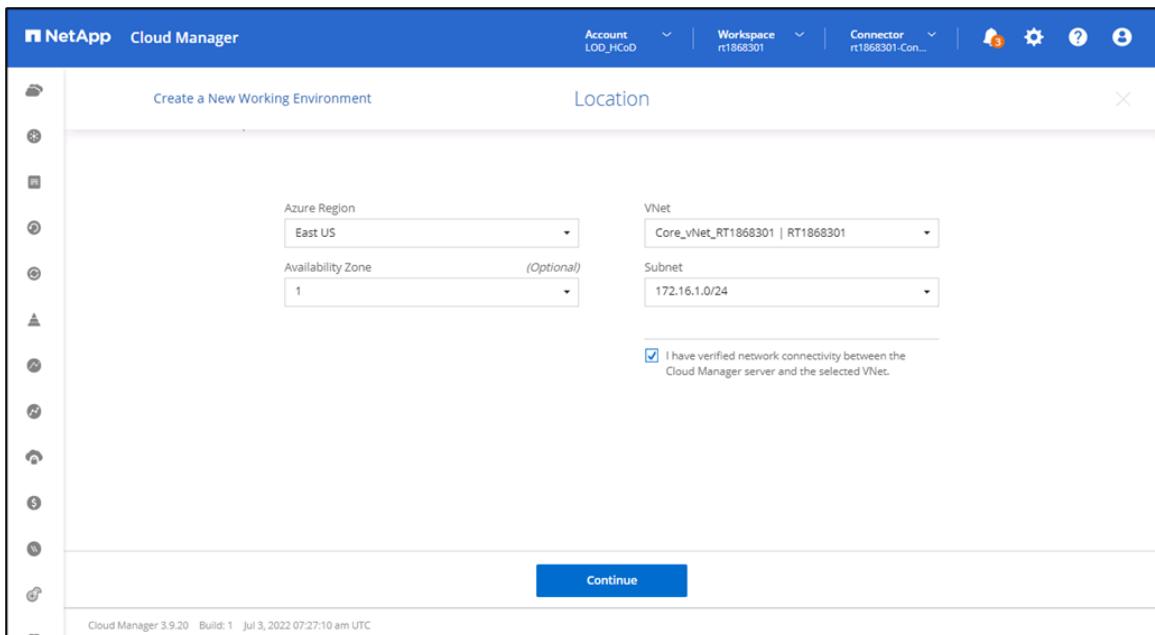
- **Data Sense**とコンプライアンス：オプション：パフォーマンスに大きな影響を与えます。をオフにします。
- クラウドへのバックアップ：オプション：パフォーマンスへの影響は最小限です。これはクラウドへの SnapMirrorサービスです。パフォーマンスへの影響はごくわずかですが、ほとんどの場合は不要です。
 - Cloud Volumes ONTAPは、スケールアウトツールおよびライブラリFlexCacheボリュームのI/Oを最大化するように最適化されています。[Backup to Cloud] - [Off]を選択します。
 - Cloud Volumes ONTAPは、スケールアウトスクラッチボリュームのI/Oを最大化するように最適化されています。[Backup to Cloud]-[Off]を選択します。
 - 高可用性と最大限のI/O、スケールアウトのスクラッチおよびリリースボリューム向けに最適化された Cloud Volumes ONTAP。[Backup to Cloud]を選択します（[On]はバックアップが必要な場合にのみ選択します。スクラッチボリュームは通常バックアップされません）。
- **監視**：オプション：監視によるパフォーマンスへの影響は最小限です。オフを選択します。この機能により、Cloud Insightsモニタリングが有効になります。



場所と接続の設定

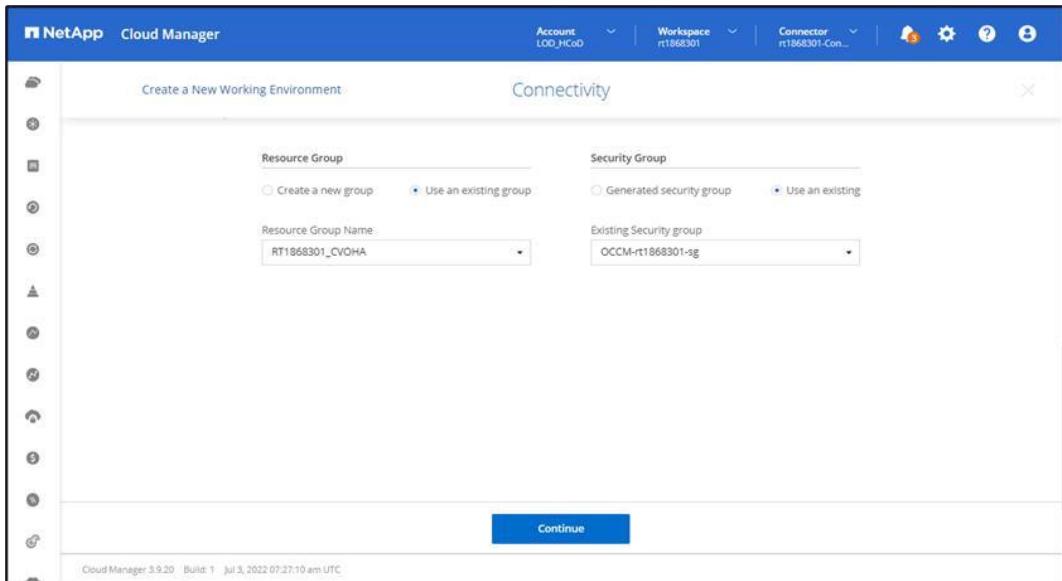
Azure Cloud Volumes ONTAPでは単一のアベイラビリティゾーン構成が使用されるため、書き込みパフォーマンスを最大限に高めることが推奨されます。

- **Azureリージョン** : EDAのベストプラクティスはありません。適切なAzureリージョンを選択します。
 - **アベイラビリティゾーン (AZ)** : このオプションは、Cloud Volumes ONTAPノードを設定するAzure AZを指定します。このオプションは、環境シングルノード環境のみです。
- 注：この設定はオプションです。優先AZを選択するか、ボックスを空のままにします。
- **VNet (VNet)** : EDAのベストプラクティスはありません。適切な指定Azure VNETを選択します。
 - **サブネット** : EDAのベストプラクティスはありません。指定した適切なAzureサブネットを選択します。



接続

NetAppでは、EDA固有のベストプラクティス要件は推奨していません。



NetApp Support Siteアカウントの管理

詳細については、次の章のクイックスタート実装の詳細 → NetApp Support Siteアカウントの管理を参照してください。



この機能を使用すると、NetApp Support SiteとのCloud Volumes ONTAP通信が可能になります。

1. [Manage NetApp Support Site Account]をクリックします。

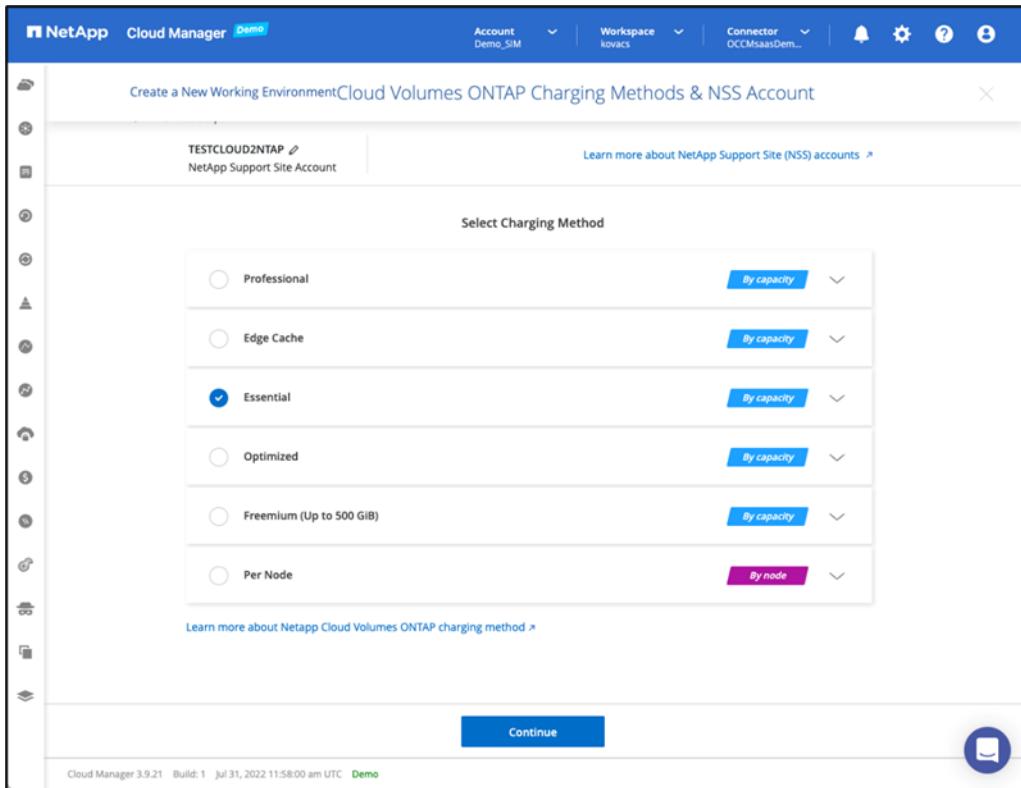
注：BlueXPも登録する必要があります。

2. BlueXP UIの上部にあるリンクをクリックします。

充電（ライセンス）方法を選択します。

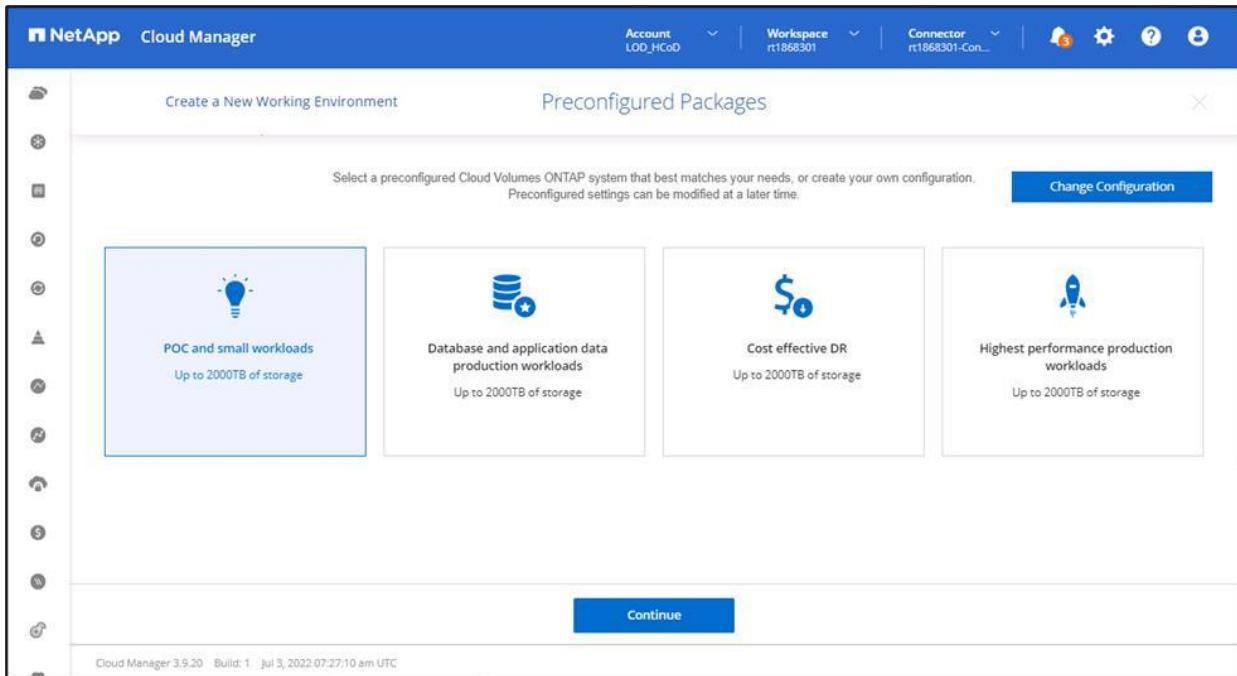
この問題の詳細については、「課金(ライセンス)方法の選択」セクションを参照してください。

クラウドへのバックアップは必要ないため、すべてのEDAユーザーがEssentialライセンスを使用しているわけではありません。Essentialsを選択します。



Cloud Volumes ONTAPインスタンスの設定

1. Cloud Volumes ONTAP設定変更構成を選択するには、[Change Configuration]をクリックします。これにより、特定のVMとディスクタイプを選択できます。



- 導入するCloud Volumes ONTAPバージョンを選択するには、[Change Version]を選択し、最新のONTAPバージョンを選択します。
- 導入するVMタイプを選択するには、次の表に基づいてVMタイプを選択します。

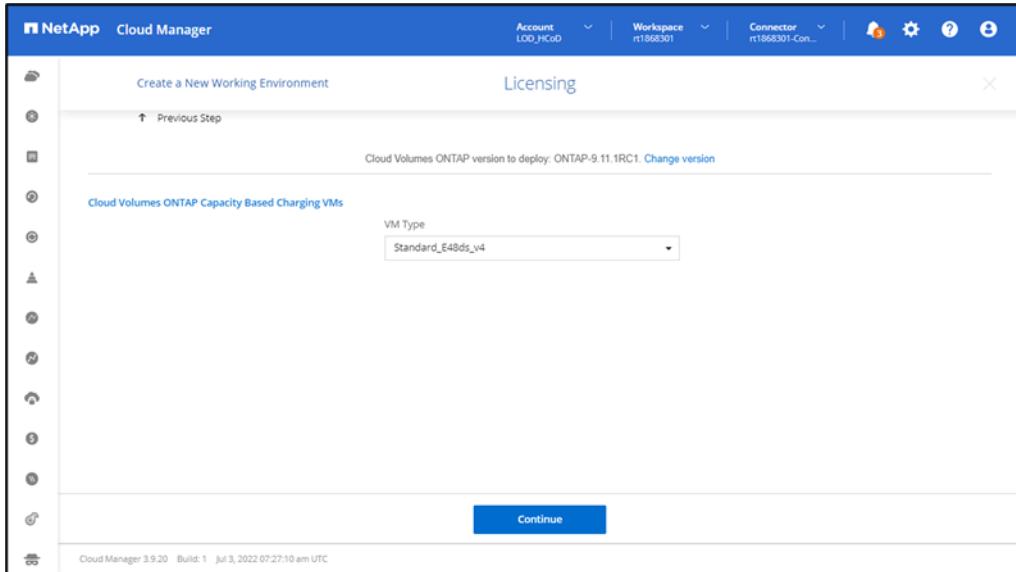


表2) シングルノードCloud Volumes ONTAP構成（VMタイプe48ds_v4に基づく）

必要な実効（使用可能）容量（TB）	VMタイプ	Azureのディスクサイズ	Cloud Volumes ONTAPインスタンスあたりの合計アグリゲート数	アグリゲートあたりのディスク数
40	E48ds_v4	4TB	1つのアグリゲート	11
80	E48ds_v4	8TB	1つのアグリゲート	11
120	E48ds_v4	8TB	2つのアグリゲート	11
180	E48ds_v4	8TB	2つのアグリゲート	12

次の表に、VMタイプe48ds_v4に基づくCloud Volumes ONTAP HAペアの構成を示します。

表3) アクティブ/アクティブHAペアCloud Volumes ONTAP構成（VMタイプe48ds_v4に基づく）

必要な実効（使用可能）容量（TB）	VMタイプ	Azureのディスクサイズ	Cloud Volumes ONTAPインスタンスあたりの合計アグリゲート数	アグリゲートあたりのディスク数
40	E48ds_v4	2TB	2つのアグリゲート	11
80	E48ds_v4	4TB	2つのアグリゲート	11
120	E48ds_v4	8TB *	2つのアグリゲート	11
180	E48ds_v4	8TB *	2つのアグリゲート	12

*このサイズには2つのデータアグリゲートが必要です。

注： さらに詳細な設定が必要な場合は、NetAppアカウントチームにお問い合わせください。

基盤となるストレージリソースを選択

詳細については、「EDAワークロードのCloud Volumes ONTAPサイジングに関する考慮事項」を参照してください。

- Cloud Volumes ONTAPに使用する基盤となるディスクタイプを選択します（上記の表を使用）。
- Premium SSDを選択します。

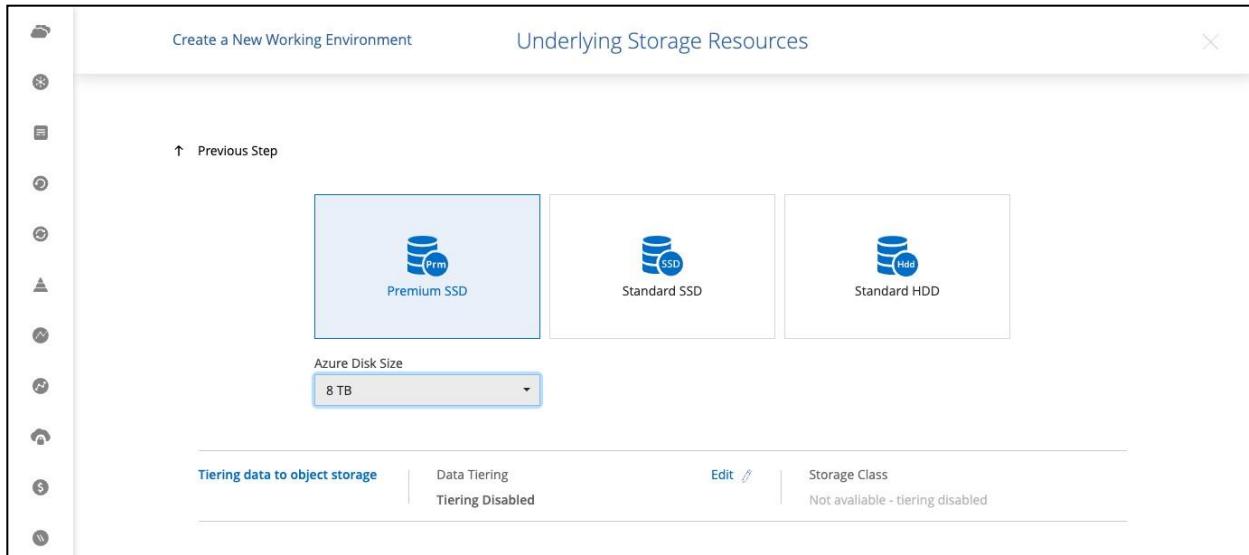
3. 単一環境またはHA環境に基づいて、対応する表からディスクサイズの値を選択します。

注：この手順では、1台のディスクを導入します。Cloud Volumes ONTAPの導入後にアグリゲートにディスクが追加されます。

を参照してください。

「POST Cloud Volumes ONTAP creation setup」を参照してください。

【続行】、【アグリゲートの変更】を選択する前に、次のようにBLOB階層化設定を編集してください。



オブジェクトストレージへのデータの階層化

使用期間ベースのコールドデータをオブジェクトストレージ階層に自動で階層化します。

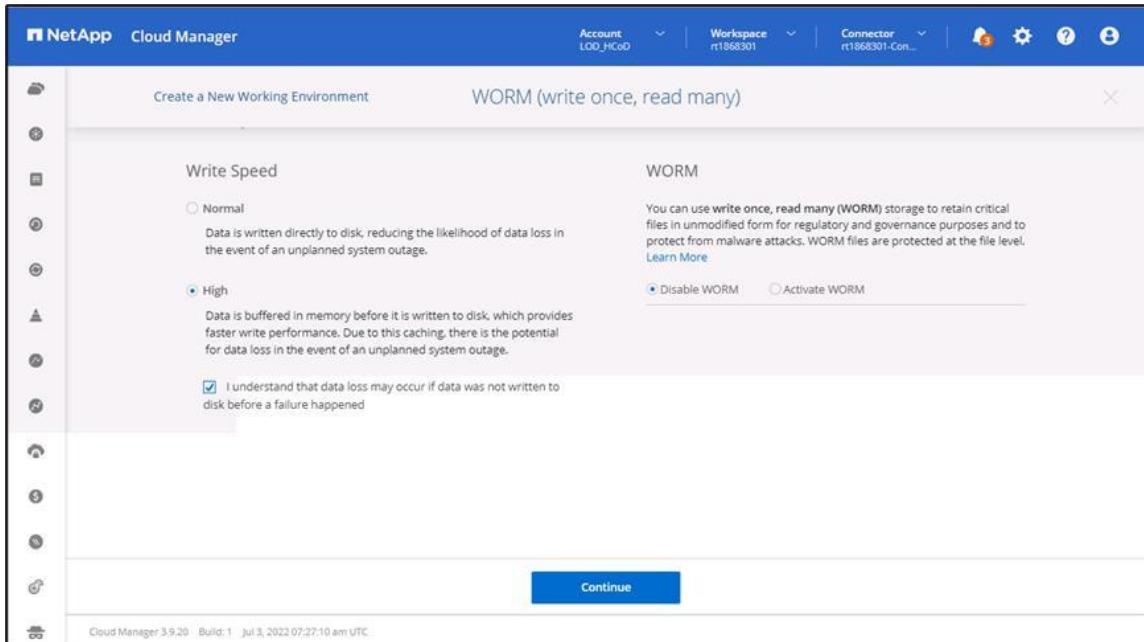
- スケールアウトツールとライブラリの**FlexCache**ボリュームのI/Oを最大化するように最適化された**Cloud Volumes ONTAP**
 - 【階層化が不要】を選択します。
 - ほとんどの場合、FlexCachedツールとライブラリボリュームの階層化は不要です。ツールボリュームやライブラリボリュームには60%を超えるコールドデータが含まれていることが多いため、ソースボリュームをオンプレミスで階層化することは理にかなっています。
- スケールアウトスクラッチボリュームのI/Oを最大化するように最適化された**Cloud Volumes ONTAP**
 - 【階層化が不要】を選択します。
 - 多くの場合、スクラッチボリュームには60~80%のコールドデータが含まれています。スクラッチワーククロード用にプロビジョニングされたボリュームがクラウドに長期間保持される場合は階層化を推奨します。それ以外の場合は階層化は不要です。これは、Cloud Volumes ONTAPの作成後に有効にすることができます。
- 高可用性と最大限のI/O、スケールアウトのスクラッチ/リリースボリューム向けに最適化された**Cloud Volumes ONTAP**
 - データのエージングデータに基づいて階層化オプションを選択します。
 - 多くの場合、スクラッチボリュームには60~80%のコールドデータが含まれています。スクラッチワーククロード用にプロビジョニングされたボリュームがクラウドに長期間保持される場合は、階層化を推奨します。



高速書き込みモードとWORMモードを有効にする

1. [Write Speed]で [High]を選択します。
2. WORMモードの場合は、[WORMを無効にする]を選択します。

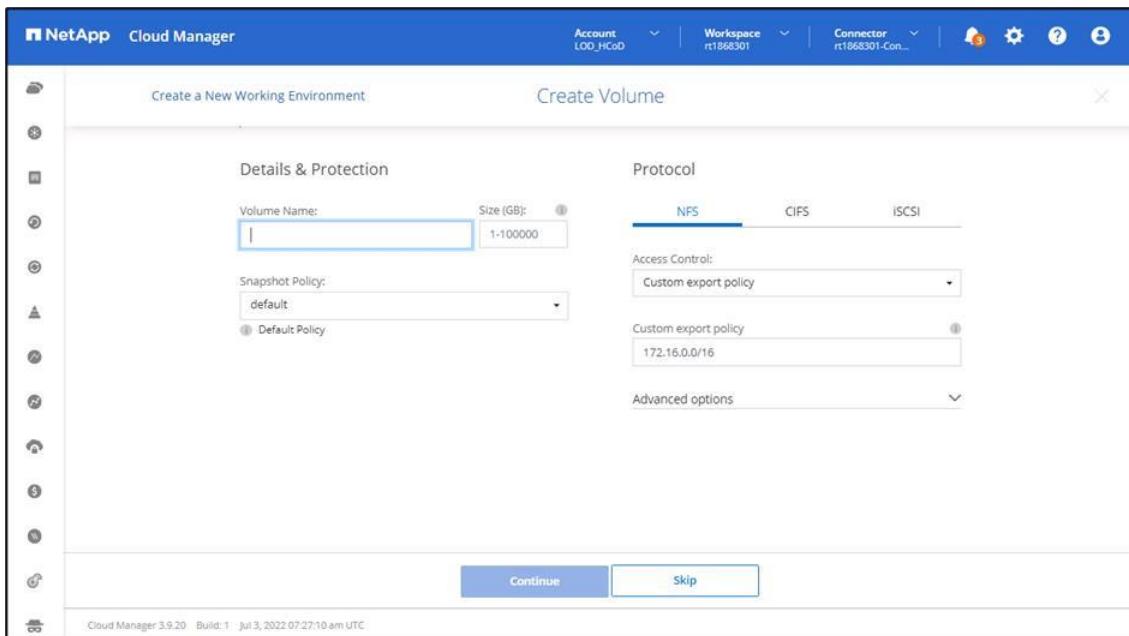
注：EDAワークロードには通常、Write-Once Read-Many (WORM) は必要ありません。通常、この機能はデータコンプライアンス要件に対して有効にします。



ボリュームを作成

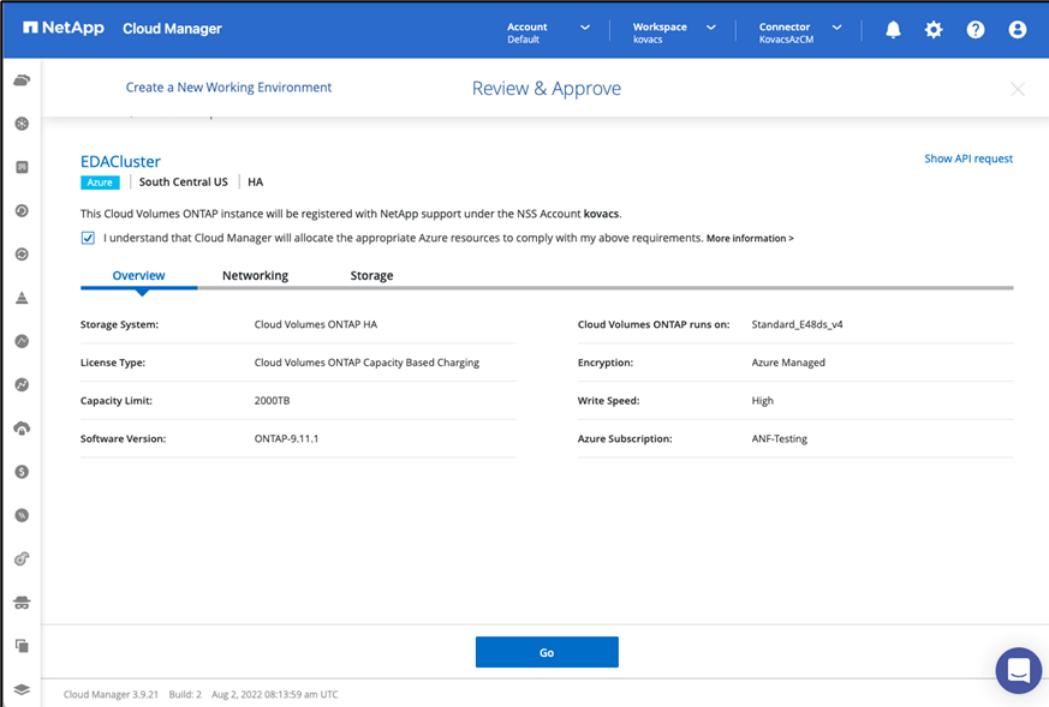
ボリュームプロビジョニングは、「Cloud Volumes ONTAP作成後のセットアップ」セクションで設定します。

1. [Skip]をクリックします。



確認と承認

設定の表を確認します。値が正しい場合は、[Go]をクリックしてCloud Volumes ONTAPファイルシステムの作成プロセスを開始します。



The screenshot shows the 'Review & Approve' step in the NetApp Cloud Manager. The configuration details are as follows:

Category	Value	Category	Value
Storage System	Cloud Volumes ONTAP HA	Cloud Volumes ONTAP runs on	Standard_E48ds_v4
License Type	Cloud Volumes ONTAP Capacity Based Charging	Encryption	Azure Managed
Capacity Limit	2000TB	Write Speed	High
Software Version	ONTAP-9.11.1	Azure Subscription	ANF-Testing

At the bottom, there is a 'Go' button and a 'Show API request' link.

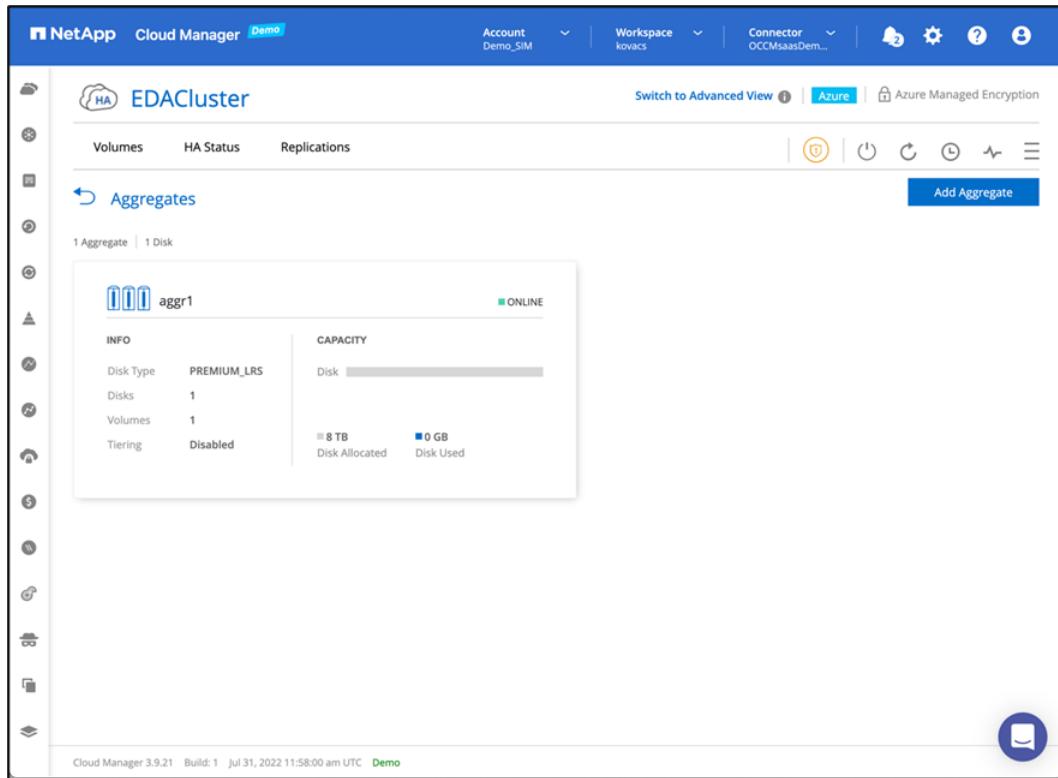
Cloud Volumes ONTAP作成後のセットアップ

次の手順は、各Cloud Volumes ONTAPファイルシステムでファイルシステムが利用可能になり、Cloud Managerでオンラインになったあとに実行します。

アグリゲートの変更

アグリゲート設定を、対応する構成テーブル（表2）に一致するように変更します（Cloud Volumes ONTAP VMタイプe48ds_v4に基づく）。および表3）アクティブ/アクティブHAペアCloud Volumes ONTAP構成（VMタイプe48ds_v4に基づく））を参照してください。「Cloud Volumes ONTAPインスタンスの設定」を参照してください。

1. [メニュー] > [詳細設定] > [詳細割り当て]を選択します。



NetApp Cloud Manager

Account: Demo_SIM | Workspace: kovacs | Connector: OCCMsaaSDem...

EDACluster

Volumes HA Status Replications

Switch to Advanced View | Azure | Azure Managed Encryption

Aggregates

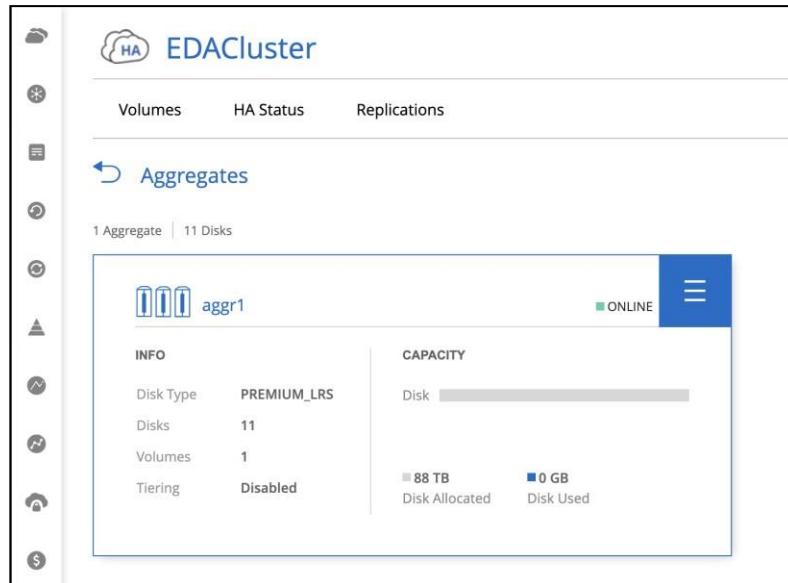
1 Aggregate | 1 Disk

aggr1 ONLINE

INFO		CAPACITY	
Disk Type	PREMIUM_LRS	Disk	<div style="width: 100%;">8 TB</div>
Disks	1	Disk Allocated	<div style="width: 0%;">0 GB</div>
Volumes	1	Disk Used	<div style="width: 0%;">0 GB</div>
Tiering	Disabled		

Cloud Manager 3.9.21 Build: 1 Jul 31, 2022 11:58:00 am UTC Demo

2. [Aggregates Menu] > [Add Azure disk]を選択します。



EDACluster

Volumes HA Status Replications

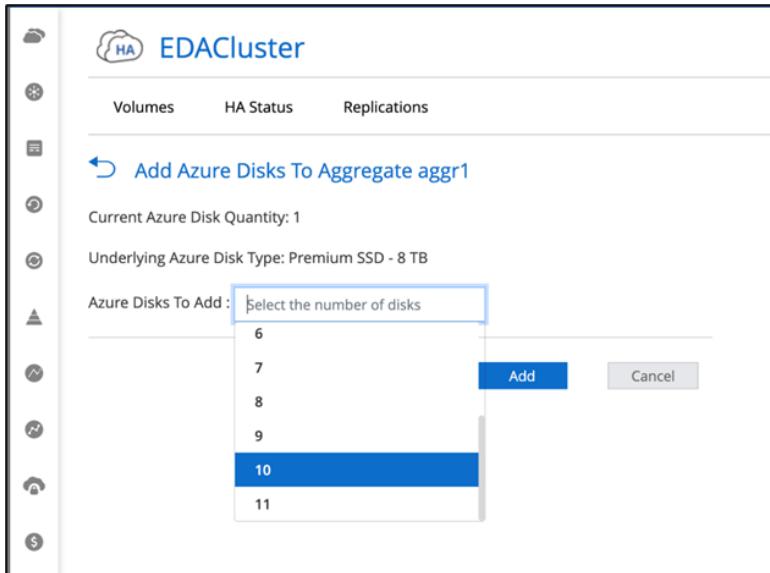
Aggregates

1 Aggregate | 11 Disks

aggr1 ONLINE

INFO		CAPACITY	
Disk Type	PREMIUM_LRS	Disk	<div style="width: 100%;">88 TB</div>
Disks	11	Disk Allocated	<div style="width: 0%;">0 GB</div>
Volumes	1	Disk Used	<div style="width: 0%;">0 GB</div>
Tiering	Disabled		

3. [Azure Disks to Add per Aggregate]を選択し、ディスク数を10に変更します。合計11本のディスクで構成されるアグリゲートが1つ作成されます。



アクティブ/パッシブまたはアクティブ/アクティブ

詳細については、「アクティブ/アクティブとアクティブ/パッシブの比較」を参照してください。

Cloud Volumes ONTAPのデフォルトはアクティブ/パッシブですが、NetAppでは、HAペア構成のEDA / HPCのパフォーマンスを最大化するために、Azureに対してアクティブ/アクティブを推奨しています。

1. [メニュー] > [詳細設定] > [詳細割り当て]を選択します。
2. [Add Aggregate] ボタンをクリックします。

3. アグリゲート名を指定してください。たとえば、aggr2を使用し、最初のアグリゲートはaggr1です。
4. 2番目のノードを選択し、[Continue]をクリックします。

Create new aggregate in azurecvoha

Aggregate Details

Aggregate Name:

Home Node:

Continue

- 前のセクション「オブジェクトストレージへのデータの階層化」に基づいて、目的の**BLOB**階層化設定を選択し、[Continue]を選択します。
- 表3に示すHA構成のディスクサイズとディスク数を選択します。

Aggregate Disks

what are storage tiers? Disabled

Premium SSD

An aggregate will be created using a Premium SSD.

Azure Disk Size Number of Disks:

推奨されるEDA固有のオプション（ボリューム前プロビジョニング）

NetAppでは、次のように64ビットIDを有効にすることを推奨します。

```
ONTAP cli> ONTAP cli> set -privilege advance
ONTAP cli> vserver nfs modify -vserver [VSERVER_NAME] -v3-64bit-identifiers enabled
```

ボリュームを作成する

「ボリュームのプロビジョニング（新しいボリュームの作成）」のEDAワークフローでは、拡張性とパフォーマンスを最大限に高めるために、FlexGroupボリュームから大きなメリットが得られます。

本ドキュメントの発行時点では、BlueXPではWeb UIからのFlexGroupまたはFlexCacheボリュームタイプの作成はサポートされていません（新しいリリースのBlueXPでは変更される可能性があります）。

注： FlexGroupボリュームとFlexCacheボリュームは、Cloud Volumes ONTAP CLIまたはONTAP REST APIを使用してプロビジョニングする必要があります。

ここでは、ONTAP CLIコマンドについて説明します。

ツールボリュームおよびライブラリボリューム用のFlexCacheボリュームの作成

ツールとライブラリボリュームは、FlexCacheを介してオンプレミスのONTAPボリュームからAzureに自動的にレプリケート（キャッシング）されます。クラウドにレプリケートする必要があるツールボリュームとライブラリボリュームの数によっては、複数のFlexCacheボリュームが必要になる場合があります。

[NetApp ONTAPのドキュメント](#)には、2つのONTAPインスタンス間にONTAPクラスタおよびSVMピア関係を設定する手順が記載されています。

クラスタピア関係とSVMピア関係のセットアップと検証が完了したら、次のコマンドを使用してFlexCacheボリュームを作成できます。

```
ONTAP cli> volume flexcache create -origin-vserver [ORIGIN_SVM_NAME] -origin-volume [ORIGIN_VOLUME_NAME] \
           -vserver [DESTINATION_SVM_NAME] -volume [DESTINATION_VOLUME_NAME] \
           -size [VOLUME_SIZE] -junction-path /[DESTINATION_VOLUME_NAME] \
           -aggr-list [AGGR_NAME] -aggr-list-multiplier 8 \
           -policy [EXPORT_POLICY]
```

FlexCacheボリュームは、単一のノード、単一のアグリゲート、および単一のLIF構成に制限する必要があります。FlexCacheボリュームは、複数のアグリゲートにまたがることはできません。これにより、ノード間のレイテンシとスループットのパフォーマンスがボリューム全体のパフォーマンスに制限されることはありません。

HA構成の場合は、FlexCacheをプロビジョニングして、2つのアグリゲート/ノード/ LIFセットにボリュームが分散されるようにします。

大容量ファイルの問題に対処するためにボリュームの容量をサイジングします（下記を参照）。

詳細については、「大容量ファイルに関する考慮事項」のセクションを参照してください。

オリジンまたはソースのatime-updateをFalseに設定する（ONTAP 9.11以前）

元の（またはソースボリュームの）atime-updatesをに設定します False。元のボリュームで読み取りのみが発生したときにキャッシングされたファイルが無効にならないようにするには、元のボリュームで最終アクセス日時の更新をオフにします。元のボリュームまたはソースボリュームでONTAP 9.6以降が実行されている場合は、FlexCacheインスタンスを作成するとatime-updateが自動的ににに設定されます False。

元のボリュームまたはソースボリュームのONTAP CLI（9.6より前のONTAPバージョン）から、次のコマンドを実行します。

```
ONTAP cli> volume modify -vserver origin-svm -volume vol_fc_origin -atime-update false
```

NetApp ONTAP 9.11.1より前のリリースでは、FlexCache元のボリュームでatime-updateを無効にして、読み取りアクセスが原因でキャッシング内で更新されるデータの量を減らすことを推奨します。ソースボリュームでatime-updateを無効にしないと、キャッシングが過度にアグレッシブに削除される可能性があります。たとえば、ソースボリュームでの単純読み取りでは、キャッシングからデータが削除されます。

NetApp ONTAP 9.11.1以降ではワークフローが改善され、atime-updateを環境全体に伝播できます。キャッシングでの削除には影響しません。キャッシングでの読み取りでも、元のデータのatimeを更新できます。キャッシングはatimeの権限ではなく、元のキャッシングだけが権限です。

元のボリュームまたはソースボリュームをONTAP9.11.1以降にアップグレードしたあと、atime-updateを使用できない場合はオフのままにします。atimeに基づいてワークフローの決定が行われる場合は、9.11.1で作成されたノブを使用します。

読み取り/書き込みスクラッチボリューム用のFlexGroupボリュームを作成する

スクラッチワークロード用のFlexGroupボリュームを1つ以上作成します。

SPECstorage 2020 EDAベンチマークテストに基づくと、FlexGroupボリュームで次の推奨事項を使用すると、FlexGroupボリュームで最高のパフォーマンスが達成されます。

- FlexGroupボリュームは、単一のノード、単一のアグリゲート、および単一のLIF構成に制限する必要があります。FlexGroupボリュームは、複数のアグリゲートにまたがることはできません。この設定により、ノード間のレイテンシとスループットのパフォーマンスがボリューム全体のパフォーマンスに制限されることがなくなります。
- 最大のスケールアウトパフォーマンスを実現するには、小規模なFlexGroupを複数プロビジョニングするのではなく、1つのFlexGroupボリュームにqtreeを配置することを推奨します。HA構成の場合は、2つのFlexGroupをプロビジョニングします。それぞれのボリュームが専用のアグリゲート/ノード/LIFに制限されます。
- 大容量ファイルの問題に対処するためにボリュームの容量をサイジングします（下記を参照）。複数の小さいボリュームをプロビジョニングするよりも、qtreeで管理される単一の大容量FlexGroupボリュームをプロビジョニングすることを推奨します。

FlexGroup CLIコマンド

```
ONTAP cli> volume create -vserver [SVM_NAME] -volume [VOLUME_NAME] \
    -size [VOLUME_SIZE] -junction-path [JUNCTION_PATH] \
    -aggr-list [aggr1 | aggr2] -aggr-list-multiplier 8 \
    -policy [EXPORT_POLICY] \
    -files-set-maximum true [-supporttiering true]
```

大容量ファイルに関する考慮事項

には aggr-list-multiplier、ノードまたはアグリゲートペアあたりのコンステイチュエントボリュームの数を指定します。EDAワークロードには8個のコンステイチュエントボリュームが推奨されます。80TBのボリュームをプロビジョニングすると、各コンステイチュエントのサイズは10TBになります。大容量ファイルとは、メンバーコンステイチュエントボリュームの1~5%の任意のファイルのことです。この例では、大容量ファイルとは、(80TB / 8 x 5%) または0.5TB (500GB) を超える任意のファイルを指します。一般に、シンプロビジョニングされたFlexGroupボリュームのサイズが小さいボリュームよりも大きいボリュームをプロビジョニングすることを推奨します。

読み取り用に最適化されたツールボリュームとライブラリボリュームのマウントコマンド

nocto an ac m o op onを使用してサーバ側キャッシングを最大化するには、次のNFS v3 mountコマンドを使用することを推奨します。SS S o noc o an ac m o a S CAN C のn ns EDAワークロードでは、Filer上のI/Oを80~90%も削減できます。

このベストプラクティスを確認し、「o’ ’’ s q」に基づいて適用してください。M n s Y A S S のバーサイドキャッシングが原因で、キャッシングの更新に時間がかかることがあります。

パフォーマンスを最大限に高めるには、ボリュームを提供するノード/アグリゲートに関連付けられたデータLIFのIPアドレスを使用します。

```
%> mount -t nfs \
    -o "nocto,actimeo=600,hard,rsize=262144,wsize=262144,vers=3,tcp,mountproto=tcp" \
    <ontap data lif ip>:</volname> /mnt/<volname>
```

スクラッチボリュームでの読み取り/書き込み混在パフォーマンスを最適化するmountコマンド

EDAスクラッチボリュームのワークロードに推奨されるNFS v3マウントコマンドは次のとおりです。

パフォーマンスを最大限に高めるには、ボリュームを提供するノード/アグリゲートに関連付けられたデータLIFのIPアドレスを使用します。

```
%> mount -t nfs \
    -o "hard,rsize=262144,wsize=262144,vers=3,tcp,mountproto=tcp" \
    <ontap data lif ip>:</volname> /mnt/<volname>
```

ツールとライブラリのスケールアウトパフォーマンスの向上

ツールボリュームやライブラリボリュームは、多くの異なるワークフロー、プロジェクト、さらにはビジネスユニットで共有されることが多く、接続されたLinuxクライアントが多数発生し、読み取りレイテンシのボトルネックにつながる可能性があります。

ツールとライブラリボリュームごとに複数のFlexCacheボリュームを使用し、サーバマウントのDNSロードバランスシングを使用してボリュームのロードを改善することで、I/Oとレイテンシのボトルネックを最小限に抑えることができます。

ツールとライブラリのパフォーマンスを拡張する方法の詳細については、次の章を参照してください。

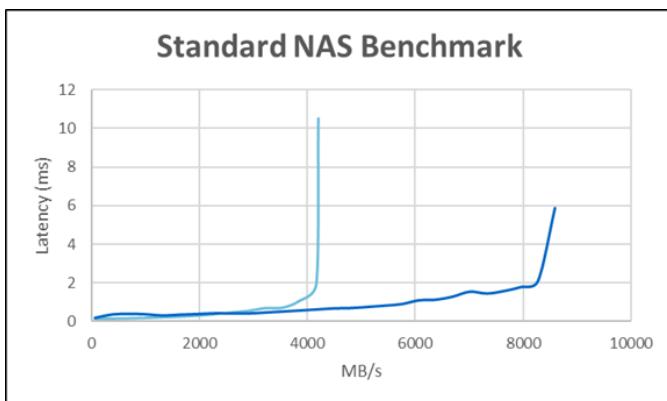
パフォーマンス監視

EDAワークロードにはパフォーマンスの監視が欠かせません。最も重要な指標はボリュームのレイテンシです。SSDベースのCloud Volumes ONTAPファイルシステムのレイテンシは2ミリ秒未満である必要があります。レイテンシが3ミリ秒を超える場合は、Cloud Volumes ONTAPファイルシステムとボリュームで大量のI/Oワークロードが実行されていることを示します。

並行して実行されているEDAジョブからのI/O負荷が高い場合、レイテンシが高くなる可能性があります。Cloud Volumes ONTAPファイルシステムのI/Oプロファイルは、NetApp A800のようなオンプレミスのONTAPアプライアンスとは異なる場合があります。NetAppでは、Cloud Volumes ONTAPで実行されるEDAワークロードを大規模にテストしてI/Oの制限を決定することを推奨しています。

次のグラフは、NASワークロードがパフォーマンス制限に押し込まれている様子を示しています。パフォーマンス曲線の膝は、コントローラがパフォーマンスプロファイルの範囲外で動作しており、レイテンシが急激に増加し始めるポイントです。この状態で長時間動作すると、EDAジョブ実行時のパフォーマンス（実測クロック）に影響します。

図7) 標準NASベンチマーク/



注：このグラフは、説明のみを目的とした一般的なNASパフォーマンスプロファイルグラフであり、実際のCloud Volumes ONTAPパフォーマンスを表しているわけではありません。

ONTAP CLIによる迅速なレイテンシ監視

リアルタイムのONTAPパフォーマンスを測定する方法は複数あります。POCでは、ONTAP CLIを使用してレイテンシを単純にレポートする方法の1つがあります。次のコマンドは、Ctrl+Cキーを押してレポートを停止するまで実行されます。EDAジョブの送信前または送信中にコマンドを開始して、リアルタイムのボリュームレイテンシを確認します。3ミリ秒を超える長時間のレイテンシは、負荷の高いストレージファイルシステムの兆候である可能性があります。

```
ontap cli> qos statistics volume latency show
```

リアルタイムのレイテンシを監視およびレポートするためのその他のアプローチは次のとおりです。

- Cloud Manager内のPerformance Manager

- クラウド固有の監視

容量レポート

FlexCacheボリュームのサイズと使用率に関するレポートを作成するには `volume show -fields <fields to report>`、ONTAP CLIからコマンドを使用します。RESTfulに相当するものもあります。

ここでは、いくつかのフィールドが入力されている例を示します。次の例では、Azurecvo proj ボリュームが、オンプレミスの proj ソースボリュームを含むFlexCacheインスタンスであることがわかります。

```
ONTAP:::> volume show -volume <volname> -fields size,used,available,percent-used,files-used,files
```

次の例では、オンプレミス proj ボリュームはFlexGroup、Azurecvo proj ボリュームはFlexCacheです。Azurecvoキャッシュ内のファイルの一部が読み取られました（事前にウォームアップされています）。

```
onPrem:::> volume show -volume proj -fields size,used,available,percent-used,files-used,files
vserver      volume size available used      percent-used files      files-used
-----
svm onPrem proj    1TB  46.47GB    15.85GB  1%            31876696  226372

Azurecvo:::> volume show -volume proj -fields size,used,available,percent-used,files-used,files
vserver      volume size available used      percent-used files      files-used
-----
svm Azurecvo proj   150GB 147.0GB    2.96GB  1%            4669368  792
```

Cloud Volumes ONTAPインスタンスの起動と停止

BlueXPでCloud Volumes ONTAPを一度停止して起動すると、クラウドコンピューティングのコストを管理できます。これは、コンセプトの実証（POC）テストで特に役立ちます。

開始/停止操作を実行する方法は複数あります。

- **BlueXPのUI**：メニューオプションを使用すると、手動スタート/トップ操作が可能になります。
- **BlueXPのUI**：事前にスケジュールされたスタート/ストップオペレーション。コンピューティングコストを削減するために、一定の期間にCloud Volumes ONTAPをシャットダウンすることもできます。手動で行う代わりに、特定の時間にシステムを自動的にシャットダウンして再起動するようにBlueXPを設定できます。
- **BlueXPのAPI**：

詳細については、[のドキュメント](#)を参照してください。

クイックスタート実装の詳細

このセクションでは、各オプションが指定された理由と、EDAワークフローのベストプラクティスの背後にあらる考慮事項を詳しく説明します。場合によっては、これらの推奨事項を特定の設計ワークフローやお客様の要件に適用する際に考慮すべきトレードオフが記載されています。

Cloud Volumes ONTAPインスタンスタイプ

ここでは、シングルペア構成とHAペア構成のトレードオフ、およびそれぞれを使用する状況について説明します。

Cloud Volumes ONTAPはクラウドベースの共有ストレージアーキテクチャを採用

AzureのCloud Volumes ONTAPで使用されるクラウドベースの共有ストレージアーキテクチャは、従来のオンプレミスのAFFクラスタやFASクラスタと似ていますが、異なります。2つの独立したコンピューティングノードは、シングルシャーシまたはデュアルシャーシアプライアンスではなくIaaSインスタンスとして動作します。Cloud Volumes ONTAPは、SASバックプレーン上の物理的に接続されたディスクではなく、LRSまたはZRS構成でAzureが提供する共有ディスクを利用します。

Cloud Volumes ONTAP HAへの書き込み処理は、パートナーノードのメッセージにミラーリングされてから、クライアントに確認応答が返されます。これらのトランザクションは、専用の物理バックプレーン（直接接続またはクラスタインターコネクト）ではなく、Azureネットワーク経由でミラーリングされます。ノードで障害が発生した場合、稼働しているノードが障害が発生したノードのディスクとアグリゲートをテイクオーバーし、データの提供を継続します。

Cloud Volumes ONTAPのパフォーマンスサイジングでは、クラウドリソースのネットワーク使用権、キャッシュされていない最大ディスク数、接続（仮想）ディスク容量の累積使用権が考慮されます。人為的なボトルネックが発生しないように、コンピューティングノードインスタンスのサイジングは、ディスクアグリゲートの総帯域幅制限に合わせて帯域幅を調整する必要があります。本ドキュメントの執筆時点での**Cloud Volumes ONTAP**ストレージのサイジングの目標は、利用可能なクラウドリソース構成からできるだけ多くのIOPSを提供し、**Software-Defined**アーキテクチャの機能が（部分的に）定義されていることを理解することです。お客様が使用するインフラの使用権によって異なるため、**A800 NetApp**ストレージコントローラなどのデータセンタークラスの最上位ハードウェアアプライアンスとは直接一致しない場合があります。

信頼性と耐久性に関する考慮事項

FlexCacheのツールボリュームとライブラリボリュームの場合、**HA**インスタンスを使用すると、複数のワークフロー、プロジェクト、チームで共有される可能性のあるボリュームの信頼性と耐久性が向上します。信頼性と耐久性が向上するため、共有のワークフロー、プロジェクト、チームが障害の影響を受ける可能性が低くなります。

FlexCacheのツールボリュームとライブラリボリュームは主に読み取り専用であるため、**HA**構成の書き込みパフォーマンスはそれほど重要ではありません。

注： **FlexCache**ツールボリュームとライブラリボリュームには、**Cloud Volumes ONTAP HA**ファイルシステムを推奨します。

耐久性が最適化されたスクラッチおよびリリースボリュームについては、アプリケーションのパフォーマンスと信頼性、コストのトレードオフ要件を評価するかどうかは、お客様次第です。これは主に、信頼性よりも並列ジョブの拡張性が高いというトレードオフです。

注： 読み取りと書き込みのパフォーマンスが最適化されたスクラッチボリュームには、**Cloud Volumes ONTAP HA**インスタンスが最適です。お客様は、特定のワークロードのパフォーマンスとコストのトレードオフについて議論する必要があります。

コストに関する考慮事項

シングルノードの**Cloud Volumes ONTAP**インスタンスに必要なコンピューティングノードは1つだけで、大規模環境での書き込みパフォーマンスが向上するため、**EDA**ジョブの実行時のパフォーマンスが向上する可能性があります。**HA**ノードに必要なクラウドインフラはシングルノードのファイルシステムよりも多く、場合によっては書き込みパフォーマンスが低下することがあります。トレードオフは高い信頼性とアップタイムの向上です。

FlexCacheのツールボリュームとライブラリボリュームには**HA**を選択しました。これは、**FlexCache**ボリュームの容量がオンプレミスのソースボリュームの約10%であり、容量コストの点で低コストであることが前提となっているためです。コストが懸念される場合は、**FlexCache**ツールボリュームとライブラリボリュームをシングルノードインスタンスとして構成できます。

追加されたサービス

Data Senseとコンプライアンス（オプション）

データセンスとコンプライアンスは、サイバーセキュリティ機能です。**EDA**固有のパフォーマンステストは実施されていません。**Data Sense**とコンプライアンスには、既知のパフォーマンスオーバーヘッドがあります。**NetApp**では、パフォーマンスを最大化するために、このオプションをオフにすることを推奨します。

追加のセキュリティが必要な場合は、この機能をオンとオフの両方でテストして、パフォーマンスへの影響を評価します。

重要: ランサムウェア対策機能は、スループットとピークIOPSの点でパフォーマンスへの影響を最小限に抑えますが、状況によってはNetAppの設定を制限することを推奨します。特に、読み取り処理が大量に発生するワークロードやデータを圧縮できるワークロードの場合、NetAppでは、ストレージノードあたりのボリューム数が50未満でこの機能を有効にすることを推奨しています。大量の書き込み処理が発生し、データを圧縮できないワークロードについては、ランサムウェア対策をストレージノードあたりのボリューム数を20以下に制限することを推奨します。あまり理想的ではないシナリオでは、特定のボリュームに対してランサムウェア対策分析の実行頻度が低下し、一般的なパフォーマンスに影響が及ぶ可能性があります。

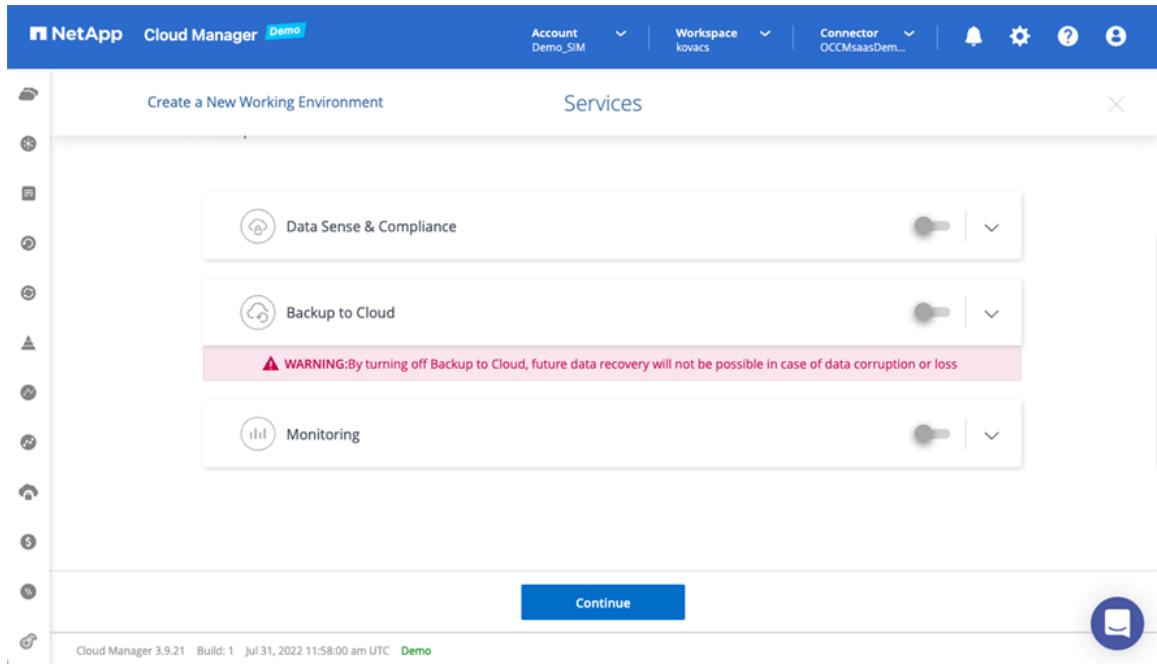
クラウドへのバックアップ（オプション）

クラウドへのバックアップは、SnapMirrorからクラウドへのサービス機能です。この機能を有効または無効にするかどうかは、Cloud Volumes ONTAPファイルシステムのタイプに基づいて次のように決定してください。クラウドへのバックアップはパフォーマンスへの影響を最小限に抑える必要があるため、この機能を有効にすることについてお客様と話し合う必要があります。

- **スケールアウトツールとライブラリのFlexCacheボリュームのI/Oを最大化するように最適化されたCloud Volumes ONTAP**
 - クラウドへのバックアップ（オプション）-オフにします。
 - ツールとライブラリのCloud Volumes ONTAPファイルシステムは、主にFlexCacheボリュームに対応しているため、バックアップの必要はありません。ソースボリューム（通常はオンプレミス）をバックアップする必要があります。FlexCacheボリュームはバックアップしないでください。
- **Cloud Volumes ONTAPは、スケールアウトスクラッチボリュームのI/Oを最大化するように最適化されています。**
 - クラウドに戻る（オプション）-オフにします。
 - スクラッチボリュームは、急速に変化する設計データとEDAツールからの出力を含む、変更率の高いボリュームです。オンプレミスのスクラッチボリュームのほとんどはバックアップされません。EDA設計データは、Perforce、ICManage、SOS、ClearCase、Gitなどのツールを使用してバージョン管理されます。そのため、重要な設計データはすでに安全にバックアップおよびバージョン管理されています。
- **高可用性と最大限のI/O、スケールアウトのスクラッチ/リリースボリューム向けに最適化されたCloud Volumes ONTAP**
 - クラウドに戻る（オプション）-オフにします。
 - このオプションは、カスタマーコンサルティングに基づいて必要に応じて有効にすることができます。特にリリースワークロード用のスクラッチボリュームは、災害発生時にバックアップが必要になる場合があります。スクラッチボリュームは通常FlexGroupボリュームであるため、現在はサポートされていません。

監視（オプション）

この機能を有効にすると、Cloud Insights監視が有効になりますEDAワークロードについては、監視によるパフォーマンスへの影響は評価されていません。NetAppでは、コンセプトの実証テスト中にこの機能をオフにすることを推奨しています。フローとフローのパフォーマンスを評価した後、パフォーマンスへの影響を評価するためにこのオプションを有効にすることができます。



アベイラビリティゾーン (AZ) を選択

Azureクラウドでは、Cloud Volumes ONTAPファイルシステムが導入されている場所を「AZ」と呼びます。Cloud Volumes ONTAPのセットアップ時にAZが1つ、2つ、3つ、または4つある場合は、Cloud Volumes ONTAPファイルシステムを作成するAZを決定する必要があります。他のクラウドでは、マルチAZとは、1つのCloud Volumes ONTAPノードを1つのAZに配置し、もう1つのAZに2つ目のノードを配置して、可用性の高いHAペア環境を構築することを指します。AzureのCloud Volumes ONTAPは現在、マルチAZ環境をサポートしていないため、シングルノードまたはHAペアのCloud Volumes ONTAPファイルシステムは単一のAZに存在します。

従来のEDA（半導体）データセンターでは、コンピューティングとストレージが同じ建物内に配置されています。データセンターで障害が発生すると、コンピューティングとストレージの両方に影響が及ぶ可能性があります。これは、ネットワークの停止、停電、またはその他の問題が原因である可能性があります。NFSファイルに格納されたEDAデータは、通常、リアルタイムでセカンダリデータセンターにレプリケートされません。一部のEDAボリュームは、1日に数回ディザスタークリアのために別のデータセンターにレプリケートまたはバックアップされますが、継続的にはレプリケートされません。

単一のAZは、基本的に前述の従来のデータセンターです。Cloud Volumes ONTAP作業環境はすべて1つのAZに存在するため、2つ目のAZへの書き込みに関連するレイテンシは発生しません。Azureでは、Cloud Volumes ONTAP HA作業環境は可用性セットの一部です。可用性セットは、各ノードを別々の障害ドメインと更新ドメインに配置することで、単一のAZに固有の冗長性リスクを軽減するのに役立ちます。

NetApp Support Siteアカウントの管理

このオプションを使用すると、Cloud Volumes ONTAPのファイルシステム通信とNetApp Support Site（Active IQとも呼ばれます）へのレポートが可能になります。Active IQでは、すべてのNetApp製品、パフォーマンス、契約更新などを単一のコンソールで監視できます。

この機能を使用するには、BlueXPも登録する必要があります。Cloud Managerを登録するには、BlueXP UIの上部にあるリンクをクリックするか、[Help]メニュー（?）に移動してください。>[Support]をクリックし、[NSS Management]タブを選択します。



96 38
Account serial number

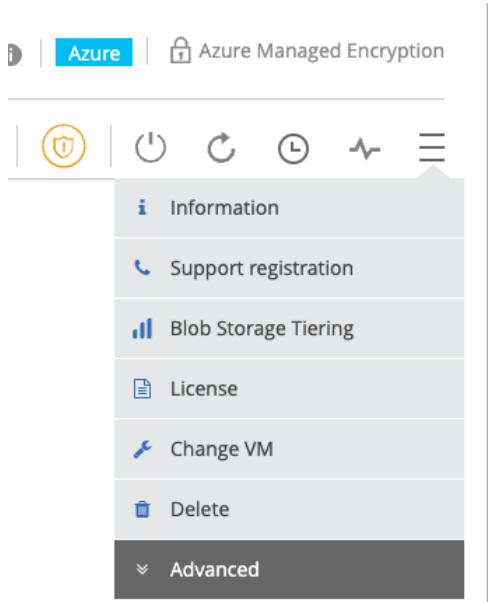


Not Registered
Support Registration

[Register for Support](#)

Active IQの接続には、外部との通信を有効にするAzureネットワーク構成が必要です。[このリンク](#)は、Active IQ (NetApp Support Siteとも呼ばれます) のセットアップと設定に役立ちます。

Cloud Volumes ONTAP作業環境を登録するには、**Cloud Volumes ONTAP**をダブルクリックし、メニューから[Support Registration]を選択します。NSSアカウントを選択し、[Register]ボタンをクリックします。



課金方法の選択（ライセンス）

すべてのEDAアカウントではないにしても、ほとんどのEDAアカウントは、より高価なProfessional by Capacityライセンスではなく、Essential by Capacityライセンス機能を選択しています。2つのライセンスの唯一の違いは、Professionalにはクラウドへのバックアップライセンスが含まれている点です。

FlexCacheツールとライブラリのCloud Volumes ONTAPインスタンスはバックアップを必要としません。ツールボリュームとライブラリボリュームのバックアップは、ソースボリュームで実行する必要があります。

EDAスクラッチボリュームは通常、オンプレミスではバックアップされないため、ほとんどの場合、クラウドへのバースト対応モデルではバックアップされません。デザインソースコードは、Perforce Helix、ICManage、SOS、ClearCase、Gitなどのツールでバージョン管理されています。一般的に、EDAジョブの出力は設計の反復が頻繁に行われるため、スナップショットやバックアップは不要であり、意味のあるものではありません。

スクラッチボリュームHAスクラッチまたはリリース作業のために、クラウドへのバックアップが必要になる場合があります。POCの場合、クラウドへのバックアップは不要なため、Essential by Capacityライセンスが適しています。本番環境（POC以外）の場合は、お客様に相談して、これらのボリュームにクラウドへのバックアップが必要かどうかを判断する必要があります。この場合、お客様は、「Essential by Capacity」と「Professional by Capacity」のどちらを選択するのが適切かを判断する必要があります。

EDAワークコードのCloud Volumes ONTAPサイジングに関する考慮事項

最新のEDAワークコードでは、1つの設計で20~100kの並列コンピューティングコアが使用されています。並列処理能力、ファイル数、メタデータ比率の高いワークコードは、スケールアウトパフォーマンスの最適化に欠かせません。

EDAワークコードは1つではありません。設計仕様の初期段階から、設計、タイミング、消費電力、面積の最適化、チップ仕上げまで、50種類以上のツールを使用してチップを設計できます。以下の推奨事項は、最大IOP、最小レイテンシ、および最大並列コンピューティングパフォーマンスを実現するためのCloud Volumes ONTAPの設定に重点を置いています。

インフラコストは常に考慮する必要がありますが、エンジニアの生産性と高価なEDA設計ツールライセンスの使用は、ITやクラウドインフラのコストよりもはるかに大きなコストです。ただし、パフォーマンス要件を決定したら、パフォーマンスとコストのトレードオフを評価できます。

EDA VMのサイジングに関する考慮事項

- 大規模なプロセッサコアを搭載したONTAPシステムは、高度な並列EDAワークロードにメリットをもたらします。コア数の多い仮想マシン（VM）を選択すると、読み取り/書き込み要求の並列処理が可能になります。
- EDAデータはファイル数が非常に多く、ディレクトリ構造の深い場所に小さなファイルと大きなファイルが混在しています。システムRAMを大量に搭載することで、ONTAPのパフォーマンスが向上します。
- NetApp FlashCacheをサポートするVM（ローカルのNVMeストレージを使用するVM）は、ONTAPのパフォーマンスを大幅に向上させます。
- ディスクアグリゲートへの最大IOPSをサポートするVMサイズを選択します。VMのIOPSパフォーマンスがアグリゲートよりも低い場合は、VMによってディスクパフォーマンスが制限されます。
- HAノード構成では、VMのスループット制限を半分に分割する必要があります。これは、書き込みはすべて最初のノードが所有するディスクアグリゲートに書き込まれる一方で、2番目のノードとそのディスクアグリゲートにも書き込まれるためです。

ストレージシステムは、I/O（またはストレージメディア）の制限になる前に、ノード（またはコンピューティング）の制限になる可能性があります。そのため、フロントエンドのコンピューティングI/O帯域幅とバックエンドストレージの総帯域幅のバランスを取ることが重要です。

共有インスタンスと専用インスタンス

共有VMはコストが低く、専用VMと同等のパフォーマンスを発揮するため、EDAでは共有VMを使用することを推奨します。

専用ファイルシステムと共有ファイルシステムの違いは何ですか。

共有テナントとは、異なる顧客の複数のVMインスタンスを同じ物理ハードウェア上に配置できることを意味します。専用モデルでは、導入した他のインスタンスがあるハードウェアでのみVMが実行され、他のお客様がお客様と同じハードウェアを使用することはありません。

センヨウインスタンス

専用インスタンスとは、単一のお客様専用のハードウェア上で仮想プライベートクラウド（VPC）内で実行されるAzure VMのことです。異なるAzureアカウントに属する専用インスタンスは、単一の支払いアカウントにリンクされている場合でも、ハードウェアレベルで物理的に分離されます。ただし、専用インスタンスは、専用でない同じAzureアカウントの他のインスタンスとハードウェアを共有する場合があります。

専用ホストは共有テナントよりもコストが高く、バースト可能なファイルシステムタイプをサポートしません。

Cloud Volumes ONTAPディスクの選択とスループット要件 EDAディ

スクのパフォーマンス要件

オールフラッシュ（SSDベース）のディスクパフォーマンスは、要件の厳しいEDAワークロードとHPCワークロードに対してその価値を実証しています。レイテンシを1ミリ秒未満に抑えることでEDAワークロードのパフォーマンスが向上し、10k SASドライブとSATAドライブの一般的な手法であったように、ストレージのパフォーマンス調整の必要がほぼ完全になくなりました。

EDAではSATAドライブがまだ一般的ですが、バックアップとコールドデータストレージにのみ使用されます。アクティブワークロードは、オンプレミスのハイパフォーマンスなオールフラッシュストレージプラットフォーム上ではほぼ実行されます。クラウドベースのNFSストレージについても同様です。

ディスクのパフォーマンスを最大化するには、ディスクストレージのスループットパフォーマンスをAzure VMのスループットと一致させることが重要です。

基盤となるディスクの選択

現在NetAppではSSDディスクが推奨されていますが、他にも使用可能なディスクオプションがあり、そのほかにも評価と認定が行われています。この文書では、一般化された方法でディスクがどのように選択されるかの背後にある思考プロセスを説明しようとします。新しいディスクオプションが使用可能になると、このセクションの内容が古くなる可能性があります。例として以下を使用します。

マネージドSSDおよびPremiumページBLOB

AzureのCloud Volumes ONTAP作業環境では、導入する作業環境のタイプに基づいて、2種類のストレージが使用されます。どちらのタイプも、高スループットと低レイテンシでI/O負荷の高いワークロードをサポートするように設計された、ハイパフォーマンスSSDストレージです。シングルノード作業環境では、ディスクあたり最大32TBのPremiumマネージドSSDを使用し、SSDあたり最大20K IOPSと900MBpsをサポートします。シングルノードの作業環境に接続できる管理対象ディスクの数は、VMのタイプによって決まります。

Cloud Volumes ONTAP HA作業環境では、Blobあたり最大8TBのPremiumページBLOBを使用します。PremiumページBLOBは、ページBLOBあたり最大7、500 IOPS、250MBpsのパフォーマンスを提供します。同様に、HA作業環境に接続できるページプロブの数は、選択したVMのサイズによって異なります。

[Cloud Volumes ONTAPリリースノート](#)のVMサイズ別のディスクと階層化の制限を参照してください。

アグリゲートとディスクのサイジング-パフォーマンス

各クラスタは、ONTAPにSCSIとして提供されたSSDボリュームのグループを基に構築され、アグリゲートにグループ化されます。アグリゲートとSSDボリュームを効果的にサイジングするには、選択したVMサイズのパフォーマンスとクラスタあたりのワークロードのサイズに基づいた計算式が必要です。

パフォーマンスに関するONTAPのベストプラクティスでは、ディスクへの書き込み並列処理を促進するため、アグリゲート内に少なくとも4本のディスクで構成されるアグリゲートを作成することを推奨します。

EDAワークロードのテストによると、パフォーマンスの結果から、最適なディスク数はアグリゲートあたり6本であることがわかりました。

アグリゲートとディスクのサイジング-容量

次に、ONTAPのワークロードサイズと容量オーバーヘッドに基づいて、アグリゲートあたりの各ディスクのサイズを検討します。計算は簡単で、次の4つの要素で構成されています。

- ディスク数
- アクティブなワークロードのサイズ
- アグリゲートSnapshotリザーブ
- WAFLオーバーヘッド

EDAのディスク数は6のままで、アクティブなワークロードのサイズはワークロードと個々のキャッシュサイズによって異なります。アグリゲートのSnapshotリザーブサイズはデフォルトで5%に設定され、WAFLオーバーヘッドは合計アグリゲートスペースの12~14%になります。

次の計算例を参照してください。

ディスクサイズ= (ワークロードサイズ* (1 + (アグリゲートのスナップリザーブ+ WAFLオーバーヘッド) / 100)) / 6

たとえば、アクティブなワークロードのサイズが4096GiBの場合、ディスクサイズ= (4096GiB * (1+ (5+14) /100)) / 6 = (4096 * 1.19) / 6 = 4875/6 = 813GiB

この方法でアグリゲートをサイジングする目的は、ディスク帯域幅がワークロードのボトルネックにならないように、コンピューティングインスタンスタイプに基づいてEBS帯域幅のIOPSとスループットを飽和させることです。

コスト対パフォーマンス

ストレージのコストは、選択したストレージタイプ（StandardまたはPremium）とSSDまたはページプロブのサイズに基づきます。ページプロブは、Premiumストレージタイプとしてのみ使用できます。

高速書き込みモード

High Write Speed (HWSM ; 高速書き込み速度) モードでは、書き込みレーテンシーがAC NO n 利用規約 データがONTAPメモリ (RAM) に書き込まれるとすぐに、OpAがオンになります。これは、書き込みスループットを向上させ、書き込みレーテンシを最大90%削減するのに適しています。これは、高スループットと低レーテンシの書き込みワークロードに非常に効果的です。

Normal Write Speed (通常書き込み速度) モード（Cloud Volumes ONTAPのデフォルト）では、データが永続的ディスクに書き込まれ、揮発性メモリに存在しなくなった場合にのみ、書き込みを確認します。

トレードオフとして、HWSMはデータ損失の可能性がある優れたパフォーマンスを提供しますが、通常の書き込みモードではデータ損失のない書き込みパフォーマンスが低いのに対し、データ損失の可能性がある優れたパフォーマンスを提供します。選ぶ方法は？

この設定では、ONTAPが整合ポイントをメモリからディスクにフラッシュするのにかかる時間に基づいて、データの耐久性を低下させることができます。これは、シングルノードシステムが本番環境のデータや永続的なデータに使用され、他のソースから置き換えることができない場合に発生するリスクです。ただし、ハイアベイラビリティ構成では、データが失われるためペアの両方のノードが最大CPフラッシュ時間である10秒以内に失われる必要があるため、リスクは比較的低くなります。

Cloud Volumes ONTAP作成後のセットアップ

アグリゲートのセットアップの詳細の変更

EDAワークロードのテストでは、アグリゲートあたり6本のディスクがパフォーマンスに最適な構成であることが確認されています。Azureでは、各ディスクのIOPS /スループットが集約され、アグリゲートのIOPS /スループットの合計値が算出されます。

例：

- ディスクが1つのアグリゲート=1×（ディスクあたりの最大IOPS）
- 6本のディスクで構成されるアグリゲート=6*（ディスクあたりの最大IOPS）（その結果、アグリゲートあたりのIOPSが6倍になります）。

Cloud Volumes ONTAPアグリゲートでは、サイズとタイプが同じディスクを最大12本使用できます。

VMtypeネットワークパフォーマンスは、最大アグリゲートIOPSをサポート（または一致）する必要があります。osn'の場合、アグリゲートではなくVMtypeスループットがボトルネックになります。

アクティブ/アクティブとアクティブ/パッシブ

Cloud Volumes ONTAPでは、HAペアを使用したデータアクセスの構成として、アクティブ/パッシブとアクティブ/アクティブの2つが考えられます。

BlueXPで作成されるすべてのクラスタはデフォルトでアクティブ/パッシブであるため、ホームノード（ノード1）となるアグリゲートが1つ作成されます。つまり、すべてのデータは data_1 LIF（論理インターフェイス）を介してノード1からネイティブにアクセスされ、ノード2はフェイルオーバー時にクライアントによってのみ使用されます。

アクティブ/アクティブとは、ノード2に少なくとも1つのアグリゲートを設定し、data_LIFを介してデータにアクセスすることです²。これにより、クライアントは、HAペアの両方のノードを読み取りとローカルキャッシング（FlashCache）で最大限に活用できます。ただし、どちらかのノードの使用率が50%を超えると、フェイルオーバー時のパフォーマンスが低下し、単一のローカルキャッシングを持つ単一のノードから発信されるすべてのトラフィックに対するサービスが自然に低下します。

アクティブ/アクティブにより、パフォーマンスが向上する可能性があります。ただし、アクティブ/アクティブモードで実行する場合は、フェイルオーバーに影響する問題に注意する必要があるため、各ノードのパフォーマンス利用率は50%以下にする必要があります。

詳細については、「[LIF（論理インターフェイス）とは](#)」を参照してください。

NetAppでは、Azure構成にはアクティブ/アクティブ構成を使用することを推奨しています。

ボリュームプロビジョニング（新しいボリュームの作成）

ツールおよびライブラリボリューム用のCloud Volumes ONTAPインスタンス

ツールボリュームおよびライブラリボリューム用のFlexCacheボリュームを作成します。キャッシングする必要があるオンプレミスのソースボリュームの数によっては、複数のボリュームが必要になる場合があります。

[このリンク](#)では、ONTAPクラスタ間にクラスタとSVMのピア関係を設定する方法、およびFlexCacheボリュームのプロビジョニング方法について説明します。

FlexCache CLIコマンド

```
ONTAP cli> volume flexcache create -origin-vserver [ORIGIN_SVM_NAME] -origin-volume [ORIGIN_VOLUME_NAME] \
-vserver [DESTINATION_SVM_NAME] -volume [DESTINATION_VOLUME_NAME] \
-size [VOLUME_SIZE] -junction-path /[DESTINATION_VOLUME_NAME] \
-aggr-list [AGGR_NAME] -aggr-list-multiplier 8 \
-policy [EXPORT_POLICY]
```

高パフォーマンスのスクラッチボリューム用のCloud Volumes ONTAPインスタンス

スクラッチワーカロード用のFlexGroupボリュームを1つ以上作成します。

SPECstorage 2020 EDAベンチマークテストに基づくと、FlexGroupボリュームで次の推奨事項を使用すると、FlexGroupボリュームのパフォーマンスが最高になります。

- FlexGroupボリュームは、シングルノード、シングルアグリゲート、およびシングルLIF構成に制限する必要があります。これにより、ノード間のレイテンシが実現し、スループットパフォーマンスによってボリューム全体のパフォーマンスが制限されることはありません。
- 複数のボリュームをプロビジョニングする場合は、パフォーマンスと容量の負荷が最も低いノード、アグリゲート、またはLIFのペアに新しいボリュームを配置します。
- 大容量ファイルの問題に対応できるサイズにボリュームの容量を設定してください。

FlexGroup CLIコマンド

```
ONTAP cli> volume create -vserver [SVM_NAME] -volume [VOLUME_NAME] \
-size [VOLUME_SIZE] -junction-path / [VOLUME_NAME] \
-aggr-list [AGGR_NAME] -aggr-list-multiplier 8 \
-policy [EXPORT_POLICY] \
-files-set-maximum true [-supporttiering true]
```

には aggr-list-multiplier、ノードまたはアグリゲートペアあたりのコンステイチュエントボリュームの数を指定します。NetAppでは、EDAワーカロード用に8個のコンステイチュエントボリュームを推奨80TBのボリュームをプロビジョニングすると、各コンステイチュエントのサイズは10TBになります。大容量ファイルとは、メンバーコンステイチュエントボリュームの1~5%のファイルのことです。この例では、大容量ファイルとは、(80TB / 8 x 5%) または0.5TB (500GB) を超える任意のファイルを指します。一般に、シンプルプロビジョニングされたFlexGroupボリュームのサイズが小さいボリュームよりも大きい方をプロビジョニングすることを推奨します。

inode数を増やす

EDA環境では数百万個のファイルが存在する場合があるため、inode値を最大レベルに引き上げなければならぬ場合があります。NetAppでは、ファイルの最大値を簡単に設定できるように、ONTAP 9.9.1で新しいボリュームオプションが導入されました。

ボリュームのfiles-set-maximumの値を **true**に設定すると、ONTAPは **maxfiles**を可能な限り大きい値に自動的に調整します。**true**に設定すると、値をリセットできません。**maxfiles**を最大値に設定する場合にのみ、値を**true**に設定します。

表4) FlexVolサイズ別のinodeのデフォルトと最大値

FlexVolボリュームサイズ	デフォルトのinode数	最大inode数
20 MB *	566	4,855
1 GB *	31,122	249,030
100GB*	3,112,959	24,903,679
1TB	21,251,126	255,013,682
7.8TB	21,251,126	2,040,109,451
100TB	21,251,126	2,040,109,451

* FlexGroupメンバーボリュームのサイズは100GB未満にする必要があります。

ボリュームのマウント

NFSv3はNFSv4.xよりもパフォーマンスが高いため、引き続きEDAワークロードの標準となっています。次の2つのマウントコマンドは、読み取り専用（またはほとんど）ツールボリュームとライブラリボリュームを最適化し、読み取り/書き込みスクラッチボリュームのパフォーマンスを可能な限り高速化するためのものです。

EDA / HPCワークロード向けにNFSマウントを最適化する方法の詳細については、追加情報の検索先

ツールボリュームとライブラリボリュームのマウントコマンド

次のスクリプトは、ツールボリュームとライブラリボリュームに推奨されるNFSマウントオプションを示しています。

```
%> mount -t nfs \
    -o "nocto,actimeo=600,hard,rsize=262144,wszie=262144,vers=3,tcp,mountproto=tcp" \
    <ontap svm>:/<volname> /mnt/<volname>
```

ファイルとメタデータのロード時間を短縮し、ONTAPストレージシステムの負荷を軽減するために、NetAppでは、本質的に読み取り専用のツールボリュームとライブラリボリュームに対して、アグレッシブなコンピューティングサーバキャッシングを使用することを推奨しています。NFSマウントオプションを使用nocto,actimeo=600すると、EDAワークロードのメタデータI/Oが最大90%削減されます。また、このオプションを使用すると、ツールやライブラリのロード時間を大幅に短縮することができます。これは、業界が7nm、5nm、3nmの設計ノード（トランジスタサイズ）を採用するにつれて、半導体テクノロジライブラリの数とサイズが劇的に増加しているため、ますます大きな課題となっています。

EDAワークフローでは、**Makefile**や**Makefile**に似た依存関係の追跡が頻繁に使用されます。その結果、ファイルシステムから日付とタイムスタンプ、ファイルの存在チェックなどが返されるため、メタデータI/Oの60~80%が表示されることも珍しくありません。**Makefile**のような依存性チェックに加えて、EDAツールやPerlやPythonのようなスクリプトツールは、ロードするモジュールからの長いインクルードファイルやディレクトリ検索パスをスキヤンすることがよくあります。その結果、ファイルおよびディレクトリの検索で使用される追加のメタデータI/Oトランザクションが生成されます。これにより、NFSファイルに不要な負荷がかかる可能性があります。

NFSサーバ側のアグレッシブなキャッシングは、ツールやライブラリのボリュームが静的または変更されていないという前提に基づいて機能します。新しいツールやライブラリは通常、古いバージョンの隣のディレ

トリにインストールされ、上書きされたり変更されたりすることはめったにありません。分析中に最新のライブラリを取り込むことは、EDAワークロードにとって不適切な手法と考えられています。EDAフローは通常、特定のジョブに対して一連のツールとライブラリを指定し、それらのバージョンが実行全体を通じて使用されます。

この問題の詳細については、ブログ記事を参照してください。[LinuxでのNFSマウント中にactimeoが意味するものは何ですか？](#)

高パフォーマンスのスクラッチボリューム用のmountコマンド

次のマウントオプションは、ハイパフォーマンスなNFS読み取り/書き込みトランザクション用に最適化されています。読み取り専用のツールボリュームやライブラリボリュームとは異なり、アグレッシブなファイルキャッシュは通常推奨されません。

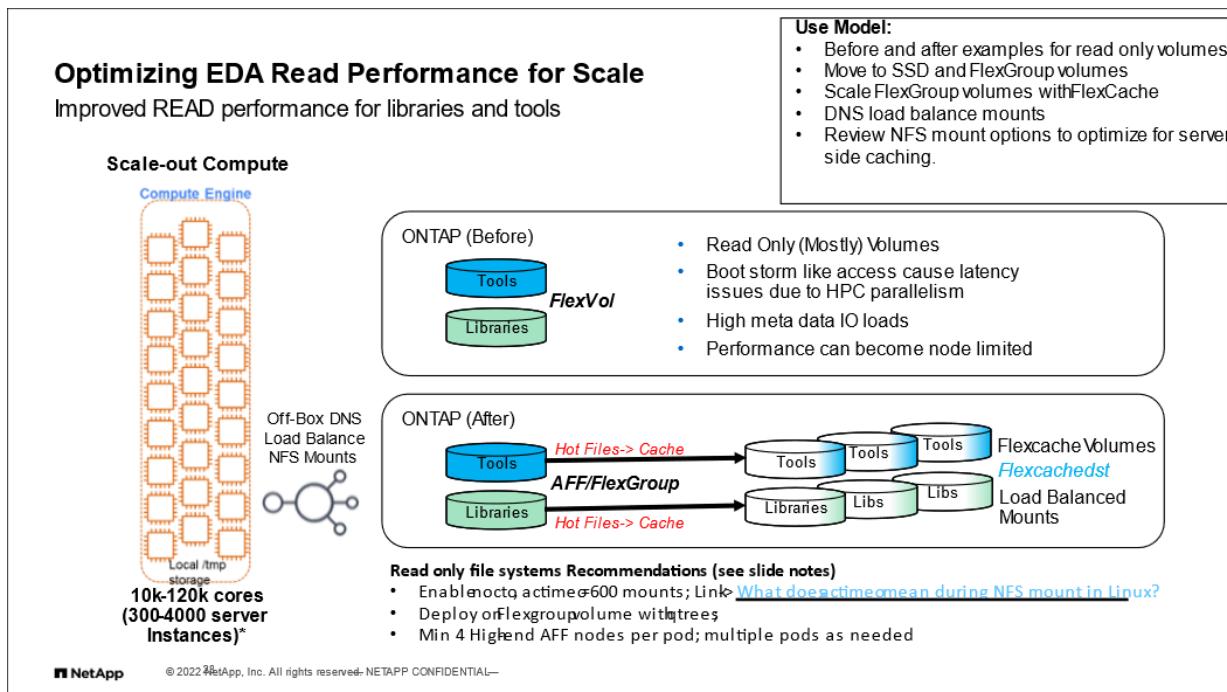
次のスクリプトは、スクラッチボリュームに推奨されるNFSマウントオプションです。

```
%> mount -t nfs \
-o "hard,rsize=262144,wszie=262144,vers=3,tcp,mountproto=tcp" \
<ontap svm>:/<volname> /mnt/<volname>
```

DNSマウントバランシングによるネットワーク負荷の向上

拡張とDNSロードバランシングNFSマウントは、大規模環境での読み取りパフォーマンスを向上させることができます。複数のツールおよびライブラリFlexCacheボリュームを作成し、使用可能なすべてのFlexCachedツールおよびライブラリボリュームにNFSマウントを分散させることができます。

図8) 拡張性に応じたEDA読み取りパフォーマンスの最適化



詳細情報の入手方法

このドキュメントに記載されている情報の詳細については、以下のドキュメントやWebサイトを確認してください。

- 電子設計自動化のベストプラクティス
<https://www.netapp.com/media/19368-tr-4617.pdf>

- Azure Cloud Volumes ONTAPオンラインドキュメント
<https://docs.netapp.com/us-en/cloud-manager-cloud-volumes-ontap/task-getting-started-azure.html>
- NetApp ONTAP FlexGroupボリュームのベストプラクティスと実装ガイド
<https://www.netapp.com/pdf.html?item=/media/12385-tr4571pdf.pdf>
- NFS in NetApp ONTAPベストプラクティスおよび実装ガイド
<https://www.netapp.com/pdf.html?item=/media/10720-tr-4067.pdf>

バージョン履歴

バージョン	日付	リリース ノート
1.0	2022年11月	初版リリース

本ドキュメントに記載されている製品や機能のバージョンがお客様の環境でサポートされるかどうかについては、NetApp サポートサイトで [Interoperability Matrix Tool \(IMT\)](#) を参照してください。NetApp IMT には、NetApp がサポートする構成を構築するために使用できる製品コンポーネントやバージョンが定義されています。サポートの可否は、お客様の実際のインストール環境が公表されている仕様に従っているかどうかによって異なります。

機械翻訳に関する免責事項

原文は英語で作成されました。英語と日本語訳の間に不一致がある場合には、英語の内容が優先されます。公式な情報については、本資料の英語版を参照してください。翻訳によって生じた矛盾や不一致は、法令の順守や施行に対していかなる拘束力も法的な効力も持ちません。

著作権に関する情報

Copyright © 2024 NetApp, Inc. All Rights Reserved. Printed in the U.S.このドキュメントは著作権によって保護されています。著作権所有者の書面による事前承諾がある場合を除き、画像媒体、電子媒体、および写真複写、記録媒体、テープ媒体、電子検索システムへの組み込みを含む機械媒体など、いかなる形式および方法による複製も禁止します。

NetApp の著作物から派生したソフトウェアは、次に示す使用許諾条項および免責条項の対象となります。

このソフトウェアは、NetApp によって「現状のまま」提供されています。NetApp は明示的な保証、または商品性および特定目的に対する適合性の暗示的保証を含み、かつこれに限定されないいかなる暗示的な保証も行いません。NetApp は、代替品または代替サービスの調達、使用不能、データ損失、利益損失、業務中断を含み、かつこれに限定されない、このソフトウェアの使用により生じたすべての直接的損害、間接的損害、偶発的損害、特別損害、懲罰的損害、必然的損害の発生に対して、損失の発生の可能性が通知されていたとしても、その発生理由、根拠とする責任論、契約の有無、厳格責任、不法行為（過失またはそうでない場合を含む）にかかわらず、一切の責任を負いません。

NetApp は、ここに記載されているすべての製品に対する変更を隨時、予告なく行う権利を保有します。NetApp による明示的な書面による合意がある場合を除き、ここに記載されている製品の使用により生じる責任および義務に対して、NetApp は責任を負いません。この製品の使用または購入は、NetApp の特許権、商標権、または他の知的所有権に基づくライセンスの供与とはみなされません。

このマニュアルに記載されている製品は、1つ以上の米国特許、その他の国の特許、および出願中の特許により保護されている場合があります。

本書に含まれるデータは市販の製品および/またはサービス (FAR 2.101 の定義に基づく) に関するもので、データの所有権は NetApp, Inc. にあります。米国政府は本データに対し、非独占的かつ移転およびサブライセンス不可で、全世界を対象とする取り消し不能の制限付き使用権を有し、本データの提供の根拠となつた米国政府契約に関連し、当該契約の裏付けとする場合にのみ本データを使用できます。前述の場合を除き、NetApp, Inc. の書面による許可を事前に得ることなく、本データを使用、開示、転載、改変するほか、上演または展示することはできません。国防総省にかかる米国政府のデータ使用権については、DFARS 252.227-7015(b) 項で定められた権利のみが認められます。

商標に関する情報

NetApp、NetAppのロゴ、<https://www.netapp.com/company/legal/trademarks/>に記載されているマークは、NetApp, Inc.の商標です。その他の会社名と製品名は、それを所有する各社の商標である場合があります。

TR-4946-JP