AFF/ FASシステム上のネットワーク構成

Enterprise applications

NetApp
February 10, 2026

目次

AFF/ FASシステム上のネットワーク構成

論理インターフェイス

Oracleデータベースにはストレージへのアクセスが必要です。Logical Interface（LIF；論理インターフェイス）は、Storage Virtual Machine（SVM）をネットワークに接続し、さらにデータベースに接続するネットワーク配管です。各データベースワークロードに十分な帯域幅を確保し、フェイルオーバーによってストレージサービスが失われないようにするには、LIFを適切に設計する必要があります。

このセクションでは、LIFの主な設計原則の概要を説明します。より包括的なドキュメントについては、"ONTAPネットワーク管理に関するドキュメント"。データベースアーキテクチャの他の要素と同様に、Storage Virtual Machine（SVM、CLIではVserver）と論理インターフェイス（LIF）の設計に最適なオプションは、拡張要件とビジネスニーズに大きく依存します。

LIFの戦略を立てる際は、主に次の点を考慮してください。

- *パフォーマンス。*ネットワーク帯域幅は十分か。
- *耐障害性。*設計に単一点障害はありますか？
- *管理性。*ネットワークを無停止で拡張できますか？

これらのトピックは、ホストからスイッチ、ストレージシステムまで、エンドツーエンドの解決策に適用されます。

LIFタイフ

LIFには複数のタイプがあります。 "LIFタイプに関するONTAPのドキュメント" このトピックのより包括的な情報を提供しますが、機能的にはLIFを次のグループに分類できます。

- *クラスタおよびノードの管理LIF。*ストレージクラスタの管理に使用するLIF。
- * SVM管理LIF。* REST APIまたはONTAPI（ZAPIとも呼ばれます）を使用してSVMへのアクセスを許可するインターフェイス。Snapshotの作成やボリュームのサイズ変更などの機能に使用できます。SnapManager for Oracle（SMO）などの製品では、SVM管理LIFにアクセスする必要があります。
- データ**LIF**。 FC、iSCSI、NVMe/FC、NVMe/TCP、NFS、 またはSMB / CIFSデータ。

> ファイアウォールポリシーを `data` 終了： `mgmt` または、HTTP、HTTPS、SSHを許可する別のポリシー。この変更により、NFSデータLIFと別の管理LIFの両方にアクセスするように各ホストを設定する必要がなくなるため、ネットワーク設定が簡易化されます。iSCSIトラフィックと管理トラフィックの両方にIPプロトコルを使用しているにもかかわらず、インターフェイスを設定することはできません。iSCSI環境では、個別の管理LIFが必要です。

SAN LIFの設計

SAN環境でのLIFの設計は、マルチパスという1つの理由で比較的簡単です。最新のSAN実装では、クライアントは複数の独立したネットワークパス経由でデータにアクセスし、アクセスに最適なパス（複数可）を選択できます。その結果、SANクライアントは使用可能な最適なパス間でI/Oの負荷を自動的に分散するため、パフォーマンスに関してはLIFの設計は容易に対処できます。

あるパスが使用できなくなった場合、クライアントは自動的に別のパスを選択します。その結果、設計がシンプルになるため、一般にSAN LIFの管理性が向上します。だからといって、SAN環境の方が常に簡単に管理できるわけではありません。SANストレージには、NFSよりもはるかに複雑な要素が多数あるからです。単純に、SAN LIFの設計が容易であることを意味します。

パフォーマンス

SAN環境でLIFのパフォーマンスを考慮する際に最も重要な考慮事項は、帯域幅です。たとえば、2ノードONTAP AFFクラスタの各ノードに16Gb FCポートを2つ搭載すると、各ノードとの間で最大32Gbの帯域幅を確保できます。

耐障害性

AFFストレージシステムでは、SAN LIFはフェイルオーバーしません。コントローラのフェイルオーバーが原因でSAN LIFに障害が発生すると、クライアントのマルチパスソフトウェアがパスの損失を検出し、I/Oを別のLIFにリダイレクトします。ASAストレージシステムでは、LIFは短時間でフェイルオーバーされますが、もう一方のコントローラにすでにアクティブなパスがあるためIOが中断されることはありません。フェイルオーバープロセスは、定義されたすべてのポートでホストアクセスをリストアするために実行されます。

管理性

NFS環境では、クラスタ内でのボリュームの再配置にLIFの移行が伴うことが多いため、LIFの移行ははるかに一般的なタスクです。SAN環境でHAペア内でボリュームを再配置しても、LIFを移行する必要はありません。ボリュームの移動が完了すると、ONTAPはパスの変更をSANに通知し、SANクライアントは自動的に再最適化します。SANを使用したLIFの移行は、主に物理ハードウェアの大幅な変更に関連しています。たとえば、コントローラの無停止アップグレードが必要な場合は、SAN LIFを新しいハードウェアに移行します。FCポートで障害が検出された場合は、LIFを未使用のポートに移行できます。

設計上の推奨事項

NetAppの推奨事項は次のとおりです。

- 必要以上の数のパスを作成しないでください。パスの数が多すぎると管理全体が複雑になり、一部のホストでのパスのフェイルオーバーで原因の問題が発生する可能性があります。さらに、一部のホストでは、SANブートなどの構成でパスが予期せず制限されます。
- ごく少数の構成では、LUNへのパスが4つ以上必要です。LUNにパスをアドバタイズするノードが3つ以上あると、LUNを所有するノードとそのHAパートナーに障害が発生した場合、LUNをホストしているアグリゲートにアクセスできなくなるため、その価値には制限があります。このような状況では、プライマリHAペア以外のノードにパスを作成しても役に立ちません。
- 参照可能なLUNパスの数はFCゾーンに含めるポートを選択することで管理できますが、一般には、ターゲットとなるポイントをすべてFCゾーンに含め、LUNの可視性をONTAPレベルで制御する方が簡単です。
- ONTAP 8.3以降では、選択的LUNマッピング（SLM）機能がデフォルトです。SLMを使用すると、新しいLUNはすべて、基盤となるアグリゲートを所有するノードとノードのHAパートナーから自動的に通知されます。これにより、ポートのアクセス性を制限するためにポートセットを作成したりゾーニングを設定したりする必要がなくなります。各LUNは、最適なパフォーマンスと耐障害性の両方を実現するために必要な最小限のノードで利用できます。
- LUNを2台のコントローラの外部に移行する必要がある場合は、 `lun mapping add-reporting-nodes` コマンドを実行して、新しいノードでLUNがアドバタイズされるようにします。これにより、LUNの移行用にLUNへの追加のSANパスが作成されます。ただし、新しいパスを使用するには、ホストで検出処理を実行する必要があります。

- 間接トラフィックを過度に気にしないでください。I/Oが大量に発生する環境ではレイテンシがマイクロ秒単位で重要になるため、間接トラフィックは避けることを推奨しますが、一般的なワークロードではパフォーマンスに目に見える影響はごくわずかです。

NFS LIFの設計

NFSでは、SANプロトコルとは異なり、データへの複数のパスを定義する機能に制限があります。NFSv4に対するParallel NFS（pNFS）拡張ではこの制限に対応していますが、イーサネットの速度が100GB以上に達しているため、パスを追加する価値があることはほとんどありません。

パフォーマンスと耐障害性

SAN LIFのパフォーマンスを測定することは、主にすべてのプライマリパスの合計帯域幅を計算することですが、NFS LIFのパフォーマンスを判断するには、正確なネットワーク構成を詳しく調べる必要があります。たとえば、2つの10Gbポートを物理ポートとして構成することも、Link Aggregation Control Protocol（LACP）インターフェイスグループとして構成することもできます。インターフェイスグループとして設定されている場合は、複数のロードバランシングポリシーを使用できます。ロードバランシングポリシーの動作は、トラフィックがスイッチングされるかルーティングされるかによって異なります。最後に、Oracle Direct NFS（dNFS）は、現時点ではどのOS NFSクライアントにも存在しないロードバランシング設定を提供します。

SANプロトコルとは異なり、NFSファイルシステムにはプロトコルレイヤでの耐障害性が必要です。たとえば、LUNは常にマルチパスを有効にして設定されるため、ストレージシステムではFCプロトコルを使用する複数の冗長チャネルを使用できます。一方NFSファイルシステムは、物理レイヤでのみ保護できる単一のTCP/IPチャネルの可用性に依存します。このような理由から、ポートフェイルオーバーやLACPポートアグリゲーションなどのオプションが用意されています。

NFS環境では、パフォーマンスと耐障害性の両方がネットワークプロトコルレイヤで提供されます。その結果、両方のトピックが絡み合っており、一緒に議論する必要があります。

ポートグループへの**LIF**のバインド

LIFをポートグループにバインドするには、LIFのIPアドレスを物理ポートのグループに関連付けます。物理ポートを1つに集約する主な方法はLACPです。LACPのフォールトトレランス機能は非常に簡単です。LACPグループ内の各ポートは監視され、障害が発生した場合はポートグループから削除されます。ただし、パフォーマンスに関してLACPがどのように機能するかについては、多くの誤解があります。

- LACPでは、エンドポイントと一致するようにスイッチで設定する必要はありません。たとえば、ONTAPにIPベースのロードバランシングを設定し、スイッチにMACベースのロードバランシングを使用することができます。
- LACP接続を使用する各エンドポイントは、パケット送信ポートを個別に選択できますが、受信に使用するポートは選択できません。これは、ONTAPから特定の宛先へのトラフィックが特定のポートに結び付けられ、リターントラフィックが別のインターフェイスに到達する可能性があることを意味します。ただし、これは原因の問題ではありません。
- LACPでは、常にトラフィックが均等に分散されるわけではありません。多数のNFSクライアントを含む大規模な環境では、通常はLACPアグリゲーションのすべてのポートが均等に使用されます。ただし、環境内の1つのNFSファイルシステムの帯域幅は、アグリゲーション全体ではなく、1つのポートの帯域幅に制限されます。
- ONTAPではロビンベースのLACPポリシーを使用できますが、スイッチからホストへの接続には対応していません。たとえば、ホストで4ポートのLACPトランクを、ONTAPで4ポートのLACPトランクを使用する構成でも、ファイルシステムの読み取りには1つのポートしか使用できません。ONTAPは4つのポートすべてを介してデータを送信できますが、4つのポートすべてを介してスイッチからホストに送信するス

イッチテクノロジは現在使用できません。使用されるのは1つだけです。

多数のデータベースホストで構成される大規模な環境で最も一般的なアプローチは、IPロードバランシングを使用して、適切な数の10Gb（またはそれよりも高速）インターフェイスでLACPアグリゲートを構築する方法です。このアプローチにより、ONTAPはクライアントが十分に存在する限り、すべてのポートを均等に使用できます。LACPトランキングでは負荷が動的に再分散されないため、構成内のクライアント数が少なくなるとロードバランシングが機能しません。

接続が確立されると、特定の方向のトラフィックは1つのポートにのみ配置されます。たとえば、あるデータベースがNFSファイルシステムに対してテーブルのフルスキャンを実行し、接続に4ポートのLACPトランクを使用している場合、データの読み取りには1枚のネットワークインターフェイスカード（NIC）のみが使用されます。このような環境にデータベースサーバが3台しかない場合は、3台すべてが同じポートから読み取りを行い、他の3つのポートはアイドル状態になる可能性があります。

物理ポートへの**LIF**のバインド

物理ポートにLIFをバインドすると、ネットワーク構成をきめ細かく制御できるようになります。これは、ONTAPシステム上の特定のIPアドレスは、一度に1つのネットワークポートにのみ関連付けられるためです。フェイルオーバーグループとフェイルオーバーポリシーを設定することで耐障害性が実現します。

フェイルオーバーポリシーとフェイルオーバーグループ

ネットワーク停止時のLIFの動作は、フェイルオーバーポリシーとフェイルオーバーグループによって制御されます。設定オプションは、ONTAPのバージョンによって変更されました。を参照してください "フェイルオーバーグループとポリシーに関するONTAPのネットワーク管理に関するドキュメント" を参照して、導入するONTAPのバージョンの詳細を確認してください。

ONTAP 8.3以降では、ブロードキャストドメインに基づいてLIFのフェイルオーバーを管理できます。そのため、特定のサブネットにアクセスできるすべてのポートを管理者が定義し、ONTAPが適切なフェイルオーバーLIFを選択できるようにすることができます。このアプローチは一部のお客様にも使用できますが、予測性がないため、高速ストレージネットワーク環境では制限があります。たとえば、ファイルシステムへの日常的なアクセスに使用する1Gbポートと、データファイルI/Oに使用する10Gbポートの両方を環境に含めることができます。両方のタイプのポートが同じブロードキャストドメインにあると、LIFのフェイルオーバーによって、データファイルI/Oが10Gbポートから1Gbポートに移動される可能性があります。

要約すると、次の方法を検討してください。

1. ユーザ定義のフェイルオーバーグループを設定します。
2. フェイルオーバーグループにストレージフェイルオーバー（SFO）パートナーコントローラのポートを含め、ストレージフェイルオーバー時にLIFがアグリゲートに従って移動するようにします。これにより、間接トラフィックの作成が回避されます。
3. パフォーマンス特性が元のLIFと一致するフェイルオーバーポートを使用します。たとえば、1つの物理10Gbポート上のLIFには、1つの10Gbポートを含むフェイルオーバーグループを含める必要があります。4ポートLACP LIFは、別の4ポートLACP LIFにフェイルオーバーする必要があります。これらのポートは、ブロードキャストドメインに定義されているポートのサブセットになります。
4. SFOパートナーのみにフェイルオーバーポリシーを設定します。これにより、フェイルオーバー時にLIFがアグリゲートに従うようになります。

自動リバート

を設定します `auto-revert` 必要に応じてパラメータを指定する。ほとんどのお客様は、このパラメータを `true` LIFをホームポートにリバートします。ただし、場合によっては、想定外のフェイルオーバーを調査し

てからLIFをホームポートに戻すように、このパラメータを「false」に設定することもできます。

LIFとボリュームの比率

よくある誤解の1つは、ボリュームとNFS LIFの間には1：1の関係が必要であるということです。この構成は、ボリュームをクラスタ内の任意の場所に移動する際に必要ですが、インターコネクトトラフィックが増えることはありません。ただし、この構成は必須要件ではありません。クラスタ間トラフィックは考慮する必要がありますが、クラスタ間トラフィックが存在するだけでは問題は発生しません。ONTAP用に作成された公開済みのベンチマークの多くには、主に間接I/Oが含まれています。

たとえば、パフォーマンスが重視されるデータベースの数が比較的少なく、合計で40個のボリュームしか必要としないデータベースプロジェクトの場合、ボリューム対LIFの戦略は1：1で、必要なIPアドレスは40個です。これにより、すべてのボリュームを関連付けられたLIFと一緒にクラスタ内の任意の場所に移動でき、トラフィックは常に直接送信されるため、レイテンシのすべてのソースをマイクロ秒レベルでも最小限に抑えることができます。

反対の例として、大規模なホスト環境では、お客様とLIFが1：1の関係にある場合、より簡単に管理できます。時間が経つにつれて、ボリュームを別のノードに移行しなければならない場合があり、間接トラフィックが原因になることがあります。ただし、インターコネクトスイッチのネットワークポートが飽和状態になっていないかぎり、パフォーマンスへの影響は検出されません。懸念がある場合は、ノードを追加して新しいLIFを設定し、次回のメンテナンス時間にホストを更新して、構成から間接トラフィックを取り除くことができます。

TCP/IPおよびイーサネット構成

Oracle on ONTAPをご利用のお客様の多くは、NFS、iSCSI、NVMe/TCPのネットワークプロトコルであるイーサネットを使用しており、特にクラウドを使用しています。

ホストOSの設定

ほとんどのアプリケーションベンダーのドキュメントには、アプリケーションが最適に動作することを確認するためのTCPおよびイーサネットの設定が含まれています。これらの設定は通常、IPベースのストレージパフォーマンスを最適化するのに十分です。

イーサネットフロー制御

このテクノロジを使用すると、クライアントは送信者にデータ転送を一時的に停止するように要求できます。これは通常、受信側が受信データを十分に迅速に処理できないために行われます。一時期、送信者に送信の中止を要求しても、バッファがいっぱいになったために受信者がパケットを破棄するよりも、中断が少なくて済みました。現在OSで使用されているTCPスタックでは、これは当てはまりません。実際、フロー制御は解決するよりも多くの問題を引き起こします。

近年、イーサネットフロー制御に起因するパフォーマンスの問題が増加しています。これは、イーサネットフロー制御が物理レイヤで動作するためです。ネットワーク構成で、任意のホストOSからストレージシステムへのイーサネットフロー制御要求の送信が許可されていると、接続されているすべてのクライアントのI/Oが一時停止します。1台のストレージコントローラで対応するクライアントの数が増えているため、1台以上のクライアントがフロー制御要求を送信する可能性が高くなります。この問題は、OSの仮想化が広範に行われているお客様のサイトで頻繁に発生しています。

NetAppシステム上のNICは、フロー制御要求を受信しないでください。この結果を得る方法は、ネットワークスイッチの製造元によって異なります。ほとんどの場合、イーサネットスイッチのフロー制御は次のように設定できます。 `receive desired` または `receive on`これは、フロー制御要求がストレージコントローラに

転送されないことを意味します。それ以外の場合は、ストレージコントローラのネットワーク接続でフロー制御の無効化が許可されないことがあります。このような場合は、ホストサーバ自体またはホストサーバが接続されているスイッチポートのNIC設定に変更して、フロー制御要求を送信しないようにクライアントを設定する必要があります。

* NetAppでは* NetAppストレージコントローラがイーサネットフロー制御パケットを受信しないようにすることを推奨しています。これは通常、コントローラが接続されているスイッチポートを設定することで実行できますが、一部のスイッチハードウェアには制限があり、代わりにクライアント側の変更が必要になる場合があります。

MTUサイス

ジャンボフレームを使用すると、CPUとネットワークのオーバーヘッドが軽減され、1Gbネットワークのパフォーマンスがある程度向上することが示されていますが、通常はそれほど大きなメリットはありません。

* NetAppでは、可能な限りジャンボフレームを実装することを推奨しています。これは、パフォーマンス上のメリットを実現し解決策、将来のニーズにも対応するためです。

10Gbネットワークではジャンボフレームの使用がほぼ必須です。これは、ほとんどの10Gb環境では、ジャンボフレームを使用しないと10Gbに達する前に1秒あたりのパケット数が制限されるためです。ジャンボフレームを使用すると、OS、サーバ、NIC、およびストレージシステムで処理できるパケットの数は少なくても大きいため、TCP / IP処理の効率が向上します。パフォーマンスの向上はNICによって異なりますが、大幅に向上します。

ジャンボフレームの実装では、接続されているすべてのデバイスでジャンボフレームがサポートされている必要があり、MTUサイズがエンドツーエンドで同じである必要があるという誤った考えがよくあります。代わりに、2つのネットワークエンドポイントは、接続を確立するときに、相互に許容可能な最大フレームサイズをネゴシエートします。一般的な環境では、ネットワークスイッチのMTUサイズは9216、NetAppコントローラは9000、クライアントは9000と1514が混在するように設定されています。MTU 9000をサポートできるクライアントはジャンボフレームを使用でき、1514しかサポートできないクライアントは低い値をネゴシエートできます。

完全にスイッチが接続された環境では、この構成に問題が生じることはほとんどありません。ただし、ルーティングされた環境では、中間ルータが強制的にジャンボフレームをフラグメント化しないように注意してください。

- NetAppでは*次の設定を推奨しています。
- ジャンボフレームの使用を推奨しますが、1Gbイーサネット（GbE）の場合は必須ではありません。
- 10GbE以上の速度で最大のパフォーマンスを実現するには、ジャンボフレームが必要です。

TCPパラメータ

TCPタイムスタンプ、選択的確認応答（SACK）、TCPウィンドウスケーリングの3つの設定が誤って設定されることがよくあります。インターネット上の古いドキュメントの多くは、パフォーマンスを向上させるために、これらのパラメータの1つまたは複数を無効にすることを推奨しています。CPU能力がはるかに低く、TCP処理のオーバーヘッドを可能な限り削減できるというメリットが何年も前にあったこの推奨事項には、いくつかのメリットがありました。

ただし、最新のOSでは、これらのTCP機能のいずれかを無効にしても、通常は検出できるメリットはなく、パフォーマンスも低下する可能性があります。特に仮想ネットワーク環境では、パケット損失やネットワーク品質の変化を効率的に処理するためにこれらの機能が必要になるため、パフォーマンスが低下する可能性があります。

* NetAppでは、ホストでTCPタイムスタンプ、SACK、TCPウィンドウスケーリングを有効にすることを推奨しています。現在のOSでは、これら3つのパラメータはすべてデフォルトでオンにする必要があります。

FC SAN構成

Oracleデータベース用にFC SANを構成する主な目的は、日常的なSANのベストプラクティスに従うことです。

これには、ホストとストレージシステムの間のSANに十分な帯域幅があることを確認したり、必要なすべてのデバイス間にすべてのSANパスが存在することを確認したり、FCスイッチベンダーが必要とするFCポート設定を使用してISLの競合を回避したりするなど、一般的な計画方法が含まれます。 SANファブリックを適切に監視します。

ゾーニング

FCゾーンに複数のイニシエータを含めることはできません。このような配置は最初は機能しているように見えるかもしれませんが、最終的にはイニシエータ間のクロストークがパフォーマンスと安定性の妨げになります。

マルチターゲットゾーンは一般に安全とみなされますが、まれに、ベンダーが異なるFCターゲットポートの動作が問題を引き起こすことがあります。たとえば、NetAppとネットアップ以外のストレージアレイのターゲットポートを同じゾーンに配置することは避けてください。また、NetAppストレージシステムとテープデバイスを同じゾーンに配置すると、原因の問題が発生する可能性がさらに高くなります。

直接接続ネットワーク

ストレージ管理者は、構成からネットワークスイッチを削除してインフラを簡易化したいと考える場合があります。これは一部のシナリオでサポートされます。

iSCSIとNVMe/TCP

iSCSIまたはNVMe/TCPを使用するホストは、ストレージシステムに直接接続して正常に動作することができます。その理由はパス設定です。2つの異なるストレージコントローラに直接接続すると、データフローが2つの独立したパスになります。パス、ポート、またはコントローラが失われても、他のパスの使用が妨げられることはありません。

NFS

直接接続されたNFSストレージも使用できますが、フェイルオーバーには大きな制限があります。スクリプト作成にはお客様の責任が伴います。

直接接続されたNFSストレージで無停止フェイルオーバーが複雑になるのは、ローカルOSで発生するルーティングが原因です。たとえば、ホストのIPアドレスが192.168.1.1/24で、IPアドレスが192.168.1.50/24

のONTAPコントローラに直接接続されているとします。フェールオーバー中、192.168.1.50アドレスはもう一方のコントローラにフェールオーバーでき、ホストが使用できるようになりますが、ホストはそのアドレスの存在をどのように検出しますか。元の192.168.1.1アドレスは、動作中のシステムに接続されていないホストNICに残っています。192.168.1.50宛てのトラフィックは、動作不能なネットワークポートに引き続き送信されます。

2番目のOS NICは19に設定できます。 2.168.1.2およびは、192.168.1.50経由でフェールオーバーされたアドレスと通信できますが、ローカルルーティングテーブルのデフォルトでは、192.168.1.0/24サブネットと通信するために1つの*および1つの*アドレスのみを使用することになります。システム管理者は、失敗したネットワーク接続を検出し、ローカルルーティングテーブルを変更したり、インターフェイスをアップ/ダウンしたりするスクリプトフレームワークを作成できます。正確な手順は、使用しているOSによって異なります。

実際にはNetAppを使用していますが、通常はフェイルオーバー中のIO一時停止が許容されるワークロードのみが対象です。ハードマウントを使用する場合は、一時停止中にIOエラーが発生しないようにしてください。ホスト上のNIC間でIPアドレスを移動するためのフェイルバックまたは手動操作によって、サービスが復元されるまでIOはハングします。

FC直接接続

FCプロトコルを使用してホストをONTAPストレージシステムに直接接続することはできません。その理由はNPIVの使用です。FCネットワークへのONTAP FCポートを識別するWWNは、NPIVと呼ばれる仮想化タイプを使用します。ONTAPシステムに接続されているすべてのデバイスがNPIV WWNを認識できる必要があります。現在、NPIVターゲットをサポートできるホストにインストールできるHBAを提供しているHBAベンダーはありません。

著作権に関する情報

商標に関する情報