



テクニカル レポート

NVMe-oFを使用した最新SANの実装と設定

NetApp
Michael Peppers / Martin George
2023年2月 | TR-4684

概要

このドキュメントでは、NVMe-oFトランスポート（NVMe/FCおよびNVMe/TCP）を実装および設定する方法について説明します。NVMeプロトコルとトランスポートを使用して可用性とパフォーマンスに優れた最新SANソリューションを構築するための設計、実装、設定、管理のガイドライン、ベストプラクティスが記載されています。

<<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

NVMeとは	5
NVMe標準はどこから来たのか?	5
NVMe-oFは高速.....	6
ストレージ接続アーキテクチャとしてのNVMe	8
業界初のエンドツーエンドNVMe - NetApp AFF A800	8
NVMe-oF	10
NVMeとデータファブリック	11
NVMeと高可用性.....	16
NVMe over Fibre Channel	18
NVMe over TCP (NVMe/TCP)	20
NVMeの使用	21
NVMe/FCとNVMe/TCPのどちらの導入を選択すべきか.....	21
ONTAP機能のサポートと共存	21
相互運用性.....	23
LUNからネームスペースまたはネームスペースへの変換.....	25
NVMe-oFのベストプラクティス	26
NVMe / FCのベストプラクティス.....	26
NVMe-oFのセットアップと設定.....	27
セットアップと設定の詳細手順リファレンス.....	27
パフォーマンス	28
NVMe / TCPのベストプラクティス	31
NVMe-oFの機能拡張	31
ONTAP 9.6	31
ONTAP 9.9.1	31
ONTAP 9.10.1	32
ONTAP 9.11.1	33
ONTAP 9.12.0	33
ONTAP 9.12.1	33
付録A : ONTAP System Managerを使用したONTAP NVMe/FCおよびNVMe/TCPオブジェクトの作成 ..	34
付録B : ONTAP NVMe/FCおよびNVMe/TCPのCLIコマンド-初期セットアップと検出	38
ONTAPコントローラの場合.....	38
ホスト上.....	40

ONTAPコントローラ	40
付録C：ホスト設定情報	41
付録D：LUNとネームスペース間の変換	41
LUNのネームスペースへの変換	41
ネームスペースのLUNへの変換	41
付録E：NVMe/FCの拡張性と制限	42
付録F：トラブルシューティング	42
NVMe/FCのIPFCの詳細ログ	42
NVMe-CLIの一般的なエラーとその回避策	43
デバッグに必要なファイルとコマンド出力	44
付録G：MCC IPでのNVMe/FCの設定とセットアップ	44
付録H：NVMe/TCP経由のセキュアな認証のセットアップ	45
追加情報の入手方法	45
ホストOSのセットアップと設定	45
標準文書	45
SUSE Enterprise Linuxのリンク	46
Brocadeリンク	46
ビデオ、Webキャスト、ブログ	46
ホワイトペーパー、製品の発表、および分析	46
NetAppのドキュメント、テクニカルレポート、その他のNVMe関連資料	46
ONTAPクラウド	47
NetApp検証済みアーキテクチャ	47
NVMeのその他のドキュメント	47
バージョン履歴	47
お問い合わせ	48
表一覧	
表1) SCSIとNVMeの用語	21
表2) NVMeでサポートされる機能、またはNVMeと共存可能なONTAP機能	22
表3) 現在NVMeでサポートされていないONTAP機能	22
表4) LUN <->ネームスペース変換ユーティリティ機能のサポート	25
表5) AFF A700の4KランダムリードNVMe/FCとFCPの比較	29

図一覧

図1) NVMeの定義	5
図2) NVMe/FCの高速性	7
図3) NVMeプロトコルが非常に効率的な理由	8
図4) NetApp AFF A800	9
図5) NetAppエンドツーエンドNVMe	9
図6) AFF A800 FCPとNVMe/FCのパフォーマンスの比較	10
図7) SCSIとNVMe-oF転送の比較	11
図8) NVMeは複数のネットワークトランスポートを使用可能	13
図9) NVMe over Fabrics (NVMe-oF)	13
図10) FCPとNVMe/FCフレームの比較	14
図11) iSCSIとNVMe/TCPデータグラムの比較	14
図12) TCPデータグラム内のNVMe	15
図13) SCSIとNVMeのスタックアーキテクチャの比較	15
図14) TP 4004 : ANAベース提案 (2018年3月批准)	16
図15) TP 4028 : ANAのパスとトランスポート (2018年1月承認)	17
図16) INCITS FC-NVMe-2 Rev 1.08 T11-2019-00210-v004 : FC標準を使用したNVMeコマンドとデータ転送の定義	17
図17) NVMe/FCストレージフェイルオーバー : ONTAP 9.5でANAが導入	18
図18) NVMe / FCとANA非対応の比較	18
図19) システムを停止することなく最新テクノロジーを導入	19
図20) NVMe/FCホスト構成ページ	
図21) 新しいIMT NVMe/FCプロトコルとNVMe/TCPプロトコルのフィルタ	24
図22) NVMe/FCによる超ハイパフォーマンス設計	28
図23) AFF A700ハイアベイラビリティ (HA) ペア、ONTAP 9.4、8KランダムリードFCPとNVMe/FCの比較	29
図24) AFF A700 HAペア、ONTAP 9.4、4KランダムリードのFCPとNVMe/FCの比較	30
図25) FCPからNVMe/FCへの移行とONTAPのアップグレードごとのパフォーマンスの向上	30
図26) リモートI/OをサポートしないNVMe-oF	31
図27) リモートI/Oを使用するNVMe-oF	32
図28) OnCommand System Manager-SVMの作成	34
図29) OnCommand System Manager-SVMの作成 : NVMe転送の設定-NVMe/FCおよびNVMe/TCP	35
図30) OnCommand System Manager-SVMの作成 : NVMe/FCの設定	35
図31) OnCommandシステムマネージャ-SVMの作成 : NVMe/TCPの設定	36
図32) OnCommand System Manager-SVMの作成 : 管理の詳細の設定	36
図33) 新しく作成したSVMの表示	37
図34) OnCommandシステムマネージャ-新しいNVMe名前空間の作成	37
図35) OnCommandシステムマネージャ-新しく作成したNVMe名前空間の表示	38
図36) 新しく作成したNVMeサブシステムの表示	38

NVMe とは

NVM Expressのデータストレージ標準であるNVMeは、新しいストレージインフラの構築や最新のストレージインフラへのアップグレードを進めている企業にとって、コアテクノロジーとして登場しつつあります。

NVMeは、ソリッドステートストレージデバイス向けに最適化されたプロトコルであり、NVMeのコンポーネントとシステムに関するオープンソースのアーキテクチャ標準でもあります。

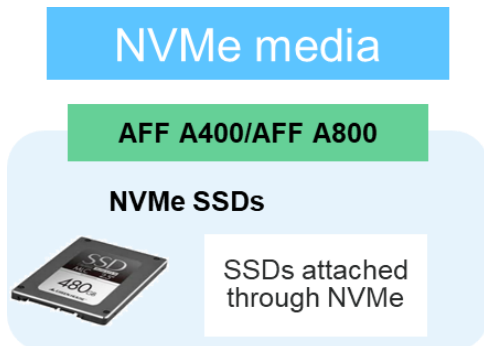
NVMe規格は、現在だけでなく将来のメモリテクノロジーにも、広帯域幅と低レイテンシのストレージアクセスを提供するように設計されています。NVMeは、SCSIコマンドセットをNVMeコマンドセットに置き換え、SCSI、Serial-Attached SCSI (SAS)、SATAなどの古い標準よりも大幅に高速で広帯域のハードウェアプロトコルであるPCIeに依存しています。

SCSIは約40年前に発表され、8インチフロッピードライブやファイルキャビネットサイズのHDDなど、当時普及していたストレージテクノロジー向けに設計されました。また、はるかに低速なシングルコアCPUと、当時利用可能だった少量のDRAMを中心に設計されていた。

一方NVMeは、マルチコアCPUとギガバイトのメモリを搭載した不揮発性フラッシュドライブで動作するように開発されました。また、1970年代以降の計算機科学の大幅な進歩を利用して、効率的にデータを解析および操作するコマンドセットを合理化できます。

NVMeはオーバーロードされています。NVMe over Fabrics (NVMe-oF) の導入に伴い、メディアタイプの置き換えを検討しているかどうかを定義することが重要になります。1つのメディアタイプをNVMe接続ディスクに交換することを指す場合によく使用されます (図1)。このユースケースでは、ラップトップ、デスクトップ、サーバ、さらにはストレージレイなど、従来のメディアタイプをNVMe接続ディスクに置き換えることができます。

図1) NVMeの定義



NVMe標準はどこから来たのか？

NVMeバージョン1.0標準は、2011年3月にNVM Express, Inc.によって承認されました。以降、この標準にはいくつかの更新があり、最新のNVMeバージョン1.3aが2017年11月に批准されました。NVMe NVM Expressは2015年11月に補完的仕様であるNVMe Management Interface (NVMe-MI) を発表した。NVMe-MIはインバンド管理とアウトオブバンド管理の両方に重点を置いています。NVMe-oF仕様は2016年6月に追加されました。NVMe-oFは、ネットワークまたはファブリック上でNVMeプロトコルを使用して定義されません。

NVMe 2.0ファミリーの仕様は、NVMe仕様の最新バージョンです。NVMe-oF、NVMe-MI、NVMe-kVなどのNVMeの機能強化と、新しいNVMe 2.0ファミリーの仕様承認されたいくつかの技術提案 (TPS) が組み合わされています。NVMe標準は、ソフトウェアスタックから最新のストレージデバイスへのデバイスアクセスまでをカバーします。NVMe 2.0仕様について知っておくべきすべての要素と新しい技術提案：

<https://nvmexpress.org/everything-You-need-to-know-the-NVMe-2-0-specifications-and-new-technology-proposals/>をご覧ください。

注：技術提案は、Internet Engineering Task Force (IETF) Request for Comment (RFC) に相当する NVMe NVM Expressです。NVM Expressワークグループは、技術提案が重点を置いている分野に関する技術提案をレビューのために提出します。最終的には、技術提案が投票で承認され、NVMeの仕様やプロセスの延長として承認される可能性があります。

NVMeプロトコルの開発と管理を担当しているのは誰ですか？

NVMeは、Peripheral Component Interconnect PCI標準化組織であるPeripheral Component Interconnect Special Interest Group (PCI-SIG) から生まれた標準化組織によって開発されました。NVMeの標準化組織は、[Non-Volatile Memory Express, Inc](#)です。

NVMe標準のドキュメントの参照先

NVMeの仕様、ホワイトペーパー、プレゼンテーション、ビデオなどの資料は、[NVMe Express, Inc.のWebサイト](#)からダウンロードできます。

NVMe/FC標準は、International Committee for Information Technology Standards (INCITS) T11 Committee FC-NVMe標準、[T11-2017-00145-v004 FC-NVMe](#)でFibre Channel Industry Association (FCIA) によってさらに定義されています。

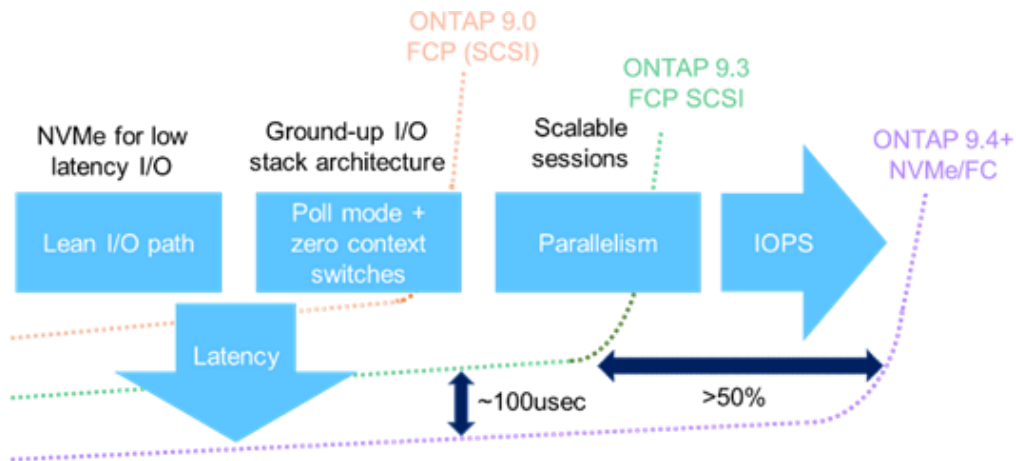
NVMe-oFは高速

NVMeは、データストレージのパフォーマンスに関して、IOPS、スループット、レイテンシ、CPU利用率という4つの重要な属性に対処できるため、最新のデータセンターに欠かせない要素になります。

- **IOPS** は、デバイスが1秒あたりに実行できる読み取りまたは書き込み処理の数です。これは、読み取りと書き込み、シーケンシャル処理とランダム処理で異なり、格納または読み出されるブロックのサイズによって異なります。ほとんどのデバイスでは、4KBブロックや8KBブロックなどの小さなブロックI/Oサイズを使用している場合、IOPS値が高くなります。しかし、実際のアプリケーションでは、32KBブロックや64KBブロックなど、より大きなI/Oサイズが必要になることが多いため、関連するI/O特性に基づいて評価することが重要です。
- **スループット** とは、ストレージデバイスがデータの読み取りや書き込みをどのくらいの速さで実行できるかを示す指標で、多くの場合、1秒あたりのギガビット数で指定されます。通常、I/Oサイズが大きいほど高く、I/Oの方向やアクセスタイプ（ランダムまたはシーケンシャル）によって異なります。したがって、実際の運用環境に応じて評価する必要があります。
- **レイテンシ** は、読み取り処理または書き込み処理が開始されてから完了するまでの時間です。ストレージのレイテンシは、転送されるデータのサイズ、シーケンシャルかランダムか、読み取りか書き込みか、ネットワークの速度によって異なります。応答性に優れた魅力的なエクスペリエンスをユーザに提供するには、低ストレージレイテンシ、特に低読み取りレイテンシが不可欠です。
- **CPU利用率** は、レイテンシ内で、特定のワークロードをサポートするのに十分なスループットでI/Oを生成するために使用されるCPUサイクルの数です。上記の他の方法とは異なり、NVMeの効率性によるCPU利用率の削減は、FCPやiSCSIなどのSCSIベースのプロトコルからNVMeに移行するメリットとしてあまり知られていません。CPU利用率を下げるメリットとしては、特定のストレージコントローラにより多くのワークロードを統合できることや、特定のワークロードをホストするために必要なサーバの数を削減できることが挙げられます。ストレージとホスト側の削減によってIT投資が増加し、NVMeに移行するプロジェクトのROI（投資回収率）を短期間で達成できます。

図2 と 図3は、NVMeが非常に効率的で高速である理由を示しています。

図2) NVMe/FCの高速性



IOPSと帯域幅の向上は、主にNVMeの柔軟性と、高速転送テクノロジーを活用してNVMeのコマンドとデータを移動できることによる成果です。次のトランスポートが含まれます。

- **FCP** : 現在、32Gbpsと64Gbpsの速度で提供されており、まもなく128Gbpsにもなります。
- **Remote Direct Memory Access (RDMA) プロトコル** :
 - データセンターのファストイーサネット : 現在、25、40、50、100、および200Gbpsで提供されています。
 - InfiniBand (IB) : 現在、最大100Gbpsの速度で提供されています。
- **PCI Express 3.08ギガトランザ/秒 (GTps)** をサポートします。これは約6.4Gbpsに相当します。
- **NVMeコマンドとペイロードでユビキタスイーサネットネットワークを使用できるTCP。**
現在、より高速なイーサネットネットワークは、10、25、40、50、100、および現在は200Gbpsで利用できます。

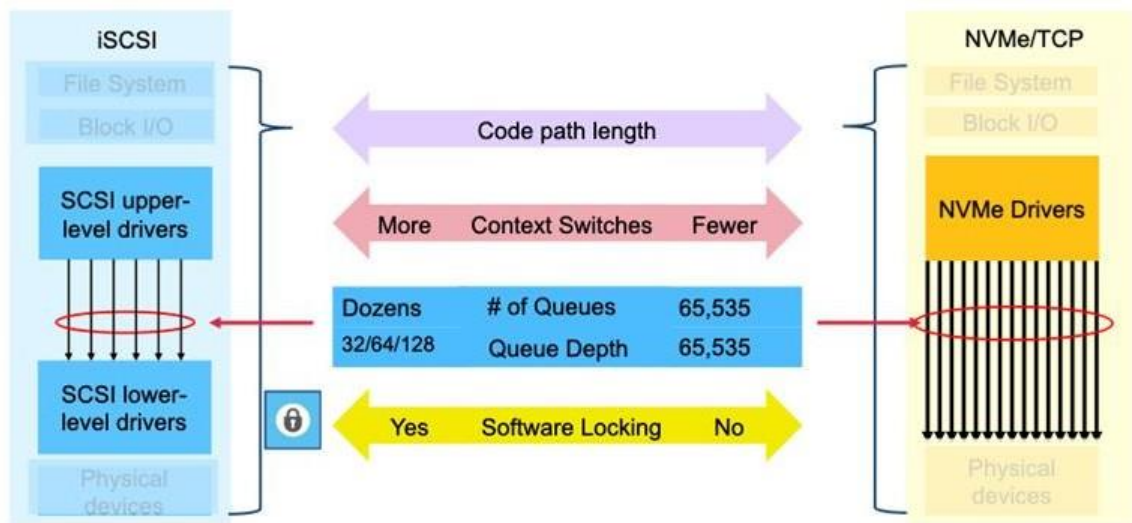
NVMeで可能な大規模な並列化の結果、パフォーマンスが向上しました。この並列化により、プロトコルは複数のスレッドを同時に処理するために、複数のコアに処理を分散させることができます。

レイテンシの改善は、次のような要因の組み合わせによってもたらされます。

- 高い並列性。I/O送信キューと完了キューのペアは、ホストCPUコアにアライメントされます。各ホスト/コントローラペアには、それぞれ独立したNVMeキューのセットがあります。
- NVMeコマンドセットの合理化
- ハードウェア割り込みに代わるポーリングモードドライバ
- ソフトウェアロックの排除
- コンテキストスイッチの削除

これらの要素が連携して、企業のビジネスクリティカルなアプリケーションの主要な指標であるスループットの向上とレイテンシの低減を実現します。

図3) NVMeプロトコルが非常に効率的な理由



ストレージ接続アーキテクチャとしてのNVMe

NVMeは現在、ディスクとディスクシェルフの接続に主に使用されています。多くのストレージベンダーやサプライヤが、NVMeをストレージ接続アーキテクチャおよび標準として使用するソリューションを導入しています。技術的には、ほとんどの場合、I/Oの実行にはNVMeが使用されますが、物理的な転送は主にPCIeです。

このシナリオでは、NVMeはSCSIコマンドセットをNVMeコマンドセットに置き換え、ドライブをストレージコントローラに接続するためにSATAまたはSASをPCIeに置き換えます。NVMeは物理的な接続と転送に依存します。PCIeをトランスポートとして使用します。

NVMe接続フラッシュは、より多くの帯域幅を提供し、レイテンシを低減します。理由は次のとおりです。

- より多くのより深いキューを提供します。64k (65,535) のキューがあり、それぞれのキューの深さは64kです。
- NVMeコマンドセットは合理化されているため、従来のSCSIコマンドセットよりも効率的です。

NVMeコマンドセットを使用してSAS 12GBバックエンドおよびSCSIコマンドセットをPCIe接続ドライブに変更すると、あらゆるバックエンドプロトコルのパフォーマンス（スループット）が向上し、レイテンシが低減します。この改善はディスクアクセスによるものです。ディスクアクセスはより効率的で、プロセッサの消費電力が少なく、並列化が可能です。理論的には、パフォーマンスの向上によってスループットが約10~15%向上し、レイテンシが10~25%短縮されます。ワークロードプロトコルやコントローラで実行される他のワークロードの違い、I/Oを実行するホストの相対的なビジー状態でさえも、これらの数値は大幅に原因されます。

業界初のエンドツーエンドNVMe - NetApp AFF A800

[AFF A800オールフラッシュアレイ](#)は、NVMe接続のソリッドステートドライブ（SSD）を使用する初のNetAppアレイです。

図4) NetApp AFF A800

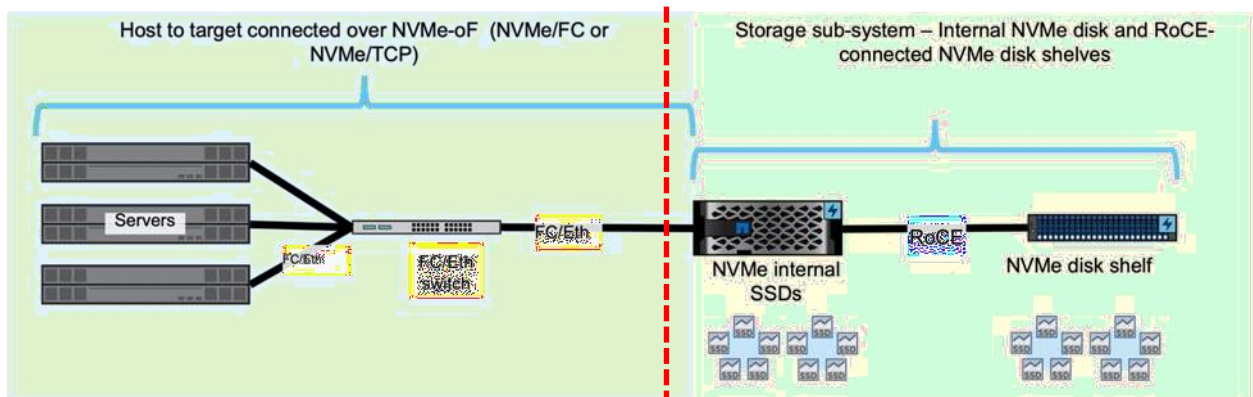


もし、2018年以降に業界をリードするハイエンドオールフラッシュアレイの設計に着手したとしたら、AFF A800のようなシステムを思いつくでしょう。その一部をご紹介します。

- 32Gbps FCを介した業界初のエンドツーエンドNVMe/FCホストツーフラッシュアレイ。AFF A800 FCホストバスアダプタ (HBA) は、16Gbpsまたは8Gbps (N-2) まで自動ネゴシエートできます。
- NASおよびiSCSIプロトコルのフロントエンド接続に対応した業界初の100GbE接続
- 15.3TB NVMe SSDを搭載した4Uシャーシで2.5PiBの実効容量を実現 (近日提供予定)
- 100GbE NetApp MetroCluster IP (MCC IP) で最高レベルのパフォーマンスを実現

ハイパフォーマンスと低レイテンシを実現する内蔵NVMe SSDは、あらゆるアプリケーションを高速化します。さらに、AFF A800をNVMe/FCと組み合わせることで、エンドツーエンドのNVMe接続が実現します (図5)。これにより、より低いレイテンシで高いパフォーマンスを実現できます。

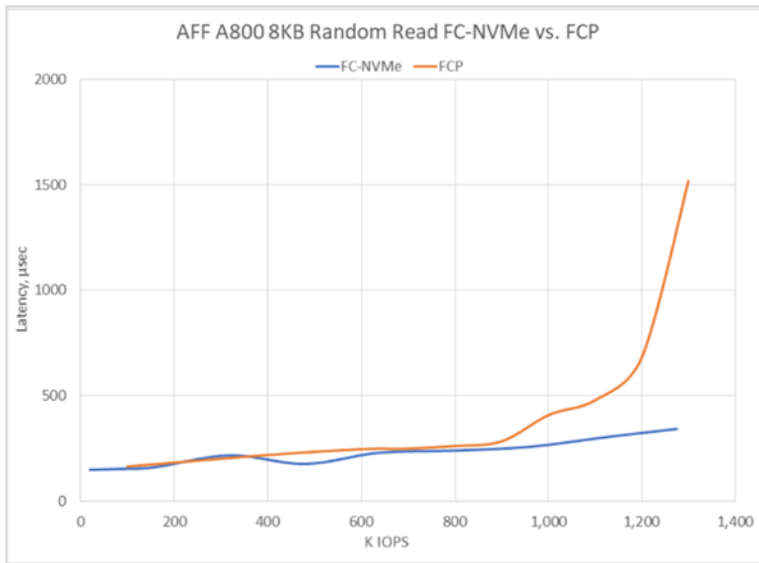
図5) NetAppエンドツーエンドNVMe



一部のお客様は、100GbEをサポートするストレージ解決策を待っていました。また、NASプロトコルとiSCSIプロトコルで100GbEを有効にし、FCPとNVMe/FCの両方で32Gbps FC (2016年にNetAppがリリース) を有効にして、優れたアプリケーション帯域幅を提供するシステムにすることもできます。

AFF A800は、SANのみの環境、NASのみの環境、またはその両方を組み合わせた環境に最適です。詳細については、この[短いライトボドビデオ](#)を参照してください。図6は、FCPまたはNVMe/FCをAFF A800に導入することで期待できるパフォーマンスの向上を示しています。青いNVMe/FCのラインはFCPと比較して横ばいです。これは、NVMe/FCが提供するパフォーマンス容量が、AFF A800のFCPに比べて大幅に大きいことを示しています。つまり、AFF A800でNVMe/FCを使用すると、スループットが向上します。FCPの場合、十分なヘッドルーム (FCPと比較して利用可能なパフォーマンス容量) があるNVMe/FCと比較すると、FCPは約120万IOPSで500ミリ秒強となります。

図6) AFF A800 FCPとNVMe/FCのパフォーマンスの比較



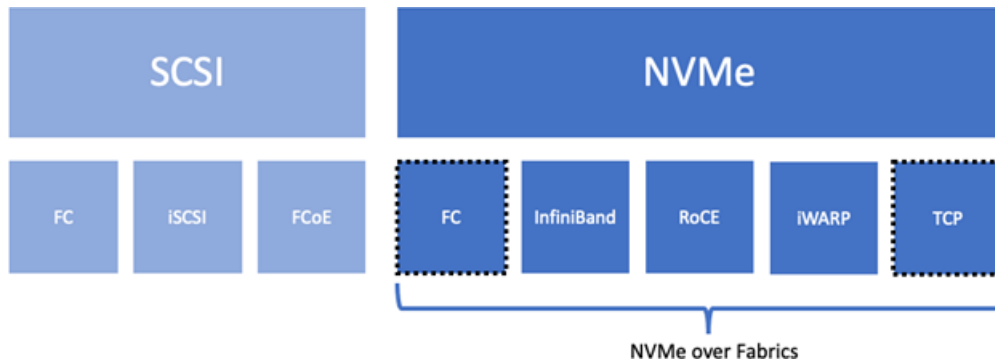
NVMe-oF

NVMeは単なるストレージ仕様ではありません。NVMe-oFプロトコル拡張機能には、サーバからネットワーク、ストレージシステムまでのデータパス全体が含まれます。HDDをフラッシュSSDに置き換えることでHDDのボトルネックが解消されたあと、もう1つのボトルネックが現れました。ローカルとSANの両方でデータへのアクセスに使用されるストレージプロトコルです。

NVMe over Fabrics委員会は、ローカルストレージアクセスプロトコルのアップグレードにNVMeを導入しました。NVMe-oFのリリースに伴い、委員会は、さまざまなネットワークプロトコルとファブリックプロトコルを使用して、NVMeプロトコルとコマンドセットを長距離にわたって使用するための仕様とアーキテクチャを追加しました。その結果、FCPやiSCSIからNVMe/FCなどのNVMe-oF仕様に移行するワークロードのパフォーマンスが大幅に向上し、レイテンシが低減します。

エンドツーエンドのNVMeを実装するには、NVMe接続のソリッドステートメディアだけでなく、ストレージコントローラからホストサーバへのNVMe転送も必要です。元々のNVMe仕様では、アタッチメント用のコマンドセットとアーキテクチャが設計されていましたが、伝送と伝送はPCIe（または別の物理トランスポート仕様）に依存しています。主に、不揮発性フラッシュストレージテクノロジーをローカルサーバに接続することに重点を置いていました。NVMeはSCSIコマンドに代わるもので、それぞれの処理キューの数とキュー深度の両方が増加します。NVMeはコンテキストスイッチを削減し、ロックレスです。これらの機能強化により、PCIバスを使用するディスクなど、NVMe接続ディスクのアクセス時間と応答時間が大幅に短縮されます。図7は、SCSIとNVMe-oFのプロトコルと転送の組み合わせを比較したものです。

図7) SCSIトランスポートとNVMe-oFトランスポートの比較



NVMeとデータファブリック

NVMeは、ローカルの不揮発性ストレージをコンピュータやサーバに接続するためのアクセスプロトコルとアーキテクチャを定義します。NVMe-oFは拡張性と範囲が強化され、元のNVMe仕様を強化します。NVMe-oFはNVMeの拡張機能であり、NVMeをSAN市場に効果的に提供します。NVMe-oFは、FC、イーサネット、IBなどのさまざまなネットワークストレージトランスポート経由でNVMeを転送する方法を定義および作成するための仕様です。

NVMe-oFは、最終的にNVMeを新しいブロックストレージプロトコルタイプとして追加します。一般的に、ベンダーが特定のNVMe-oFトランスポート（NVMe/FCなど）を開発する場合は、そのプロトコルとアーキテクチャの仕様に従う必要があります。

NVMe-oFは、NVMeがFCやイーサネットなどの既存の転送テクノロジーを使用してNVMeプロトコルを距離を超えて転送する方法を定義し、スイッチやルータなどのネットワークテクノロジーを使用できるようにします。NVMe-oFはこれらの転送プロトコルをサポートすることで、大規模なストレージレイのパフォーマンスを大幅に向上させると同時に、次のような他のプロトコルに代わるストレージプロトコルの並列化を促進します。

- **FCP** : FCフレーム内にカプセル化されたSCSIコマンド。
- **iSCSI** : IP/イーサネットフレームにカプセル化されたSCSIコマンド。
- **FCoE** : カプセル化されたSCSIコマンドを含むFCフレームは、イーサネットフレーム内にカプセル化されます。
- NetAppは、NetApp ONTAP 9.4で、NVMe/FCを使用する初のNVMe-oF実装を発表しました。実際、ONTAP 9.4は、NVMe/FCを使用した完全なエンドツーエンド（ホストからストレージ）解決策の業界初のバージョンです。
- ONTAP 9.10.1では、イーサネットベースのNVMe-oF転送としてNVMe/TCPが追加されました。
- 図5は、NVMe-FCと新しいAFF A800でエンドツーエンドのNVMeフラッシュを実現するONTAP 9.4を示しています。

NVMe-oFは主に、NVMeプロトコルをデータネットワークやファブリックに拡張することを目的としています。コンピューティングをブロックベースのストレージに接続するために使用するアクセスアーキテクチャとプロトコルを定義します。これは、次のような現在のブロックプロトコルの更新と考えるのが最も簡単です。

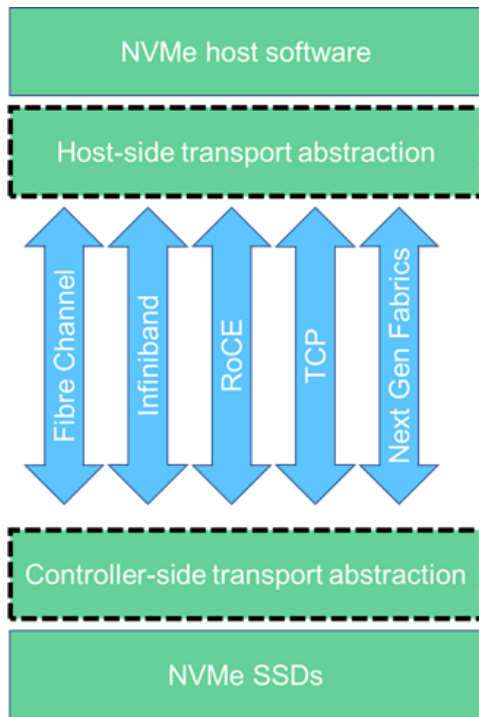
- **FCP** : FCPは、SCSI Command Descriptor Block（CDB;コマンド記述子ブロック）をFCフレーム内にカプセル化します。FCは転送方法を定義しますが、FCPはFCプロトコルを使用してSCSI（CDB）をカプセル化することを特に意味します。現在、最も一般的なSANプロトコルはFCPです。FCPファブリック（ネットワーク）の速度は1~32Gbpsです。最も高速なのは、8Gbps、16Gbps、32Gbpsです。
- **iSCSI** : iSCSIはInternet Engineering Task Force（IETF）によって[RFC 3270 Internet Small Computer Systems Interface](#)で最初に定義された。このドキュメントは、[RFC 7143 Internet Small Computer System Interface \(iSCSI\) Protocol \(Consolidated\)](#)に置き換えられました。RFC 7143は、2004年に導入された元の仕様を更新および最新化しました。

NVMe-oFは、さまざまな現在のトランスポート内でNVMeを転送するために使用できる仕様、アーキテクチャ標準、およびモデルを提供します。NVMe-oF転送には次のものがあります。

- **NVMe/FC**。転送にFCを使用するNVMe。詳細については、NVMe over Fibre Channel 「チャンネル」の項を参照してください。
- **NVMe/TCP** NVMeはTCPデータグラムにカプセル化されます。NVMe/TCPは、最も一般的なNVMe over Ethernetバリエーションである可能性があり、最終的にはiSCSIの論理的な置き換えになる可能性があります。iSCSIと同様に、NVMe/TCPは標準のNICとイーサネットスイッチを使用します。そのため、RDMA over Converged Ethernet (RoCE) をサポートするために必要なRDMA NIC (RNIC) やデータセンターブリッジング (DCB) スイッチなどの特殊なビルディングブロックを使用せずに、イーサネット上でNVMe-oFを導入したい環境にとって魅力的なオプションです。
- **RDMAを使用したNVMe転送**。RDMAをサポートするトランスポートは次のとおりです。
 - **NVMe over InfiniBand (NVMe/IB)** : この解決策は、現在100Gbpsに対応可能なIBを超高速トランスポートとして使用しています。IBは信じられないほど高速ですが、高価であり、距離と拡張性の両方に制限があります。NVMe-oF (NVMe/IBターゲットを使用) を提供した業界初のエンタープライズクラスのストレージレイはNetApp EF570アレイです。2Uプラットフォームで100万IOPSと21GBpsを実現できます。詳細については、[NetApp EF570オールフラッシュアレイのデータシート](#)を参照してください。
 - **RDMA over Converged Ethernet (RoCE)** :
 - Internet Wide-Area RDMA Protocol (iWARP) は、Direct Data Placement Protocol (DDP) を使用してRDMAを転送します。DDPは、TCPまたはSecure TCP (STCP) を使用して転送されます。DDPはストリームでデータを送信し、TCPプロトコルのデータ単位に合わせてセグメント化することはありません。
 - RoCEはTCPを必要としないため、レイテンシが低くなります。RoCEには、DCBとPriority Flow Control (PFC ; 優先フロー制御) をサポートするイーサネットスイッチが必要です。DCBスイッチは標準のイーサネットスイッチとは異なり、高価になる傾向があります。ホストおよびストレージコントローラにはRNICがインストールされている必要があります。これらの要件により、RoCEのクラウド導入が制限される可能性があります。RoCEには次の2種類があります。
 - 元々のRoCE仕様であるRoCE v1では、同じサブネット内のイニシエータとターゲット間の通信を可能にするデータリンク層プロトコルが定義されています。RoCEは、サブネット間でルーティングできないリンクレイヤプロトコルです。
 - RoCE v2は、IPv4またはIPv6でユーザデータグラムプロトコル (UDP) を使用するインターネットレイヤプロトコルです。これは、サブネット間でルーティングできるレイヤ3インターネットレイヤプロトコルです。UDPでは順序どおりの配信が強制されず、RoCE v2仕様では順序どおりのパケット配信が許可されないため、DCBネットワークはパケットを送信順に配信する必要があります。RoCE v2では、[明示的輻輳通知 \(ECN\)](#) ビットを使用してフレームと[輻輳通知パケット \(CNP\)](#) をマークするフロー制御メカニズムも定義されています。領収書を確認します。
 - **iWARPを介したRDMA**。RDMAの使用をサポートするための機能拡張。

図8は、NVMe-oFスタックと、トランスポートとして使用できる一部のネットワークとファブリックを示しています。

図8) NVMeは複数のネットワークトランスポートを使用可能



NVMeは、NVMe-oFとも呼ばれます。NVMe-oFは、NVMeコマンドセットとデータペイロードをさまざまな一般的なネットワークプロトコルやファブリックプロトコルにカプセル化する方法を規定した標準的な機能強化です。図9は、NVMe-oFスタックとこれらのプロトコル/トランスポートの一部を示しています。図9は、NVMe-oFを使用して、NVMe-oFとネットワーク接続できるオブジェクトの直径と数を拡張する方法を示しています。

図9) NVMe over Fabrics (NVMe-oF)

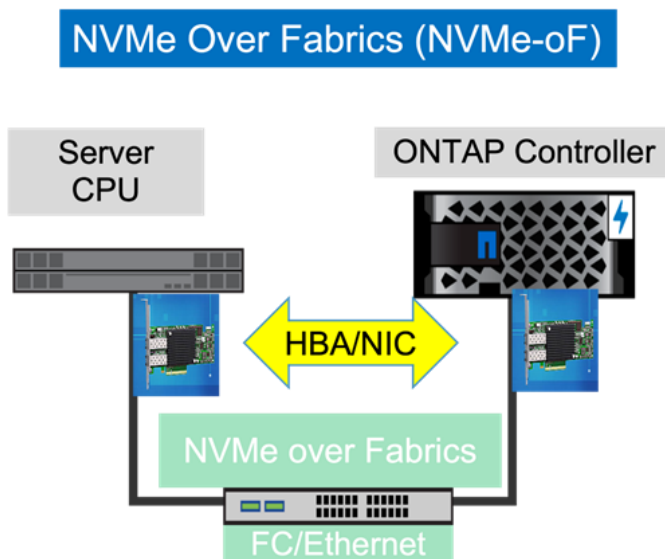


図10は、ネイティブのFCフレームと、FCフレーム内にカプセル化されたNVMeの違いを示しています。ご覧のように、NVMe/FCはSCSI-3コマンド記述子ブロック (CDB) をNVMeプロトコルコマンドに置き換えます。

このシンプルな置き換えにより、NVMeプロトコルをFCファブリック内で転送できるようになり、ネットワーク接続可能なオブジェクトの直径と数が増加し、非常に信頼性の高い転送が実現します。

図10) FCPとNVMe/FCフレームの比較

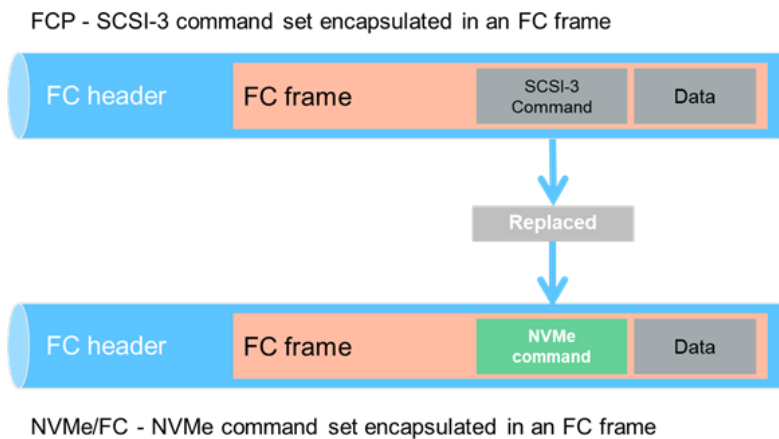
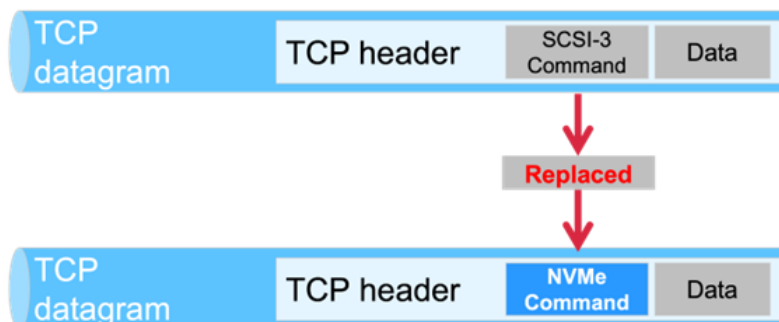


図11は、NVMe/TCPを示しています。NVMe/TCPは、TCPデータグラム内でカプセル化および転送されるNVMeコマンドセットです。

図11) iSCSIデータグラムとNVMe/TCPデータグラムの比較

- iSCSI: SCSI-3 command set encapsulated in a TCP Datagram



- NVMe/TCP: NVMe command set encapsulated in an TCP Datagram

図12は、NVMe/TCPの別のビューで、NVMeプロトコルとペイロードの両方がTCPデータグラムのデータ/ペイロード部分にカプセル化されていることを示しています。ご覧のとおり、NVMe-oFは、NVMeコマンドとデータペイロードをさまざまな既存のネットワークプロトコルやファブリックプロトコルにカプセル化し、NVMeを拡張して長距離間で転送します。

図12) TCPデータグラム内のNVMe

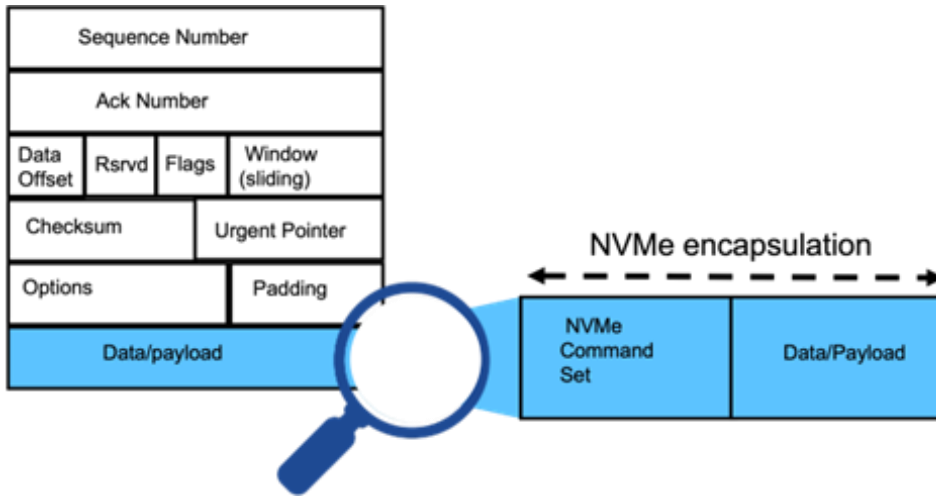
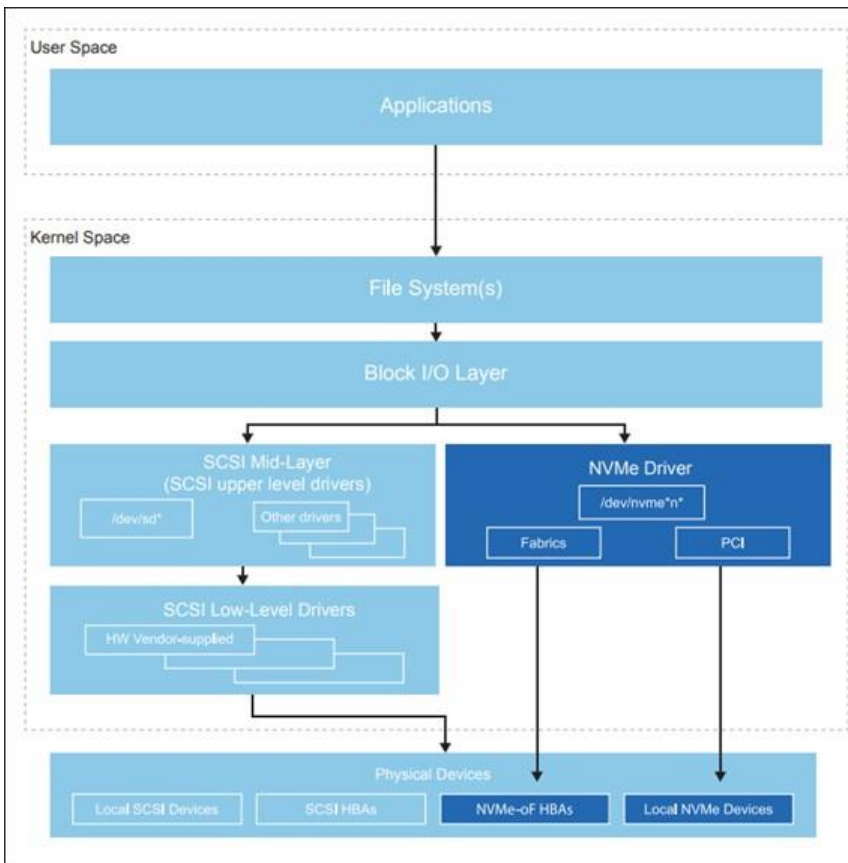


図13は、SCSIとNVMeのプロトコルスタックを比較したものです。SCSIスタックと比べてNVMeスタックがどれくらい短いかに注目してください。

図13) SCSIとNVMeのスタックアーキテクチャの比較



オープンソースフレームワークの活用

NetAppエンジニアリングは、いくつかのオープンソースアーキテクチャ仕様を利用してONTAP NVMe-oFターゲットを迅速に開発し、他のNVMeハードウェアおよびソフトウェアとのシームレスな相互運用性をサポートしました。NetAppはこれらの仕様を採用し遵守する一方で、NVM Express, Inc.やINCITSなどのいくつかの標準化団体でも活動しており、NetAppは設計、仕様、ガイダンスをオープンスタンダードコミュニティに提供しています。NetAppは、NVM ExpressおよびINCITSにコードとデザインを提供するだけでなく、次のような多数のリファレンスデザインやアーキテクチャ標準を採用し、貢献してきました。

- [データプレーン開発キット \(DPDK\)](#)
- [ストレージパフォーマンス開発キット \(SPDK\)](#)

NVMeと高可用性

NetApp NVM Express委員会の代表であるフレッド・ナイトは、機能的な高可用性エラー報告とフェイルオーバープロトコルを定義するTP 4004とTP 4028の技術提案を提出した筆頭著者であった。新しいプロトコルであるAsymmetric Namespace Access (ANA)は、2018年3月に批准されました。

全日空

ALUAと同様に、ANAはイニシエータ側とターゲット側の両方の実装で、ホスト側のマルチパス実装で各オペレーティングシステムスタックで使用されるストレージHAマルチパスソフトウェアと連携するために必要なすべてのパスとパスの状態情報を提供できるようにする必要があります。ANAが機能するには、ターゲット側とイニシエータ側の両方がANAを実装し、サポートしている必要があります。

NVMe/FCは、パスとターゲットの両方のフェイルオーバーに必要なマルチパスとパス管理機能を、ANAプロトコルを使用して提供します。ANAプロトコルは、NVMeサブシステムがパスとサブシステムのエラーをホストに伝える方法を定義します。これにより、ホストはパスを管理し、あるパスから別のパスへのフェイルオーバーを実行できます。ANAは、FCP (ファイバチャネルプロトコル) とiSCSIの両方のプロトコルに対してALUAが果たしているのと同じ役割をNVMe/FCで果たします。MPIOやDevice Mapper Multipathing (DM-Multipath) などのホストオペレーティングシステムパス管理機能を備えたANAは、NVMe/FCのパス管理機能とフェイルオーバー機能を提供します。図14 および 図15 は、NVM Express, Inc.に提出された技術提案書の表紙を示しています。図16 は、T11-2017-00145-v004 FC-NVMe仕様のINCITS T11カバーシートを示しています。INCITS T11委員会は、Intelligent Peripheral Interface (IPI)、High-Performance Parallel Interface (HIPPI)、FCの各分野の標準開発を担当しています。

図14) TP 4004 : ANAベース提案 (2018年3月批准)

NVM Express Technical Proposal for New Feature	
Technical Proposal ID	TP 4004
Change Date	02/26/2018
Builds on Specification	NVM Express 1.3 or later; does not apply to versions earlier than 1.3.

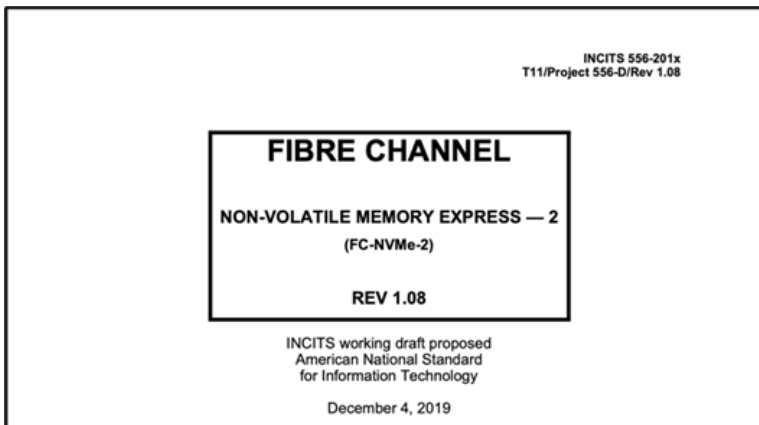
Technical Proposal Author(s)	
Name	Company
Fred Knight	NetApp
David Black	Dell EMC
Curtis Ballard	HPE
Christoph Hellwig	WDC

図15) TP 4028 : ANAのパスとトランスポート (2018年1月批准)

NVM Express Technical Proposal for New Feature	
Technical Proposal ID	TP 4028
Change Date	01/09/2018
Builds on Specification	NVM Express 1.3

Technical Proposal Author(s)	
Name	Company
Fred Knight	NetApp
David Black	Dell EMC
Sagi Grimberg	LightBits Labs

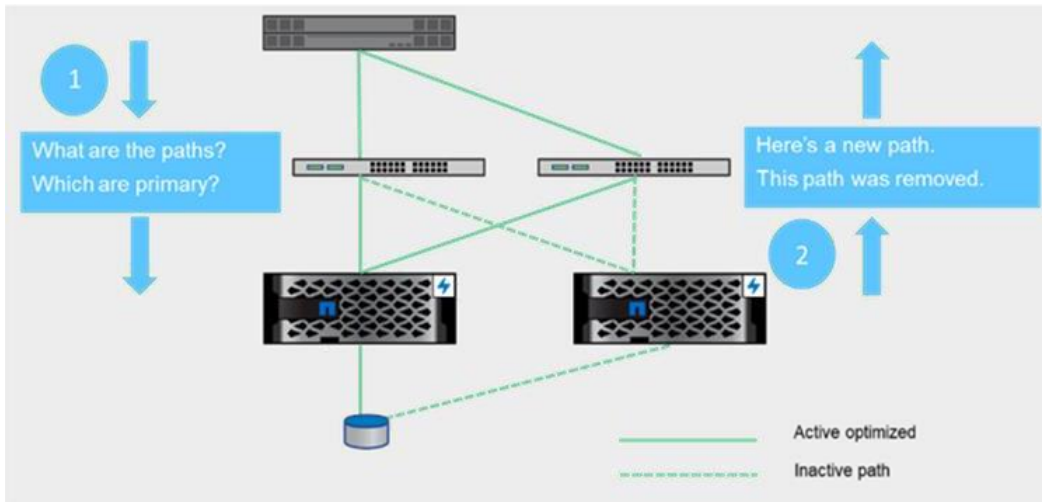
図16) INCITS FC-NVMe-2 Rev 1.08 T11-2019-00210-v004では、FC標準を使用したNVMeコマンドとデータ転送が定義されています。



ANAには次の2つのコンポーネントがあります。

- イニシエータ側のANAは、ターゲットに対してプライマリやセカンダリなどのパス属性を照会します。このデータは、イニシエータMPIOスタックによってパスの最適化に使用されます。
- ターゲット側のANAは、パスの状態の変化を通知します。

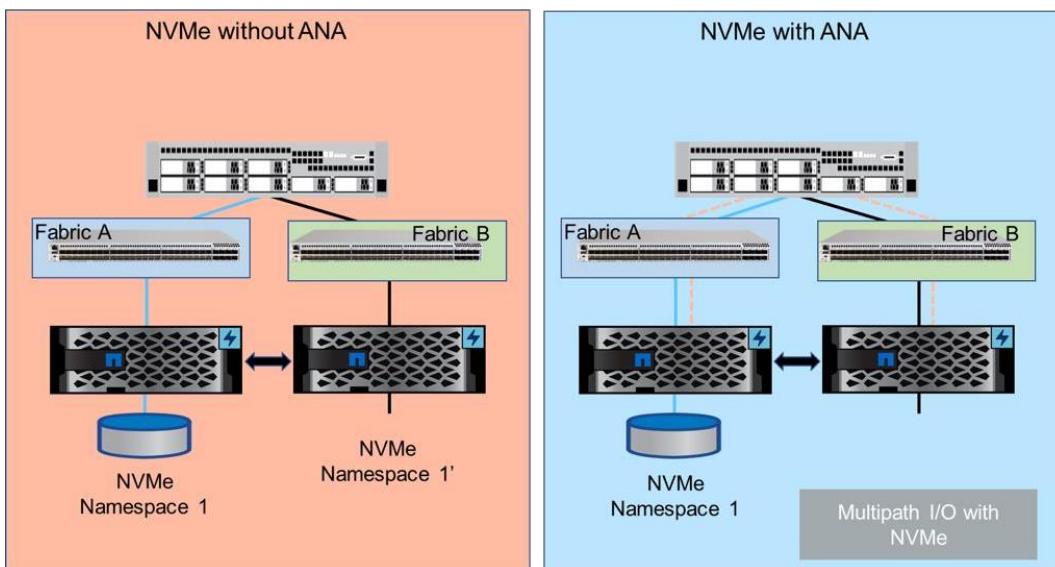
図17) NVMe/FCストレージフェイルオーバー : ONTAP 9.5でANAが導入



ONTAPはリモートI/Oをサポートしていますが、これは当社のALUAの実装に似ています。

図18は、ANAを使用した場合と使用しない場合のNVMe/FCの比較を示しています。

図18) NVMe/FCとANA非対応の比較



NVMe over Fibre Channel

NetAppのお客様は、世界最大級のデータセンターをいくつか構築、管理しています。今後数年間で、データストレージシステムをNVMeにアップグレードする予定です。では、どのNVMe-oF転送を選択するのでしょうか。

RDMA転送は重要ですが、データセンターファブリックでは当初NVMe over Fibre Channel (NVMe/FC) が主流になると思われます。FCをトランスポートとして使用すると、次のようなメリットがあります。

- 現在、ハイパフォーマンスブロックワークロードのほぼすべてがFCPで実行されています。
- これらの組織のほぼすべて（約70%）が、FCPでSANを使用しています。

- パフォーマンス重視のワークロードのほとんどは、現在、ファブリックに第5世代または第6世代（16Gbpsまたは32Gbps）のスイッチを使用しています。
- 現在、データセンターに設置されている25 / 50 / 100Gbpsイーサネットスイッチは、設置面積が小さいため、RDMA over IP、TCP、RoCE、またはその他の同様のトランスポートのバックボーンインフラとなります。
- FCPとNVMe/FCは、同じ物理コンポーネントを使用してSCSI-3とNVMeを同時に転送できます。
- 多くのNetAppユーザは、NVMe/FCの実行に必要なハードウェアをすでに所有しており、NetApp ONTAP 9.4以降にソフトウェアをアップグレードするだけでNVMe/FCの使用を開始できます。

FCPとNVMeは、共通のハードウェアおよびファブリックコンポーネントをすべて共有し、同じワイヤ（技術的には光ファイバ）、ポート、スイッチ、およびストレージコントローラ上に共存させることができます。そのため、組織は独自のペースでNVMeに移行できるため、簡単に移行できます。実際最近スイッチやダイレクタ（Gen 5/6）をアップグレードした場合は無停止でONTAP 9.4以降にアップグレードできます

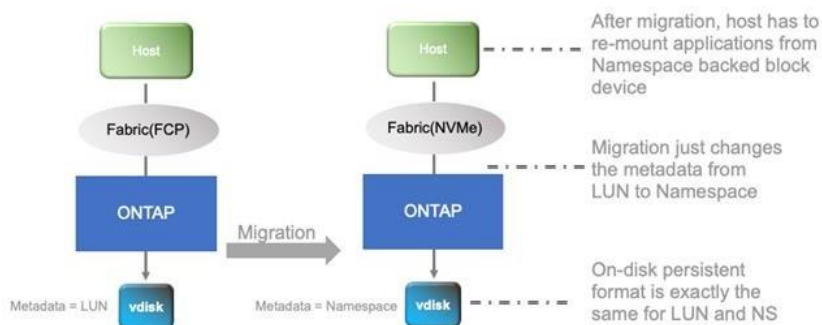
NVMe/FCはFCPによく似ています。FCPは、SCSI-3 CDBをFCフレーム内にカプセル化することで定義されています。どちらも似ているように見えるのは、NVMe/FCが旧式のSCSI-3 CDBを合理化された新しいNVMeコマンドセットに交換するためです。NVMe / FCは、このシンプルな置き換えにより、スループットとレイテンシが大幅に向上します。

NetAppは、NVMe / FCを最初にリリースしました。NVMe / FCは主要なSANプロトコルであり、次の最大プロトコルであるiSCSIの約3倍の採用実績があります。これは、企業がFCインフラストラクチャとスキルセットにすでに多額の投資を行っていることを意味します。さらに、パフォーマンスを最重要視する場合は、FC SANが選択されることがほとんどです。

NVMe/FCはコマンドセットをSCSIからNVMeにスワップするだけなので、簡単に移行できます。NVMe/FCは同じFC転送を使用するため、ホストからスイッチ経由でストレージアレイ上のNVMe/FCターゲットポートに至るまで、同じハードウェアを使用します。そのため、NVMe/FC環境では、HBA、スイッチ、ゾーン、ターゲット、およびケーブル配線。

ONTAPはFCP LIFとは別のNVMe/FC LIFを使用しますが、両方のLIFをホストイニシエータとストレージターゲットの両方の同じ物理HBAポートでホストできます。NVMe/FCとFCPでは同じ物理インフラを同時に共有できるため、同じ物理ポート、ケーブル、スイッチポート、およびターゲットポートでFCPとNVMe/FCの両方のフレームを同時にホストして送信できます。この2つのプロトコルは物理レイヤではなく論理レイヤで分離されているため、FCPからNVMe/FCへの移行を簡単かつシームレスに行うことができます。FCPからNVMe/FCにワークロードを自由なペースで移行できます。本番環境の運用を中断したり、複数のインフラを並行して実行したりする必要はありません。NVMe/FCとFCPは同じ物理インフラを使用するため、NetAppのお客様は、図19に示すように、新しいテクノロジーを無停止で実装して、パフォーマンスの向上、新しいワークフローの導入、移行される既存のワークフローのパフォーマンスの向上を実現できます。

図19) システムを停止することなく最新のテクノロジーを導入



NetApp NVMe/FCリリースの発表

ONTAP 9.4リリースノートには、NVMe/FCのサポートに関する次の情報が記載されています。ONTAP 9.4では、業界初のNVMe/FCが導入されています。

NVMe/FCは、SCSIではなくNVMeコマンドセットを使用して、FCPやiSCSIなどのホストにブロックを提供する新しいブロックアクセスプロトコルです。NVMeアーキテクチャは無駄のないコマンドセットと拡張性に優れたセッションを構築するため、レイテンシの大幅な低減と並列処理能力の向上が可能です。これにより、インメモリデータベースや分析など、低レイテンシで高スループットのアプリケーションに最適です。

NVMe/FCは、標準搭載のNetApp OnCommand® System Managerソフトウェア（Webブラウザでクラスタ管理ポートまたはいずれかのノード管理ポートのIPアドレスを指定）またはCLIを使用してプロビジョニングおよび設定できます。

この新しいプロトコルを使用して最大限のパフォーマンスを実現するには、ホストからSANファブリックを介してNetApp AFFコントローラへのエンドツーエンドのNVMe/FC接続が必要です。ONTAPでサポートされている最新の解決策スタックについては、[NetApp Interoperability Matrix Tool \(IMT\)](#) を参照してください。

注： ONTAPにNVMe/FCを実装するには、アプリケーションレベルの高可用性が必要です。コントローラの損失やパス障害が発生した場合は、（アプリケーション）HAパートナーへのパスフェイルオーバーをアプリケーションホストで管理する必要があります。これは、SCSIプロトコルのALUAに相当するANAと呼ばれるNVMeマルチパス仕様がまだ開発中であるためです。

NVMe / FCを実装する一方で、NetAppはNVMeフォーラムでANAプロトコルの設計を支援し、最近批准されました。ONTAPの今後のリリースでは、この機能強化をサポートする予定です。

注： 一部のNetAppのドキュメントやUIでは、NVMe over Fibre Channelが現在標準で商標登録されているNVMe/FCではなく、FC-NVMeと表記される場合があります。FC-NVMeとNVMe/FCは、どちらもNVMe over Fibre Channelの略で、同じ意味で使用されます。

NVMe over TCP (NVMe/TCP)

ONTAP 9.10.1で、ONTAP初のNVMe-oFイーサネットベースプロトコルNVMe/TCPが追加されました。NVMe/TCPは、NVMe/FCをONTAPでサポートされる2番目のNVMeブロックプロトコルとして追加します。NVMe/TCPは、Transmission Control Protocol (TCP) over Ethernetを使用します。転送用にTCPとイーサネットを組み合わせることで、TCPとイーサネットが存在する任意の場所にNVMe/TCPを導入できます。ほとんどの場合、所有するデータセンターに限定されるFCとは異なり、TCPとイーサネットはユビキタスです。さらに、NVMe/TCPにはネットワークハードウェアの実質的な制限はありません。ほぼすべてのイーサネットネットワーク機器でサポートされています。たとえば、1Mbpsのネットワークインターフェイスカード (NIC) と1MbpsのスイッチやルータでNVMe/TCPを実行できます。解決策は機能しますが、パフォーマンスには不満があるかもしれません。重要なのは、実際のハードウェア要件がなく、イーサネットとTCPにはほぼ普遍的な性質があるため、NVMe/TCPはほぼどこでも実行できるということです。これは、クラウドが急速に拡大する中で特に重要です。NVMe/TCPは、iSCSIをSANプロトコルまたはブロックプロトコルとして追加し、企業のデータセンター、サードパーティのホスティング、およびさまざまなクラウドエンドポイント間を接続できます。NVMe/TCPの最大の強みは、FCやiSCSIなどのSCSIベースのプロトコルからNVMe/FCやNVMe/TCPなどのNVMeベースのプロトコルに移行すると得られる効率化のすべて、TCPとイーサネットの柔軟性とモビリティです。本書の執筆時点（2023年1月）では、Red Hat Enterprise Linux、SUSE Enterprise Linux、Oracle Linux、VMware ESXiがNVMe/TCPをサポートしています。今後は他のオペレーティングシステムでもサポートが追加される予定です。そのため、[NetApp IMT](#)で現在NVMe/TCPをサポートしているOSの一覧を確認してください。

注： SCSIと同様、NVMeプロトコルは転送に依存しません。つまり、1つ以上のNVMe-oFプロトコルを使用して、NVMeネームスペースに個別にまたは同時にアクセスできます。

NVMeの導入

NetAppは、NVMe-oFの初期導入のほとんどは、本番環境のワークロードを移行する前に、組織でNVMe/FCやNVMe/TCPのテストと認定を行っている組織からのものと予測しています。ほとんどのエンタープライズストレージチームはリスクを回避しており、新しいプロトコルを本番環境に導入する前に、綿密な検証とテストを実施したいと考えています。さらに、NetAppでは、ONTAP NVMe/FCターゲットにANAが追加され、NVMe/FCサポートの一環として希望するホストOSがANAをサポートするまでは、ほとんどの早期導入者が待たなければならないと予測しています。ANAの不足を懸念しない可能性が高いのは、ストレージレイヤではなくアプリケーションレイヤで高可用性を管理するアプリケーションを使用している場合だけです。前述したように、これらのアプリケーションにはMongoDBやOracle ASMが含まれる場合があります。

NVMe/FCとNVMe/TCPのどちらの導入を選択すべきか

ほとんどの場合、新しいNVMe-oFの導入が新規に導入された環境であることはほとんどありません。ほとんどは既存のデータセンター運用への追加です。つまり、ある輸送手段を他の輸送手段よりも有利に働く可能性のある既存のインフラが存在する可能性があることを意味します。もちろん、すでにインストールされているさまざまなワークフローやインフラの要件に基づいて、複数のトランスポートを組み合わせることができないという要件もありません。

最初に理解しておくべきことは、NVMe-oFトランスポートは、iSCSIやFCPなどのSCSIベースのブロックプロトコルと非常によく共存していることです。さらに、NVMe/FCはFCPとまったく同じコンポーネントを同時に使用できます。これは、iSCSIとNVMe/TCPの2つのイーサネットプロトコルにも当てはまります。実際、同じホストからFCPやiSCSIを使用して12個のLUNに簡単にアクセスでき、NVMe/FCとNVMe/TCPのどちらかまたは両方を使用してさらに12個のネームスペースにも、まったく同じ物理HBA / NICポート、ケーブル、スイッチ、ONTAPコントローラを使用して簡単にアクセスできます。これにより、4つのブロックプロトコル間の共存と移行が非常に簡単になります。

特定のワークフローとともに導入するNVMe-oFトランスポートを検討する場合は、何を最適化したいかという疑問が生じます。自社所有のデータセンターで卓越したパフォーマンスを求めている場合は、NVMe/FCが第一の選択肢になるでしょう。柔軟性とモビリティを最適化する場合、特にFCファブリックが存在しない場合や、所有データセンターとクラウドエンドポイント間の接続が必要な場合は、まずNVMe/TCPを選択します。

ONTAP機能のサポートと共存

表2と表3に、現在NVMe-oFでサポート、共存、またはサポートされないONTAPのツールと機能を示します。これらの表に記載されている項目は本ドキュメントの執筆時点のものです。ただし、ONTAP NVMe-oFターゲットで引き続き機能とパフォーマンスが向上するため、サポート対象外の表に記載されている機能の多くがサポート対象に移行される可能性が高くなります。

表1) SCSIとNVMeの用語

FCP / iSCSI	NVMe-oFに関する注意事項	備考
FCP - Worldwide Port Name (WWPN) iSCSI-iSCSI Qualified Name (IQN)	NVMe Qualified Name (NQN)	一意の識別子
SCSIターゲット	NVMe サブシステム	ストレージオブジェクトとアクセスポイントを含むエンティティ
ポート	ポート	通信用アクセスポイント
I_Tネクスス	NVMeコントローラ (複数のキューペアを使用)	イニシエータとターゲット間のセッション
LUN	ネームスペース	ストレージオブジェクト

FCP / iSCSI	NVMe-oFに関する注意事項	備考
Asymmetric Logical Unit Access (ALUA ; 非対称論理ユニットアクセス)	全日空	ヒタイシヨウアクセストクセイ

表2) NVMeでサポートされる機能、またはNVMeと共存可能なONTAP機能

ONTAPリリース	NVMe-oFの機能
9.4	<ul style="list-style-type: none"> NVMe / FC シングルノードのみ (高可用性がないなど) 4Kブロックサイズ
9.5	<ul style="list-style-type: none"> Asymmetric Namespace Access (ANA) を使用した2ノードHA NVMe-oFライセンス
9.6	<ul style="list-style-type: none"> コピーと書き込み (CAW) 512Bブロックサイズ 読み取り専用ネームスペース 変更されたNSリストやベンダー固有のログページなどの追加のログページ
9.7	<ul style="list-style-type: none"> NVMe-oF同期SnapMirror
9.8	<ul style="list-style-type: none"> 同じSVM内でのLUNとネームスペースの共存
9.9	<ul style="list-style-type: none"> ASAのNVMe/FC ASAで大規模なネームスペース (最大128TBまで拡張可能) NVMe/FC VMware vSphere Virtual Volumes (vVol) NVMe-oF中止 4ノードアレイのNVMe/FC
9.10.1	<ul style="list-style-type: none"> NVMe/TCP ネームスペースのサイズ変更 NVMe-oFのキャンセル FASのNVMe-oF NVMe/TCPでのIPSec
9.11.1	<ul style="list-style-type: none"> NVMe/TCPパフォーマンスの機能拡張 LUN <->ネームスペース双方向インプレース変換 NVMe-oF Scaleの機能拡張 (最大12ノードのサポートなど) <p>メモ : 詳細については、NetApp HWUを参照してください。</p>
9.12.0	<ul style="list-style-type: none"> Cloud Volumes ONTAP (Amazon Web Services [AWS]およびAzure) とFSxでのNVMe/TCP
9.12.1	<ul style="list-style-type: none"> GCP Cloud Volumes ONTAPでのNVMe/TCP NVMe-oFインバンド認証 MCC IPのNVMe/FC NVMe-oF Sync SMとNDO

表3) 現在NVMeでサポートされていないONTAP機能

項目	メモ
NetAppオールSANアレイ (ASA) のNVMe-oF対称アクティブ/アクティブ構成	ASAではFCPとiSCSIのみがA/Aであり、NVMe-oFではない
ネームスペースのサービス品質 (QoS)	QoSはボリュームレベルとSVMレベルでのみサポートされ、NSレベルではサポートされない

項目	メモ
ネームスペースの移動	許可されていません（強制）
NVMe/TCP VVOL	サポート対象外（現在サポートされているのはNVMe/FC VVolのみ）
MCC IPのNVMe/TCP	サポート対象外（MCC IP 4パックでは現在NVMe/FCがサポートされています）
NVMe/TCPでのTLSのサポート	サポート対象外
NVMe-oF in NetApp SnapMirror ビジネス継続性	SnapMirrorビジネス継続性で現在サポートされているのはFCPとiSCSIのみです。
SVMの移行	ブロックのサポートなし
Foreign LUN Import（FLI）	FLIではFCPを使用してすべての移行を実行

注：インポートの完了後に、組み込みの双方向SCSI LUNからNVMeネームスペースへのインプレース変換ユーティリティを使用して、インポートしたLUNをネームスペースに変換できます。ユーティリティはLUNまたはネームスペースに関するメタデータのみを変更するため、非常に高速です。

相互運用性

NetAppは、SUSEおよびBroadcomとのパートナーシップにより、本番環境に対応したNVMe/FC製品を市場に初めて投入しました。NetApp ONTAP 9.4リリースでは、NVMe/FCを市場に投入するために提携した4社のベンダーから次のコンポーネントが認定されています。

- SUSE Enterprise Linuxバージョン12 SP3
- Broadcom/Emulex LPe3200X HBA
- Fabric OS 8.1.0A以降を実行しているBroadcom Brocade第5世代または第6世代（16Gbpsまたは32Gbpsスイッチ）
- NetApp AFF A300、AFF A700、AFF A700s、またはAFF A800システムとONTAP 9.4（各ノードに少なくとも1つの32Gbpsターゲットアダプタを搭載）

それ以降、複数の新しいONTAPリリースとホストOSベンダーの数によって、OSの新しいリリースでNVMe/FCとNVMe/TCPのサポートが追加されています。サポートされる構成の最新のリストについては、[IMT](#)を参照してください。OSの設定手順については、[ONTAP SANホストの設定ページ](#)（[NVMe/FCホストの設定]セクション）を参照してください。[NVMe/FC Host Configurations]ページの例を図 20に示します。

図20) [NVMe/FC Host Configurations]ページ

メモ：特定のバージョンのソフトウェアファームウェアとドライバについては、NetApp [IMT](#)を参照してください。コンポーネントがテストされて認定されると、IMTにさらにアイテムが追加されます。相互運用性の最新情報については、必ずNetApp [IMT](#)を参照してください。

NetApp Interoperability Matrix Tool

計画した構成または現在の構成全体が、NetApp [IMT](#)のいずれかの構成と一致していることを確認します。NVMe-oFは次の3つの軸で急速に発展しているため、これは重要です。

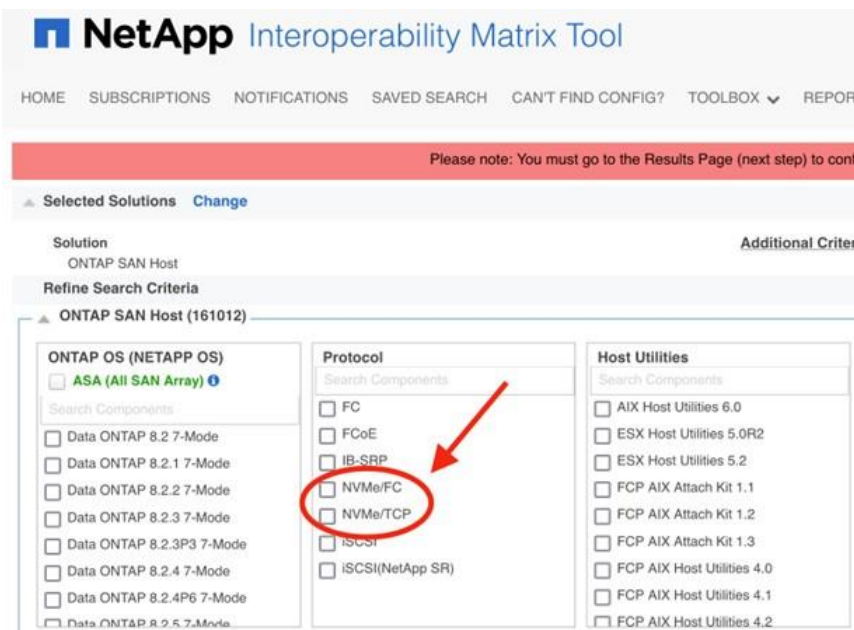
- NetApp NVMe-oFターゲット
- 関連付けられているスイッチ、HBAなど
- NVMe-oFを現在サポートしているホストOSは次のとおりです。
 - Linux (RHEL、Oracle Linux、およびSLES)
 - VMware ESXi
 - Windows

NVMe/FCのみ

- HBA (32G ~ Gen6以降) 対応ドライバ、ファームウェア：
 - Broadcom (Emulex)
 - Marvel (Qlogic)
- FCスイッチ (16G Gen 5以降) スイッチOS：
 - Broadcom (Brocade)
 - Cisco

[IMT](#)に記載されている構成と異なると、一貫性のない予期せぬパフォーマンスが発生する可能性があります。[IMT](#)に、NVMe/FCという新しいプロトコルフィルタが追加されました (図21を参照)。このフィルタを使用し、NVMe/FCで認定された構成を確認できます。

図21) 新しいIMT NVMe/FCプロトコルフィルタとNVMe/TCPプロトコルフィルタ



LUNからネームスペースまたはネームスペースからLUNへの変換

ONTAP 9.11.1では、インプレース双方向LUNがネームスペース変換ユーティリティに追加されています。LUNとネームスペースの変換は非常に高速で（メタデータのみが変更される）、非常に簡単です。これにより、NVMeの採用が容易になり、SCSIプロトコルとNVMeプロトコル間の移行がよりシームレスになります。

特長

- 同じONTAPボリューム内でのインプレースLUN <->ネームスペース変換
- データコピーは不要。実際のユーザデータはそのまま保持され、メタデータの更新のみが必要
- 既存のSnapshotコピーへのアクセス（移行前に作成）が失われなため、データの管理と保護の継続性が可能
- 変換時にブロックサイズと識別子（シリアル番号やUUIDなど）を保持
- 変換後のパフォーマンス低下なし（同様のプラットフォーム構成で新規作成した場合と比較）
- エンドツーエンドのLUN <->ネームスペース変換プロセスを完了するにはホストの修正が必要

キノウノセイケン

- LUNまたはネームスペースを1つのコマンドで一括変換することはできません。
- マッピングされたLUNをネームスペースに変換できない（またはその逆）
- FLI関係のLUNをネームスペースに変換できない
- SnapMirrorビジネス継続性関係のLUNをネームスペースに変換することはできません
- MetroCluster構成のLUNをネームスペースに変換できない
- VVOLバインドが設定されているLUNやPEとして機能しているLUNはネームスペースに変換できません。
- プレフィックスまたはサフィックスストリームが0以外のLUNはネームスペースに変換できません（ONTAPではネームスペースのプレフィックス/サフィックスストリームがサポートされないため）。
- ブロックサイズが4KのネームスペースはLUNに変換できません（ONTAPでは4KのLUNがサポートされないため）。
- `os_types` 変換できるのは、ネームスペースでサポートされている有効なLUNのみです

表4) LUN <->ネームスペース変換ユーティリティ機能のサポート

LUNのOS_type	プレフィックスサイズ	サフィックスサイズ	LUN <-> NSは許可されていますか？
vmware	0	0	はい
hyper_v	0	0	はい
windows_2008	0	0	○（「Windows」へ）
windows_gpt	17（ONTAP 9.8より前） 0（ONTAP 9.8以降）	0	○（prefixが0の場合のみ、 'windows'）
windows	31.5（ONTAP 9.8より前） 0（ONTAP 9.8以降）	0	○（プレフィックスが0の場合のみ）
linux	0	0	はい
xen	0	0	はい
solaris	1 * cylinder_size（ONTAP 9.6以前） 0（ONTAP 9.6以降）	2 * cylinder_size（ONTAP 9.6以前） 0（ONTAP 9.6以降）	いいえ（サポート対象外のNVMe OS_TYPE）
solaris_efi	17（ONTAP 9.8より前） 0（ONTAP 9.8以降）	0	いいえ（サポート対象外のNVMe OS_TYPE）

LUNのOS_type	プレフィックスサイズ	サフィックスサイズ	LUN <-> NSは許可されていますか？
hpux	0	0	いいえ (サポート対象外のNVMe OS_TYPE)
aix	0	0	いいえ (サポート対象外のNVMe OS_TYPE)
netware	0	0	いいえ (サポート対象外のNVMe OS_TYPE)
openvms	0	0	いいえ (サポート対象外のNVMe OS_TYPE)
Image (デフォルト)	0	0	いいえ (サポート対象外のNVMe OS_TYPE)

LUNとネームスペース間の変換の詳細については、「[付録D : LUNとネームスペース間の変換](#)」を参照してください。

NVMe-oFのベストプラクティス

NVMe / FCのベストプラクティス

本番環境でNVMe/FCワークロードのテストと認定を行うか、使用するにかかわらず、すべてのチームがFCP SANの一般的なベストプラクティスに従う必要があります。NVMe/FCでは転送にFCを使用するため、これらのベストプラクティスが適用されます。NetApp SANのベストプラクティスについては、[TR-4080 : 『Best Practices for Scalable SAN』](#)を参照してください。

ファブリックとスイッチの構成と運用のベストプラクティス

NVMe/FCでは、BrocadeやCiscoのFCスイッチおよびファブリックに関する一般的なベストプラクティスと異なる特別な設定やベストプラクティスは必要ありません。シングルイニシエータゾーンニングを使用することを推奨します。もう1つのベストプラクティスは、(スイッチポートベースのゾーンメンバーシップやハードゾーンニングではなく) WWPNを使用してゾーンメンバーシップを割り当てることです。

ベストプラクティス

BrocadeのSAN Healthツールを実行してファブリック構成の詳細を収集SAN Healthでは、ファブリックが正常でエラーがないことを確認します。SAN Healthでは、ファブリックを自動的に文書化して可視化することもできます。Brocadeスイッチが存在するかどうかに関係なく、すべてのファブリックでSAN Healthを使用できます。

Healthツールのコピーやビデオとガイドについては、担当のNetAppまたはパートナーアカウントチームにお問い合わせください。

NVMe-oFのベストプラクティス：パス

インターフェイスの単一点障害を回避するために、NetAppでは、SVMごと、ノードごと、ファブリックごとに2つのパスをプロビジョニングすることを強く推奨します。ONTAPでは、ストレージ管理者がノードまたはSVMごとに作成できるLIFは2つまでです。これは、NetApp SANのターゲットエンジニアが、LIFの作成を制限してイニシエータに提供するパスの数を減らすことを意図的に選択したことです。この制限を設定することで、NetAppは十分な冗長性を確保して単一点障害を排除すると同時に、イニシエータに提供されるパスの数を制限するためにSCSIベースの構成で必要なSelective LUN Map (SLM ; 選択的LUNマップ) などの機能の必要性を軽減または排除できます。特定のノードに追加のLIFが必要な場合は、SVMを追加で作成し、ノードごとに2つのLIFを作成してネームスペースとのI/Oを分散させることができます。

NVMeでは、通信、アラート、およびパスとパスの状態の変化を管理するために、ANAプロトコルが追加されました。ANAは2つのコンポーネントで構成されています

- ターゲット (ONTAPノード) に現在のパス状態情報を照会するホスト側の実装。
- ストレージノードの実装。パスの状態が変わったときにアラートを送信し、使用可能なすべてのパスの列挙に対するイニシエータ側のクエリに応答します。

ホスト側のANA実装では、受け取ったすべてのパス情報をホストのマルチパススタックに渡します。マルチパススタック (dm-multipathなど) をホストし、パスの設定と使用方法を管理します。

注： ONTAP 9.9.1ではNVMe-oFリモートI/Oのサポートが追加され、ユニファイドシステム (AFF / FAS) ではSCSIベースのプロトコルを使用するため、NVMe-oFパスが変更されてパスがAOまたはANOとしてアドバタイズされます。

注： ASAで対称なSCSIプロトコルiSCSIおよびFCとは異なり、NVMe-oFプロトコルの非対称AO/ANO特性はASAではASA以外のプロトコルと同じです。これは、ローカルパス処理とリモートパス処理でNVMe-oFプロトコルが異なるためです。

NetAppでは、Linux NVMeマルチパスをONTAPネームスペース専用を使用することを推奨しています。

マルチパスの推奨事項

NetAppでは、ONTAP LUNにのみLinux dm-multipathを使用し、ONTAPネームスペースにはNVMeマルチパスを使用することを推奨しています。

以前のバージョンのSUSE 15 (SLES15) では、NVMeマルチパスでラウンドロビンによるロードバランシングを使用できなかったため、パフォーマンスが低下していました。この制限は、新しいバージョンのSLES15カーネル/NVMe-CLIで解決されました。詳細については、『ONTAP [IMT](#)』を参照してください。

NVMe-oFのセットアップと設定

セットアップと設定のクイックリスト

NVMe/FC / NVMe/TCPをセットアップする前に、次の要件が満たされていることを確認してください。

1. 使用する構成が、IMTに記載されている認定構成と完全に一致していることを確認します。そうしないと、ストレージの実装が最適ではなく、適切に構成されていない可能性があります。
2. NetAppおよびスイッチベンダーのSANのベストプラクティスに準拠するように、物理インフラを導入、ケーブル接続、設定します。「追加情報の検索場所」の項を参照してください。
3. (NVMe/FCのみ) すべてのファブリックスイッチでN_Port ID Virtualization (NPIV) を有効にします。
4. (NVMe/FCのみ) シングルイニシエータゾーニングを使用し、WWPNを使用してゾーンメンバーシップを指定します。ゾーンメンバーシップまたはハードゾーニングを示すためにスイッチポートの接続を使用しないでください。
5. ONTAP System ManagerまたはONTAP CLIを使用して、NVMe-oFオブジェクト (SVM、ボリューム、ネームスペース、サブシステム、およびLIF) を作成します。詳細については、「[付録A： ONTAP System Managerを使用した ONTAP NVMe/FCおよびNVMe/TCPオブジェクトの作成](#)」および「[付録B： ONTAP NVMe/FCおよび NVMe/TCPのCLIコマンド-初期セットアップおよび検出](#)」を参照してください。
6. NetApp Active IQ Unified Managerを使用して、新しく作成したNVMeオブジェクトの健全性とパフォーマンスを監視し、レポートのしきい値とアラートを作成します。

セットアップと構成の詳細な手順リファレンス

NVMeのセットアップと設定のタスクの実行の詳細については、次の参考資料を参照してください。

- ONTAP System Manager (組み込みのUI) を使用してONTAPでNVMe/FCオブジェクトをセットアップおよび設定する方法については、「[付録A： ONTAP System Managerを使用したONTAP NVMe/FCおよび NVMe/TCPオブジェクトの作成](#)」を参照してください。
- ONTAPでCLIを使用してNVMe/FCをセットアップおよび設定する方法については、「[付録B： ONTAP NVMe/FCおよび NVMe/TCPのCLIコマンド-初期セットアップと検出](#)」を参照してください。
- サポートされるホストオペレーティングシステムの構成情報については、『[ONTAP SANホスト 構成ガイド](#)』を参照してください。
- NVMeオブジェクトを表示してI/Oを実行する方法については、「[付録B： ONTAP NVMe/FCおよび NVMe/TCPのCLIコマンド-初期セットアップと検出](#)」を参照してください。

パフォーマンス

NVM Express, Inc.は当初、フラッシュが原因のボトルネックに対処するためにNVMeの仕様とアーキテクチャを開発しました。フラッシュがハードドライブを交換したとき、ハードドライブが作成した主なストレージのボトルネックは取り除かれましたが、原因もう1つのボトルネックはコマンドセットとコントロールプレーンです。NVMeの仕様とアーキテクチャは、SCSIコマンドセットを次のよりスリムな新しいコマンドセットに置き換えます。

- コマンドを合理化 (SCSIコマンドは40年近く前に最初に作成された元の標準と下位互換性があります)
- ハードウェア割り込みではなくポーリングモードを使用する
- コンテキストスイッチの削減
- ロックレス
- キューを64k (65, 535) に増やし、それぞれのキュー深度を64kにします。

図22は、前述の各ポイントがスループットとレイテンシにどのように影響するかを示しています。リーンコマンドセット (I/Oパス)の方が効率的であり、同じ作業セットとインフラストラクチャで同じ時間内により多くのI/Oを可能にします。コンテキストスイッチを減らし、ハードウェアの割り込みではなくポーリングモードを使用することで、NVMeはプロセッサの強制アイドル時間を大幅に短縮します。ソフトウェアロックを解除すると、アイドル状態でのプロセッサ時間も短縮されます。最大のパフォーマンス向上は、膨大な数のキューとそれに関連付けられたキュー深度によるものです。これにより、各キューが個別のプロセッサコアを使用してI/Oを同時に処理できるようになります。

これらの機能強化により、パフォーマンスが大幅に向上します。このような増加は、スループットやIOPSの増加とレイテンシの低下が最も顕著に見られます。最初のテストでは、NetAppワークロードのエンジニアリングチームとパフォーマンスチームは、パフォーマンスが50%以上向上することを確認しました。IOPSを測定しながらレイテンシを80~100低減しました。

図22) NVMe/FCの超ハイパフォーマンス設計

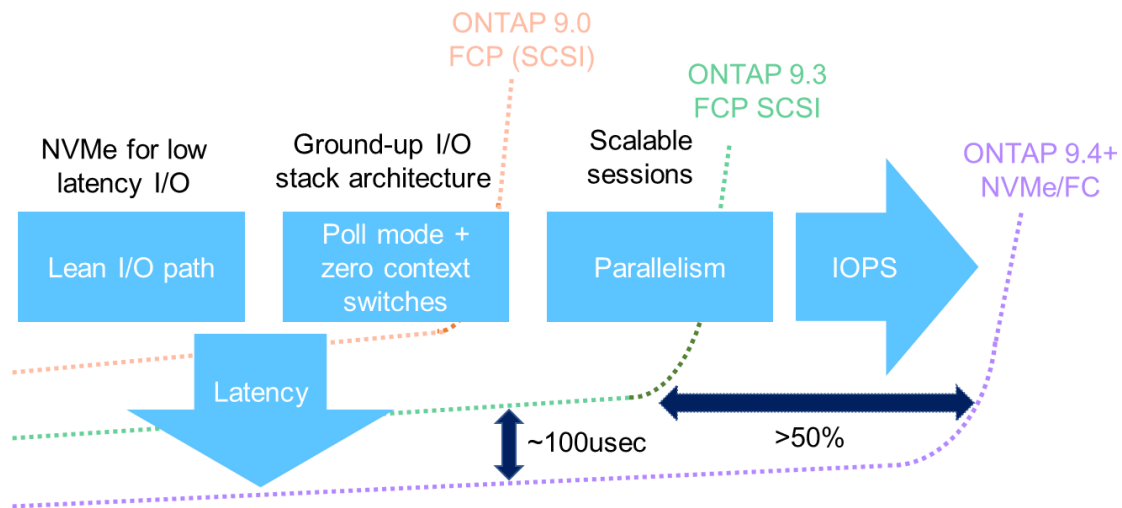


表5に、NetAppによる一部の内部テストの結果を示します。単一のLUNと単一のネームスペースアクセスを比較すると、大幅に増加しています。この比較では、並列化の増加がパフォーマンスにどの程度影響するかを強調しています。NetAppでは、パフォーマンスが必要な場合に単一のLUNを使用することは推奨されません。ストリームのI/O処理が単一のCPUコアに制限されるためです。

表5) AFF A700の4KランダムリードNVMe/FCとFCPの比較

	NVMe / FC	デルタとFCPの比較 (パーセンテージ)
シングルポート、IOPS	619,000	207%
グローバルネームスペース/LUN、IOPS	54万	八八〇パーセント
ピーク IOPS	865K	+51%

図23 と 図24 は、NetAppのパフォーマンス特性評価ラボで実施した8kと4kのランダムリードのパフォーマンスを示しています。このテストでは、NetApp、Brocade、Broadcomが、いくつかの主要な重要なエンタープライズアプリケーションについて検証したリファレンスアーキテクチャを定義、構築、テスト、文書化しました。作成されたドキュメントへのリンクは、このドキュメントの「[NetApp Verified Architectures](#)」セクションにあります。

図23) AFF A700ハイアベイラビリティ (HA) ペア、ONTAP 9.4、8KランダムリードFCPとNVMe/FCの比較

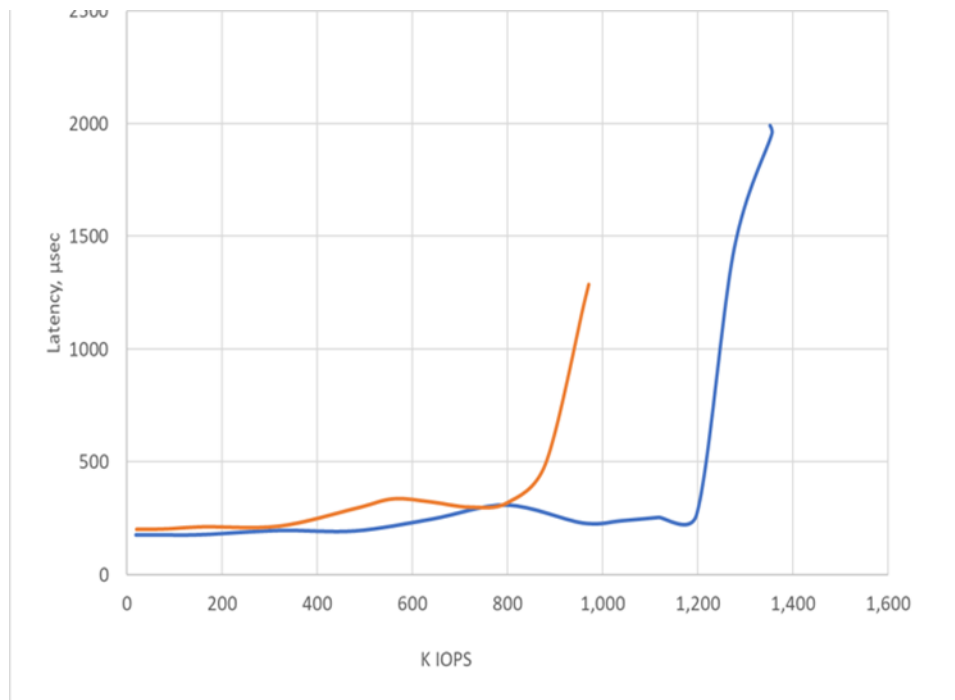


図24) AFF A700 HAペア、ONTAP 9.4、4KランダムリードのFCPとNVMe/FCの比較

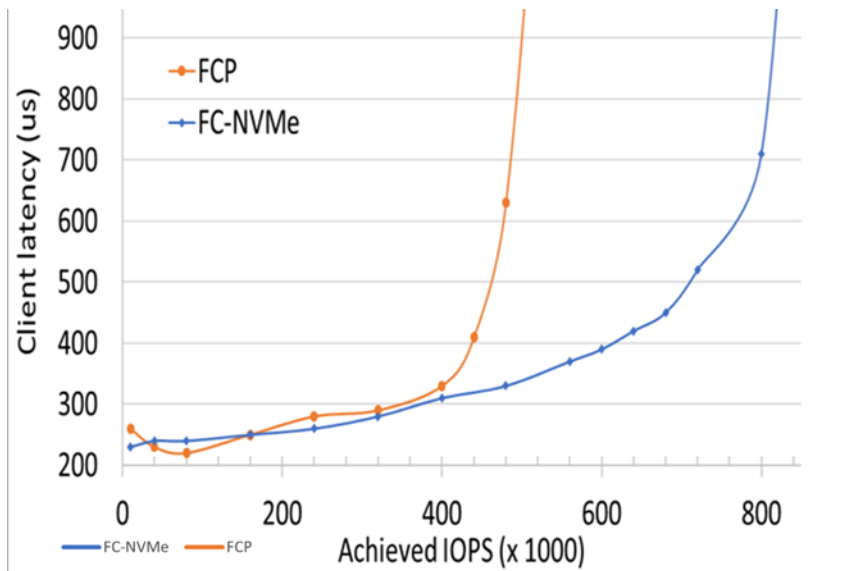
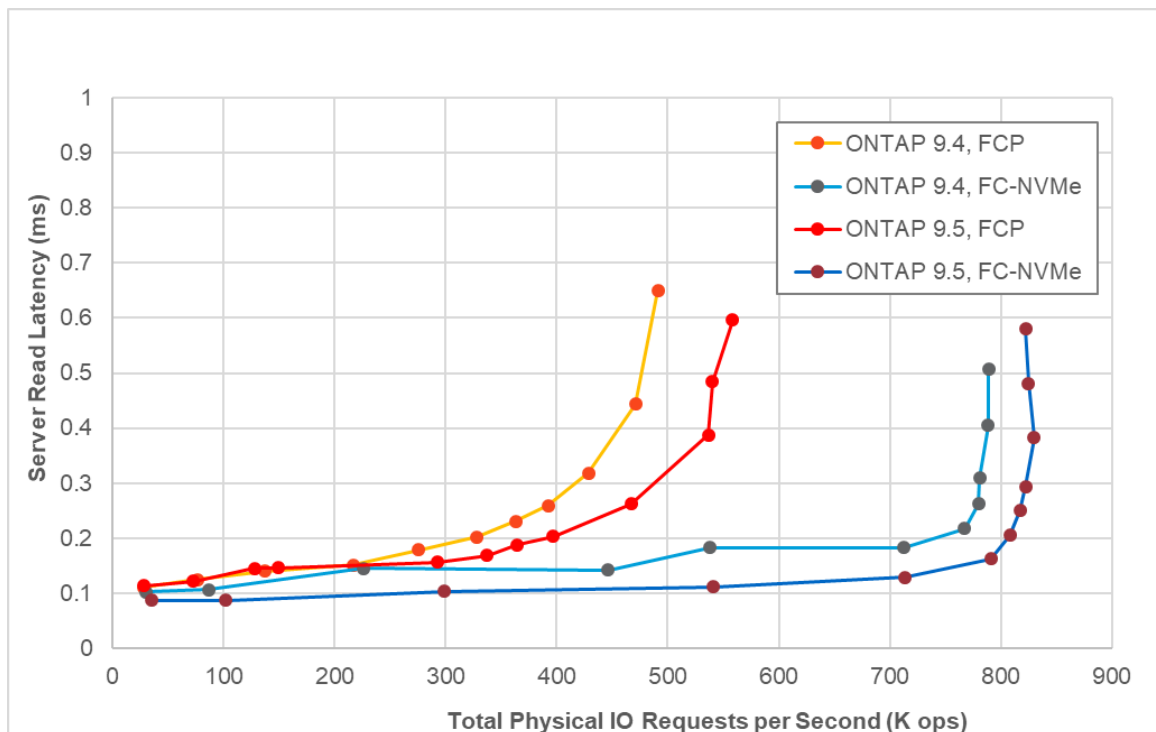


図25 は、FCPとNVMeのレイテンシあたりのIOPS数の比較です。これは、コントローラが一定期間内に完了できるI/O数が大幅に増加したことを示しています。興味深いことに、NVMe/FCは達成されるIOPSを増加させ、それらのIOPSを完了するのに必要な時間を短縮します。

図25) FCPからNVMe/FCに移行し、ONTAPをアップグレードするたびにパフォーマンスが向上



サポートされるホストオペレーティングシステムの構成情報については、『[ONTAP SANホスト構成ガイド](#)』を参照してください。

NVMe / TCPのベストプラクティス

NVMe/TCPはイーサネットとTCP/IPを使用するため、ネットワークのパフォーマンスを向上させるものは、iSCSIの場合と同様にNVMe/TCPのパフォーマンスにプラスの影響を与える可能性があります。一般に、ネットワークベンダーのスイッチとルータに関するベストプラクティスに従うことをお勧めします。

NVMe-oFの機能拡張

ONTAP 9.6

512バイトブロック

NVMeのブロックサイズは、すべてのOSで4096または4kです。NetAppでは、VMwareの仮想マシンファイルシステム (VMFS) をより簡単にサポートできるように、ESXi用のデフォルトのブロックサイズがONTAP 9.6でのみ導入されました。ブロックサイズを512バイトブロックに縮小することで、NetAppは複数のESX 512bを4KのONTAPブロックに集約するのではなく、共通のブロックサイズを提供することで、ESXiとの相互運用が容易になります。512バイトのブロックのサポートにより、ESXiのコピーと書き込み/ Atomic Test and Set (ATS) をサポートするONTAPの機能も強化されます。

ONTAP 9.9.1

NVMe-oFリモートI/Oのサポート

NVMe-oFでは、リモートI/Oのサポートが追加されています。これにより、NVMe-oFパスの設定がアクティブ/非アクティブモデルから、他のすべてのONTAPブロックプロトコルで使用されるAO /アクティブ非最適化 (ANO) モデルに変更されます。

図26は、リモートI/OをサポートしないNVMe-oFを示しています。図27は、リモートI/OをサポートするNVMe-oFを示しています。

図26) リモートI/OをサポートしないNVMe-oF

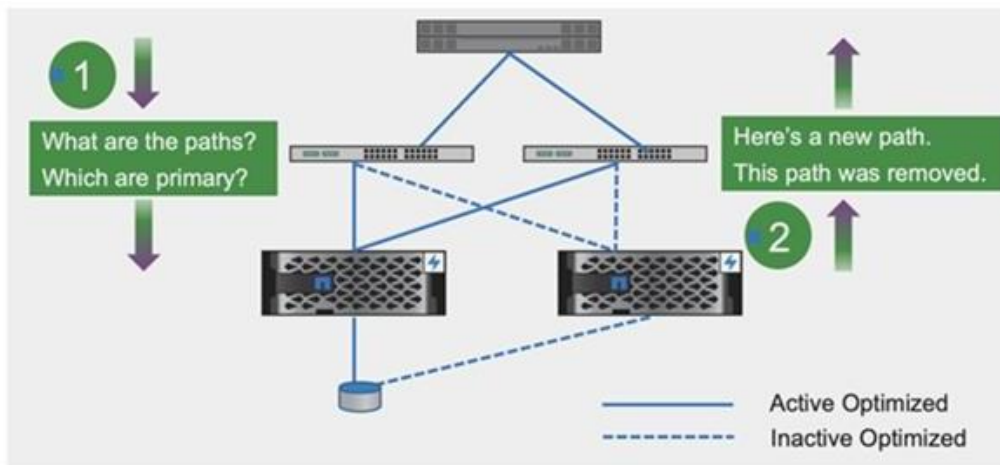


図27) リモートI/Oを使用するNVMe-oF

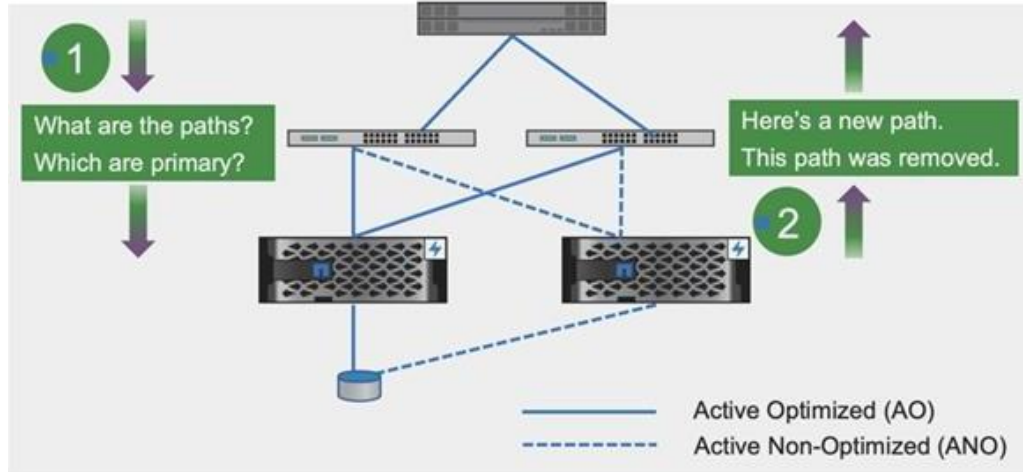


図26 と 図27を比べると、大した違いはないようですが、微妙な違いがあります。リモートI/Oでは、NVMe-oFのすべてのパスのサポートがアクティブになります。つまり、これらのパスのいずれかに送信されたI/Oが確認応答されて応答または応答されます。これまでは、リモートI/Oがなければ非アクティブなパスは使用できず、使用できませんでした。

All SAN Array adds NVMe/FC support
ONTAP 9.9.1 adds NVMe/FC as an additional block protocol. Unlike either FC or iSCSI, NVMe/FC on All SAN Array (ASA) will continue to be asymmetric (AO/ANO). This assignment is due to differences in how NVMe-oF works with remote versus local paths.

ONTAP 9.10.1以降

NVMe/TCPが導入された

ONTAP初のイーサネットベースおよび2番目のNVMe-oFプロトコルトランスポート。

ネームスペースのサイズ変更

ネームスペースのサイズを変更できるようになりました。

大規模なネームスペース

ASAでネームスペースの最大サイズが128TBに拡張されました。大規模なネームスペースのサポートがFASコントローラとAFFコントローラに追加され、ONTAP 9.12.1でプレビューが公開されます。また、ONTAPの今後のリリースで一般提供される機能になります。

検出コントローラでの非同期イベント要求 (AER、opcode OC) のサポート

新しいマップまたはマップを持つサブシステムが追加されると、非同期イベント要求 (AER) がトリガーされます。AERは、マップが追加されるサブシステムの一部であるすべてのアクティブなセッションに対して発行されます。UNMAPのサポートは、ONTAPの今後のリリースで追加される予定です。AERは、他の管理キュー要求と同様に処理されます。AERの完了には長い時間がかかることがあります。

ONTAP 9.11.1

NVMe/TCPのパフォーマンスの強化

ONTAP 9.11.1では、次のNVMe/TCPパフォーマンスの強化が導入されました。

- 書き込み用のカプセル内データなどのファーストバースト
- 読み取りの読み取り応答の折りたたみ

これらの機能強化により、NVMe/TCPのパフォーマンスが大幅に向上し、複数のI/OワークロードでiSCSIを上回るレベルにまで向上しました。

LUNからネームスペースへの双方向変換ユーティリティ

使いやすく、非常に高速なLUNからネームスペース、ネームスペースからLUNへの変換ユーティリティ。ユーティリティはONTAPに組み込まれています。変換は、LUN/ネームスペースに関するメタデータのみを変更するインプレース変換であるため、非常に高速です。

LUNとネームスペース間の変換の詳細については、「[付録D：LUNとネームスペース間の変換](#)」を参照してください。

ONTAP 9.12.0

AWS FSxとCloud Volumes ONTAPでNVMe/TCPのサポートが導入されました

- FSx HAにNVMe/TCPのサポートを追加
- Cloud Volumes ONTAP AWS HAにNVMe/TCPのサポートを追加
- Cloud Volumes ONTAP Azure HAへのNVMe/TCPのサポートの追加

ONTAP 9.12.1

その他のNetAppクラウドサービスでNVMe/TCPクラウドをサポート

- FSxシングルノードにNVMe/TCPのサポートを追加
- Cloud Volumes ONTAP AWSシングルノードでNVMe/TCPのサポートを追加
- Cloud Volumes ONTAP AzureシングルノードでNVMe/TCPのサポートを追加
- Cloud Volumes ONTAP GCPシングルノードおよびHA

NVMe/TCPに対する双方向のインバンド認証

ONTAP 9.12.1以降では、DH-HMAC-CHAP認証プロトコルを使用したNVMe-TCP経由でNVMeホストとコントローラの間でセキュアな双方向認証がサポートされます。

各ホストまたはコントローラには、DH-HMAC-CHAPキーを関連付ける必要があります。DH-HMAC-CHAPキーは、NVMeホストまたはコントローラのNQNと、NVMeホストまたはコントローラがピアを認証するために管理者が設定した認証シークレットを組み合わせたものです。ピアに関連付けられているキーを認識する必要があります。SHA-256がデフォルトのハッシュ関数です。

リポート前の確認事項

NVMe/TCPプロトコルを実行していて、DH-HMAC-CHAPを使用したセキュアな認証を確立している場合は、リポート前にDH-HMAC-CHAPを使用するホストをNVMeサブシステムから削除する必要があります。ホストを削除しないと、リポートは失敗します。

MCC IPに対するNVMe/FCのサポート

ONTAP 9.12.1では、MCC IPにNVMe/FCがサポートされるようになりました。詳細については、「[付録I: MetroCluster IPでのNVMe/FCの設定とセットアップ](#)」を参照してください。

付録A : ONTAP System Managerを使用したONTAP NVMe/FCオブジェクトとNVMe/TCPオブジェクトの作成

System Managerを使用して、次の手順を実行してONTAPオブジェクトを作成します。

1. NVMeをサポートするSVMを作成

注： この手順では、このワークフローの残りの部分で作成したNVMeストレージオブジェクトをすべて含むSVMが作成されます。

- ONTAPシステムマネージャで、[ストレージ]>[SVM]に移動します。[作成]をクリックします。
- NVMeを選択すると、サブシステム、NVMe Qualified Name (NQN)、およびネームスペース情報を作成および定義して、[SVMセットアップ]ダイアログボックスでNVMeを設定するように求められます。[送信して続行]をクリックします。

注： ホストのNQNを表示するには、次のコマンドをLinuxで実行します。

```
# cat /etc/nvme/hostnqn
```

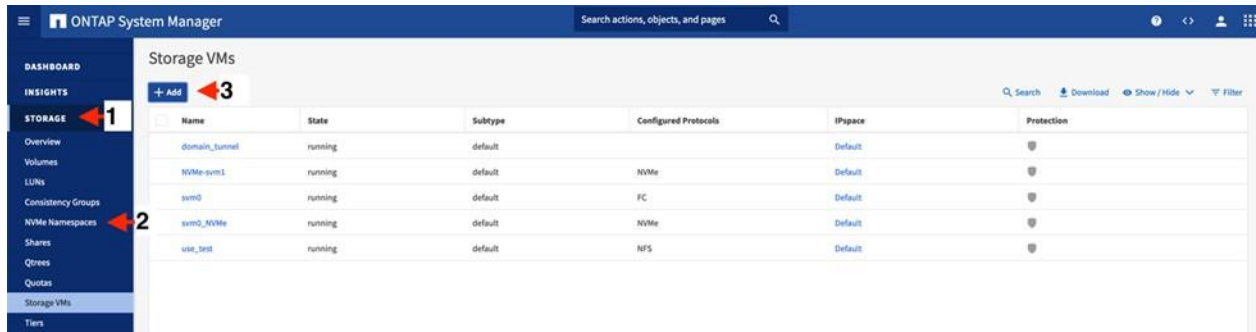
注： SVMのNQNを表示するには、次のコマンドを実行します。

```
vserver nvme show -vserver <vserver_name>
tme-a800::> nvme show -vserver NVMe-svm1
(vserver nvme show)

Vserver Name: NVMe-svm1
Administrative Status: up
Discovery Subsystem NQN: nqn.1992-08.com.netapp:sn.70c0b1366f3611edacf100a098e22473:discovery
```

注： ONTAP System Managerで[ストレージ]>[NVMe]>[NVMeネームスペース]に移動して、サブシステムのNQNを確認することもできます。次に、NQNを表示するネームスペースのリンクをクリックします。

図28) OnCommandシステムマネージャ-SVMの作成



- 次のいずれかを実行します。
 - [SVM administrator]ダイアログボックスで、SVM管理者の詳細を設定します。
 - 特定のSVM管理アカウントの追加を省略するには、[Skip]をクリックします。

図29) OnCommand System Manager-SVMの作成 : NVMe転送の設定-NVMe/FCおよびNVMe/TCP

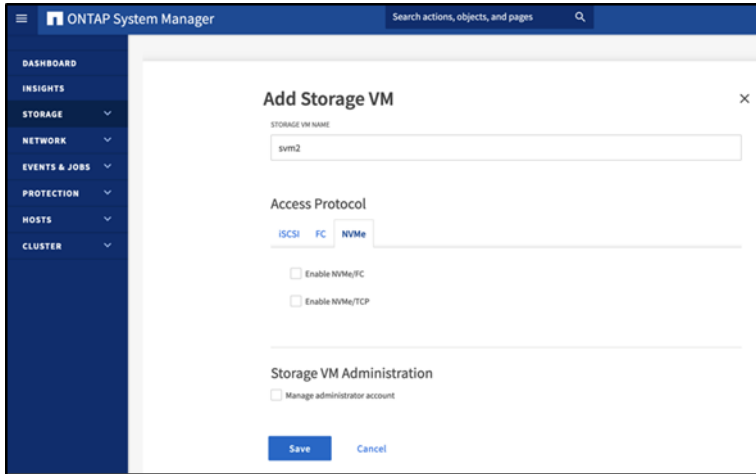


図30) OnCommand System Manager-SVMの作成 : NVMe/FCの設定

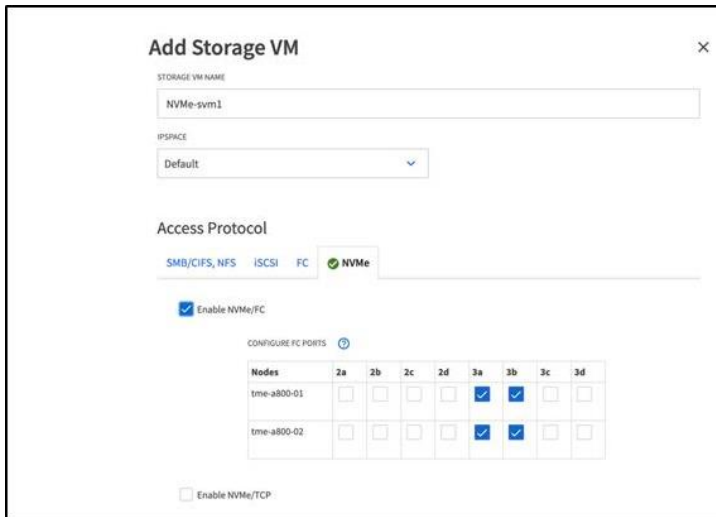
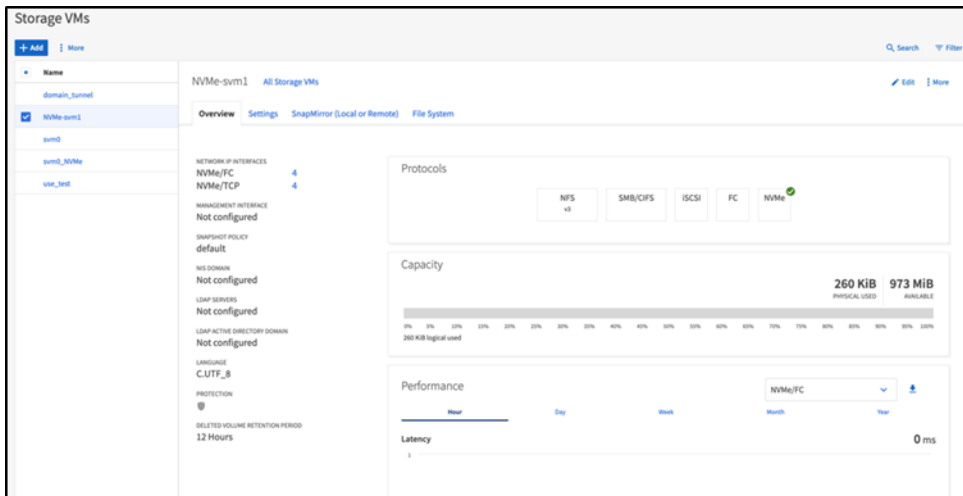


図31) OnCommandシステムマネージャ-SVMの作成 : NVMe/TCPの設定

図32) OnCommandシステムマネージャ-SVMの作成 : 管理の詳細を設定

2. 作成したSVMの概要を確認し、[OK]をクリックします。
3. 新しく作成したSVMを選択します。すべてのプロトコル設定とサービスステータスを確認するには、トップメニューの[SVM Settings]をクリックします。

図33) 新しく作成したSVMの表示



4. SVMダッシュボードページに戻るには、SVM設定ページの右上にある[Back]をクリックします。[SVM Dashboard]ページにNVMeステータスが緑で表示されます。
5. クラスタ内のすべてのネームスペースの詳細が表示された[ネームスペース管理]ウィンドウを起動します。左側のメニューペインで、[ストレージ]>[NVMe]>[NVMeネームスペース]に移動します。次のようにネームスペースを作成します。
 - a. [作成]をクリックします。
 - b. 作成したSVMを選択します。
 - c. [アドバンスト]オプションを使用して、すべてのネームスペース名の前に付ける命名パターンを作成します。
 - d. [命名パターン (Naming Pattern)]ダイアログボックスに関連する詳細を入力する。
 - e. [適用]をクリックします。
 - f. [送信]をクリックしてネームスペースを作成します。
6. [Close]をクリックします。

図34) システムマネージャ-新しいOnCommandネームスペースの作成

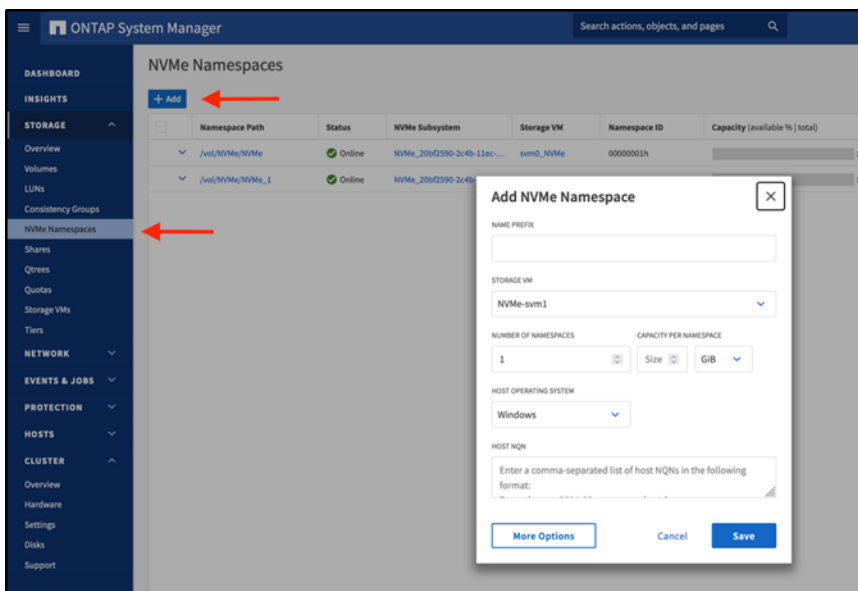


図35) OnCommandシステムマネージャ-新しく作成されたNVMeネームスペースの表示

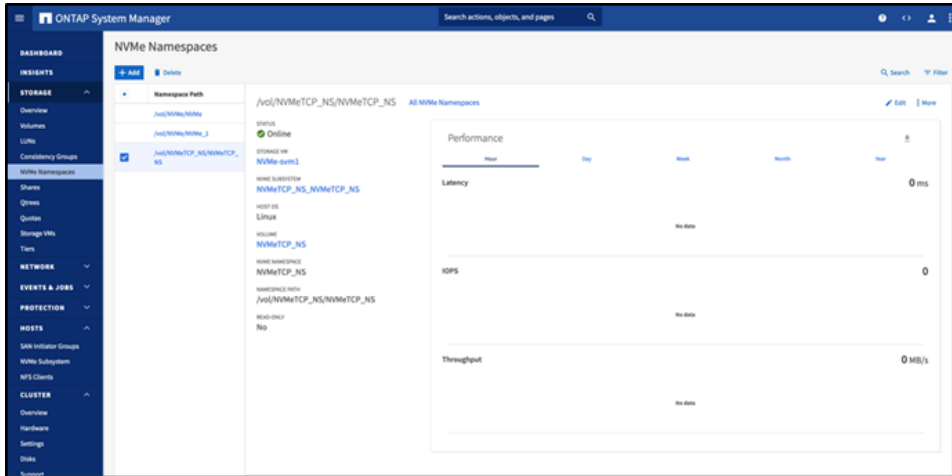


図36) 作成したNVMeサブシステムの表示



付録B : ONTAP NVMe/FCおよびNVMe/TCPのCLIコマンド-初期セットアップと検出

ONTAPコントローラ

1. クラスタにNVMe / FC対応アダプタがインストールされていることを確認します。

注 : これらはNVMe/FCでは必要ですが、NVMe/TCPでは不要です。

```
AFF::> fcp adapter show -data-protocols-supported fc-nvme
(network fcp adapter show)
Connection Port Admin Operational
Node Adapter Established Address Status Status
-----
AFF_1 1a true 10100 up online
AFF_1 1b true 10200 up online
2 entries were displayed.
```

2. NVMe/TCP LIFをホストするイーサネットポートを表示する

```
AFF::> net port show
(network port show)

Node: tme-a800-01

Port      Ipspace      Broadcast Domain  Link MTU      Speed (Mbps) Health
Admin/Oper Status
-----
e0M      Default     Default           up 1500 auto/1000 healthy
e0a      Cluster     Cluster           up 9000 auto/100000 healthy
```

```

e1a      Cluster      Cluster      up    9000  auto/100000
healthy
e4a      Default      Default      down 1500  auto/-
-
e4b      Default      Default      down 1500  auto/-
-
e4c      Default      Default      down 1500  auto/-
-
e4d      Default      Default      down 1500  auto/-
-
<Output omitted>
e5b      Default      NVMeT       up    9000  auto/100000
healthy
19 entries were displayed.

```

3. NVMeトラフィックをホストするSVMを作成します。

```

AFF::> vserver create -vserver nvme1
[Job 2831] Job succeeded:
Vserver creation completed.

```

4. 作成されたSVMを表示します。

```

AFF::> vserver show
          Admin  Operational Root
Vserver  Type Subtype State      State      Volume  Aggregate
-----
AFF      admin -      -      -      -      -
AFF_1    node -      -      -      -      -
AFF_2    node -      -      -      -      -
nvme1    data default running running   svm_root AFF_2_SSD_1
4 entries were displayed.

```

5. SVMで許可されているプロトコルを表示します。

```

AFF::> vserver show -vserver nvme1 -fields allowed-protocols
vserver allowed-protocols
-----
nvme1 -

```

6. NVMeプロトコルを追加します。

```

AFF::> vserver add-protocols -vserver nvme1 -protocols nvme

```

7. SVMで許可されているプロトコルを表示します。

```

AFF::> vserver show -vserver nvme1 -fields allowed-protocols
vserver allowed-protocols
-----
nvme1 nvme

```

8. NVMeサービスを作成します。

```

AFF::> vserver nvme create -vserver nvme1

```

9. NVMeサービスのステータスを表示します。

```

AFF::> vserver nvme show -vserver nvme1

Vserver Name: nvme1
Administrative Status: up

```

10. NVMe/FC LIFを作成する

```

AFF::> network interface create -vserver nvme1 -lif fcnvme-nodel-1a -role data -data-protocol fc-
nvme -home-node nodel -home-port 1a

```

11. NVMe/TCP LIFを作成します。

```

network interface create -vserver nvme1 -lif lif_nvme1_182 -role data -data-protocol nvme-tcp -
home-node tme-a800-02 -home-port e5b -address 192.168.1.20 -netmask 255.255.255.0

```

12. 新しく作成したLIFを表示します。

```

AFF::> net interface show -vserver nvme1

```

```
(network interface show)
Logical Status Network Current Current Is
Vserver Interface Admin/Oper Address/Mask Node Port Home
-----
nvmel
  fcnvme-nodel-1a
    up/up 20:60:00:a0:98:b3:f7:a7
      AFF_1 1a true
  lif_nvmel_182
    up/up 192.168.1.20/24
      AFF_1 e5b true
```

13. LIFと同じノードにボリュームを作成します。

```
AFF::> vol create -vserver nvmel -volume nsvol1 -aggregate AFF_2_SSD_1 -size 50gb
```

```
Warning: You are creating a volume that, based on the hardware configuration, would normally have
the "auto" efficiency policy enabled. Because the effective cluster version is not 9.3.0 or
later, the volume will be created without the "auto" efficiency policy. After upgrading, the
"auto" efficiency policy can be enabled by using the "volume efficiency modify" command.
[Job 2832] Job succeeded: Successful 1
```

注：この警告は無視してかまいません。volume efficiency modify コマンドを使用して、作成するボリュームに「auto」効率化機能を追加する必要があることが説明されています。

14. ネームスペースを作成します。

```
AFF::> vserver nvme namespace create -vserver nvmel -path /vol/nsvol1/ns1 -size 1GB -ostype linux
Created a namespace of size 1GB (1073741824).
```

15. サブシステムを作成します。

```
cluster1:> vserver nvme subsystem create -vserver nvmel -subsystem mysubsystem -ostype linux
```

16. 新しく作成したサブシステムを表示します。

```
AFF::> vserver nvme subsystem show -vserver nvmel
Vserver Subsystem Target NQN
-----
nvmel
  mysubsystem nqn.1992-08.com.netapp:sn.a6f7f76d40d511e8b3c900a098b3f7a7:subsystem.mysubsystem
```

ホスト

1. ホストからNQNを取得します。

注：この hostnqn 文字列は、/etc/nvme/hostnqn NVMe-CLIパッケージのインストール時に自動的に入力され、永続的に設定されます。その文字列はすでに一意です。したがって、hostnqn Linux nvme gen-hostnqn コマンドを使用して個別に文字列を生成する必要はありません。ホストNQNを削除した場合は、Linux nvme get-hostnqn ユーティリティを使用して生成できます。LinuxホストのNQNを永続的に設定するには、/etc/nvme/hostnqn ファイルにNQNを追加します。

2. ホストのNQNを表示します。

```
SLES_host:~ # cat /etc/nvme/hostnqn
nqn.2014-08.org.nvmexpress:fc_lif:uuid:2cd61a74-17f9-4c22-b350-3020020c458d
```

ONTAPコントローラ

1. hostnqn サブシステムに文字列を追加します。

```
AFF::> vserver nvme subsystem host add -vserver nvmel -subsystem mysubsystem -host-nqn nqn.1992-08.com.netapp:sn.a6f7f76d40d511e8b3c900a098b3f7a7:subsystem.mysubsystem
```

2. ネームスペースをサブシステムにマッピングします。

```
AFF::> vserver nvme subsystem map add -vserver nvmel -subsystem mysubsystem -path /vol/nsvol1/ns1
```

```
AFF::> vservers nvme namespace show -vservers nvme1 -instance

Vserver Name: nvme1
Namespace Path: /vol/nsvol1/ns1
  Size: 1GB
  Block Size: 4KB
  Size Used: 0B
  OS Type: linux
  Comment:
  State: online
  Is Read Only: false
  Creation Time: 4/15/2018 18:09:09
  Namespace UUID: 567fb229-a05e-4a57-aec9-d093e03cdf44
  Restore Inaccessible: false
Node Hosting the Namespace: AFF_1
  Volume Name: nsvol1
  Qtree Name:
Attached Subsystem: mysubsystem
  Namespace ID: 1
  Vserver ID: 89
```

付録C : ホスト構成情報

ホストの設定手順については、<https://docs.netapp.com/us-en/ontap-sanhost/>を参照してください。

付録D : LUNとネームスペース間の変換

LUNのネームスペースへの変換

1. LUNを表示します。

```
tme-a700s-clus::> lun show
Vserver Path State Mapped Type Size
-----
svm0 /vol/testLUN/testLUN online mapped linux 1GB
```

2. 変換するLUNのマッピングを解除します。

```
tme-a700s-clus::> lun unmap -vservers svm0 -path /vol/testLUN/testLUN -igroup
new_15Mar21_tif5_igroup
```

3. LUNを変換します。

```
tme-a700s-clus::> vservers nvme namespace convert-from-lun -vservers svm0 -lun-path
/vol/testLUN/testLUN
```

4. ネームスペースをNVMeサブシステムにマッピングします。

```
vservers nvme subsystem map add -vservers svm0 -subsystem svm0_subsystem_909 -path
/vol/testLUN/testLUN
```

5. 新しいネームスペースを表示します。

```
tme-a700s-clus::> vservers nvme namespace show
Vserver Path State Size Subsystem NSID
-----
Svm0 /vol/testLUN/testLUN online 1GB svm0_subsystem_909 00000001h
```

ネームスペースのLUNへの変換

1. 新しいネームスペースを表示します。

```
tme-a700s-clus::> vservers nvme namespace show
Vserver Path State Size Subsystem NSID
-----
```

```
Svm0 /vol/testLUN/testLUN online 1GB svm0_subsystem_909 00000001h
```

2. ネームスペースのマッピングを解除します。

```
vserver nvme subsystem map remove -vserver svm0 -subsystem svm0_subsystem_909 -path /vol/testLUN/testLUN
```

3. ネームスペースをLUNに変換します。

```
lun convert-from-namespace -vserver svm0 -namespace-path /vol/testLUN/testLUN
```

4. LUNをigroupにマッピングします。

```
lun map -vserver svm0 -path /vol/testLUN/testLUN -igroup new_15Mar21_tif5_igroup -lun-id 20
```

5. LUNを表示します。

```
tme-a700s-clus::> lun show -vserver svm0
Vserver Path State Mapped Type Size
-----
svm0 /vol/testLUN/testLUN online mapped linux 1GB
```

付録E : NVMe/FCの拡張性と制限

本書の執筆時点では、NVMe/FCの拡張性と制限は <https://hww.netapp.com/>に記載されています。

付録F : トラブルシューティング

NVMe/FCの障害のトラブルシューティングを行う前に、必ず [IMT](#) 仕様に準拠した構成を実行していることを確認してください。次に、次のセクションの手順に進み、ホスト側の問題をデバッグします。

NVMe/FCのIPFC詳細ログ

```
Here is a list of lpfc driver logging bitmasks available for NVMe/FC, as seen in
drivers/scsi/lpfc/lpfc_logmsg.h:
#define LOG_NVME 0x00100000 /* NVME general events. */
#define LOG_NVME_DISC 0x00200000 /* NVME Discovery/Connect events. */
#define LOG_NVME_ABTS 0x00400000 /* NVME ABTS events. */
#define LOG_NVME_IOERR 0x00800000 /* NVME IO Error events. */
```

lpfc_log_verbose lpfc /etc/modprobe.d/lpfc.conf lpfc ドライバ側からNVMe/FCイベントを記録する場合は、ドライバ設定（の行に追加）を上記のいずれかの値に設定します。次に initiramfs、を実行してを再作成 dracut -f し、ホストをリブートします。リブート後、LOG_NVME_DISC 例として前のビットマスクを使用して次の出力を確認し、詳細ロギングが適用されていることを確認します。

```
# cat /etc/modprobe.d/lpfc.conf
options lpfc lpfc_enable_fc4_type=3 lpfc_log_verbose=0x00200000
# cat /sys/module/lpfc/parameters/lpfc_log_verbose
2097152
```

NetAppでは、一般的な問題に対して次のlpfcロギングビットマスク値を推奨しています。

- 一般的なNVMe検出/接続イベント : 0x00200000
- リンクバウンス時のFC-LS検出の問題に関連するLPFCドライバイベント（LIF /ポートトグルイベントなど） : 0xf00083

NVMe-CLIの一般的なエラーとその回避策

ここでは `nvme discover`、`nvme connect`、および `nvme connect-all` の処理中にNVMe-CLIユーティリティによって表示されるいくつかのエラーメッセージについて説明します。これらのエラーの考えられる原因とその回避策について説明します。

エラーメッセージ

```
Failed to write to /dev/nvme-fabrics: Invalid argument.
```

- **Probable原因**：このエラーメッセージは、通常、構文が間違っている場合に表示されます。
- **回避策**：前の回避策コマンドの正しい構文を使用してください。

エラーメッセージ

```
Failed to write to /dev/nvme-fabrics: No such file or directory.
```

- **Probable原因**：原因このエラーが発生する可能性のあるいくつかの問題。一般的な原因には次のようなものがあります。

以前のNVMeコマンドに間違った引数が渡されました。

- **回避策**：前のコマンドに対して適切な引数（正しい回避策文字列、WWPN文字列など）が渡されていることを確認します。

引数が正しいにもかかわらずエラーが表示される場合は、`/sys/class/scsi_host/host*/nvme_info` 出力が正しいかどうかを確認します。Enabled [Remote Ports]セクションの下にNVMeイニシエータが、NVMe/FCターゲットLIFが正しく表示されていることを確認します。

例：

```
# cat /sys/class/scsi_host/host*/nvme_info
NVME Initiator Enabled
NVME LPORT lpfc0 WWPN x10000090fae0ec9d WWNN x20000090fae0ec9d DID x012000 ONLINE
NVME RPORT WWPN x200b00a098c80f09 WWNN x200a00a098c80f09 DID x010601 TARGET DISCSRVC ONLINE
NVME Statistics
LS: Xmt 00000000000000006 Cmpl 00000000000000006
FCP: Rd 00000000000000071 Wr 0000000000000005 IO 0000000000000031
Cmpl 00000000000000a6 Outstanding 0000000000000001
NVME Initiator Enabled
NVME LPORT lpfc1 WWPN x10000090fae0ec9e WWNN x20000090fae0ec9e DID x012400 ONLINE
NVME RPORT WWPN x200900a098c80f09 WWNN x200800a098c80f09 DID x010301 TARGET DISCSRVC ONLINE
NVME Statistics
LS: Xmt 00000000000000006 Cmpl 00000000000000006
FCP: Rd 00000000000000073 Wr 0000000000000005 IO 0000000000000031
Cmpl 00000000000000a8 Outstanding 0000000000000001
```

回避策：ターゲットLIFが上記のように `nvme_info` 出力に表示されない場合は `/var/log/messages`、および `dmesg` の出力に不審なNVMe/FC障害がないかどうかを確認し、必要に応じてレポートまたは修正します。

エラーメッセージ

```
Failed to write to /dev/nvme-fabrics: Operation already in progress
```

Probable原因：このエラーメッセージは、コントローラの関連付けまたは指定した処理がすでに作成されているか、または作成中の場合に表示されます。これは、インストールされている自動接続スクリプトの一部として発生する可能性があります。

回避策：なし。の場合 `nvme discover`は、しばらくしてからこのコマンドを実行してください。`nvme connect` およびの場合は `connect-all`、を実行して、`nvme list` ネームスペースデバイスがすでに作成されてホストに表示されていることを確認します。

エラーメッセージ

```
No discovery log entries to fetch
```

Probable原因： 通常、このエラーメッセージは、`/etc/nvme/hostnqn` NetAppアレイの対応するサブシステムに文字列が追加されていない場合に表示されます。このエラーは、それぞれのサブシステムに誤った`hostnqn`文字列が追加されている場合にも発生することがあります。

回避策： `/etc/nvme/hostnqn` NetAppアレイの対応するサブシステムに正確な文字列が追加されていることを確認します。`vserver nvme subsystem host show` コマンドを実行して確認します。

デバッグに必要なファイルとコマンド出力

引き続き問題が発生する場合は、次のファイルとコマンド出力を収集してNetAppに送信し、さらに優先順位を付けてください。

```
cat /sys/class/scsi_host/host*/nvme_info
/var/log/messages
dmesg
```

`nvme discover` 出力 (を参照) :

```
nvme discover --transport=fc --traddr=nn-0x200a00a098c80f09:pn-0x200b00a098c80f09 --host-
traddr=nn-0x20000090fae0ec9d:pn-0x10000090fae0ec9d
nvme list
```

付録G : MCC IPでのNVMe/FCの設定とセットアップ

ONTAP 9.12.1で、MCC IP 4パッククラスタでのNVMe/FCのMCC IPサポートが追加されました。

MCC NVMeホストタイムアウトセツテイ

ONTAP MCIPスイッチオーバー中に、プライマリクラスタサイトがセカンダリ/ピアクラスタサイトにスイッチオーバーしている間、**All Paths Down (APD)** ウィンドウが表示されます。すべてのパスがここにあるため、このAPDウィンドウの実行中、ホスト/クライアントはすべてのNVMeネームスペースデバイスにアクセスできなくなります。このAPDウィンドウが指定した間隔を超えると、Linux NVMe/FCホストは停止し、上位のアプリケーションへのI/Oエラーが発生します。この動作は、リンク損失時のLinux NVMe/FCホストの動作が原因で発生します。この動作は、と呼ばれるトランスポートレイヤパラメータで制御されます `NVMe/FC dev_loss_tmo`。

次の技術情報アーティクルを参照して、NVMe/FC `DEV_LOSS_TMO`を調整してください。

[NetApp MetroCluster IPでのNVMe/FCプロトコルのサポートに関するLinuxホストの推奨事項。](#)

NVMeプラットフォームのサポートおよび設定の制限事項

NVMe-oFプロトコルのサポートは、ONTAPのバージョンによってプラットフォームと構成によって異なります。

NVMe FC LIFの制限事項

両方のクラスタのFCポートを同じファブリックに接続し、ソフトゾーニングを使用する必要があります。これにより、接続されたポートにLIFが配置され、スイッチオーバー後にホストがLIFにログインできるようになります。

詳細については、次のMetroClusterテクニカルレポートを参照してください。

- TR-4689 : 『MetroCluster IP -解決策Architecture and Design』
<https://www.netapp.com/pdf.html?item=/media/13481-tr4689.pdf>
- TR-4705 : NetApp MetroClusterソリューションのアーキテクチャと設計
<https://www.netapp.com/media/13480-tr4705.pdf>

付録H : NVMe/TCP経由のセキュアな認証のセットアップ

手順

1. NVMeサブシステムにDH-HMAC-CHAP認証を追加します。

```
vserver nvme subsystem host add - vserver svm_name
...
-subsystem subsystem
-host-nqn host_nqn
-dhchap-host-secret authentication_host_secret
-dhchap-controller-secret authentication_controller_secret
-dhchap-hash-function {sha-256|sha-512}
-dhchap-group {none|2048-bit|3072-bit|4096-bit|6144-bit|8192-bit}
```

2. DH-HMAC CHAP認証プロトコルがホストに追加されていることを確認します。

```
vserver nvme subsystem host show
...
[ -dhchap-hash-function {sha-256|sha-512} ] Authentication Hash
Function [ -dhchap-group {none|2048-bit|3072-bit|4096-bit|6144-
bit|8192-bit} ]
Authentication Diffie-Hellman
Group
[ -dhchap-mode {none|unidirectional|bidirectional} ]
```

3. DH-HMAC CHAP認証プロトコルがコントローラに追加されていることを確認します。

```
vserver nvme subsystem controller show
...
[ -dhchap-hash-function {sha-256|sha-512} ] Authentication Hash Function
[ -dhchap-group {none|2048-bit|3072-bit|4096-bit|6144-bit|8192-bit} ]
Authentication Diffie-Hellman
Group
[ -dhchap-mode {none|unidirectional|bidirectional} ]
Authentication Mode
```

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。

ホストOSのセットアップと設定

- ホストのセットアップと設定の情報
<https://docs.netapp.com/us-en/ontap-sanhost/>

標準ドキュメント

- 提案されるT10 SCSIブロックコマンド
<http://t10.org/ftp/t10/document.05/05-344r0.pdf>
- RFC 3270 Internet Small Computer Systems Interface (iSCSI)
<https://tools.ietf.org/html/rfc3270>
- RFC 7143 Internet Small Computer System Interface (iSCSI) プロトコル (統合)
<https://tools.ietf.org/html/rfc7143>
- RFC 5041信頼性の高いトランスポート上での直接データ配置
<https://tools.ietf.org/html/rfc5041>
- INCITS T11-T11-2017-00145-v004 FC-NVMe仕様
https://standards.incits.org/apps/group_public/document.php?document_id=92164

SUSE Enterprise Linuxのリンク

- SUSEのダウンロード
<https://www.suse.com/download-linux/>
- SUSE Linuxのマニュアル
<https://www.suse.com/documentation/>

Brocadeリンク

- SANの健全性 :
<https://www.broadcom.com/support/fibre-channel-networking/tools/san-health/diagnostics-capture>
- サポートSAN健全性診断キャプチャのダウンロード
<https://www.broadcom.com/support/fibre-channel-networking/tools/san-health/diagnostics-capture>
- Fibre ChannelでNVMeのパワーを最大限に引き出す
<https://docs.broadcom.com/docs/12395453>
- 『NVMe over Fibre Channel for Dummies Book』（第3版） :
<https://docs.broadcom.com/docs/nvme-over-fibre-channel-for-dummies-book>
- 『Planning for the Transition of Production-Ready NVMe over Fabrics Deployments in the Enterprise :
<https://docs.broadcom.com/docs/planning-for-the-transition-to-production-ready-nvme>
- Broadcom NVMe over Fibre ChannelとVMware vSphere 7.0のサポート
<https://docs.broadcom.com/docs/12399210>
- IDC : 『Native NVMe/FC Support Provides a Performance Growth Path for Virtual Infrastructure』
<https://docs.broadcom.com/docs/12398958>
- Tollyテストレポート- Emulex Gen 7 LPe35002 VMware ESXi 7.0、NVMe/FCとSCSI
<https://docs.broadcom.com/docs/Emulex-Gen-7-LPe35002-VMware-ESXi-7.0-NVMe-FC-SCSI>

ビデオ、Webキャスト、ブログ

- FCIA : Introducing Fibre Channel NVMe (Bright Talk Webキャスト)
https://www.brighttalk.com/webcast/8615/242341?utm_campaign=webcasts-search-results-feed&utm_content=FCIA&utm_source=brighttalk-portal&utm_medium=web
- FCIA : Dive Deep into NVMe over Fibre Channel (FC-NVMe) (Bright talk Webキャスト)
https://www.brighttalk.com/webcast/14967/265459?utm_campaign=webcasts-search-results-feed&utm_content=FCIA&utm_source=brighttalk-portal&utm_medium=web

ホワイトペーパー、製品の発表、分析

- RoCEとiWARPの競合分析
http://www.mellanox.com/related-docs/whitepapers/WP_RoCE_vs_iWARP.pdf
- 超低レイテンシのSamsung Z-NAND SSD
https://www.samsung.com/us/labs/pdfs/collateral/Samsung_Z-NAND_Technology_Brief_v5.pdf

NetAppのドキュメント、テクニカルレポート、その他のNVMe関連資料

- ONTAPのNVMeプロトコルのライセンス情報
https://kb.netapp.com/Advice_and_Troubleshooting/Data_Storage_Software/ONTAP_OS/Licensing_information_for_NVMe_protocol_on_ONTAP
- TR-4080 : 『Best Practices for Scalable SAN』
<https://www.netapp.com/pdf.html?item=/media/10680-tr4080pdf.pdf>
- NetAppのNVMeソリューション : お客様重視のテクノロジーで業界をリード
www.netapp.com/us/info/nvme.aspx
- NVMeCLIに関する素敵なブログ-NVMeコマンドセット
<https://nvmexpress.org/open-source-nvme-management-utility-nvme-command-line-interface-nvme-cli/>

- Tech OnTapポッドキャストエピソード72 : NVMeを徹底解剖
https://soundcloud.com/techontap_podcast/episode-72-demystifying-nvme?utm_source=clipboard&utm_medium=text&utm_campaign=social_sharing
- データシート : NetApp EF570オールフラッシュアレイ
<https://www.netapp.com/pdf.html?item=/media/81117-ds-3893.pdf>
- NetAppのドキュメント
<https://mysupport.netapp.com/documentation/productsatoz/index.html#O>
- NetAppソフトウェアのダウンロード
<https://mysupport.netapp.com/NOW/cgi-bin/software/>

ONTAP Cloud

- Amazon FSx for NetApp ONTAPとは
<https://docs.aws.amazon.com/fsx/latest/ONTAPGuide/what-is-fsx-ontap.html>
- Azure NetApp Filesのドキュメント
<https://learn.microsoft.com/en-us/azure/azure-netapp-files/>
- Cloud Volumes ONTAPのドキュメント
<https://docs.netapp.com/us-en/cloud-manager-cloud-volumes-ontap/>

NetApp検証済みアーキテクチャ

また、ベスト・プラクティス・リファレンス・アーキテクチャ、テストおよびテストされた構成、予想されるパフォーマンス結果を文書化した、NetApp Verified Architectureレポートを1つ以上参照することもできます。本ドキュメントの執筆時点では、NVMe/FCを使用したソリューションを対象とするNetApp Verified Architectureは2つありました。

- [NVA-1126 -設計 : NetApp and Broadcom Modern SAN Cloud-Connected Flash解決策Oracle and SUSE NetApp Verified Architecture Design Edition](#)
- [NVA-1127 -設計 : 『NetApp and Broadcom Modern SAN Cloud-Connected Flash解決策MongoDB and SUSE NetApp Verified Architecture Design Edition』](#)
- [NVA-1145 - 『Design : Modern SAN Cloud-Connected Flash解決策 NetApp、VMware、and Broadcom Verified Architecture Design Edition : with MS Windows Server 2019 and MS SQL Server 2017 workloads』](#)
- [NVA-1147 -設計 : 『SAP HANA on NetApp All SAN Array Modern SAN、データ保護、ディザスタリカバリ』](#)
- [NVA-1159 -設計 : 『Epic on Modern SAN NetApp Broadcom Healthcare解決策』](#)

追加のNetApp検証済みアーキテクチャが作成および計画されています。

NVMeに関するその他のドキュメント

NVMeとNVMe/FCによってパフォーマンスが向上し、レイテンシが短縮される仕組みの詳細については、次の資料を参照してください。

- [『NVMeによる最新SAN入門』](#)
- [Demartek評価版 : NVMe over Fibre Channelのパフォーマンス上のメリット-新しいパラレルで効率的なプロトコル](#)

バージョン履歴

バージョン	日付	ドキュメントの改訂履歴
バージョン1.0	2018年6月	初版リリース
バージョン2.0	2018年11月	ONTAP 9.5およびANA用に更新され、ホストのインストールと設定の付録が追加されました。
バージョン3.0	2019年4月	マイナーな更新と誤植訂正。

バージョン	日付	ドキュメントの改訂履歴
バージョン4.0	2019年6月	ONTAP 9.6、ANA、ラウンドロビン、およびマルチパスの推奨事項用に更新。
バージョン 4.1	2020年4月	ESXi 7.0のNVMe構成を追加して更新。
バージョン5.0	2020年12月	OS設定リンクおよびその他のマイナーアップデートを追加して更新。
バージョン6.0	2021年6月	ONTAP 9.9.1の新機能と推奨事項を追加して更新。
バージョン7.0	2023年2月	NVMe/TCP転送およびその他のONTAP更新を追加して更新。ホストおよびターゲット側の構成に関する推奨事項の最適化に関するセクションを追加しました。

お問い合わせ

本テクニカル レポートの品質向上について、ご意見をお寄せください。ご連絡先：decom-ng-doccomments@netapp.com、件名に「Technical Report 4684」を添えてください。

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 © 2023 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な保証を含み、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、その発牛理由、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随時、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）に関係し、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权を有し、本データの提供の根拠となっ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权については、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-4684-0223-JP