

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

NVMe over Fabrics のサポート

ETERNUS AB/HB series システムの InfiniBand、RoCE、
およびファイバチャネルを経由する Nonvolatile Memory Express

目次

1.	NVMe-oF の導入	6
2.	トランスポート層：InfiniBand、RoCE、FC.....	7
3.	NVMe と NVMe-oF：プロトコルと概念	8
3.1	コマンドセットとトランスポートプロトコル	8
3.2	NVMe の概念	9
3.2.1	NVMe	9
3.2.2	NVMe-oF	9
3.2.3	コントローラ	9
3.2.4	ネームスペース	10
3.2.5	ネームスペース ID	10
3.2.6	キューペア	11
3.3	コマンドセット	13
3.3.1	admin コマンド	13
3.3.2	NVM コマンド	13
3.3.3	ファブリックコマンド	14
3.4	ホストドライバおよびツール	14
3.4.1	ドライバスタック	14
3.4.2	マルチパスとフェイルオーバー	15
3.4.3	NVMe/FC と FC の共存	16
3.4.4	NVMe/IB、iSER、SRP の共存	16
3.4.5	NVMe/RoCE と iSCSI の共存	17
3.4.6	NVMe CLI	17
A.	よくある質問	18

目次

図 1.1	AB5100/HB5000 システムの NVMe-oF フロントエンド	6
図 1.2	AB3100/AB6100 システムの NVMe エンドツーエンド	6
図 3.1	論理 NVMe コントローラを使用した NVMe-oF ホスト / アレイのトポロジ	10
図 3.2	ホストグループへのネームスペース ID のマッピング	11
図 3.3	NVMe コントローラのキューペア	12
図 3.4	Linux OS ドライバの構造	15
図 3.5	NVMe/FC と FC の共存例.....	16
図 3.6	ホスト側の NVMe/IB、iSER、SRP の共存例.....	16
図 3.7	NVMe/RoCE と iSCSI の共存例.....	17

表目次

表 3.1	ETERNUS AB/HB series のトランスポートプロトコルとコマンドセットの組み合わせの履歴.....	8
表 3.2	ETERNUS AB/HB series でサポートされている admin コマンド.....	13
表 3.3	ETERNUS AB/HB series でサポートされている NVM コマンド	13
表 3.4	ETERNUS AB/HB series でサポートされている fabrics コマンド	14
表 3.5	便利な NVMe CLI コマンド	17

はじめに

ETERNUS AB/HB series では、InfiniBand (IB)、RDMA over Converged Ethernet (RoCE)、またはファイバチャネル (FC) 接続のいずれかを使用して新しい Nonvolatile Memory over Fabrics (NVMe-oF) プロトコルをサポートします。本書は、このシステムの実装、利点、および制限に関する技術的な詳細を提供します。また、SCSI と NVMe-oF の構造も比較します。

第 2 版
2025 年 3 月

登録商標

本製品に関連する他社商標については、以下のサイトを参照してください。

<https://www.fujitsu.com/jp/products/computing/storage/trademark/>

本書では、本文中の™、®などの記号は省略しています。

本書の読み方

対象読者

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

関連マニュアル

ETERNUS AB/HB に関連する最新の情報は、以下のサイトで公開されています。

<https://www.fujitsu.com/jp/products/computing/storage/manual/>

本書の表記について

■ 本文中の記号

本文中では、以下の記号を使用しています。

注意

お使いになるときに注意していただきたいことを記述しています。必ずお読みください。

備考

本文を補足する内容や、参考情報を記述しています。

1. NVMe-oF の導入

Nonvolatile Memory Express (NVMe) は PCIe ソリッドステートドライブ (SSD) の業界標準インターフェースになりました。プロトコルとコマンドセットが合理化され、I/O あたりのクロックサイクルが減少したことで、NVMe は最大 64,000 のキューとキューあたり最大 64,000 のコマンドをサポートします。これらの属性により、FC、SAS、SATA などの SCSI ベースのプロトコルよりも効率的です。

NVMe over Fabrics (NVMe-oF) の導入は、インターフェースの特徴である低レイテンシと少ないオーバーヘッドはそのまま、NVMe をよりスケーラブルにします。

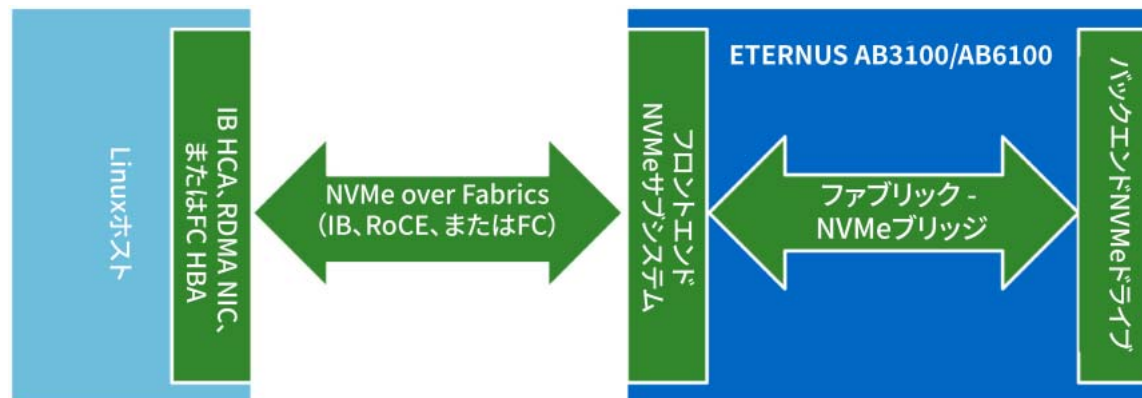
ETERNUS AB/HB series では、一部のモデルで NVMe-oF をサポートしており、トランスポートプロトコルとして NVMe/IB、NVMe/RoCE、および NVMe/FC があります。NVMe-oF は、ホストから AB5100 オールフラッシュアレイまたは HB5000 ハイブリッドアレイのフロントエンドまでサポートされますが、バックエンドは SAS ドライブを搭載した SCSI ベースのままです ([図 1.1](#) を参照)。

AB3100/AB6100 オールフラッシュアレイの登場により、ETERNUS AB/HB series では、[図 1.2](#) に示すように、フロントエンドからドライブに至るまで、NVMe に対するエンドツーエンドのサポートを提供しています。

図 1.1 AB5100/HB5000 システムの NVMe-oF フロントエンド



図 1.2 AB3100/AB6100 システムの NVMe エンドツーエンド



2. トランスポート層：InfiniBand、RoCE、FC

NVMeexpress.org の仕様では、リモートダイレクトメモリアクセス (RDMA) および FC を介した NVMe-oF のサポートの概要を説明しています。RDMA ベースのプロトコルは、IB または RDMA over Converged Ethernet version 2 (RoCE v2) です。本書では、NVMe/RoCE について言及する場合は、バージョン 2 が明示的に記載されているかどうかにかかわらず、バージョン 2 を示しています。

ETERNUS AB/HB series の実装では、NVMe/IB、NVMe/RoCE v2、および NVMe/FC がサポートされており、以下の特性などを含む利点があります。

- ETERNUS AB/HB series ストレージは、SCSI プロトコルコマンドのトランスポート層として FC をすでにサポートしています。NVMe/FC は、このように確立されたトランスポート層に新しいプロトコルを追加します。
- FC を実行する ETERNUS AB/HB series と同じハードウェアで、NVMe/FC を (同時にではありませんが) 実行できます。
- 両方のプロトコル (FC および NVMe/FC) を同じファブリック上で共存させることができ、ホスト側の同じ FC ホストバスアダプタ (HBA) ポート上でも共存させることができます。この機能により、FC を実行している既存のファブリックを使用しているお客様は、NVMe/FC を実行している ETERNUS AB/HB series を同じファブリックに接続できます。
- ファブリック内のすべての FC コンポーネント (ETERNUS AB/HB series、スイッチ、HBA) は、必要に応じて速度を下げるようネゴシエートできます。ネゴシエートできる速度は、32 Gbps、16 Gbps、および 8 Gbps です。速度を遅くすると、従来のコンポーネントへの接続が容易になります。
- IB と RoCE には RDMA が組み込まれています。
- ETERNUS AB/HB series ストレージは、RDMA (SCSI ベース) 上で iSCSI Extensions for RDMA (iSER) や SCSI RDMA Protocol (SRP) などの他のプロトコルをサポートします。
- AB5100 と HB5000 の同じホストインターフェースカードで、iSER、SRP、NVMe/IB、または NVMe/RoCE を実行できますが、同時には実行できません。
- AB3100/AB6100 上の同じホストインターフェースカードで NVMe/IB または NVMe/RoCE を実行できますが、同時には実行できません。
- 3 つのプロトコル (iSER、SRP、および NVMe/IB) はすべて、同じファブリック上で、さらにはホスト側の同じ InfiniBand ホストチャネルアダプタ (HCA) ポート上でも共存できます。この機能により、iSER や SRP を実行している既存のファブリックを使用しているお客様は、NVMe/IB を実行している ETERNUS AB/HB series を同じファブリックに接続できます。
- iSCSI と NVMe/RoCE は、ホスト側の同じファブリック上で共存できます。
- ETERNUS AB/HB series では、NVMe/RoCE に対して 100 Gbps、50 Gbps、40 Gbps、25 Gbps、10 Gbps の速度をサポートしています。
- ETERNUS AB/HB series は、NVMe/RoCE v2 (ルーティング可能) をサポートしており、RoCE v1 との下位互換性もあります。
- ファブリック内のすべての IB コンポーネント (ETERNUS AB/HB series、スイッチ、HCA) は、必要に応じて速度を下げるようネゴシエートできます。ネゴシエートできる速度は、拡張データレート (EDR) 100 Gbps、フォーティーンデータレート (FDR) 56 Gbps、またはクアドデータレート (QDR) 40 Gbps です。速度を遅くすると、従来のコンポーネントへの接続が容易になります。

3. NVMe と NVMe-oF : プロトコルと概念

3.1 コマンドセットとトランスポートプロトコル

このドキュメントでは、次の定義を使用します。

- **コマンドセット**
ストレージアレイとの間でデータを移動し、そのアレイ上のストレージを管理するために使用されるコマンドセット。
- **トランスポートプロトコル**
ストレージアレイとの間でコマンドおよびデータを伝送するために使用される、基礎となる物理接続およびプロトコル。

当初から、ETERNUS AB/HB series のストレージソフトウェアは、SCSI コマンドセットのストレージ機能を提供するように設計されました。長年にわたり、このソフトウェアは、追加の SCSI トランスポートプロトコルをサポートするために、アーキテクチャ的および機能的に拡張されてきました。しかし、いずれの場合も、基本的には SCSI 製品のままです。したがって、サポートされるコマンドセットの概念と基礎となるトランスポートプロトコルの概念とを区別する必要はありませんでした。つまり、ETERNUS AB/HB series は従来 SCSI 製品であったため、FC ホスト / ファブリック接続をサポートしているという事実は、FC トランスポートプロトコルがサポートする FCP SCSI プロトコルをサポートしていることを意味します。

[表 3.1](#) に、ETERNUS AB/HB series ストレージ製品がサポートするトランスポートプロトコルとコマンドセットの組み合わせを示します。歴史的には、基本となるトランスポートメカニズムは様々ですが、SCSI はあらゆるケースで使用されてきたコマンドです。

表 3.1 ETERNUS AB/HB series のトランスポートプロトコルとコマンドセットの組み合わせの履歴

物理接続	トランスポートプロトコル	コマンドセット
パラレル SCSI バス	SCSI	SCSI
FC	FC SCSI プロトコル (FCP)	SCSI
シリアル接続 SCSI	シリアル SCSI プロトコル (SSP)	SCSI
イーサネット	インターネット SCSI プロトコル (iSCSI)	SCSI
IB	SCSI RDMA プロトコル (SRP)	SCSI
IB	RDMA 用の iSCSI 拡張 (iSER)	SCSI
InfiniBand	NVMe-oF (NVMe/IB)	NVMe
RDMA over Converged Ethernet	NVMe-oF (NVMe/RoCE)	NVMe
FC	NVMe-oF (NVMe/FC)	NVMe

ETERNUS AB/HB series ストレージソフトウェア SANtricity OS は、SCSI 以外のコマンドセットである NVMe-oF をサポートします。SANtricity OS では、IB NVMe、RoCE v2 NVMe、FC NVMe の 3 つの NVMe トランスポートプロトコルすべてをサポートしています。

3.2 NVMe の概念

3.2.1 NVMe

NVMe は、アプリケーション (OS が提供するファイルシステムとドライバを使用) が PCI Express (PCIe) 接続を介して不揮発性メモリデータストレージと通信するための、仕様定義されたレジスタレベルインターフェースです。このインターフェースは、ストレージデバイスが同じ物理エンクロージャ内に存在し、ホストの OS とアプリケーションを、サーバやノートパソコンなどの PCIe 経由で直接接続できる場合に使用します。

3.2.2 NVMe-oF

NVMe-oF は、NVMe に対する仕様定義の拡張であり、PCIe 以外のインターコネクトを介した NVMe ベースの通信を可能にします。このインターフェースにより、基本的な通信メカニズムとして NVMe を使用しながら、「外部」ストレージエンクロージャを直接またはスイッチ経由でサーバに接続できます。

NVMe-oF プロトコルは、RoCE、IB、FC などのトランスポートプロトコルを含む、複数の異なる物理接続を経由して伝送できます。相互接続に関係なく、NVMe-oF はトランスポートに依存しません。サーバとストレージは、NVMe-oF プロトコルを伝送するために使用される基本的なトランスポートとは無関係に、NVMe-oF を介して通信します。ETERNUS AB/HB series ソフトウェアの SANtricity OS では、InfiniBand、RoCE v2、および FC トランスポートプロトコルがすべてサポートされています。

3.2.3 コントローラ

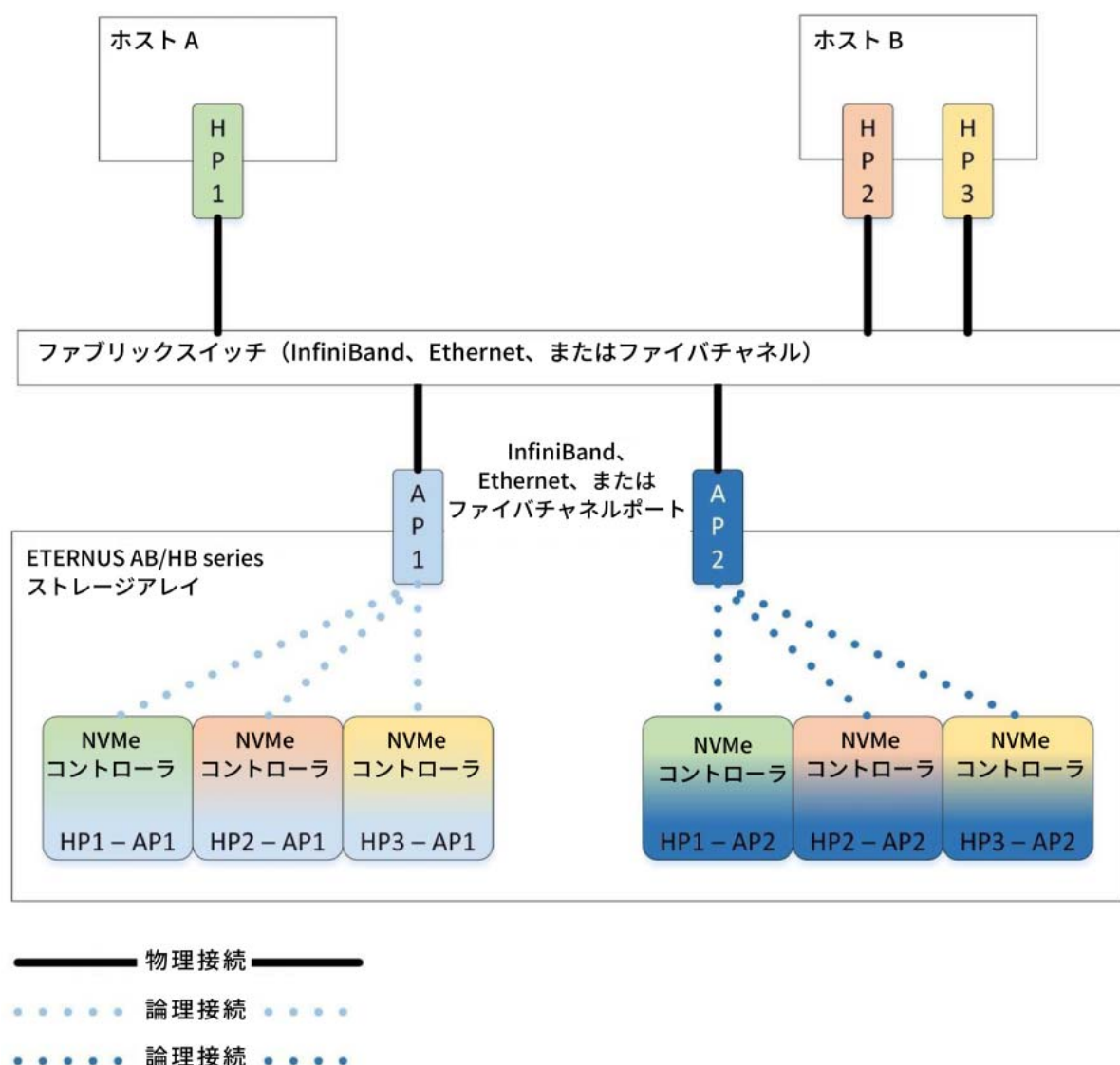
コントローラという用語は、NVMe のコンテキストでは非常に特殊な意味を持っています。NVMe コントローラは、NVMe Express を実装する PCI Express 機能として仕様で定義された具体的な実体です。本質的に、コントローラは、不揮発性データを格納するために PCIe でアクセス可能なハードウェアレジスタのセットを提供するデバイスです。

NVMe-oF 標準は、コントローラの定義を拡張しますが、一方でコントローラが一度に 1 つのホストに関連付けられることを指定します。したがって、NVMe-oF のコンテキストでは、コントローラは、単一のホストポートと単一のストレージアレイポートの間の関係を表す抽象的な実体になります。SCSI 用語では、これはイニシエータポートとターゲットポートの関係の概念に類似しています。イニシエータポートとターゲットポートの間に NVMe-oF 接続が確立されると、その接続を表すために NVMe コントローラが作成されます。

[図 3.1](#) は、典型的な NVMe-oF トポロジにおける論理 NVMe コントローラ構成を示します。ホスト A はスイッチに接続された 1 つのポート (HP 1) を提供し、ホスト B は同じスイッチに接続された 2 つのポート (HP 2 および HP 3) を提供します。ETERNUS AB/HB series コントローラには、ホストポートと同じスイッチに接続された AP 1 と AP 2 の 2 つのポートがあります。NVMe-oF 接続は、各ホストポートから各アレイポートへ確立されます。したがって、ETERNUS AB/HB series ソフトウェアは、接続ごとに論理 NVMe コントローラを作成します。

[図 3.1](#) は、構成例として 6 つの NVMe コントローラを示しており、ホストとアレイのポートの関連付けは HPx-APx と記述されています。

図 3.1 論理 NVMe コントローラを使用した NVMe-oF ホスト / アレイのトポロジ



3.2.4 ネームスペース

NVMe ネームスペースは、記憶領域を論理的に分割した各論理領域として定義されます。この定義は、論理ユニットの SCSI 概念と実質的に同じです。したがって、ETERNUS AB/HB series のストレージアレイで作成されたボリュームは、NVMe-oF インターフェースにはネームスペースとして提供されます。

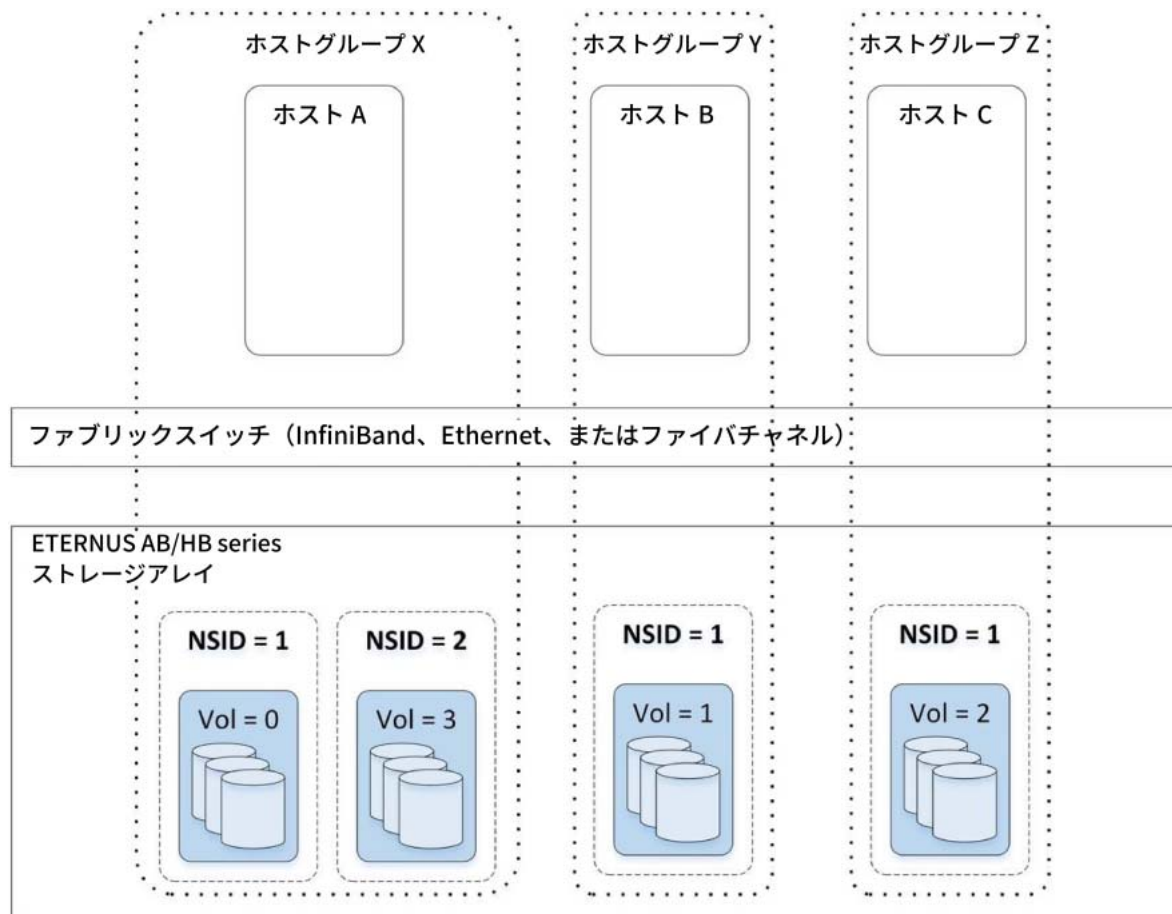
3.2.5 ネームスペース ID

ネームスペース ID (NSID) は、ネームスペースへのアクセスを提供するために ETERNUS AB/HB series コントローラによって使用される識別子です。これは、SCSI の論理ユニット番号 (LUN) とほぼ同じです。2つの例外があります。まず、NSID の値を 0 にすることはできません。第 2 に、FFFFFFFFh の NSID は、すべてのネームスペースを指定するために使用されるブロードキャスト値です。

ホストによるボリュームへのアクセスは、そのホストまたはホストグループのネームスペース ID の設定とともに、管理インターフェースから構成されます。SCSI と同様に、論理ボリュームは一度に 1 つのホストグループにのみマッピングできます。また、特定のホストグループに重複する NSID を割り当てることはできません。

図 3.2 は、複数のホストに対するストレージの論理パーティショニングを示しています。論理ボリューム 0 と 3 は、(0 が有効な NSID でないことを考慮すると、) ホストグループ X にそれぞれ NSID 1 と NSID 2 として割り当てられます。論理ボリューム 1 はホストグループ Y に NSID 1 として割り当てられ、論理ボリューム 2 はホストグループ Z に NSID 1 として割り当てられます。

図 3.2 ホストグループへのネームスペース ID のマッピング



3.2.6 キューペア

他のいくつかの RDMA ベースプロトコルと同様に、デバイス間の NVMe および NVMe-oF 通信は、キューペア (QP) の概念に依存します。QP は、送信キュー (SQ) と完了キュー (CQ) の組み合わせです。ホスト (イニシエータ) は、ストレージアレイ (ターゲット) によって読み取られる SQ にコマンドを配置します。

ターゲットは、受信したコマンドに関する完了情報を、コマンドを受信した SQ に関連付けられた CQ に配置します。NVMe-oF の場合、送信キューと完了キューの間に 1:1 の関係がなければなりません。つまり、すべての SQ には、その SQ だけに関連付けられた 1 つの固有の CQ が必要です。

NVMe コントローラが作成されると、最低 2 つの QP が作成されます。(キュー ID 0 を使用している) 最初の QP は、admin キューと呼ばれます。残りのすべてのキューは I/O キューと呼ばれます。admin キューは admin コマンドセット (「[3.3.1 admin コマンド](#)」(P.13) 参照) の処理に使用され、I/O キューは NVM コマンドセット (「[3.3.2 NVM コマンド](#)」(P.13) 参照) の処理に使用されます。

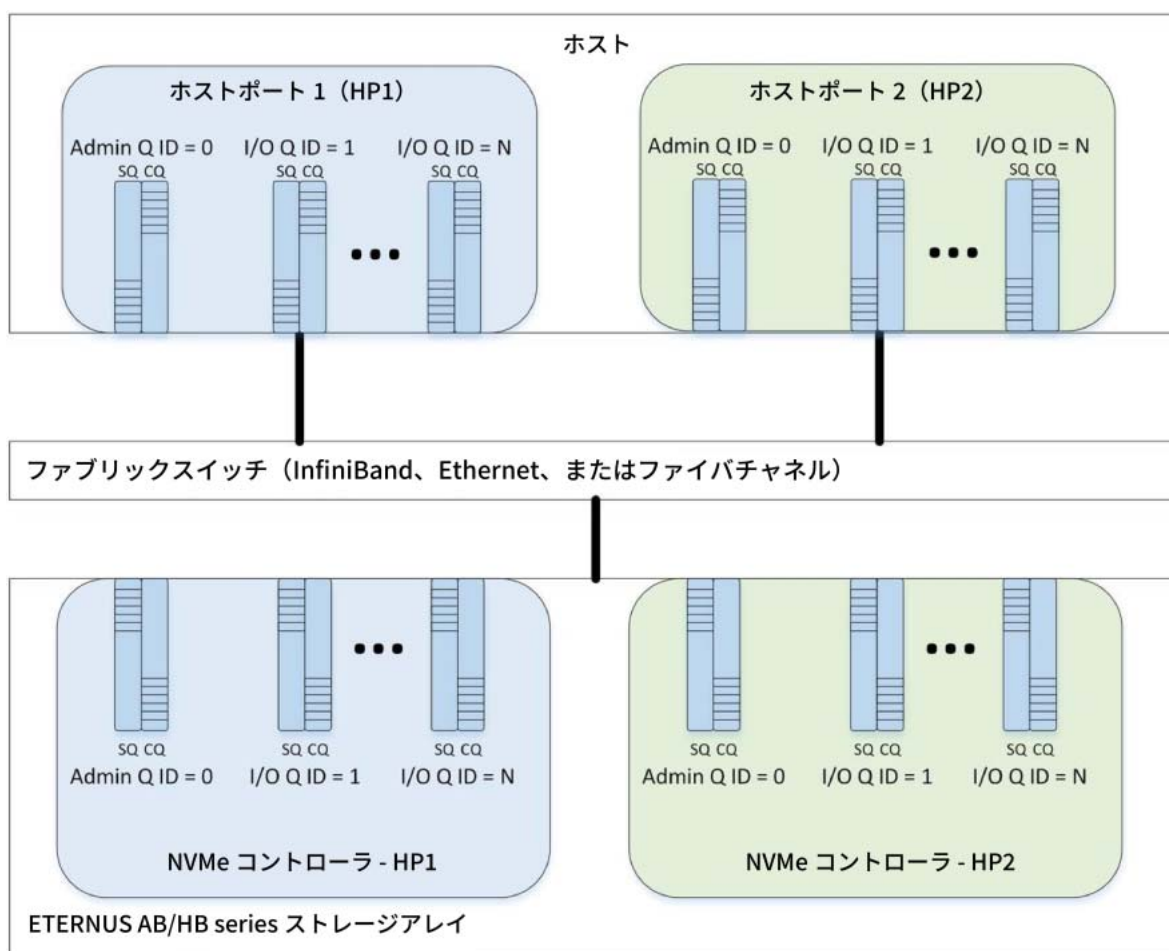
コントローラによって作成される I/O キューの数は、イニシエータとターゲットによってネゴシエートされます。

- 1 イニシエータはいくつかの I/O キューを要求します。
- 2 ターゲットは、そのコントローラで作成できる I/O キューの数で応答します。
- 3 イニシエータは、ターゲットによって示されたキュー数の I/O キュー作成要求を発行します。

ETERNUS AB/HB series の SANtricity OS ソフトウェアは、コントローラごとに 4 つの I/O キューをサポートします。

図 3.3 は、NVMe コントローラに適用されるキューペアの概念を示しています。NVMe コントローラは、ETERNUS AB/HB series ストレージソフトウェア内の論理構成要素です。これは、ホストとストレージ間の NVMe-oF 接続を形成する、ホストポートとアレイポートの組み合わせを表します。NVMe-oF 接続が確立されると、ホストは admin キュー (キュー ID が 0) を作成し、アレイとネゴシエーションを行って N 個の I/O キュー (キュー ID が 1 ~ N) を作成します。ホストキューペアとストレージアレイキューペアは同じ物理メモリシステムの一部ではありませんが、NVMe-oF は、それらを論理的に共有されたメモリキューペアにすることができ、コマンドおよび制御構造は、ホストとストレージアレイの間の RDMA を介して共有されます。したがって、ホストが I/O キュー 1 の送信キューのスロットに新しいコマンドを入れると、そのコマンドは、RDMA を介して、ストレージアレイ上の関連付けられた NVMe コントローラの I/O キュー 1 の同じスロットに置かれます。

図 3.3 NVMe コントローラのキューペア



3.3 コマンドセット

このセクションでは、ETERNUS AB/HB series ストレージレイソフトウェアでサポートされている NVMe および NVMe-oF コマンドの概要について説明します。特定のコマンドの詳細については、[NVMe および NVMe-oF の仕様](#)、ならびに ETERNUS AB/HB series の NVMe ホストインターフェースソフトウェアの仕様を参照してください。

3.3.1 admin コマンド

admin コマンドは、コントローラ admin QP を介して受信され、完了します。これらのコマンドは、特に、ホストに接続された NVMe コントローラのステータスを設定および監視するために使用されます。

特に、SCSI INQUIRY、MODE SELECT/SENSE、LOG SELECT/SENSE などの SCSI 非読み取り / 書き込みコマンドに類似しています。

表 3.2 ETERNUS AB/HB series でサポートされている admin コマンド

コマンド	仕様	必須 / オプション
Get Log Page	NVMe	必須
Identify	NVMe	必須
Abort	NVMe	必須
Set Features	NVMe	必須
Get Features	NVMe	必須
Asynchronous Event Request	NVMe	必須
Keep Alive	NVMe	オプション

3.3.2 NVM コマンド

NVM コマンドはコントローラ I/O QP を介して受信され、完了します。これらのコマンドは、データ移動やネームスペースアクセスなどの操作に使用され、READ、WRITE、SYNC CACHE、PERSISTENT RESERVE IN などの SCSI コマンドに類似しています。

表 3.3 ETERNUS AB/HB series でサポートされている NVM コマンド

コマンド	仕様	必須 / オプション
Flush	NVMe	必須
Write	NVMe	必須
Read	NVMe	必須
Reservation Register	NVMe	オプション
Reservation Report	NVMe	オプション
Reservation Acquire	NVMe	オプション
Reservation Release	NVMe	オプション

[表 3.3](#) に示すように、必須の NVM コマンドに加えて、NVMe-oF ホストインターフェースを使用する ETERNUS AB/HB series アレイは、オプションの予約機能をサポートします。この機能は、ホストがネームスペースを予約することを可能にし、他のホストがネームスペースに書き込むことを防ぎます (読み取りも制限されることがあります)。この機能は、クラスタ内のノードが誤って同じネームスペースに同時にアクセスしないようにするために、パッケージをクラスタリングする際によく使用されます。

SCSI の persistent reservation に詳しい人は、この機能が非常によく似ていると思うかもしれませんが、いくつかの違いがあります。SCSI では、登録者は 1 つのイニシエータとターゲットと LUN (I_T_L) の関係にのみ紐づけられていますが、NVMe-oF では登録者はネームスペースおよびホスト ID に紐づけられています。複数のコントローラが同じホスト ID を持つことができ、同じホスト ID を使用するすべてのコントローラが同じように扱われます。

もう 1 つの違いは、Persist Through Power Loss 設定をレジスタコマンドの外部で設定できることと、レジスタに On と Off に加えて No Change オプションを指定できることです。

NVMe-oF には、特定イニシエータまたはすべてのターゲットポートの概念がありません。特定のホストからのすべての接続に同じホスト ID を使用すると、同じ問題が根本的に解決します。NVMe-oF 標準 1.2 には、予約コマンドごとに Ignore Existing Key (IEKEY) パラメータがありますが、標準 1.3 のガイダンスを選択したため、register コマンド以外ではこのパラメータが削除されています。実際のコマンドおよびパラメータデータの形式は、SCSI とは明らかに異なります。

3.3.3 ファブリックコマンド

[表 3.4](#) に、キューペアの作成と NVMe コントローラの初期化に使用されるファブリックコマンドを示します。これらは、FC ポートログインなどのプロトコル固有の初期化コマンドにやや似ています。

表 3.4 ETERNUS AB/HB series でサポートされている fabrics コマンド

コマンド	仕様	必須 / オプション
Property Set	NVMe-oF	必須
Connect	NVMe-oF	必須
Property Get	NVMe-oF	必須

3.4 ホストドライバおよびツール

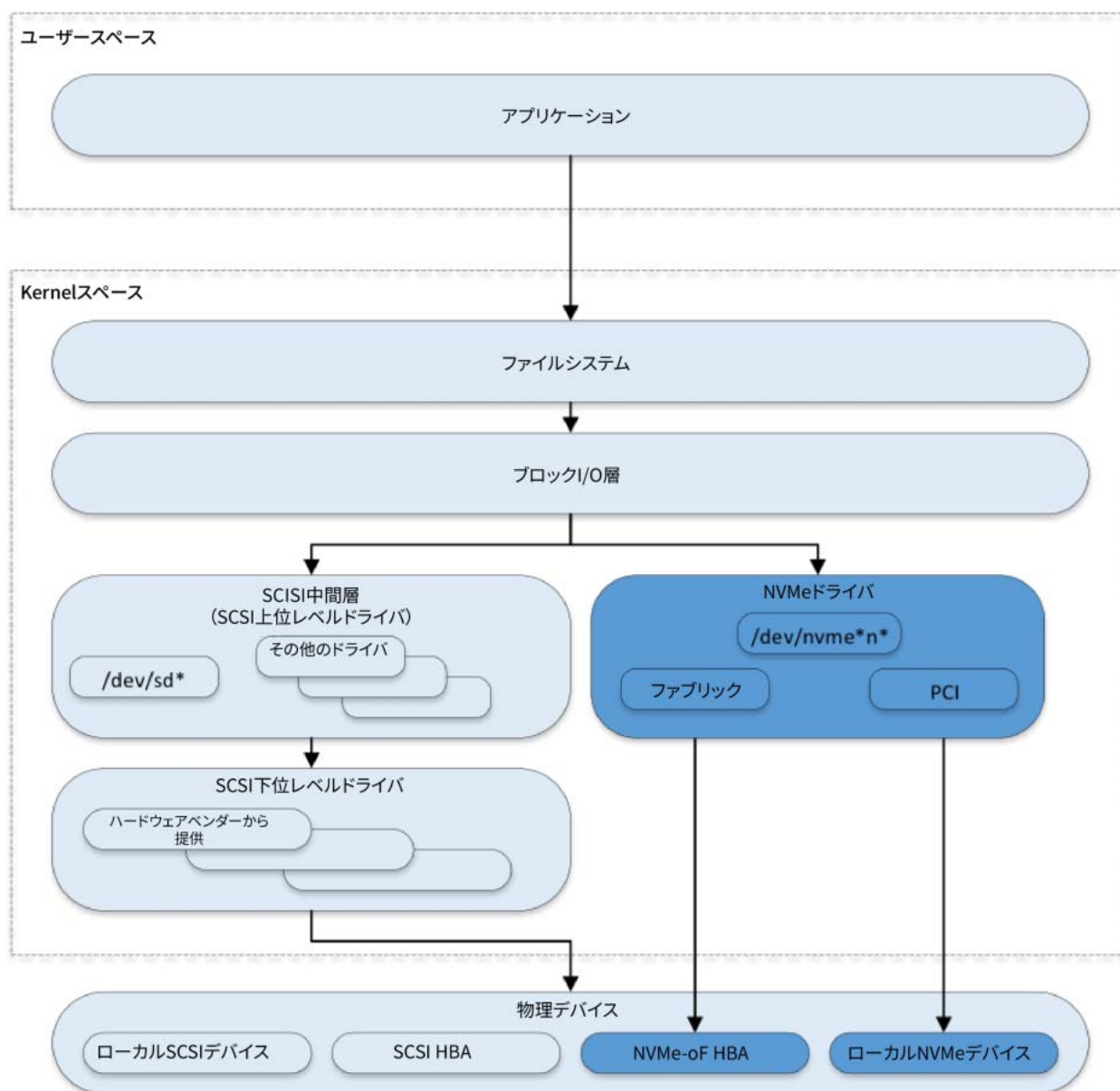
3.4.1 ドライバスタック

NVMe (と NVMe-oF) が SCSI に比べて優れている点の 1 つは、遅延の少ない I/O をサポートできることです。これは、デバイスが高速であるだけでなく、ホスト OS ドライバスタックに利点があるためです。したがって、アプリケーションからストレージへの I/O アクセスに要する時間が短縮され、レスポンスタイムが短縮されます。

[図 3.4](#) は、Linux OS ドライバスタックを単純化した図です。SCSI ドライバと同様に、NVMe ドライバはブロック I/O 層の下に位置します。ただし、NVMe ドライバは上位レベルと下位レベルに分割されません。ドライバは、NVMe デバイスをブロック I/O システムに提供し、PCIe 接続デバイスとファブリック接続デバイスの両方をサポートするドライバを含みます。ドライバのファブリック部分には、IB ベースの RDMA などの操作を処理するためのトランスポート固有のコードが含まれています。

NVMe ドライバを使用するもう 1 つの利点は、高速で不揮発性のメモリ記憶装置を考慮して設計されていることです。つまり、低レイテンシのストレージで適切に動作するように最適化されています。一方、SCSI ドライバはもともと回転式ディスクドライブのような回転メディアベースのストレージと一体になって作られ、使われていました。これらのデバイスでは、I/O 応答時間の大部分はデバイスの動作にかかっていました (たとえば、回転シーク時間)。したがって、ドライバ自体を実行時間のパフォーマンスに対して高度に最適化する必要はありませんでした。また、SCSI ドライバスタックは、ディスク、テープ、プリンターなど、複数の異なる種類の SCSI デバイスをサポートするように設計されていました。そのためドライバには、各種デバイスへの I/O を処理するために、NVMe ドライバには存在しないオーバーヘッドが組み込まれています。

図 3.4 Linux OS ドライバの構造



3.4.2 マルチパスとフェイルオーバー

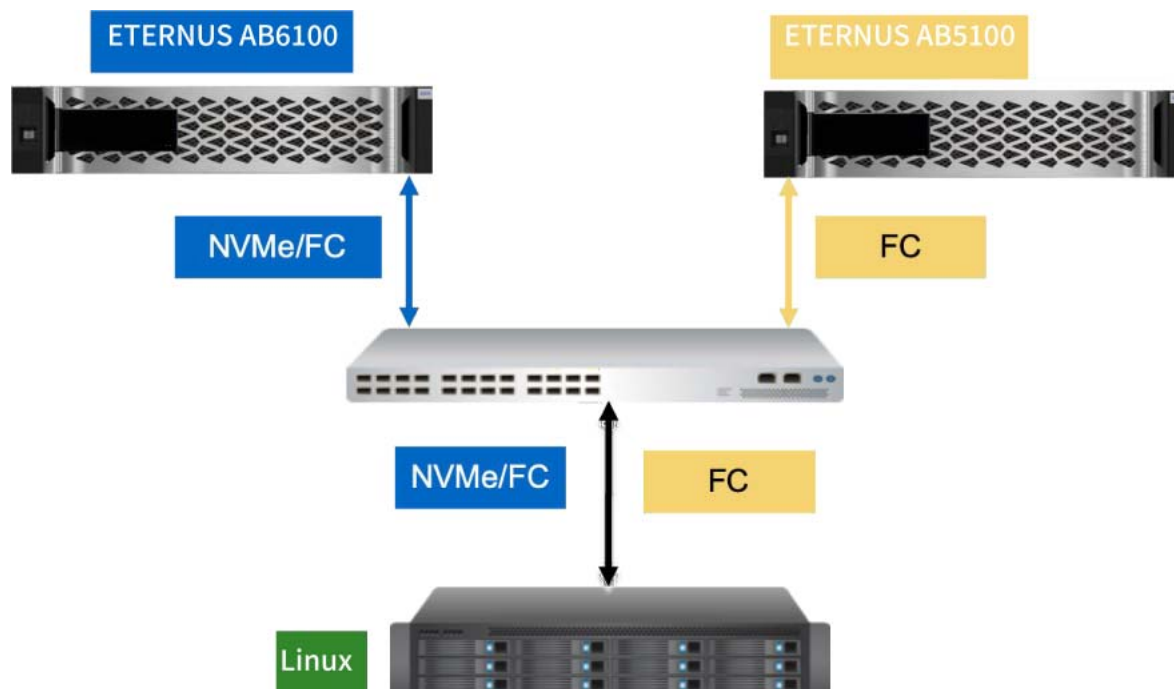
現在、マルチパスとフェイルオーバー機能は Linux Device Mapper Multipath (DM-MP) モジュールで提供されています。ETERNUS AB/HB series アレイ上のネームスペースへの複数パスは、自動的に 1 つの論理デバイスに構成されます。Linux マルチパスツールは、NVMe Asynchronous Namespace Access (ANA) 機能を使用してパスに関する情報を取得し、DM-MP と連携して I/O が最適パスを介して ETERNUS AB/HB series のネームスペースに送信されるようにして、フェイルオーバーやフェイルバックの操作を支援します。

新しい Linux リリースには、Linux NVMe ドライバに組み込まれているマルチパス / フェイルオーバー機能を使用するオプションがあります。このネイティブ NVMe マルチパスにより、ETERNUS AB/HB series のネームスペースに I/O を送信する際に DM-MP レイヤが不要になります。NVMe カーネルドライバは、NVMe ANA イベントを使用して、パス情報を更新し、I/O を最適パスにルーティングします。この情報は、フェイルオーバーやフェイルバックの操作でも使用されます。

3.4.3 NVMe/FC と FC の共存

図 3.5 に示すように、ホスト側では、FC と NVMe/FC の両方を同じホストバスアダプタ上で同時に実行できますが、異なる ETERNUS AB/HB series のターゲットに対してしか実行できません。

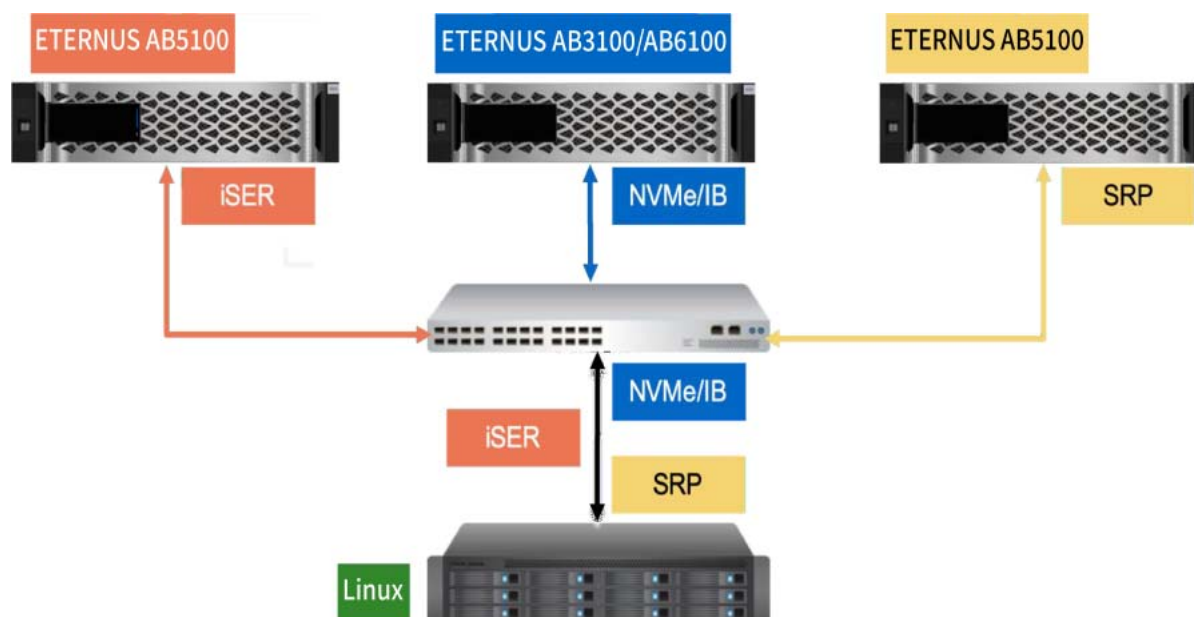
図 3.5 NVMe/FC と FC の共存例



3.4.4 NVMe/IB、iSER、SRP の共存

図 3.6 に示すように、ホスト側では、IB で ETERNUS AB/HB series デバイスがサポートするすべてのプロトコル (SRP、iSER、および NVMe/IB) を、同じホストチャネルアダプタ上で同時に実行できますが、異なる ETERNUS AB/HB series のターゲットに対してしか実行できません。

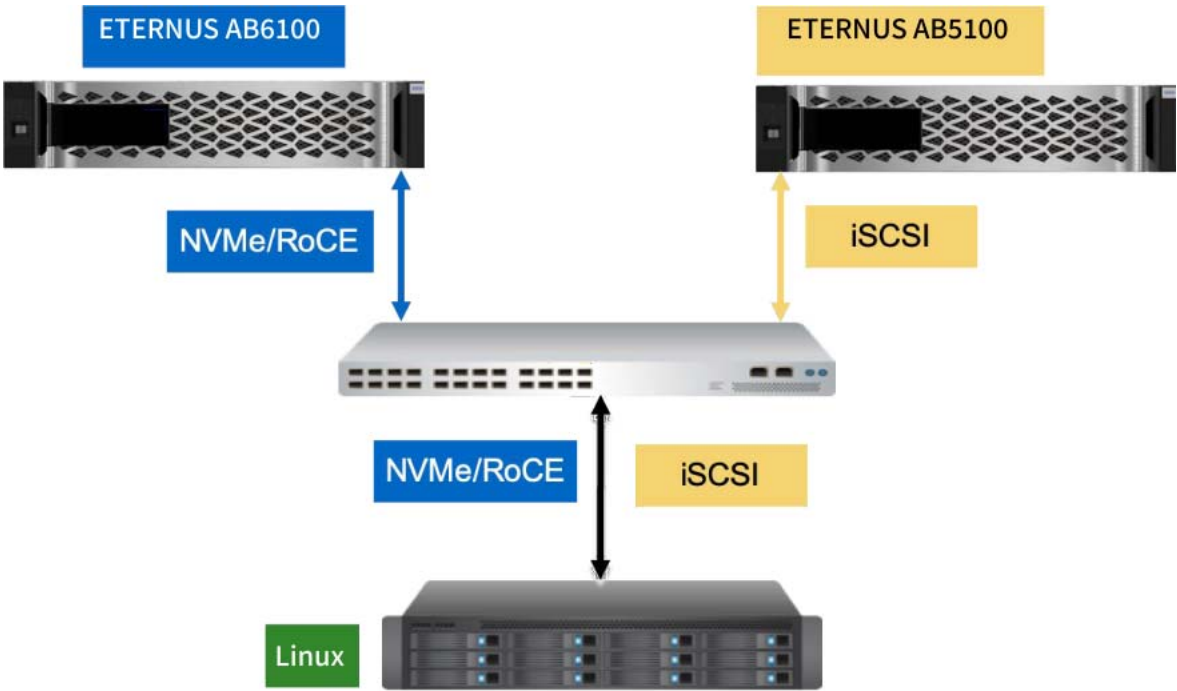
図 3.6 ホスト側の NVMe/IB、iSER、SRP の共存例



3.4.5 NVMe/RoCE と iSCSI の共存

図 3.7 に示すように、RDMA が同時にサポートされる場合、ホスト側では、iSCSI と NVMe/RoCE を同じネットワークアダプタ上で実行できますが、異なる ETERNUS AB/HB series のターゲットに対してしか実行できません。

図 3.7 NVMe/RoCE と iSCSI の共存例



3.4.6 NVMe CLI

NVMe CLI は、Linux ホスト上の NVMe デバイスの管理ツールを提供するコマンドラインユーティリティです。このユーティリティは、ホスト上でブロックデバイスとして構成されている NVMe ネームスペースをリストし、それらのネームスペースの詳細を提供できます。NVMe-oF には、ファブリック上の NVMe コントローラを検出して接続するためのサブコマンドがあります。表 3.5 に、いくつかの便利なコマンドをリストします。詳細は、Linux シェルから `man nvme` コマンドを実行して、Linux の man ページを参照してください。

表 3.5 便利な NVMe CLI コマンド

コマンド	説明
<code>nvme list</code>	システム上でブロックデバイスとして構成されている各 NVMe ネームスペースに関する情報をリストします。
<code>nvme discover</code>	ファブリック上の NVMe コントローラを検出します。
<code>nvme connect</code>	ファブリック上の NVMe コントローラに接続します。
<code>nvme connect-all</code>	<code>/etc/nvme/discovery.conf</code> に含まれる設定を使用して、ファブリック上の複数の NVMe コントローラを検出して接続します。

A. よくある質問

- 通常のネットワークインターフェースカード (NIC) は NVMe/RoCE でサポートされていますか。

回答：

いいえ。カードは RDMA をサポートしている必要があります。これらは RDMA NIC (rNIC) と呼ばれ、広く利用されています。

- 直接接続はサポートされていますか、またはスイッチが必要ですか。

回答：

NVMe/IB、NVMe/RoCE、および NVMe/FC では、直接接続とスイッチを介したファブリック接続の両方がサポートされます。

- ETERNUS AB/HB series のコントローラは、NVMe/RoCE 用の 4 x 25 GbE で 100 GbE のケーブルを使用した複数ホストへの直接接続をサポートしていますか。

回答：

いいえ。サポートしていません。

- 速度によるフォームファクタの違いは何ですか。

回答：

- NVMe/IB:QSFP28 (100 Gbps、56 Gbps、および 40 Gbps 用)
- NVMe/RoCE:QSFP28 (100 Gbps および 50 Gbps 用)、QSFP+ (40 Gbps 用)、SFP28 (25 Gbps 用)、SFP+ (10 Gbps 用)

- 100 Gbps InfiniBand ホストインターフェースカードでは、AB5100 と HB5000 はどのプロトコルをサポートしていますか。

回答：

SRP、iSER、NVMe/IB、および NVMe/RoCE プロトコルはサポートされていますが、同時にはサポートされていません。プロトコル間の移行はソフトウェアベースであり、ハードウェアの変更は必要ありません。

- ETERNUS AB/HB series のシステムは、SCSI を実行する RoCE をサポートしていますか。

回答：

いいえ。RoCE 上で動作するためにサポートされているのは NVMe プロトコルだけであり、SCSI プロトコルはサポートされていません。

- ETERNUS AB/HB series がサポートしている NVMe/RoCE のバージョンを教えてください。

回答：

NVMe/RoCE バージョン 2 をサポートし、NVMe/RoCE バージョン 1 と互換性があります。

ETERNUS AB series オールフラッシュアレイ , ETERNUS HB series ハイブリッドアレイ
NVMe over Fabrics のサポート

P3AG-6392-02Z0

発行年月 2025 年 3 月

発行責任 エフサステクノロジーズ株式会社

- 本書の内容は、改善のため事前連絡なしに変更することがあります。
- 本書の内容は、細心の注意を払って制作致しましたが、本書中の誤字、情報の抜け、本書情報の使用に起因する運用結果に関しましては、責任を負いかねますので予めご了承ください。
- 本書に記載されたデータの使用に起因する第三者の特許権およびその他の権利の侵害については、当社はその責を負いません。
- 無断転載を禁じます。