

AI/ML アプリケーションのための シスコ データセンター ネットワーキング ブループリント

Contents

概要	3
AI クラスタのトランスポートとしての RoCEv2	4
AI クラスタにはロスレス ネットワークが必要	5
明示的輻輳通知 (ECN)	5
PFC	5
AI/ML クラスタ ネットワークで輻輳を効率的に管理する方法	6
ECN の仕組み	6
PFC の仕組み	8
ECN と PFC を併用してロスレス イーサネット ネットワークを構築する	10
Approximate Fair Drop (AFD) の使用	13
ネットワークの動作を可視化することで、トランスポートとトラブルシューティングがどのように改善されるか	14
Nexusダッシュボードインサイトの使用	14
AI/ML クラスタの最高のパフォーマンスに対応するネットワーク設計	15
Cisco Nexus ダッシュボードファブリック コントローラを使用した AI/ML ネットワークの自動化	17
まとめ	17
関連資料	17

概要

人工知能と機械学習 (AI/ML) アプリケーションは、データセンターでますます一般的になりつつあります。AI のサブセットである機械学習は、最も一般的なアプリケーションの 1 つです。ML は、観察とデータから決定と予測を行うことを学習するコンピューター システムの能力です。現在、広く利用されている GPU アクセラレーション サーバーは、カスタム ディープ ニューラル ネットワークを設計およびトレーニングするための柔軟性を生み出します。より優れたサーバー ハードウェアと、Python や C/C++ などの一般的に使用されるプログラミング言語、および GPU をネイティブに活用するように構築された PyTorch、TensorFlow、JAX などのフレームワークが利用できるようになったことで、GPU アクセラレーション ML アプリケーションの構築が簡素化されました。これらのアプリケーションは、高度な医学研究、コンピューター支援による創薬、自然言語処理、自動運転車、ショッピングの推奨、ビデオ ストリーム内の画像の認識など、多くの目的に使用できます。多くの場合、ML アプリケーションの構築は、複数の反復にわたる大規模なデータセットを使用してディープ ニューラル ネットワークをトレーニングすることから始まります。ニューラル ネットワークは、通常、サーバーごとに複数の GPU を備えた数千の GPU で構成できる GPU クラスタを利用します。多くの場合、これらのサーバーには、厳密なネットワーク要件に従って、2 つの 100Gb ネットワーク インターフェイス カード (NIC) が別々のスイッチに接続されています。

ディープ ラーニング モデルには、raw データから直接学習できる非常に柔軟なアーキテクチャがあります。大規模なデータ セットを使用してディープ ラーニング クラスタをトレーニングすると、予測精度が向上します。予想どおり、これらのアプリケーションは大量のデータを生成し、リアルタイムで収集および処理する必要があり、時には数千に及ぶ複数のデバイス間で共有されます。推論フレームワークは、トレーニングされたニューラル ネットワーク モデルから知識を取得し、それらを新しいデータに適用して結果を予測します。推論クラスタにはさまざまな要件があり、パフォーマンスが最適化されています。ディープ ラーニング システムは、大量のデータを処理して結果を再評価するように最適化されています。推論システムは、より小さなデータ セットを使用する場合がありますが、多くのデバイスにハイパースケールされます。この例は、スマートフォンや自動運転車のアプリケーションです。

上で説明したラーニング サイクルの一部は、非常に大規模なデータ セットの場合、完了するまでに数日から数週間かかる場合があります。ラーニング サイクルに関するサーバー クラスタ間の通信に長い待ち時間が発生したり、パケットがドロップしたりすると、ラーニング ジョブの完了に時間がかかるか、場合によっては失敗します。このため、AI ワークロードには厳しいインフラストラクチャの要件があります。

AI アプリケーションは、低遅延のロスレス ネットワークを利用します。これを実現するには、ネットワーク管理者は、AI アプリケーションのニーズをサポートする構成とともに、適切なハードウェアとソフトウェアの機能を展開する必要があります。AI アプリケーションには、必要に応じて調整できるようにホット スポットを可視化できるネットワークも必要です。最後に、AI アプリケーションは自動化フレームワークを利用して、ネットワーク ファブリック全体が正しく構成され、構成のばらつきがないことを確認する必要があります。

Cisco Nexus 9000 スイッチは、現在利用可能なハードウェアとソフトウェアの機能を備えており、AI/ML アプリケーションの要件を満たす適切な遅延、輻輳管理メカニズム、およびテレメトリを提供します。Cisco Nexus 9000 スイッチは、可視化のための Cisco Nexus Dashboard Insights や自動化のための Nexus Dashboard Fabric Controller などのツールと組み合わせることで、高性能の AI/ML ネットワーク ファブリックを構築するための理想的なプラットフォームになります。

このドキュメントは、出荷されたハードウェアとソフトウェアの機能を使用して AI/ML ワークロードを最高の状態で実行できる最新のネットワーク環境を構築するためのベスト プラクティスの青写真を提供します。このブループリントの構成例が含まれている、[AI/ML アプリケーションのデータセンター ネットワーキング ブループリントの Cisco Validated Design](#) も参照してください。

AI クラスターのトランスポートとしての RoCEv2

リモートダイレクトメモリアクセス (RDMA) は、ハイパフォーマンスコンピューティング (HPC) およびストレージネットワーキング環境で使用されるよく知られたテクノロジーです。RDMA の利点は、CPU に負荷をかけずに、メモリ間レベルで計算ノード間で情報を高スループットで低遅延で転送できることです。この転送機能は、オペレーティングシステムソフトウェアのネットワークスタックをバイパスするために、ネットワークアダプタハードウェアにオフロードされます。この技術には、所要電力が削減されるという利点もあります。

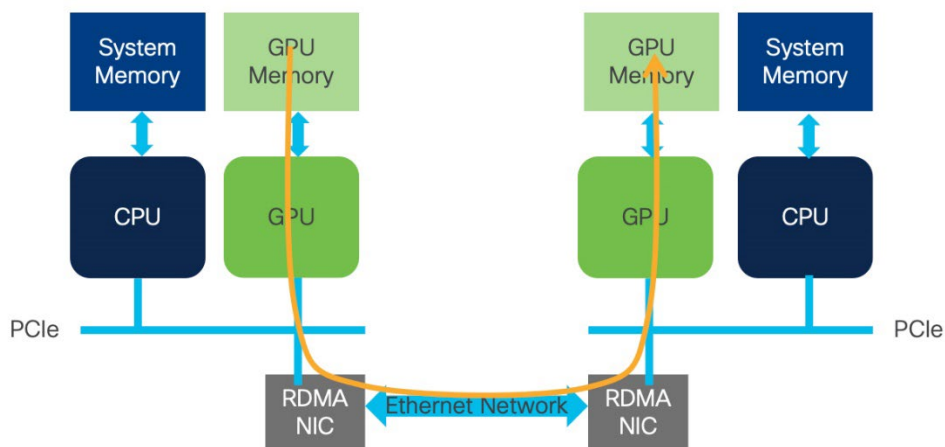


図 1. RDMA を使用した直接 GPU メモリ通信

InfiniBand (IB) は、その最初のアプリケーションで、RDMA のすべての利点を市場にもたらし、高いスループットと CPU バイパスを提供し、待ち時間を減らしました。InfiniBand は、プロトコルに輻輳管理も組み込んでいます。これらの利点により、InfiniBand はハイパフォーマンスコンピューティングトランスポートとして選ばれるようになりました。

HPC ワークロードを必要とするエンタープライズネットワークの場合、InfiniBand は、そのすべての利点を活用するための別個のネットワークの設計につながりました。これらの専用ネットワークは、企業に追加のコストと複雑さをもたらしました。RDMA は、ネットワークトランスポートのいくつかの導入を提供します。それらの 1 つはイーサネットベース、つまり RDMA over Converged Ethernet (RoCE) です。

RoCE は、2010 年に最初に導入された InfiniBand Trade Association (IBTA) 標準であり、同じレイヤ 2 ブロードキャストドメインでのみ機能します。RoCE バージョン 2 は 2014 年に導入され、トラフィックのルーティングを可能にします。RoCE は、イーサネット転送を備えた InfiniBand の拡張機能です。RoCEv2 は、イーサネット、IP、および UDP ヘッダーで IB トランスポートをカプセル化するため、イーサネットネットワーク経由でルーティングできます。

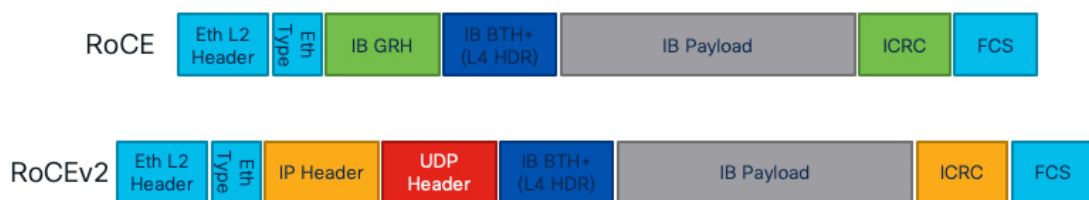


図 2. イーサネット上に RoCEv2 IP および UDP ヘッダーを使用した RoCE および RoCEv2 フレーム形式

イーサネットは、エンタープライズデータセンターのいたるところにあります。ネットワーク管理者はイーサネットに非常に精通しており、これはこのテクノロジーの大きな利点です。それに加えて、RDMA ワークロードとともに通常のエンタープライズトラフィックを伝送する「収束された」ファブリックを手頃な価格で作成

できることは、顧客にとって非常に魅力的です。これが、RoCEv2 がデータセンター ネットワークに実装されている理由の 1 つです。RoCEv2 では、ロスレス トランスポートが必要です。これは、Explicit Congestion Notification (ECN; 明示的輻輳通知) およびプライオリティ フロー制御 (PFC) 輻輳回避アルゴリズムを使用して実現できます。

AI/ML クラスタでは、RDMA を使用して、ネットワーク経由で GPU 間のメモリ間通信を行います。この実装は、GPUDirect RDMA と呼ばれます。RoCEv2 は、GPUDirect RDMA の優れたトランスポートです。

AI クラスタにはロスレス ネットワークが必要

RoCEv2 トランスポートの場合、ネットワークは、輻輳が発生した場合のトラフィックの低下を回避しながら、高スループットと低遅延を提供する必要があります。Cisco Nexus 9000 スイッチは、データセンター ネットワーク向けに構築されており、必要な低遅延を提供します。ASIC あたり最大 25.6Tbps の帯域幅を備えたこれらのスイッチは、RoCEv2 トランスポート上で実行される AI/ML クラスタを満たすために必要な非常に高いスループットを提供します。Cisco Nexus 9000 は、ECN と PFC の両方でソフトウェアとハードウェアのテレメトリを使用して、ロスレス ネットワークのサポートと可視性も提供します。

明示的輻輳通知 (ECN)

輻輳情報をエンドツーエンドで伝達する必要がある状況では、ECN を輻輳管理に使用できます。ECN は、2 つの最下位ビットの IP ヘッダー タイプ オブ サービス (TOS) フィールド内で輻輳が発生しているネットワーク ノードでマークされます。受信側は、ECN 輻輳エクスペリエンス ビットが 0x11 に設定されたパケットを受信すると、輻輳通知パケット (CNP) を生成して送信側に送信します。送信者は輻輳通知を受信すると、通知に一致するフローを遅くします。このエンド ツー エンドのプロセスはデータ パスに組み込まれているため、輻輳を効率的に管理できます。ECN がどのように機能するかの場合は、このドキュメントの後半で提供されます。

ECN ビット	ECN の動作
0x00	非 ECN 対応
0x10	ECN 対応トランスポート (0)
0x01	ECN 対応トランスポート (1)
0x11	輻輳が発生しました

表 1 ネットワーク デバイスとエンド ホストによって使用される ECN ビット値。

PFC

プライオリティ フロー制御は、ロスレス イーサネットを可能にする主要なメカニズムとしてレイヤ 2 ネットワークに導入されました。フロー制御は、レイヤ 2 フレームのサービス クラス (COS) 値によって駆動され、輻輳はポーズ フレームとポーズ メカニズムを使用して通知および管理されます。ただし、スケーラブルなレイヤ 2 ネットワークを構築することは、ネットワーク管理者にとって困難な作業になる場合があります。このため、ネットワーク設計は、ほとんどがレイヤ 3 のルーテッド ファブリックに進化してきました。

RoCEv2 をルーティングできるため、PFC は、差別化サービス コードポイント (DSCP) の優先度と連動して、ネットワーク内のルーティングされたホップ間の輻輳を通知するように調整されました。DSCP は、IP ネットワーク上のネットワーク トラフィックを分類するために使用されるメカニズムです。パケット分類の目的で、IP ヘッダーの 6 ビットの差別化サービス フィールドを使用します。レイヤ 3 マーキングを使用すると、トラフィックはルータ間で分類セマンティクスを維持できます。PFC フレームはリンク ローカル アドレス指定を使用するため、ネットワーク デバイスはルーテッド トラフィックとスイッチド トラフィックの両方に対してポーズ シグナリングを受信して実行できます。PFC は、輻輳の場所からトラフィックの送信元まで、ホップごと

に送信されます。この段階的な動作は、送信元に反映されるまでに時間がかかる場合があります。PFC は、RoCEv2 トランスポートの輻輳を管理するための主要なツールとして使用されます。PFC がどのように機能するかは、このドキュメントの後で提供されます。

Cisco Nexus 9000 スイッチは、PFC 輻輳管理と ECN マーキングの両方をサポートし、重み付けランダム早期検出 (WRED) または近似公平ドロップ (AFD) のいずれかを使用して、ネットワーク ノードの輻輳を示します。

AI/ML クラスタ ネットワークで輻輳を効率的に管理する方法

PFC と ECN は相互に補完して、最も効率的な輻輳管理を提供します。一緒に、それらは輻輳時に最高のスループットと最低の遅延ペナルティを提供します。これらはそれぞれ、ロスレスイーサネット ネットワークを構築する上で重要な役割を果たします。それらの補完的な役割を理解するために、まず ECN シグナリングがどのように機能するかを見て、次に 2 層 (スパイン スイッチとリーフ スイッチ) ネットワークでの PFC 輻輳制御の例を見ていきます。

ECN と PFC は、サービスの質 (QoS) 構成を通じて、システム内でエンドツーエンドで有効にする必要があります。エンドホストとネットワーク ノードの両方が ECN と PFC に参加して、ロスレストラフィック機能を有効にする必要があります。

ECN の仕組み

ECN は、2 つの ECN 対応エンドポイント間で使用される機能です。Nexus 9000 スイッチは、ネットワークが輻輳した場合に ECN ビットでパケットをマークできます。ECN 対応ネットワーク ノードは、輻輳回避アルゴリズムを使用して、使用されているキューの量をチェックし、指定されたしきい値に達すると、輻輳の原因となっているトラフィックをマークします。この例では、重み付けランダム早期検出 (WRED) を使用して、輻輳を示し、ECN ビットでトラフィックをマークします。

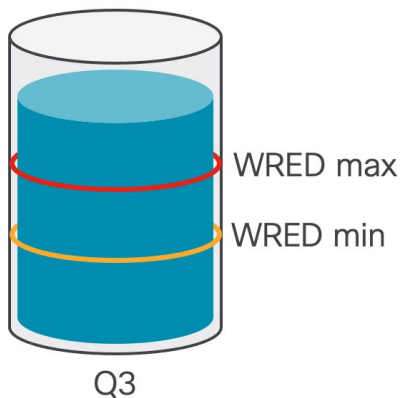


図 3. キューに設定された WRED の最小 (橙) および最大 (赤) のしきい値

Cisco Nexus 9000 の WRED は、キューごとのレベルで実行されます。キューには、2 つのしきい値が設定されています。WRED の最小しきい値は、バッファ使用率が低く、拡大する可能性のある軽度の輻輳を示しています。バッファ使用率が増加し続けると、最小しきい値に達すると、WRED はキューから出る発信パケットの量をマークします。パケットの数は、WRED 設定のドロップ確率値に依存します。Cisco Nexus 9000 では、これはすべての発信パケットの割合 (%) として表されます。たとえば、ドロップ確率パラメータが 10 に設定されている場合、すべての発信パケットの 10% がマークされることを意味します。これで輻輳が緩和されない場合は、キュー内のバッファ使用率が増大し、WRED の最大しきい値に達します。WRED の最大しきい値を超えると、スイッチはキューのすべての発信パケットに ECN をマークします。この WRED ECN メカニズムにより、エンドポイントは輻輳について学習し、輻輳に対応できます。これについては、次の例で説明します。

この例では、2層ネットワークがあり、ホスト A と B がホスト X にデータを送信しています。

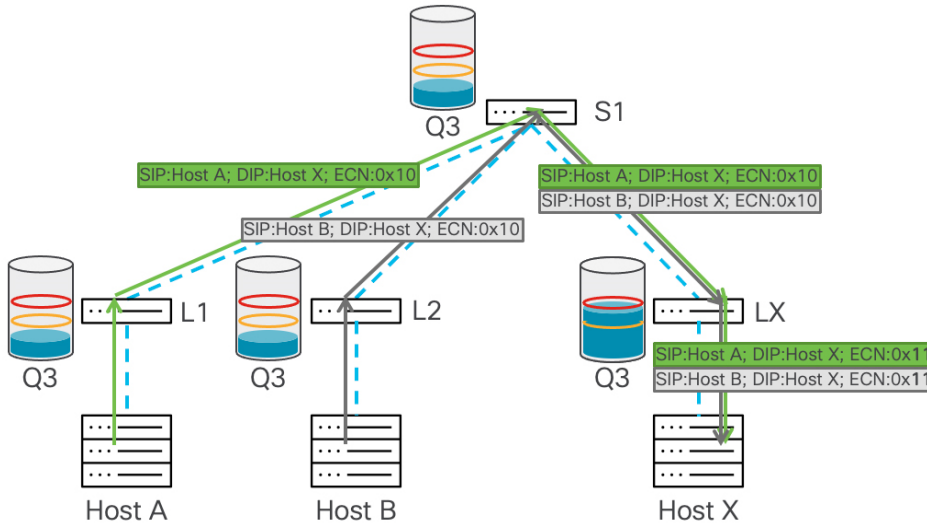


図 4. バッファ占有率が WRED の最小しきい値を超え、リーフ X が IP ヘッダーで ECN をマーク

輻輳はリーフ X で発生します。これは、スパインスイッチからの帯域幅がホスト X よりも多く、ホスト X に接続されているポートがオーバーサブスクライブされているためです。バッファ使用量はリーフ X で蓄積し始めます。バッファが WRED の最小しきい値に達すると、リーフスイッチは、データパスの輻輳を示すために、いくつかのパケットの ECN フィールドに 0x11 値でマーキングを開始します。

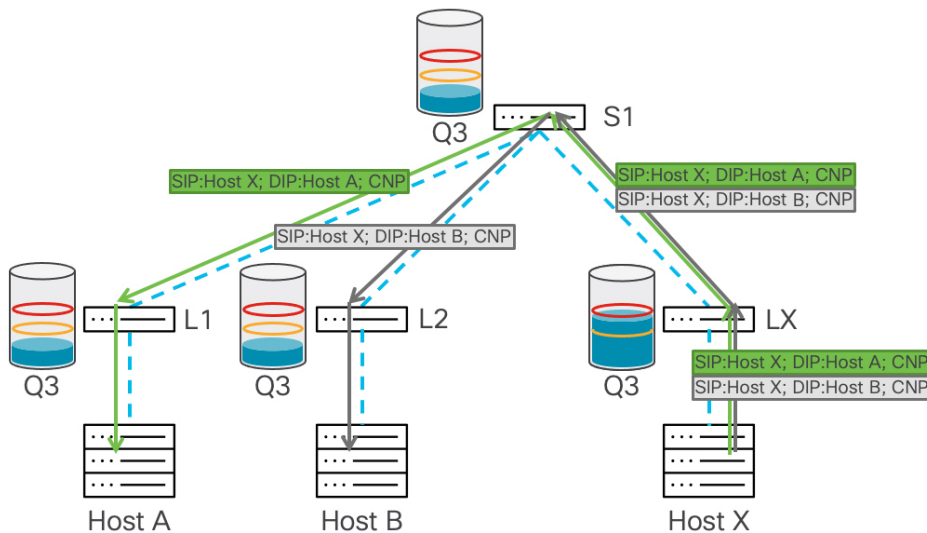


図 5. ホスト X は、CNP パケットを送信して、ネットワークの輻輳についてホスト A と B に通知します。

宛先 (ホスト X) は、マークされたパケットを受信すると、データパスで輻輳が発生していることを学習し、現在マークされているパケットを受信した送信元ホストに CNP パケットを生成します。

一部のデータパケットだけが輻輳が発生したビットでマークされているため、送信元はそのフローのトラフィックスルーputを削減し、パケットの送信を継続します。輻輳が続き、バッファ使用量が WRED の最大しきい値を超えると、スイッチはすべてのパケットに輻輳が発生したビットをマークします。これは、送信者が多くの CNP パケットを受信することを意味し、そのアルゴリズムに基づいて、接続先へのデータ転送速度を大幅に下げることがあります。これにより輻輳が軽減され、バッファの排出が開始されます。そうなれば、次に渋滞が通知されるまで、トラフィックレートは上昇するはずですが。

PFC の仕組み

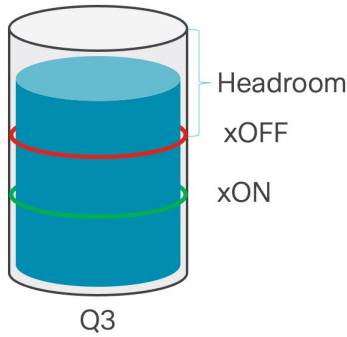


図 6. プライオリティフロー制御 **no-drop** キュー、**xOFF** (赤) および **xON** (緑) のしきい値、およびバッファヘッドルーム
Cisco Nexus 9000 スイッチで PFC が有効になっている場合、サービスクラスはロスレス トランスポート専用になります。このクラスのトラフィックは、他のクラスのトラフィックとは異なる方法で処理されます。PFC が構成された Cisco Nexus 9000 スイッチのポートには、専用の **no-drop** キューとそのキュー専用のバッファが割り当てられます。

ロスレス機能を提供するために、キューには 2 つのしきい値があります。**xOFF** しきい値はバッファで高く設定されており、これは、PFC フレームが生成され、トラフィックの送信元に向けて送信されるバッファ使用率のポイントです。バッファの排出が始まり、**xON** しきい値を下回ると、ここでポーズフレームが停止し、送信側に送信されなくなります。これは、システムが輻輳が終了したと判断した時点です。

次の例は、PFC が輻輳を管理する方法を示しています。リーフ X は、ホスト A および B から着信するトラフィックを受信します。これにより、ホスト X に向かうポートで輻輳が発生します。この時点で、スイッチは専用バッファを使用して着信トラフィックを吸収します。トラフィックはリーフ X によってバッファリングされ、**xOFF** しきい値に達すると、ポーズフレームがアップストリーム ホップ（この図ではスパインスイッチ S1）に送信されます。

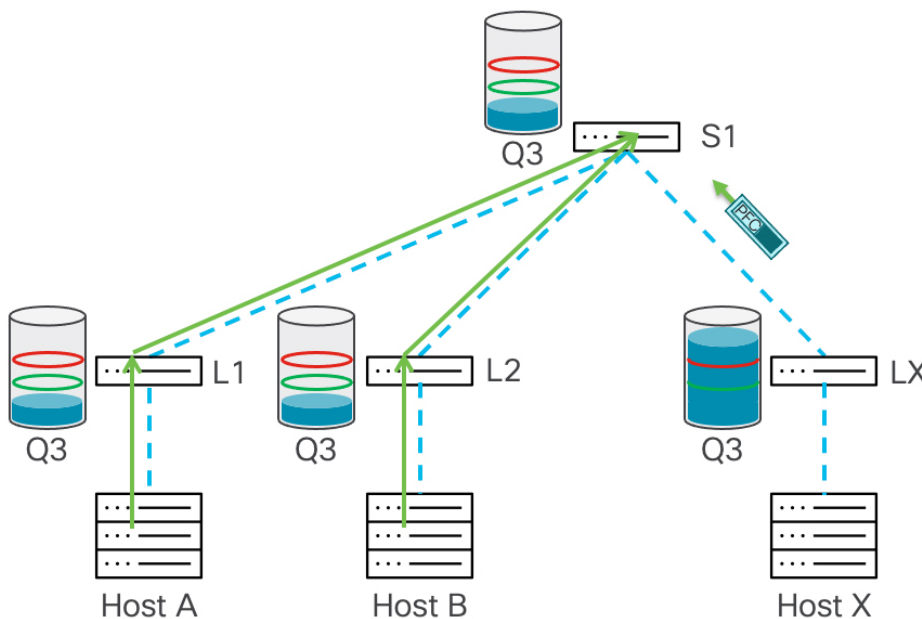


図 7. リーフ X からスパイン 1 への PFC のシグナリング、スパインスイッチからリーフスイッチへのトラフィックの一時停止

スパインスイッチ S1 がポーズフレームを受信すると、S1 はトラフィックの送信を停止します。これにより、リーフ X でのさらなる輻輳が防止されます。同時に、S1 は no-drop キューでのトラフィックのバッファリングを開始し、バッファが xOFF しきい値に達した後、両方がスパインスイッチにトラフィックを送信しているため、ポーズフレームをアップストリームデバイス、リーフ 1 およびリーフ 2 に送信します。

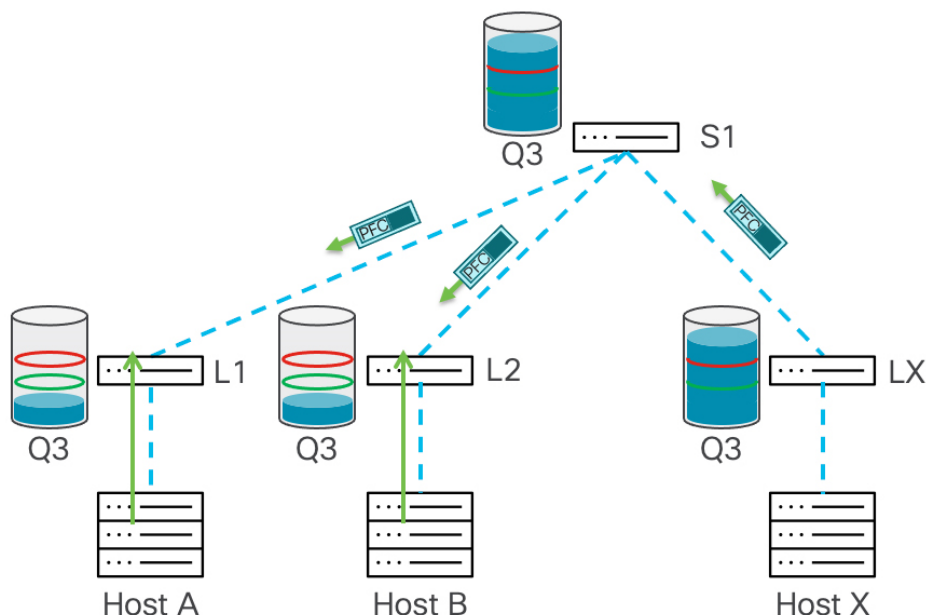


図 8. スパイン 1 からリーフスイッチ 1 および 2 への PFC シグナリングにより、リーフ 1 およびリーフ 2 からスパインスイッチへのトラフィックが一時停止

このポーズ動作のホップごとの伝達は継続し、リーフスイッチはポーズフレームを受信し、スイッチはスパインスイッチへの送信を一時停止し、トラフィックのバッファリングを開始します。これにより、リーフ 1 と 2 がそれぞれ Host A と B にポーズフレームを送信します。

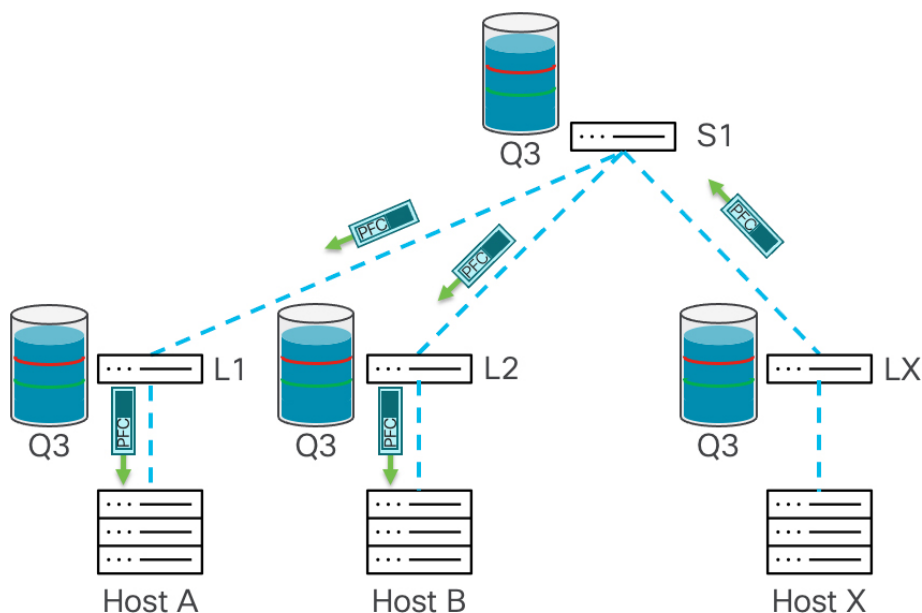


図 9. PFC はリーフ スイッチ 1 および 2 からホスト A および B に信号で送信されます。ホストからのトラフィックの一時停止は PFC の最終ホップであり、これにより輻輳を緩和するためにホストが一時停止されます。

送信側がポーズ フレームを受信すると、ストリームの速度が低下します。これにより、ネットワーク内のすべてのスイッチでバッファを空にすることができます。各デバイスが xON しきい値に達すると、そのデバイスはポーズ フレームの伝達を停止します。輻輳が緩和され、ポーズ フレームが送信されなくなった後、ホスト X はトラフィックの受信を再開します。

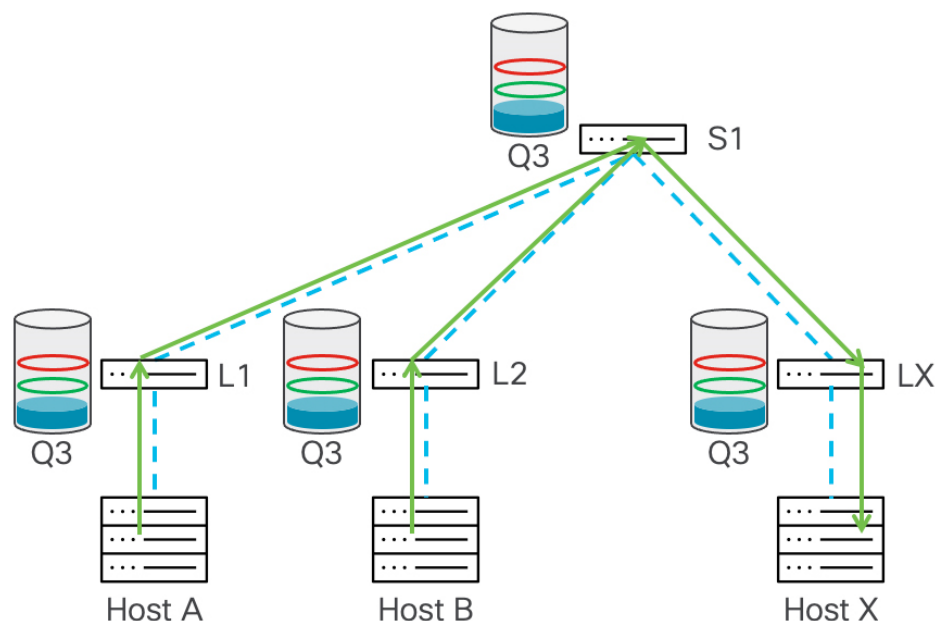


図 10. バッファが xON しきい値を下回ると、トラフィックはホスト A および B からホスト X に復元されます
この一時停止プロセスは、ネットワークで輻輳が発生するたびに、または PFC が有効になっているエンドポイントで輻輳ポイントから送信者まで繰り返されます。例に示すように、このプロセスでは、パス内のすべてのデバイスが送信を停止する前にポーズ フレームを受信する必要があります。このプロセスは、パケットのドロップを防ぐために使用されます。

まれに、動作不良のホストが継続的に PFC フレームを送信すると、PFC ストームが発生することがあります。この動作により、すべてのネットワーク ノードのバッファが飽和する可能性があり、すべてのエンド ホストに到達すると、ネットワークが完全に停止する可能性があります。PFC ストームを防止するには、PFC ウォッチドッグ機能を使用する必要があります。PFC ウォッチドッグ間隔は、no-drop キュー内のパケットが指定された時間内にドレインされているかどうかを検出するように構成できます。期間が経過すると、ネットワークの PFC デッドロックを回避するために、ドレインされていない PFC キューと一致するすべての発信パケットがドロップされます。

ECN と PFC を併用してロスレス イーサネット ネットワークを構築する

前の例で示したように、ECN と PFC はどちらもそれ自体で輻輳をうまく管理できます。併用することで、さらに効果を高めることができます。ECN は、輻輳を軽減するために最初に反応することができます。ECN が十分に速く反応せず、バッファ使用率が増加し続ける場合、PFC はフェイルセーフとして動作し、トラフィックのドロップを防ぎます。これは、輻輳を管理し、ロスレス イーサネット ネットワークを構築するための最も効率的な方法です。

PFC と ECN が一緒に輻輳を管理するこの協調プロセスは、Data Center Quantized Congestion Notification (DCQCN) と呼ばれ、RoCE ネットワーク用に開発されました。

PFC と ECN は連携して、効率的なエンドツーエンドの輻輳管理を提供します。システムでバッファ使用量が中程度の軽度の輻輳が発生している場合、ECN を使用した WRED は輻輳をシームレスに管理します。輻輳がより深刻な場合、またはマイクロバーストによってバッファの使用率が高くなるのが原因である場合、PFC がトリガーされ、その輻輳が管理されます。WRED と ECN の両方が説明どおりに機能するには、適切なしきい値を設定する必要があります。次の例では、WRED の最小しきい値と最大しきい値は、バッファ使用率を低くして最初に輻輳を軽減するように設定され、PFC しきい値は、ECN 後の輻輳を軽減するセーフティ ネットとして高く設定されています。ECN と PFC は no-drop のキューで動作し、ロスレス トランスポートを提供します。

この例では、ホスト A と B の両方がホスト X にトラフィックを送信します。リーフアップリンクは、すべてのトラフィックがリーフ X に到着するのに十分な帯域幅を提供するため、輻輳ポイントはホスト X への発信インターフェイス上にあります。

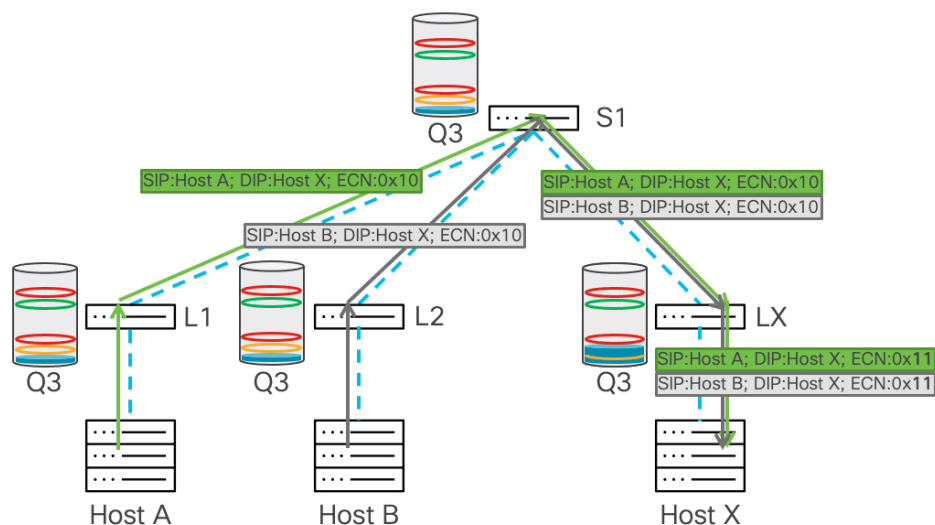


図 11. リーフ X の輻輳により、一部のパケットが ECN でマークされるようにトリガーします

この図では、WRED ECN と PFC の両方が、ネットワーク内のすべてのスイッチの no-drop キューに設定されています。リーフ X は、WRED の最小しきい値を超えるバッファの蓄積を経験し、スイッチは ECN ビットで IP ヘッダーをマークします。

トラフィックは、リーフ X の最小 WRED しきい値に達します。リーフ X の WRED は、ECN 0x11 ビットでトラフィックをマーキングすることによって反応します。ホスト X はこのトラフィックを受信すると、CNP パケットをホスト A と B に送り返します。

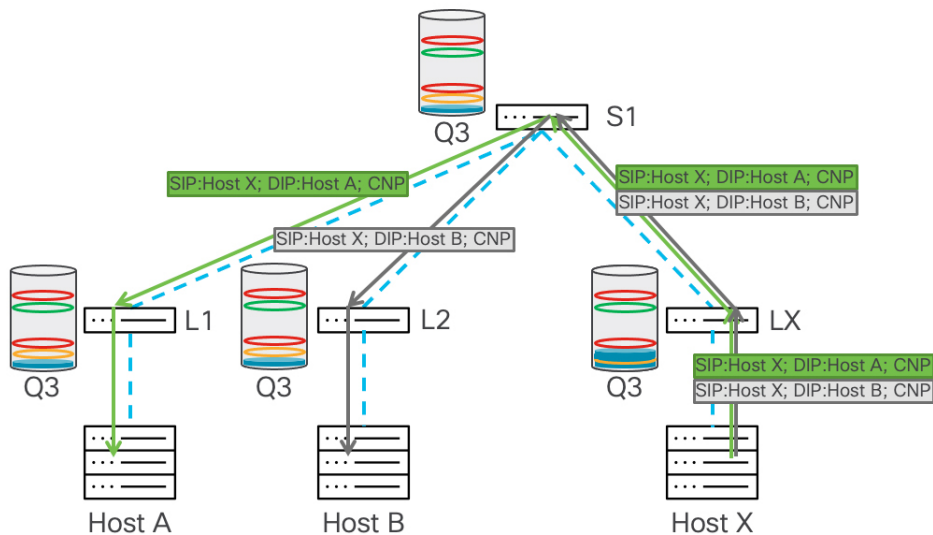


図 12. ホスト X は、**CNP** パケットを送信することによって、ホスト A と B に輻輳があることを示します。WRED ECN では不十分な場合があります、PFC しきい値を高くすると、輻輳をさらに軽減できます。トラフィックは引き続き複数のホストから来ており、前の例で説明したように **ECN** を使用した **WRED** が使用されていますが、バッファの使用量は **xOFF** しきい値に達するまで増加し続けます。この時点で、スイッチは送信側に向けてポーズ フレームを生成します。この例では、これがスパイン スイッチに送信されます。

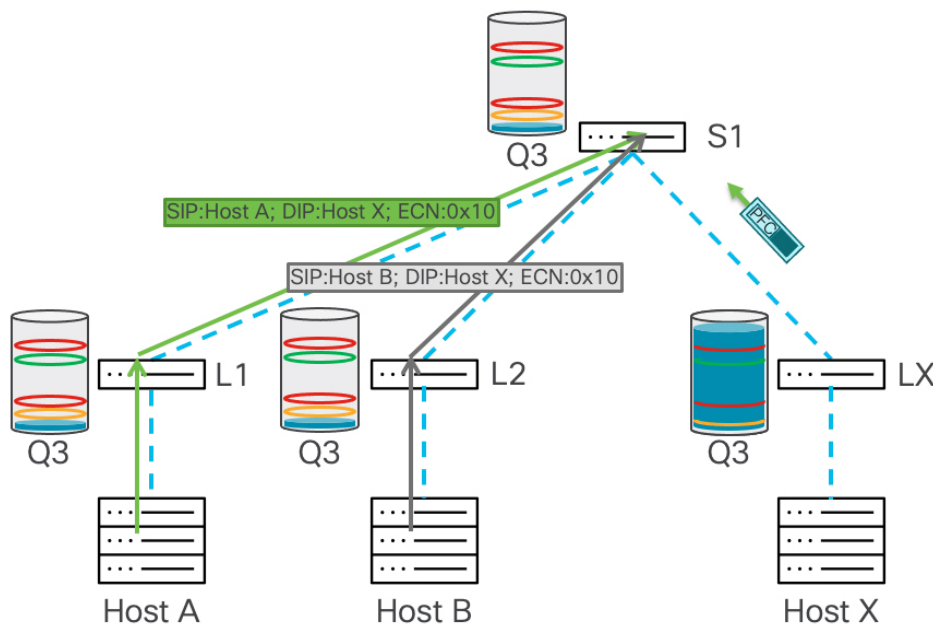


図 13. リーフ X のキューが蓄積し続けると、リーフ スイッチはスパイン 1 への **PFC** フレームを生成し、ホスト X へのトラフィックの送信を停止します。パケット ドロップを防ぐために、**PFC** はスパイン スイッチからリーフ X へのトラフィックを遅くします。これにより、リーフ スイッチがトラフィックをドロップしないようにします。これにより、スパイン スイッチでさらにバッファが使用されます。スパイン スイッチでは、**xOFF** しきい値に達するまでバッファ使用率が増加します。これにより、スパイン スイッチから送信者が接続されているリーフ スイッチまでの **PFC** フレームがトリガーされます。

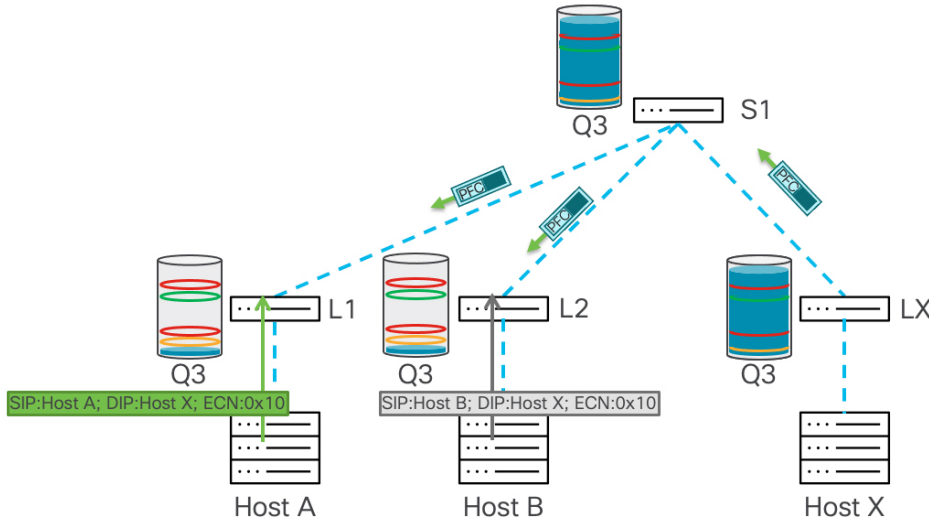


図 14. スパイン 1 は、リーフ 1 および 2 への PFC フレームを生成します。
 各送信側に最も近いリーフスイッチが PFC ポーズフレームを受信し、トラフィックのバッファリングを開始します。バッファは構築を開始し、WRED のしきい値トラフィックを超えた後、バッファは ECN でマークされますが、前の例と同様に構築を続けます。リーフスイッチの xOFF しきい値に達すると、システムは送信側に向けてポーズフレームを生成します。これにより、送信側からのレートがさらに低下し、パケットのドロップが防止されます。

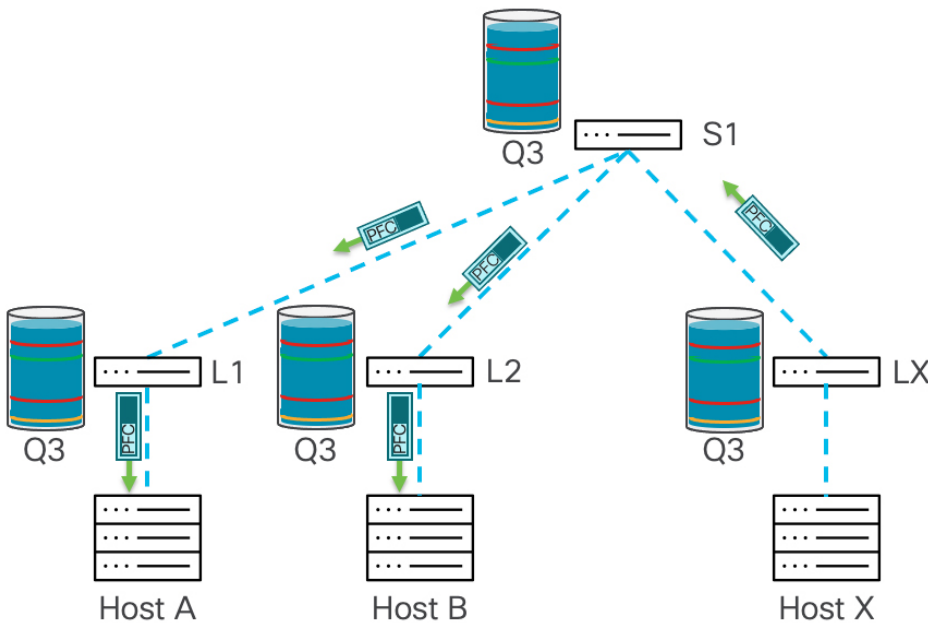


図 15. リーフ 1 とリーフ 2 は PFC フレームをホスト A と B に送信し、輻輳を k 軽減します
 前述のように、ECN と PFC を一緒に使用することは、このグループプリントで推奨されるアプローチです。次のセクションでは、近似フェアドロップ (AFD) を使用して検討します。

Approximate Fair Drop (AFD) の使用

輻輳を管理する別の方法は、高度な QoS アルゴリズムを使用することです。Cisco Nexus 9000 スイッチには、近似フェアドロップ (AFD) などのインテリジェントなバッファ機能が付属しています。AFD を使用して、高帯域幅 (エレファントフロー) を短期間および低帯域幅フロー (マウスフロー) と区別できます。AFD

がエレファント フローを構成するトラフィックに関する情報を取得した後、AFD は ECN ビットに 0x11 値をマークできますが、これは高帯域幅フローに対してのみです。フローで使用される帯域幅に基づいて、異なる数のパケットが ECN でマークされます。たとえば、1G で実行されているフローには、10G で実行されているフローよりも 0x11 ECN ビットでマークされたパケットが少ないため、AFD はフローのサイズに比例して異なる数のマークされたパケットをトリガーします。エンド ホストで適切なアルゴリズムを使用すると、この方法でマーキングを使用すると、システム全体の輻輳の原因となるフローを効率的に減速させることができます。このようにして、パフォーマンスは最小のレイテンシーに合わせて最適化されます。WRED に対する AFD の利点は、最も輻輳を引き起こしているフローのセットを区別できることです。WRED は、キュー内のすべてのトラフィックを均等にマークします。AFD はより細かく、より高い帯域幅のエレファント フローのみをマークし、マウス フローをマークしないままにして、ペナルティを課したり遅くしたりしないようにします。

AI クラスタでは、データの長い転送とその結果として生じる輻輳が速度を低下させないようにすることで、継続期間の短い通信を最後まで実行できるようにすることが有利です。パケット ドロップは引き続き回避されますが、システムは象のフローを区別して処理を遅くするだけなので、多くのトランザクションはより速く完了します。

Cisco Nexus 9000 シリーズ スイッチでのインテリジェント バッファ管理の詳細については、『[Intelligent Buffer Management on Cisco Nexus 9000 Series Switches ホワイトペーパー](#)』を参照してください。

ネットワークの動作を可視化することで、トランスポートとトラブルシューティングがどのように改善されるか

ネットワーク管理者が AI/ML ネットワークを最適化して最高のパフォーマンスを実現し、問題がサービスに影響を与える前に予測できるようにするには、ネットワーク システムが輻輳管理アルゴリズムとネットワーク システム全体の健康に対して深いレベルの可視性を提供することが不可欠です。Cisco Nexus 9000 スイッチには、ネットワークの問題を関連付け、RoCEv2 トランスポート向けに最適化するために使用できる強力なテレメトリ機能が組み込まれています。

Cisco Nexus 9000 ファミリのスイッチは、フロー テーブルおよびフロー テーブル イベントを通じてハードウェア フロー テレメトリ情報を提供します。これらの機能により、スイッチを通過するすべてのパケットを説明し、観察し、マイクロバーストやパケット ドロップなどの動作と関連付けることができます。このデータを Cisco Nexus Dashboard Insights にエクスポートして、デバイスごと、インターフェイスごと、フローごとの粒度までデータを表示できます。

Nexus ダッシュボードインサイトの使用

Cisco Nexus Dashboard Insights は、デバイスごと、インターフェイスごと、およびフロー レベルで ECN マーク カウンターを提供できます。さらに、サービス レベルのクラスごとにスイッチによって発行または受信された PFC パケットに関する情報をレポートできます。この情報により、ネットワーク管理者はリアルタイムのネットワーク輻輳統計を観察し、それらを使用してネットワークを調整し、輻輳への対応を改善することができます。

Cisco Nexus Dashboard Insights によって提供される詳細な可視性により、ネットワーク管理者はドロップを観察し、通常のトラフィック条件でドロップが停止するまで WRED または AFD のしきい値を調整できます。これは、AI/ML ネットワークが定期的な交通渋滞の発生に効果的に対処できるようにするための最初で最も重要なステップです。多くのサーバが単一の宛先と通信するマイクロバースト状態の場合、ネットワーク管理者はカウンタ データを使用して、完全にロスレスな動作が有効になるように PFC とともに WRED または AFD のしきい値を調整できます。ドロップが防止されたら、ECN マーキングと PFC RX/TX カウンターのレポートを使用して、システムをさらに調整して最高のパフォーマンスを実現できます。

Cisco Nexus Dashboard Insights の Operational Intelligence Engine には、一連の高度なアラート、ベースライン、相関、および予測アルゴリズムが組み込まれており、ネットワークおよびコンピューティング コンポーネントから取得したテレメトリデータを利用して、ネットワークの動作を詳細に把握することができます。Cisco Nexus Dashboard Insights は、トラブルシューティングを自動化し、迅速な根本原因の特定と早期修復を支援します。統合されたネットワーク リポジトリとコンプライアンス ルールにより、ネットワークの状態がオペレーターの意図に同調した状態に保たれます。

Cisco Nexus Dashboard Insights の詳細については、[Cisco Nexus Dashboard Insights for the Data Center Data Sheet](#) を参照してください。

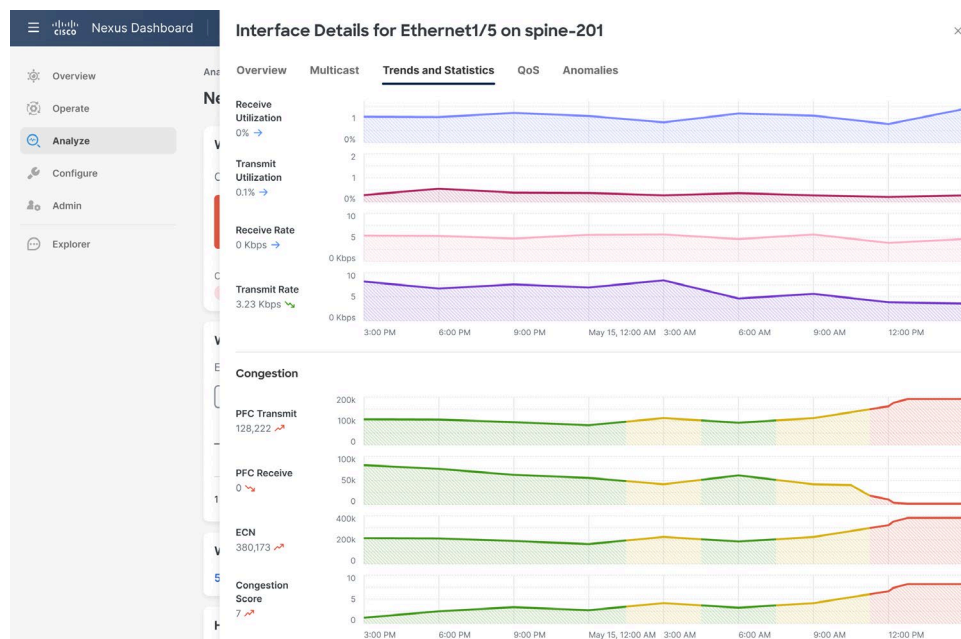


図 16. Cisco Nexus Dashboard Insights は、ECN でマークされたパケット、および PFC が送受信したポーズフレームを追跡します

AI/ML クラスタの最高のパフォーマンスに対応するネットワーク設計

ネットワーク設計は、AI/ML クラスタの全体的なパフォーマンスに大きな影響を与えます。このドキュメントで前述した輻輳管理ツールに加えて、ネットワーク設計では、GPU ワークロードが必要とするすべてのスループットに対応する非ブロッキング ファブリックを提供する必要があります。これにより、輻輳管理アルゴリズムに必要な作業が減り、AI/ML ジョブをより迅速に完了することができます。

ノンブロッキング ネットワークを構築する方法はいくつかありますが、2 層のスパイン-スイッチ-リーフ-スイッチ設計により、遅延とスケラビリティが最も低くなります。これを説明するために、1024 GPU を備えた GPU クラスタと、GPU で処理する必要のあるデータを保持するストレージ デバイスを作成したい会社の例を使用します。

この例では、サーバーごとに 8 つの GPU があり、ネットワーク I/O はサーバーごとに 100Gbps の 2 つのポートです。それぞれ 2x100G ポートを備えた 128 台のサーバーに対応するには、アクセス レイヤに 256 x 100G ポートが必要です。この例では、低遅延が不可欠であるため、Cisco Nexus 9300 スイッチで構成されるスパイン/リーフ スイッチ ネットワークが推奨されます。ノンブロッキング ネットワークにするには、スパイン スイッチからのアップリンクに、フロント パネルのサーバー側ポートと同じ帯域幅容量が必要です。リーフ (アクセス) レイヤの要件に対応するには、Cisco Nexus 93600CD-GX スイッチが最適です。

Cisco Nexus 93600CD-GX スイッチには、サーバーポートとして使用できる 100G の 28 ポートと 400G の 8 つのアップリンクがあります。このダウンリンクポートとアップリンクポートの集合により、これはノンブロッキングスイッチになります。256 個のポートすべてを接続するには、10 個のリーフスイッチが必要です。サーバーは、ネットワークの冗長性を提供するために 2 つの別個のリーフスイッチにデュアルホーム接続されます。この設計により、ストレージデバイスまたはストレージクラスタに接続し、この AI/ML サーバークラスタをエンタープライズネットワークの他の部分に接続するために使用可能なポートが確保されます。

リーフスイッチからの帯域幅に対応するには、80x400G ポートが必要です。冗長性の理由から、2 つのスパインスイッチを選択できますが、AI/ML ワークロードにはスケールと回復力が重要であるため、システムは 4 つのスパインスイッチで構築されます。このネットワークでは、スパインスイッチとして Cisco Nexus 9332D-GX2B スイッチが選択されています。スパインスイッチは、それぞれ 20 個の 400G ポートに接続します。これにより、各スパインスイッチで 12 個のポートが空くので、ネットワークの非ブロッキングの側面を危険にさらすことなく、リーフスイッチを追加してこの環境を拡張できます。ネットワークは次の図で表されます。

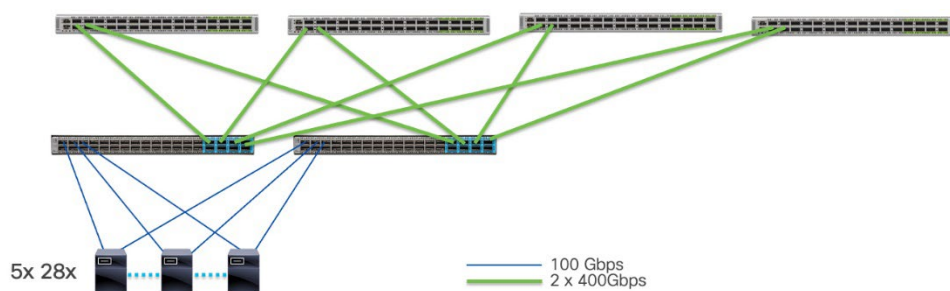


図 17. 1024 GPU をノンブロッキングネットワークで接続するスパインリーフネットワーク設計

ノンブロッキングネットワークを構築しました。ただし、最高のパフォーマンスに対応し、ロスレスイーサネットファブリックを保証するために、輻輳管理アルゴリズムを引き続き使用する必要があります。ノンブロッキングネットワークでも、2 台のサーバーが 1 台のサーバーに回線速度で送信する場合など、輻輳が発生する可能性があります。このドキュメントで前述したように、ECN と PFC の両方を輻輳管理ツールとして使用する必要があります。

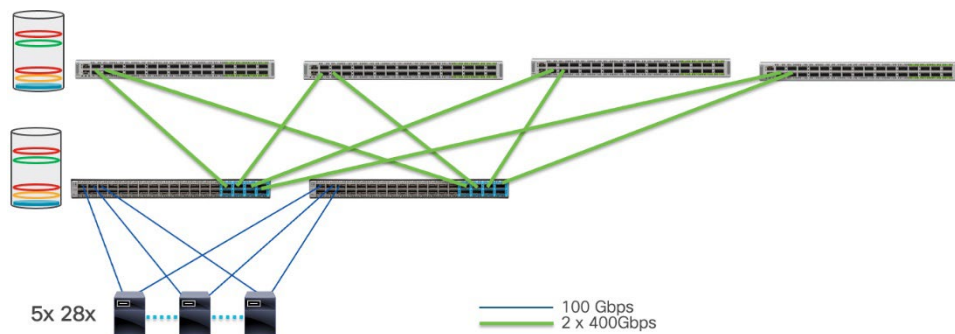


図 18. スパインリーフネットワーク設計、PFC および WRED ECN を使用したリーフレイヤおよびスパインレイヤで有効な輻輳管理

Cisco Nexus 93600CD-GX リーフスイッチおよび Cisco Nexus 9332D-GX2B スパインスイッチの遅延は 1.5 マイクロ秒です。非常に遅延の影響を受けやすいマシンを同じリーフスイッチのペアに接続できます。このネットワークファブリックのエンドツーエンドの最大遅延は、トラフィックがリーフスイッチとスパインスイッチの両方を通過して宛先に到達する必要がある場合、約 4.5 マイクロ秒です。ほとんどの場合、輻輳は WRED ECN によって管理されるため、エンドポイントの RoCEv2 トランスポートによって提供される遅延を維持できます。

上記のネットワーク例は、リーフ スイッチを追加することで簡単に拡張できます。また、64 X 400G ポートを備えた Cisco Nexus 9364D-GX2A スパイン スイッチを使用するか、ノンブロッキング ファブリックを維持するためにスパイン スイッチを追加することで、スパイン キャパシティを簡単に 2 倍にすることができます。最後に、3 層 (スーパー スパイン タイプ) 設計を使用して、複数のノンブロッキング ネットワーク ファブリックを相互接続できます。

最も単純な反復では、このネットワークは AI/ML ワークロード専用であり、レイヤ 3 リーフ スイッチへのコントロール プレーンとして BGP を実行する単純な超スケーラブル データ センター (MSDC) ネットワーク設計原則を念頭に置いて構築されています。多くのエンタープライズ環境など、このネットワークが複数のテナントと機能に対応する必要がある場合は、MP-BGP EVPN VXLAN ネットワークを利用できます。VXLAN ネットワークにより、テナント間のネットワーク分離が可能になります。このドキュメントで説明されている設計原則は、単純なレイヤ 3 または VXLAN 設計でうまく機能します。

MSDC の設計とベスト プラクティスの詳細については、[シスコの大規模なスケーラブルなデータ センター ネットワーク ファブリック ホワイト ペーパー](#)を参照してください。

Cisco Nexus ダッシュボード ファブリック コントローラを使用した AI/ML ネットワークの自動化

ネットワーク アーキテクチャの選択 (MSDC または VXLAN) に関係なく、Cisco Nexus ダッシュボード ファブリック コントローラ (ファブリック コントローラ サービスとも呼ばれる) は、ベスト プラクティスの構成と自動化機能を提供できます。ここでは、ネットワークは PFC と ECN の QoS 構成を含めて、数分で構成されます。ファブリック コントローラ サービスは、新しいリーフ スイッチまたはスパイン スイッチを追加し、アクセス ポート構成を変更するための自動化も提供します。

ファブリック コントローラ サービスの機能の詳細については、[Cisco Nexus Dashboard ファブリック コントローラ 12 データ シート](#)を参照してください。

まとめ

AI/ML クラスタには、厳しいインフラストラクチャ要件があります。ネットワークは、大規模な AI/ML ジョブをより迅速に完了するための重要な機能を果たし、正しく設計されていれば、高遅延やパケット ドロップが原因で大規模な AI/ML ジョブが失敗するリスクを軽減します。

AI/ML クラスタはイーサネットを効率的に使用できるため、イーサネットを活用して、RoCEv2 をトランスポートとして使用するトラフィックの低レイテンシーと高スループットを提供します。RoCEv2 に必要なロスレスイーサネット ネットワークを構築するには、このドキュメントで説明されているように、輻輳管理ツールを使用する必要があります。このブループリントで説明されている設計に加えて、Nexus Dashboard Insights などの監視ツールを使用してネットワーク ファブリックの動作を観察し、それに応じて調整して、可能な限り最高のパフォーマンスを提供することが重要です。

Cisco Nexus 9000 スイッチは、最新の高性能 AI/ML ネットワーク ファブリックを構築するための適切な輻輳管理メカニズム、テレメトリ機能、ポート速度、および遅延を提供するために、現在利用可能なハードウェアおよびソフトウェア機能を備えています。

関連資料

- [NX-OS VXLAN ファブリックを介した RoCE ストレージの実装](#)
- [データ センター データ シート向け Cisco Nexus Dashboard Insights](#)
- [Cisco Nexus Dashboard ファブリック コントローラ 12 データ シート](#)

米国本社
Cisco Systems, Inc.
サンノゼ (カリフォルニア州)

アジア太平洋本社
Cisco Systems (USA) Pte. Ltd.
Singapore

ヨーロッパ本社
Cisco Systems International BV Amsterdam
アムステルダム (オランダ)

シスコは世界各国 200 箇所にオフィスを開設しています。住所、電話番号、ファックス番号はシスコの Web サイト <https://www.cisco.com/go/offices> を掲載されています。

Cisco および Cisco ロゴは、Cisco Systems, Inc. またはその関連会社の米国およびその他の国における登録商標または商標です。シスコの商標の一覧は、http://www.cisco.com/web/JP/trademark_statement.html でご確認いただけます。Third-party trademarks mentioned are the property of their respective owners. 「パートナー」という言葉が使用されていても、シスコと他社の間にパートナーシップ関係が存在することを意味するものではありません。(1110R)。

翻訳について

このドキュメントは、米国シスコ発行ドキュメントの参考和訳です。リンク情報につきましては、日本語版掲載時点で、英語版にアップデートがあり、リンク先のページが移動/変更されている場合がありますことをご了承ください。あくまでも参考和訳となりますので、正式な内容については米国サイトのドキュメントを参照ください。