Arista 7050X3 シリーズ・スイッチのアーキテクチャ

仮想化とコンテナを活用するハイパフォーマンス・サーバーの普及と、広帯域化の進展に伴い、エンタープライズ、クラウド、キャリア・ネットワークのリーフとスパインの両方の階層において、高密度の 25G および 100G のイーサネット・スイッチの必要性が高まっています。これからの次世代ネットワークには、パフォーマンス、スケール、効率を高いレベルで調和させ、アーキテクチャの変更に対応できるシステムが必要です。さらに、トンネリングとルーティングの新しいオプションと、高度な監視、テレメトリ、プログラマビリティを、既存の機能を損なわずに実現することが求められます。Arista 7050X3 シリーズは柔軟性のあるデータセンター向けスイッチであり、ワイヤスピードのレイヤ 2 およびレイヤ 3 機能、低遅延、および Software Driven Cloud Networking（SDCN）に対応する包括的で一貫性のある各種機能を備え、負荷分散、ネットワーク・トレーシング、スケールのイノベーションを提供します。

- マルチパス・ネットワークを効率的に使用できる最適パス選択
- ラインレートで低遅延のネットワーク・アドレス変換（NAT）
- トリガーによるバッファ・キャプチャとフロー・トラッカーを備えた高度なネットワーク・テレメトリ
- 完全共有インテリジェント・バッファによるスループットとスケールの倍増
- VXLAN ルーティングなど高度な機能のサポート

7050X3 は、多様なインターフェイス速度（10G、25G、40G、50G、100G）をサポートし、Arista EOS®と組み合わせることで、ビッグデータ、クラウド、仮想化、従来型のネットワーク設計に必要な各種機能を提供します。さらに、最新のデータセンターの多様なアプリケーションや水平型（East-West）トラフィック・パターンに対応できます。

Arista 7050X3 スイッチは、Arista 7050X のポートフォリオを強化するもので、レイヤ 2 およびレイヤ 3 スケールの大幅な機能向上を実現する、重要な最新テクノロジーと機能を備えています。

7050X3 シリーズにより、7050X のポートフォリオで 25G と 100G のアップリンクがサポートされるようになります。100G のネットワーク運用では、帯域幅が 40G に比べて大きく増えるとともに、総帯域幅のスケーリングの効率と費用対効果が向上します。7050X3 は、一貫したケーブリングで 10G または 25G、および 40G から 100G への接続に対応し、ケーブリングおよびサーバー・インフラに対する投資を保護しながら、ネットワークを容易にアップグレードできる移行パスを提供します。そのため、お客様のあらゆるサイズのネットワークを 25G サーバー・テクノロジーに移行して、サーバー・パフォーマンスと広帯域幅の利点を最大限に活用できます。7050X3 には、他にも次のような特長があります。

- 全ポート共通の完全共有 32MB パケット・バッファ。インテリジェントな動的バッファ管理ですべてのポートにバッファを均等分配し、重要なアプリケーション・トラフィック用にバッファを予約することで、速度の変化、マイクロバースト、持続的なネットワーク輻輳に対処します。また、7050X3 は、PFC、ETS、RoCE などの機能のサポートにより、ストレージ・アプリケーション向けのロスレス・イーサネットを可能にします。
- ネットワークのスケーラビリティは、スイッチのフォワーディング・テーブルのサイズによって直接影響を受けます。多くのシステムでは、一般的に転送エントリが固定サイズのテーブルを使用して、「すべてに適合する」アプローチが採用されています。Arista 7050X3 プラットフォームは、L2 MAC、L3 ルーティング、L3 ホスト、IP マルチキャストの転送のエントリが共用する統合フォワーディング・テーブル（UFT）を採用しており、エントリの種類ごとにパーティション分割が可能です。各パーティションの理想的なサイズは、ネットワーク展開の状況ごとに異なります。7050X3 には、事前に定義されたさまざまな構成プロファイルもあります。統合フォワーディング・テーブルの柔軟性を生かすことにより、どのようなネットワーク・トポロジやネットワーク仮想化技術でも、最適なリソース割り当てを実現できます。
- Arista 7050X3 シリーズのパケット・プロセッサのアーキテクチャは、柔軟なパケット・パイプラインで強化されています。これにより、システムを変更または交換することなく、ソフトウェア・アップグレードによってスイッチの転送プレーンに新しい機能を追加できます。そのため、高い費用をかけて機器を交換したり、ネットワークのアップグレードを待ったりする必要がなくなり、迅速に新機能のテストと導入を行えます。UFT が提供する柔軟なリソース割り当てと、柔軟なパイプラインの組み合わせによって、プラットフォームの汎用性を高め、幅広いユースケースを可能にし、投資保護を確保します。
- Arista 7050X3 シリーズのアーキテクチャは、カットスルー方式とストア・アンド・フォワード方式のスイッチングをサポートします。このプラットフォームは、同じ速度の 2 つのポート間または高速ポートから低速ポートへのカットスルー・スイッチングによって、800 ナノ秒の超低遅延を実現します。
- 7050X3 は、Smart System Upgrade（SSU）、LANZ、高度なネットワーク・テレメトリ、パケット・タイムスタンプなど、Arista X シリーズ・システムで既にサポートされている、一貫性のあるEOS機能セットをサポートします。運用と機能の一貫性が維持されるため、新製品導入の際に通常行われる評価の時間を短縮できます。

7050X3 プラットフォームは、高いシステム・パフォーマンス、スケール、一貫性のある機能とイノベーションを備えているため、大規模なエンタープライズ・データセンター、ビッグデータおよび機械学習環境、従来型データセンターおよび仮想データセンター、サービス・プロバイダのエッジ・ネットワーキングを拡張させたいときに最適です。

Arista 7050X3 シリーズの各モデル

7050X3 シリーズには複数のモデルがあり、さまざまなインターフェイスの種類とシステム密度に対応しています。

図 1:Arista 7050X3 シリーズ

7050X3 シリーズの各モデルの詳細を次の表に示します。

表 1:7050X3 シリーズ – システム仕様

7050X3 のモデル	7050CX3-32S	7050CX3M-32S	7050SX3-96YC8	7050SX3-48YC12	7050SX3-48YC8	7050SX3-48C8	7050TX3-48C8
スイッチの高さ(RU)	1	1	2	1	1	1	1
10G RJ45	--	--	--	--	--	--	48
10G SFP+	2	2	2	--	--	48	--
25G SFP	--	--	96	48	48	--	48
40G QSFP+	--	--	--	--	--	--	--
100G QSFP	32	32	8	12	8	8	8
10GbE の最大ポート数*	128(129)*	128(129)*	128(129)*	96	64	64	64
25GbE の最大ポート数*	128	128	128	96	64	16	16
40GbE の最大ポート数*	32	32	8	12	8	8	8
100GbE の最大ポート数*	32	32	8	12	8	8	8
最大 HW システム スループット(Tbps)	6.4	6.4	6.4	4.8	4.0	2.56	2.56
最大転送レート(Bpps)	2	2	2	2	1	1	1
遅延	800 ナノ秒から	800 ナノ秒から	800 ナノ秒から	800 ナノ秒から	800 ナノ秒から	800 ナノ秒から	3 マイクロ秒から
パケット・バッファ・メモリ	32MB	32MB	32MB	32MB	32MB	32MB	32MB

* 最大ポート数を使うには、ブレークアウトの使用と他のインターフェイスの無効化が必要となる場合があります。これらのモデルの 10G のポート数には、追加の SFP+ポートの使用が含まれます。1 番目の SFP+ポートは常にアクティブで、2 番目は代替インターフェイスを無効化すると使用可能になる共有ポートです。

Arista 7050X3 シリーズ導入のシナリオ

7050X3 の各モデルには、スケーラブルなリーフ/スパイン設計の構築に柔軟に対応できる、複数の接続オプションが用意されています。運用の柔軟性は 7050X3 シリーズ全体に共通しており、多種多様な展開の状況に適しています。

高密度のトップオブラックで、10GbE と 25GbE の両方のシステムのサーバー・ラックのユースケースに対応します。

- **10GbE から 25GbE への移行** – 次世代イーサネットのパフォーマンスへのシームレスな移行のための IEEE 802.3by 25GbE およびコンソーシアム策定規格への準拠
- **グリッド/HPC** – 費用対効果と電力効率に優れたシステムを必要とする設計。10G および 25G サーバーでノンブロッキングのネットワークやオーバーサブスクリプションを最小限に抑えたネットワークを実現
- **リーフ/スパイン** – テレメトリおよび可視化機能を備えた、オープン規格に基づく L2 および L3 の 2 階層設計
- **セキュアなクラウド環境** – リーフ/スパイン用の 7050X3M シリーズ上で MACsec に対応
- **100GbE スケールアウト設計** – 電力効率および高密度でコンパクトなシステムが必要な中小規模の場所
- **最大 128way の ECMP 設計** – オープン・プロトコル、および Arista 7320X と 7500R の 100GbE モジュール型スパイン・スイッチを使用して、費用対効果に優れたマルチパスを実現
- **大規模な L2 環境** – UFT が提供する柔軟なリソース割り当てにより、最大限の規模を実現

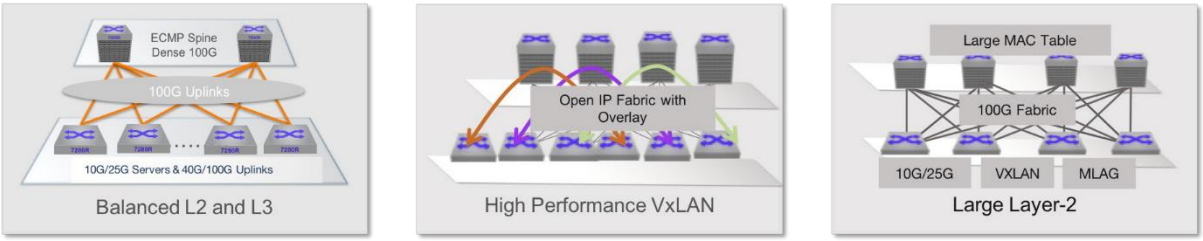

図 2:7050X3 導入のシナリオ

Arista 7050X3 スイッチのアーキテクチャ

システム設計は 7050X3 シリーズのすべてに共通で、コントロール・プレーンには高性能な x86 CPU と 8GB のシステム・メモリを搭載しています。CPU は、内部フラッシュ、ブートフラッシュ、電源、ファン、管理 I/O、周辺機器に接続されます。

この x86 CPU はさらに PCIe 経由でスイッチ・オン・チップに接続されます。このスイッチ・オン・チップが、すべてのデータ・プレーン転送を前面パネルのポートに直接接続します。

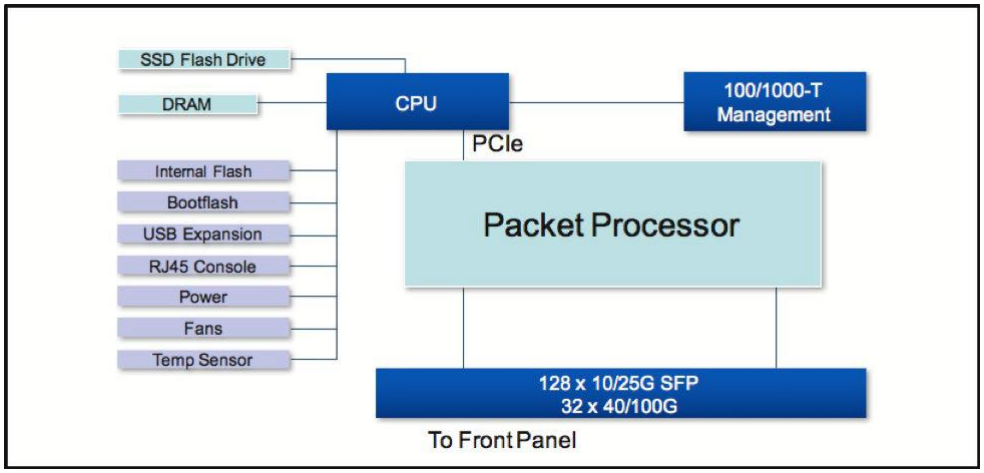

図 3:7050X3 のアーキテクチャ

ARISTA 7050CX3-32S

7050CX3-32S は、32 の 100G QSFP ポートを収めた 1RU システムで、最大ワイヤスピード・スループットは 6.4 Tbps です。このスイッチは、高密度の 100G または 40G スパイン接続、あるいは各 QSFP インターフェイスで 2x50G をサポートするハイパフォーマンス・サーバーおよびストレージ接続に最適化されています。

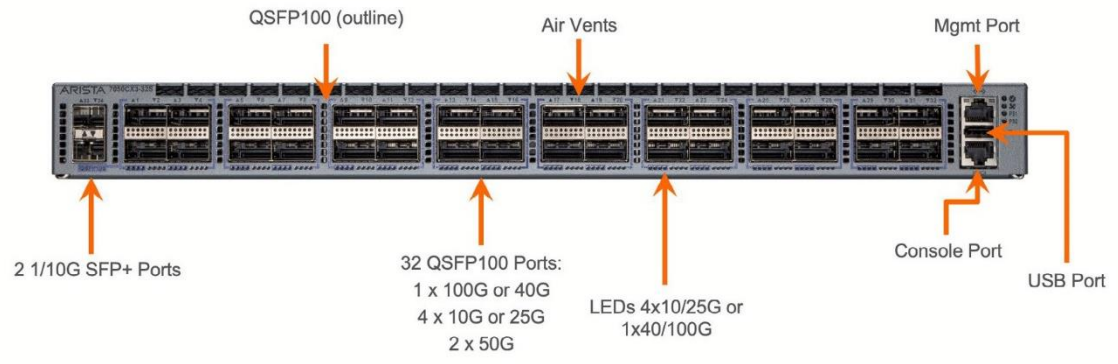

図 4：7050CX3-32S スイッチ

- 7050CX3-32S は高密度の 100GbE システムで、最大 129 の固有のインターフェイスを提供します。
- 32 の QSFP ポートで 10GbE から 100GbE の 5 種類の速度をサポートします。
  - 各ポートは、4 x 10G、4 x 25G、1 x 40G、2 x 50G、1 x 100G として構成可能です。
- 代替接続オプション用に 2 個の 1/10G SFP+ポート（33、34）が提供されています。
  - SFP+ポート 34 は QSFP ポート 32 の 2 つのレーン（32/2 と 32/4）で共有されます。
  - デフォルトでは、ポート 32/1、32/3、34 はアクティブで、2 つの 10/25 ブレークアウトと 1 つの SFP ポートを提供します（32/2 と 32/4 は無効化されています）。
  - ポート 34 を無効化すると、QSFP 32/1-4 をアクティブにして 4 つの 10/25 ブレークアウトを提供できます。
- QSFP ポートは IEEE 25GbE をサポートしています。
- ワイヤスピード性能（32MB の完全共有バッファ搭載）を提供します。
- カットスルー・スイッチングで 800 ナノ秒の低遅延をサポートしています。

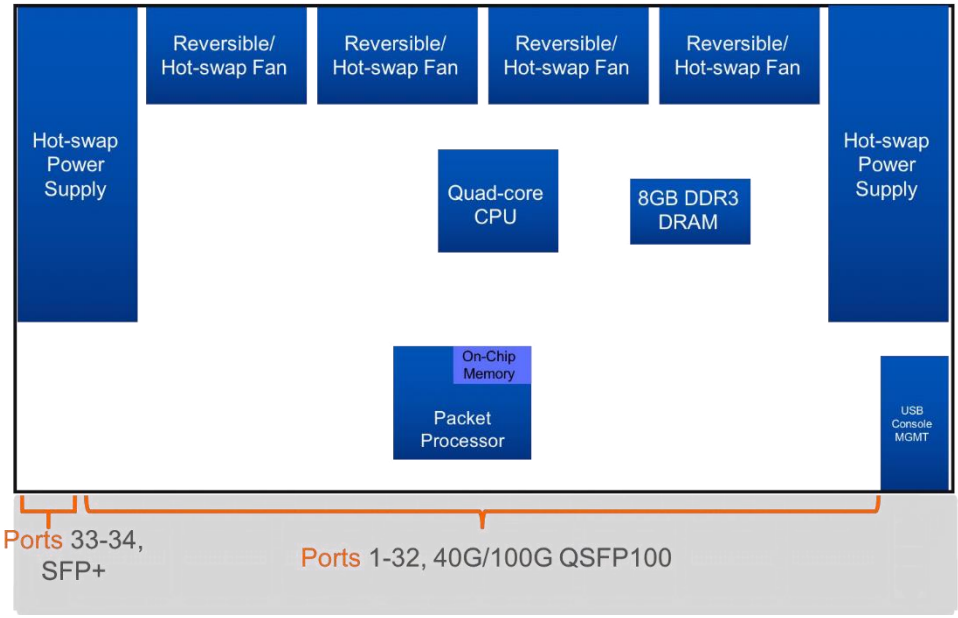

図 5：Arista 7050CX3-32S のアーキテクチャのブロック図

ARISTA 7050CX3M-32S

7050CX3M-32S は、32 の 100G QSFP ポートを収めた 1RU システムで、6.4 Tbps の最大ワイヤスピード・スループットを提供し、MACsec に対応しています。このスイッチは、高密度の 100G または 40G スパイン接続、あるいは各 QSFP インターフェイスで 2x50G をサポートするハイパフォーマンス・サーバーおよびストレージ接続に最適化されています。

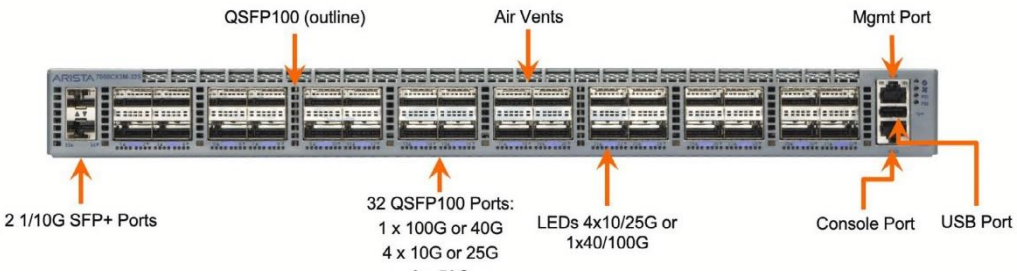

図 6：7050CX3M-32S スイッチ

7050CX3M-32S は高密度の 100GbE システムで、最大 129 の固有のインターフェイスを提供します。

- 32 の QSFP ポートで 10GbE から 100GbE の 5 種類の速度をサポートします。
  - 各ポートは、4 x 10G、4 x 25G、1 x 40G、2 x 50G、1 x 100G として構成可能です。
- 代替接続オプション用に 2 個の 1/10G SFP+ポート（33、34）が提供されています。
  - SFP+ポート 34 は QSFP ポート 32 の 2 つのレーン（32/2 と 32/4）で共有されます。
  - デフォルトでは、ポート 32/1、32/3、34 はアクティブで、2 つの 10/25 ブレークアウトと 1 つの SFP ポートを提供します（32/2 と 32/4 は無効化されています）。
  - ポート 34 を無効化すると、QSFP 32/1-4 をアクティブにして 4 つの 10/25 ブレークアウトを提供できます。
- すべての QSFP ポートでワイヤスピード MACsec を提供し、多様なインターフェイス速度に対応します。
- QSFP ポートは IEEE 25GbE 仕様をサポートしています。
- ワイヤスピード性能（32MB の完全共有バッファ搭載）を提供します。
- カットスルー・スイッチングで 800 ナノ秒の低遅延をサポートしています。

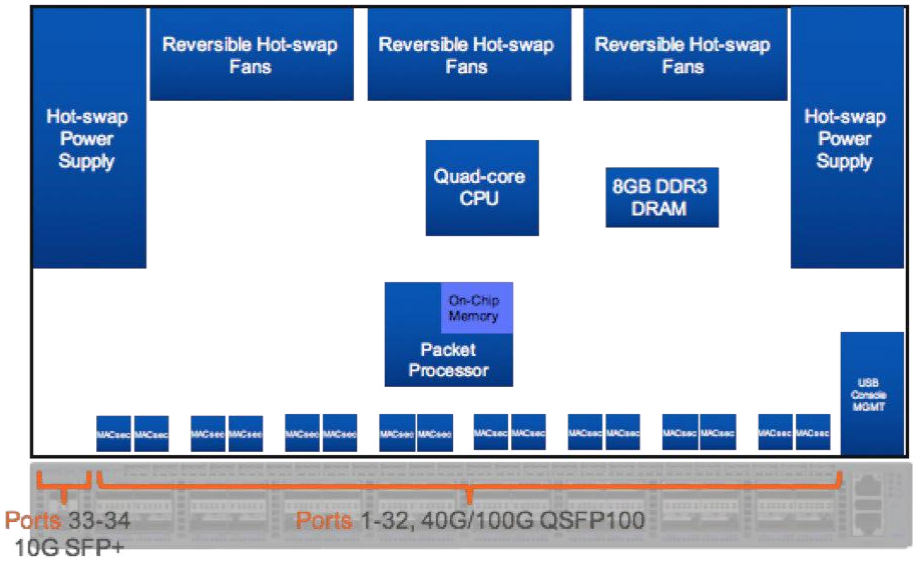

図 7：Arista 7050CX3M-32S のアーキテクチャのブロック図

ARISTA 7050SX3-96YC8

Arista 7050SX3-96YC8 は、25G の SFP ポートを 96 ポートと 100G の QSFP ポートを 8 ポート収めた 2RU システムで、合計スループットは 6.4Tbps です。このスイッチは、10/25G のリーフ・ネットワークと 40G または 100G のアップリンク・ネットワークに最適化されており、さまざまなポート速度構成に対応します。高密度スイッチを導入する必要があり、1/10G から移行しようとするコンピュートおよびストレージ・ラックに適しています。

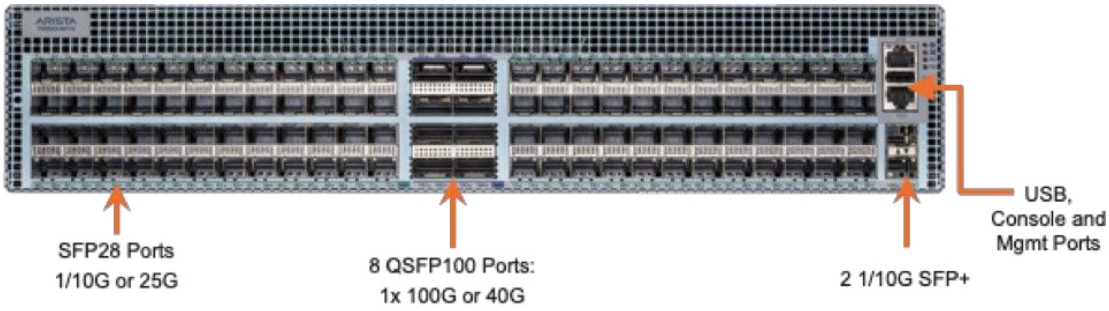

図 8：7050SX3-96YC8 スイッチ

7050SX3-96YC8 は、100GbE の QSFP ポートを備えた高密度 25GbE/10GbE システムで、最大 129 の固有のインターフェイスを提供します。

- 96 ポートの SFP28 でワイヤスピードの 1G/10G/25G をサポートします。
  - 4 ポートを 1 グループとし、グループ単位でポート速度を構成できます。
- 8 ポートの QSFP ポートで 10GbE から 100GbE の 5 種類の速度をサポートします。
  - 各ポートは、4 x 10G、4 x 25G、1 x 40G、2 x 50G、1 x 100G として構成可能です。
- 代替接続オプション用に 2 個の 1/10G SFP+ポート（105、106）が提供されています。
  - SFP+ポート 105 は SFP28 ポート 46 と 48 で共有されます。
  - SFP28 ポート 46 と 48 はデフォルトで有効になっており、ポート 105 をアクティブにするにはこれらを無効化する必要があります。
- IEEE 25GbE 仕様を完全にサポートし、10G から 25G への移行が可能です。
- 使い慣れた SFP 接続とケーブリング・ソリューションを使用して 1/10G から 1/10/25G に容易に移行できます。
- ワイヤスピード性能（32MB バッファ搭載）を提供します。
- カットスルー・スイッチングで 800 ナノ秒の低遅延をサポートしています。

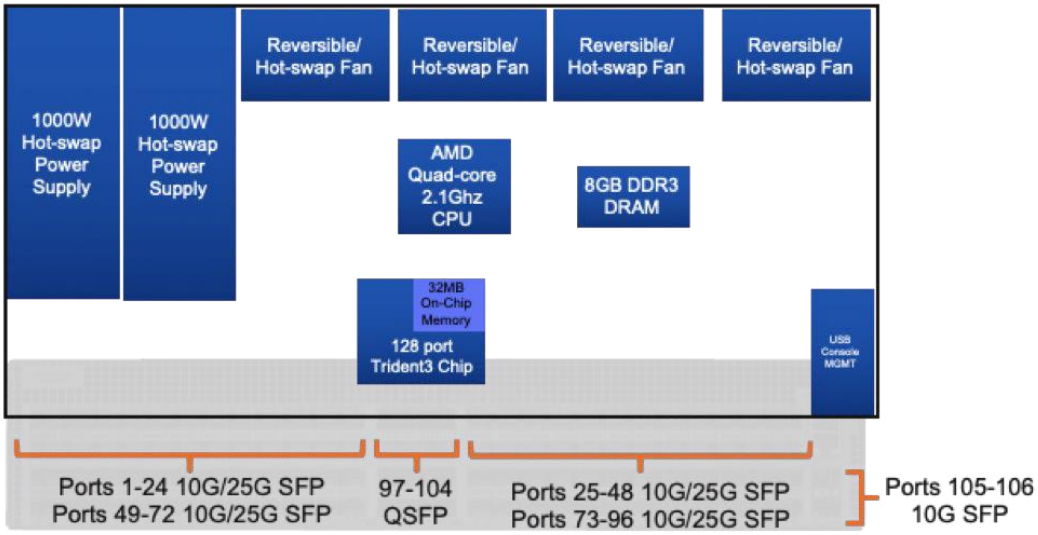

図 9：Arista 7050SX3-96YC8 のアーキテクチャのブロック図

ARISTA 7050SX3-48YC12

Arista 7050SX3-48YC12 は、25G の SFP ポートを 48 ポートと 100G の QSFP ポートを 12 ポートを収めた 1RU システムで、合計スループットは 4.8Tbps です。このスイッチはノンブロッキング設計で、さまざまなポート構成から選択でき、リーフおよびスパインの展開に加えて、高密度で高性能のコンピュートおよびストレージ・ラックに適しています。

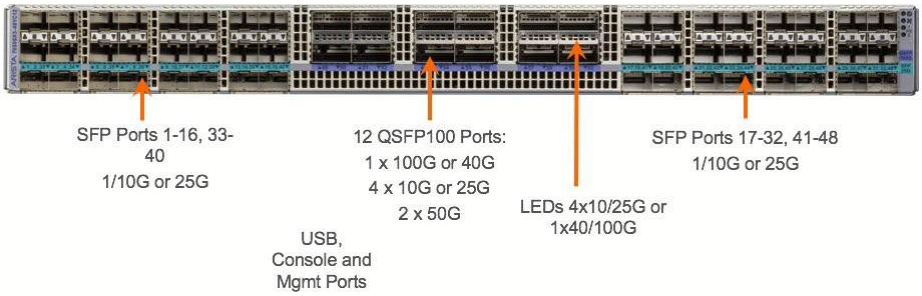

図 10：7050SX3-48YC12 スイッチ

7050SX3-48YC12 は、100GbE の QSFP ポートを備えた高密度 25GbE/10GbE システムで、次の機能を提供します。

- ワイヤスピードの 25GbE を 48 ポートと 40/100G の QSFP ポートを 12 ポートを備え、ブレークアウト・ケーブルとパラレル光ファイバーを使用する際は最大で合計 96 ポートの 25G または 10G に対応します。
- IEEE 25GbE 仕様を完全にサポートし、10G から 25G への移行が可能です。
- 48 ポートの高密度 SFP ポートを 4 ポート 1 グループの単位で有効化でき、速度 25G または 10G/1G 混在で使用できます。
- 12 ポートの QSFP ポートはそれぞれ、4 x 10G、4 x 25G、1 x 40G、2 x 50G、1 x 100G の速度構成を選択できます。
- 使い慣れた SFP 接続とケーブリング・ソリューションを使用して 1/10G から 1/10/25G に容易に移行できます。
- ワイヤスピード性能（32MB バッファ搭載）を提供します。
- カットスルー・スイッチングで 800 ナノ秒の低遅延をサポートしています。

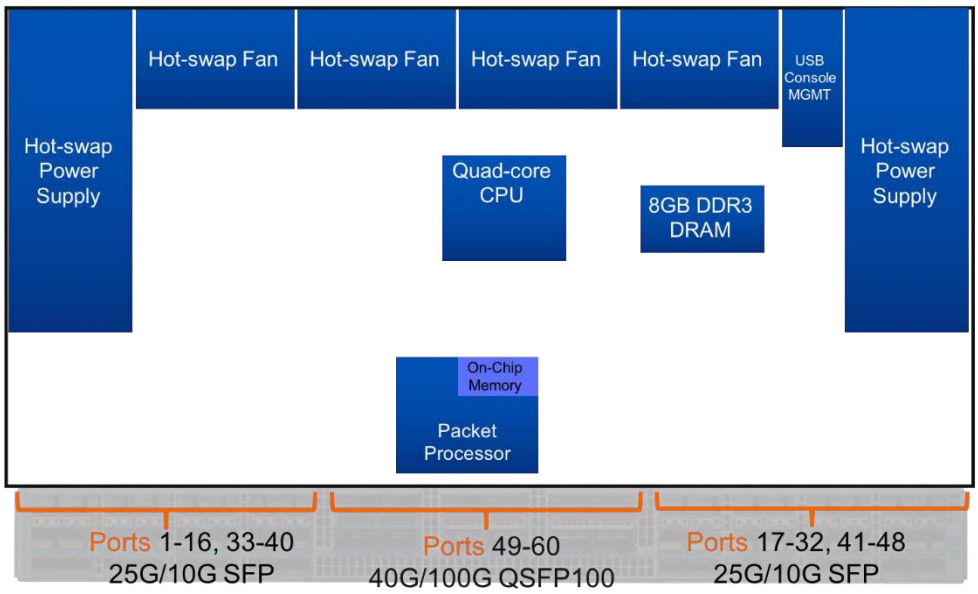

図 11：Arista 7050SX3-48YC12 のアーキテクチャのブロック図

ARISTA 7050SX3-48YC8 と 7050SX3-48C8

Arista 7050SX3-48YC8 と 7050SX3-48C8 は、SFP ポートを 48 ポートと 100G の QSFP ポートを 8 ポート収めた 1RU システムで、合計スループットは最大 4Tbps です。48YC8 プラットフォームは 1/10/25G の SFP トランシーバーをサポートしますが、48C8 モデルは 1/10G の SFP のみをサポートします。

このスイッチは、10/25G のリーフ・ネットワークと 40G または 100G のアップリンク・ネットワーク向けに設計されており、さまざまなポート速度構成を選択でき、高密度のコンピュートおよびストレージ・ラックに適しています。

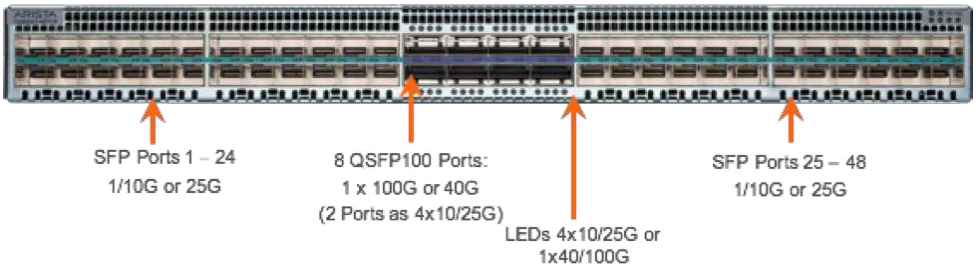

図 12：7050SX3-48YC8 スイッチ

7050SX3-48YC8 は、100GbE の QSFP ポートを備えた高密度 25GbE/10GbE システムで、次の機能を提供します。

- ワイヤスピードの 25GbE ポートを 48 ポートと 40/100G の QSFP ポートを 8 ポート備え、システムごとの最大論理ポート数は 64 ポートです。
- ブレークアウト・ケーブルとパラレル光ファイバーのいずれかまたは両方を使用する際は、最大で合計 64 ポートの 25G または 10G に対応します（フル・ブレークアウトの場合は、4 個の QSFP ポートを無効化する必要があります）。
- IEEE 25GbE 仕様を完全にサポートし、10G から 25G への移行が可能です。
- 48 ポートの高密度 SFP ポートを有効化でき、速度 25G または 10G/1G 混在で、任意の組み合わせで使用できます。
- 8 ポートの QSFP ポートにはそれぞれ、100GbE と 40GbE の速度構成を選択できます。
- 使い慣れた SFP 接続とケーブリング・ソリューションを使用して 1/10G から 1/10/25G に容易に移行できます。
- ワイヤスピード性能（32MB バッファ搭載）を提供します。

7050SX3-48C8 は、100GbE の QSFP を備えた高密度 10GbE システムで、カットスルー・スイッチングと 800 ナノ秒の低遅延をサポートします。48YC8 モデルとの相違点は次のとおりです。

- ワイヤスピード 1/10GbE の SFP ポートを 48 ポートと 40/100G の QSFP ポートを 8 ポート備え、システムごとの最大論理ポート数は 64 ポートです。
- ブレークアウト・ケーブルとパラレル光ファイバーのいずれかまたは両方を使用する際は、最大で合計 64 の 10G ポート、または 48 の 10G ポートと 16 の 25G ポートに対応します（フル・ブレークアウトの場合は、4 つの QSFP ポートを無効化する必要があります）。

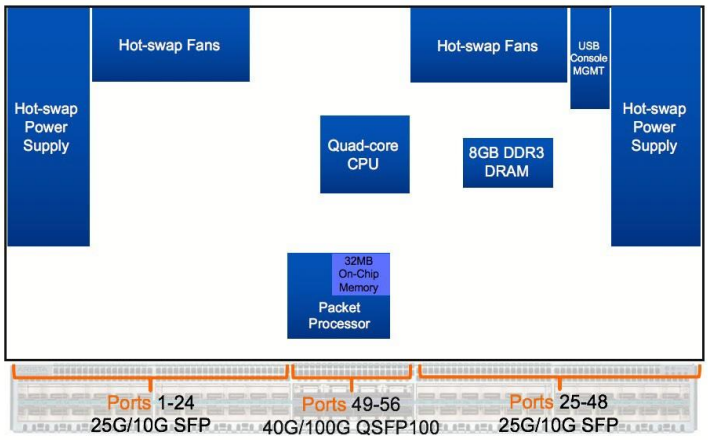

図 13：Arista 7050SX3-48YC8 / 48C8 のアーキテクチャのブロック図

ARISTA 7050TX3-48C8

Arista 7050TX3-48C8 は、10GBaseT ポートを 48 ポートと 100G の QSFP ポートを 8 ポート収めた 1RU システムで、合計スループットは 2.56Tbps です。このスイッチは、10G RJ45 のリーフ・ネットワークと 40G または 100G のアップリンク・ネットワーク向けに設計されており、さまざまなポート速度構成を選択でき、8-way ECMP でスパイン・ネットワークに接続される高密度のコンピュートおよびストレージ・ラックに適しています。

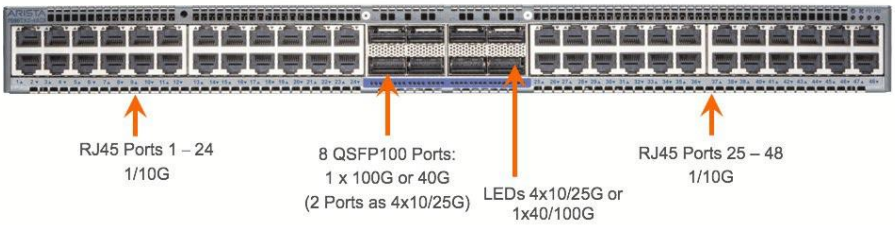

図 14：7050TX3-48C8 スイッチ

7050TX3-48C8 は、100GbE の QSFP ポートを備えた高密度 10GbE RJ45 システムで、次の機能を提供します。

- ワイヤスピードの 10GbE RJ45 ポートを 48 ポートと 40/100G の QSFP ポートを 8 ポート備え、システムごとの最大論理ポート数は 64 ポートです。
- ブレークアウト・ケーブルとパラレル光ファイバーのいずれかまたは両方を使用する際は、最大で合計 64 の 10G ポート、または QSFP を使用して 10G/25G と 40/100G の柔軟な組み合わせに対応できます（フル・ブレークアウトの場合は、4 つの QSFP ポートを無効化する必要があります）。
- 1G および 10G BaseT RJ45 は UTP カテゴリ 5、5e、6、および 6A の構内配線に対応します。
- 8 つの QSFP ポートにはそれぞれ、100GbE と 40GbE の速度構成を選択できます。
- 10G から 25G のサーバーを扱う他の 7050X3 シリーズと組み合わせることで、100G スパインに容易に移行できます。
- ワイヤスピード性能（32MB バッファ搭載）を提供します。
- 低消費電力、カットスルー・スイッチング、低遅延のサポートにより、コスト効率とスペース効率に優れた高密度のサーバー・ラックを実現します。

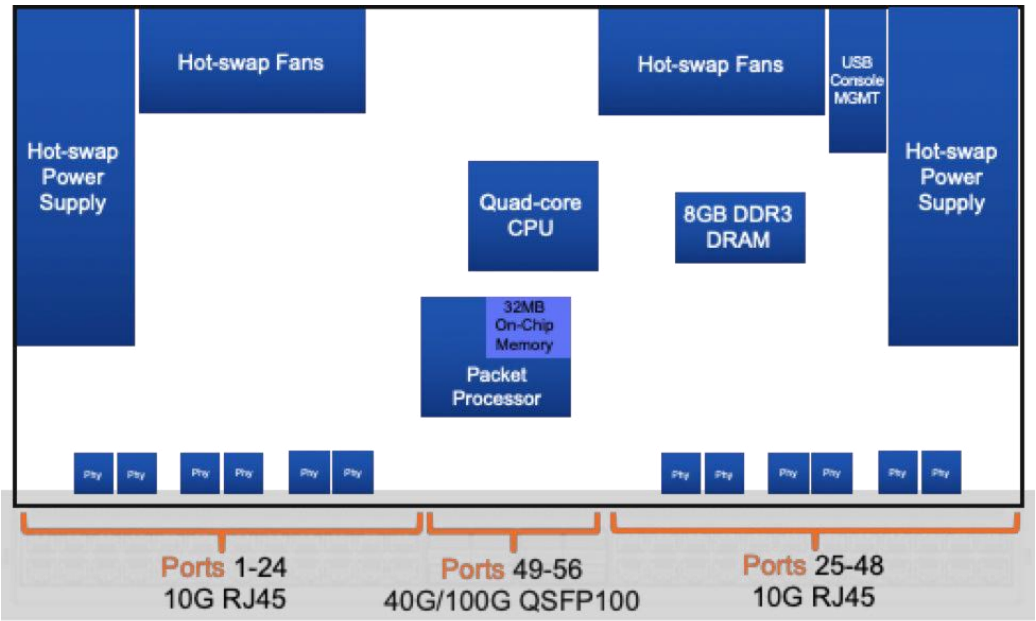

図 15：Arista 7050TX3-48C8 のアーキテクチャのブロック図

ポートの識別

40GbE QSFP+と 100GbE QSFP100 のトランシーバーは物理サイズが同じであり、どの QSFP ベース・ポートにも挿入できます。エンジニアがトランシーバーを適切な種類のポートに挿入できるようにするために、スイッチの前面パネルにポートの機能を明示することが重要です。

次の図は、2 種類の QSFP ポートを示しています。右側のポートには、QSFP100 対応を示す薄紫色の線が引かれています。これらのポートは、QSFP+または QSFP100 のトランシーバーに対応しています。薄紫色の線がない左側のポートは、QSFP+限定で、QSFP100 には対応していません。ポートの下の LED インジケーターは、4 ポートずつのグループに分けられています。左側の 1 組は上のポート、右側の 1 組は下のポートを表します。ポートが 40G または 100G に構成されている場合は、接続時に 4 個の LED インジケーターのうち左端の 1 個が緑色に点灯し、残りはオレンジ色に点灯します。ポートが 10G または 25G のブレークアウトとして構成されている場合は、4 個の LED のうち、接続されたリンクに対応する LED が緑色になります。QSFP100 ポートに 50G で接続する場合は、LED 1 と 3 がリンク状態を表し、2 と 4 はオレンジ色のままになります。

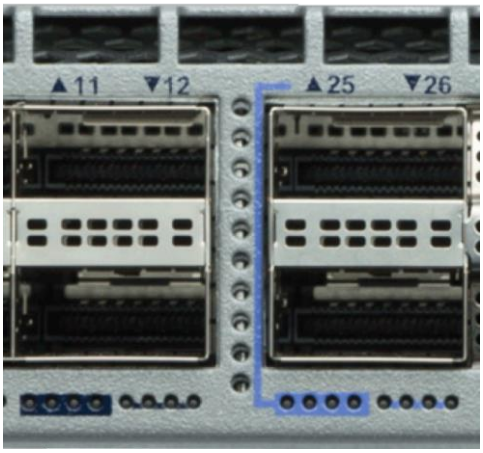

図 16：QSFP+ポートと QSFP100 ポート

25GbE に使用されるプラガブル・トランシーバーは SFP28 であり、サイズは 1GbE SFP および 10GbE SFP+と同じです。25GbE 対応ポートを識別しやすくするために、25GbE 対応 SFP ポートは次のように青緑色の線で示されています。

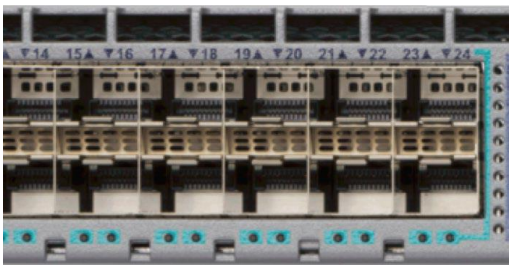

図 17：SFP28 25GbE ポート

25G SFP ポートのグループ化

7050SX3-48YC12 と 7050SX3-96YC8 の 10G/25G 対応 SFP ポートは、同じ速度構成のポート 4 つずつのグループに構成されます。同じグループの 4 つのポートは、単一のポート速度設定（1/10G または 25G）を使用します。最初に構成したポートの速度によって、そのグループの他のすべてのポートの速度が決まります。これは、100G QSFP ポートを速度 1/10G または 25G のブレークアウト・モードで使用する場合も同様です。このポート・グループ化モードは、7050SX3-48YC8 には適用されません。

データセンター・グレードの可用性と冗長性

Arista 7050X3 シリーズのスイッチは、ソフトウェアおよびハードウェアの両面において、高い可用性を発揮するように設計されています。高可用性を実現するための主な機能は次のとおりです。

- 1+1 冗長のホット・スワップ可能な電源と、N+1 冗長のホット・スワップ・ファン 4 個
- カラーコード付きの PSU とファン
- ライブ・ソフトウェア・パッチ
- ステートフル・フォールト・リペア（SFR）による自己回復ソフトウェア
- Smart System Upgrade（SSU）
- マルチシャーシ LAG によるアクティブ/アクティブ L2 マルチパス
- 128way の ECMP ルーティングによる負荷分散と冗長性

表 2：7050X3 シリーズ - 電源とファンの冗長性

7050X3 のモデル	7050CX3-32S	7050CX3M-32S	7050SX3-96YC8	7050SX3-48YC12	7050SX3-48YC8	7050SX3-48C8	7050TX3-48C8
電源（冗長）	2 個のホット・スワップ可能電源（1+1）						
ファン（冗長）	3 個のホット・スワップ可能ファン（N+1）	4 個のホット・スワップ可能ファン（N+1）			2 個のホット・スワップ可能ファン（N+1）		
エアフロー	前面吸気/背面排気および背面吸気/前面排気	前面吸気/背面排気および背面吸気/前面排気	前面吸気/背面排気および背面吸気/前面排気	前面吸気/背面排気	前面吸気/背面排気および背面吸気/前面排気		

図 18：Arista 7050SX3-48YC12 スイッチの背面

図 19：Arista 7050CX3M-32S スイッチの背面

図 20：Arista 7050SX3-48YC8、7050SX3-48C8、7050TX3-48C8 スイッチの背面

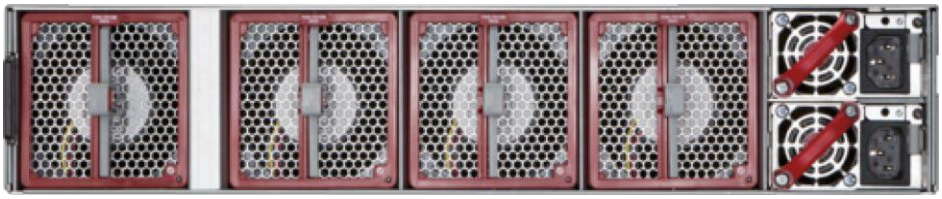

図 21：Arista 7050SX3-96YC8 スイッチの背面 - 前面吸気/背面排気（赤）

コントロール・プレーンの拡張

7050X3 シリーズ・スイッチの中央 CPU コンプレックスは、コントロール・プレーンと管理機能専用で使用され、すべてのデータ・プレーン転送ロジックは、パケット・プロセッサ・レベルで行われます。

すべての Arista スイッチのコントロール・プレーン・ソフトウェアである Arista EOS®は、マルチコア x86 CPU と数ギガバイトの DRAM で実行されます。

EOS は、マルチスレッドに対応し、Linux カーネルで実行され、拡張可能です。大容量 RAM と高速マルチコア CPU により、効率に優れたコントロール・プレーンを実行できるとともに、EOS と同じ Linux インスタンス内あるいはゲスト仮想マシンまたはコンテナ内でサードパーティ・ソフトウェアも実行できます。

シリアル・コンソール・ポートまたは 10/100/1000 イーサネットの管理インターフェイスあるいはその両方を使用して、帯域外管理を行うことができます。7050X3 シリーズは USB2.0 インターフェイスも提供しており、イメージやログの転送などさまざまな用途に使用できます。

表 3:7050X3 シリーズの CPU コンプレックス

7050X3 のモデル	7050CX3-32S	7050CX3M-32S	7050SX3-48YC12	7050SX3-48YC8	7050SX3-48C8	7050TX3-48C8
CPU	クアッドコア x86					
システム・メモリ	8GB					
フラッシュ・ストレージ	8GB					

柔軟性に優れた次世代パケット・プロセッサ

7050X3 シリーズには、単一のスイッチ・オン・チップ(SoC)デバイスが使用されています。7050X3 は、広く導入されている 7050X シリーズとの一貫性を維持しながら、新しい柔軟なパイプラインの実装で多くの利点を提供します。この柔軟性により、パイプラインに変更を加えて、EOS で実装される新機能を導入できます。

以下のセクションでは、パケット・プロセッサのさまざまな主要コンポーネントと、それらのコンポーネントが提供する機能について説明します。

表 4:7050X3 の柔軟な転送の特長

機能	利点
完全に確定的なパイプライン	ラインレート・スループットと低遅延
大容量のルックアップ・ステージ	メモリ効率の向上と消費電力の削減
インテリジェントな並列処理	少ないステージ数でのパケット処理能力強化
構成可能なパイプライン	独自の高度な機能
柔軟なプログラマブル・テレメトリ	動的ネットワーク監視ときめ細かな可視化

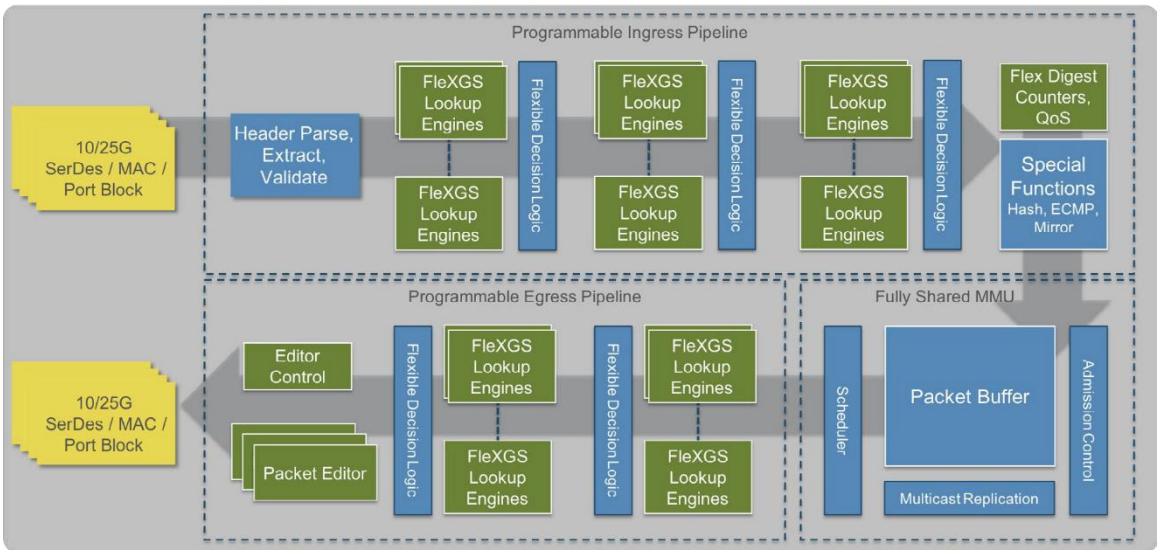

図 22:7050X3 シリーズのパケット・プロセッサ

柔軟なパイプライン

Arista 7050X3 シリーズは、柔軟なパケット・パイプラインで強化された転送アーキテクチャをサポートしており、基盤となるハードウェアを変更または交換することなく、ソフトウェア・アップグレードによってパケット・プロセッサのデータ・プレーンに新しい機能を追加できます。そのため、高い費用をかけて機器の交換や大規模なアップグレードを行う必要がなくなり、迅速に新機能のテストと導入を行えます。

7050X3 のパイプラインを、統合フォワーディング・テーブル（UFT）が提供する構成可能なルックアップ転送リソースの割り当てと組み合わせることで、幅広いユースケースに対応できる柔軟なシステムを構築し、投資保護を実現できます。

統合フォワーディング・テーブル

ネットワークのスケーラビリティは、スイッチのフォワーディング・テーブルのサイズによって直接影響を受けます。多くのシステムでは、一般的な種類の転送エントリごとに個別の固定サイズ・テーブルを使用して、「すべてに適合する」アプローチが採用されています。Arista 7050X3 スイッチは、L2 MAC、L3 ルーティング、L3 ホスト、IP マルチキャストの転送のエントリが共用する統合フォワーディング・テーブル（UFT）を採用しており、エントリの種類ごとにパーティション分割が可能です。各パーティションの理想的なサイズは、ネットワーク展開の状況ごとに異なります。7050X3 には、事前に定義されたさまざまなプロファイルもあります。統合フォワーディング・テーブルの柔軟性を生かすことにより、どのようなネットワーク・トポロジやネットワーク仮想化技術でも、最適なリソース割り当てを実現できます。

統合フォワーディング・テーブルは、最大で 32,000 のエントリが含まれるバンクを 8 つ保持し、各バンクは L2 または L3 に個別に割り当てられます。通常、バンク 2 は常に L2/MAC エントリに使用されますが、バンク 3、4、5 は L2 または L3（ホスト・ルート）に個別に割り当てることができます。

表 5：Arista 7050X3 の UFT プロファイル

UFT モード	0	1	2 デフォルト	3	4
MAC アドレス	288K	224K	160K	96K	32K
IPv4 ホスト・ルート	16K	80K	144K	168K	16K
IPv4 マルチキャスト（S、G）	8K	40K	72K	104K	8K
IPv6 ホスト・ルート	8K	40K	72K	104K	8K

また、UFT をアルゴリズム LPM（ALPM）モードで使用することもできます。この場合、すべてのバンクが LPM に割り当てられ、最大 384,000 の IPv4 LPM ルートを保存できます。

表 6：Arista 7050X3 の ALPM モード

LPM テーブル・モード	ALPM	1	2	3	4
IPv4 LPM ルート	384K	32K	32K	32K	32K
IPv6 LPM ルート・ユニキャスト（プレフィックス長 <= /64）	192K	12K	8K	4K	--
IPv6 LPM ルート・ユニキャスト（任意のプレフィックス長）	40K	2K	4K	6K	8K

ALPM モードでは、すべての UFT バンクが LPM テーブルの拡張専用となります。これにより、ホスト・ルート数は 16,000 に、MAC アドレス数は 32,000 に減少します。ALPM モードでは、L3 ホストおよび MAC テーブルの拡張はできません。

UFT と ALPM の柔軟性により、お客様は複数のユースケースに導入した 7050X3 シリーズでデータセンター・スイッチングを標準化し、それぞれで使用可能なすべてのスイッチ・リソースを最大限効率的に活用できます。

動的完全共有バッファのアーキテクチャ

25G カットスルー・モードでは、Arista 7050X3 スイッチは、一貫して 800 ナノ秒の低遅延でパケットを転送します。輻輳時には、合計サイズ 32MB のインテリジェントな完全共有パケット・メモリにパケットがバッファリングされ、優れたバースト吸収を実現します。ポートごとに固定のパケット・メモリを持つ他のアーキテクチャとは異なり、動的なしきい値によって、パケット・メモリをトラフィック・クラス、キューのサイズ、およびサービス品質ポリシーに基づいて割り当て、ロスありクラスとロスレス・クラスの両方のすべてのポートに均等分配します。バッファの利用率、占有率、およびしきい値は Arista LANZ ですべて可視化され、監視ツールにエクスポートして、デバイスおよびエンドツーエンドでのホットスポットの特定と遅延の測定に使用できます。

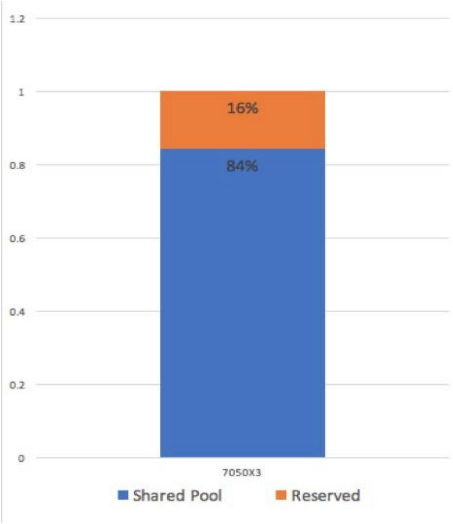

図 23:7050X3 のバッファ割り当て

パケット・バッファは、ロスありとロスレスの両方のトラフィック・クラスを処理するように設計されています。ロスレス・フレーム配信を必要とするトラフィック・クラスの場合には、PFC/PAUSE などのフロー制御の発行後に到着する転送中パケットを吸収するために、一定量のバッファが確保されます。ロスレス・バッファはすべての入力ポート共用の共有プールであり、最小および最大バッファ容量はポートごとに定義されています。これにより、全ポートに対するバッファ容量の割り当てを控えめにして、1 つのポートがバッファをオーバーフローさせる可能性を回避できます。

図中の予約済みプールは、ロスレス・トラフィック・クラスに割り当てられた合計プールを示しています。残りのバッファは、ロスありトラフィック・クラス用の共有バッファ・プールに集約されます。バッファ管理ではセル・サイズも重要です。7050X3 システムのセル・サイズは 208B なので非常に細分化したバッファリングができ、バッファの無駄を最小限に抑えて高いパフォーマンスを実現します。

ネットワーク負荷分散の自動化

7050X3 シリーズは、レイヤ 3 の ECMP およびレイヤ 2 の MLAG 環境のトラフィックに最適化されたネットワーク負荷分散を提供し、ネットワーク全体のパフォーマンスを向上させます。7050X3 の負荷分散ハッシュ・アルゴリズムの機能強化では、リンクのリアルタイムの負荷を考慮し、新しいフローを最適なリンクに動的に割り当てます。また、不均衡が検出されたときは、アクティブなフローのバランスを自動的に調整してリンク輻輳の可能性を抑制し、最大のスループットを実現します。

カットスルー・スイッチングと低遅延

7050X3 シリーズのアーキテクチャは、スイッチのすべてのポートで一貫した低遅延を実現します。パケット・プロセッサは、カットスルー・スイッチングとストア・アンド・フォワード・スイッチングの両方に対応します。カットスルー・スイッチングは、同速度の 2 つのポート間か、高速ポートから低速ポートへの切り替えでサポートされます。次の表は、さまざまな速度構成間のカットスルー・スイッチングのサポートを示しています。

表 7:7050X3 のカットスルー転送速度

同じポート速度間	異なるポート速度間
10GbE から 10GbE	40GbE から 10GbE
25GbE から 25GbE	50GbE から 10GbE
40GbE から 40GbE	50GbE から 25GbE
50GbE から 50GbE	100GbE から 40GbE
100GbE から 100GbE	100GbE から 50GbE
	100GbE から 25GbE

同じ速度間と異なる速度間のカットスルー転送により、7050X3 シリーズは幅広い環境で低遅延を実現できます。ストア・アンド・フォワード・モードは、低速から高速のパケット境界での転送で有効化され、他のフローのカットスルー動作に影響を与えません。

図 24 は、一般的な HFT および HPC のユースケースの 2 階層シナリオを示しています。これらのリーフ/スパイン設計では、2 階層のサーバー間の遅延を、100G エンドポイントで 2 マイクロ秒、25G エンドポイントで 3 マイクロ秒まで低くすることができます。

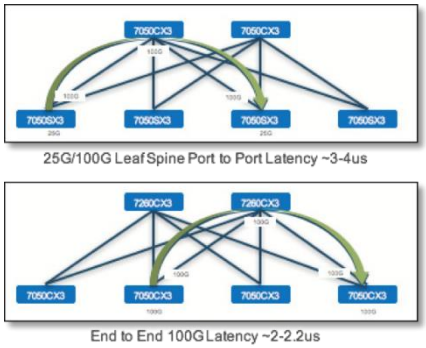

図 24：7050X3 のサーバー間の遅延

ネットワーク・アドレス変換（NAT）

ネットワーク・アドレス変換（NAT）は、一般的に、内部ネットワークをマスクするため、あるいは IP 割り当てが使用ホスト数に比べて小さすぎる問題を回避するために、ネットワーク内部で使用されます。この技術は、企業、金融取引所、市場取引など、コロケーション参加者に割り当てられる IP アドレス範囲が小さな範囲に限られている場合によく利用されます。たとえば、取引所から取引エンティティに割り当てられるアドレスが 30 で、サーバー数が 60 台ある場合に、この問題を回避する 1 つの方法は取引所から追加のアドレス 30 を取得することですが、それが無理ならば NAT を使用してサーバー・アドレスをマスクする必要があります。また別の例として、取引所が周知のマルチキャスト・アドレスを使用してフィードを公開している場合に、その送信元またはグループを、外部に周知されている情報を変更することなく内部的に変更したいことがあります。一般に、NAT はモジュール型ルーター、ファイアウォール、またはソフトウェアベースのルーターで実行され、多くの場合、数十マイクロ秒以上の非常に高い遅延が発生します。

Arista 7050X3 は、ハイパフォーマンスのコンパクトなデータセンター・スイッチで NAT 機能を提供します。これは、標準的な L2 ブリッジング、L3 ルーティング、または L4 インスペクションと同じ低遅延で NAT を実装しています。同じデバイスでアドレス変換を提供することにより、遅延を短縮するだけでなく、デバイスの統合も可能になり、投資支出と運用コストを大幅に削減できます。マルチキャスト NAT は、データのサブスクライブまたはパブリッシュを行うユーザーにとって有益な機能であり、トラフィックを変換して単一の送信元から送信されたように見せることができます。また、マルチキャスト・グループと宛先を変換することにより、IP インフラストラクチャでの競合を回避したり、複数のポイントで受信される同一のトラフィックを区別したりすることもできます。

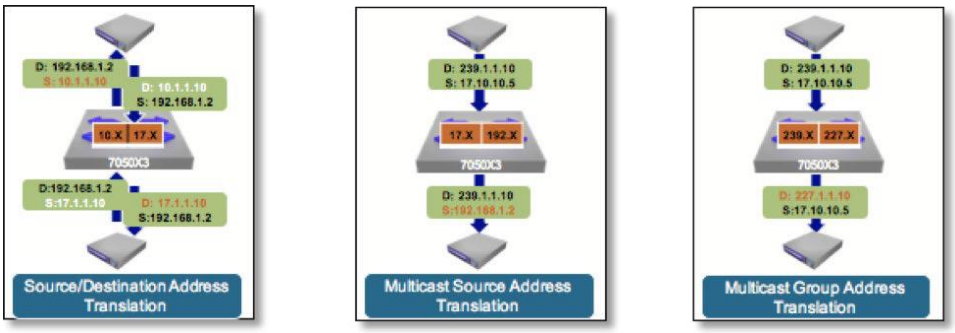

図 25：7050X3 のネットワーク・アドレス変換（NAT）

802.3by IEEE 25GbE の仕様

7050X3 シリーズは、IEEE 802.3by 25 ギガビット・イーサネット規格を完全にサポートしており、長期的な投資保護を実現します。既存の 25G デバイスとの下位互換性のために、25G および 50G コンソーシアム策定規格にも準拠しています。

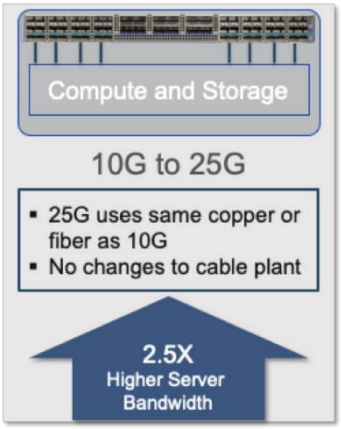

図 26：10GbE から 25GbE に移行する利点

増加し続けるトラフィックとストレージ容量の拡大により、ハイパフォーマンス・ネットワークの需要が高まっています。25GbE を導入することで、同じ使い慣れたデュプレックス・ケーブル・インフラストラクチャを使用して、10GbE の 2.5 倍のパフォーマンス向上を実現できます。10G/25GbE モードのサポートにより、高い費用をかけてネットワークをアップグレードしなくても、必要に応じた移行ができるため、これまでの投資が無駄になりません。

サーバーとストレージを 25GbE に移行すると、次のような利点があります。

- 高性能システムで使用可能なすべての帯域幅を使用して、スイッチとサーバーのスループットと効率を最大化できます。
- 複数の 10G ポートを使用する場合と比べ、少ないケーブル数とスイッチ・ポート数で帯域幅を増大でき、投資支出を抑えることができます。
- 代替の 40G よりも消費電力と冷却要件が低いため、運用コストを低減できます。
- 25GbE は同じスイッチ技術とケーブリングを使用して 2.5 倍のスループットを実現するため、パフォーマンスのビットあたりのコストを最小限に抑えることができます。

7050X3 シリーズは、25G コンソーシアム策定規格と IEEE 802.3by 25GbE 規格の両方に完全に準拠しており、多くの 10G ネットワークに容易な移行パスを提供します。25GbE インターフェイスは、多様な 10G SFP+光ファイバーおよびケーブルとの下位互換性を持つため、各ポートが他のポートから独立してデュアル・スピードに対応できます。10G から 25G へのインターフェイス速度変更は、他のポートのトラフィックを中断することなくヒットレスで実行されるため、徐々に移行を進めることができます。また、7050X3 シリーズのすべての 100GbE ポートは、パラレル光ファイバーとブレークアウト・ケーブルを使用する 10G、25G、50G に加えて、40G および 100G に対応し、合計 5 種類のインターフェイス速度を利用できます。

ワイヤスピード暗号化

7050CX3M-32S は、業界標準 IEEE 802.1AE（MAC セキュリティ標準、MACsec）の機能を提供します。組み込みの MACsec 機能により、外部暗号化デバイスが不要になり、侵入、受動的盗聴、その他のプレイバック攻撃に対するセキュリティが提供されます。MACsec 暗号化は、規制コンプライアンス要件を満たし、パフォーマンスを損なわないデータ保護を提供します。

MACsec は実績ある暗号化技術でトラフィックを保護することにより、シンプルで信頼性の高いスケーラブルなデータセンター相互接続を実現し、リーフとスパインからなるデータセンター設計の階層間リンクを保護します。

Arista EOS：スケール、安定性、拡張性のためのプラットフォーム

Arista 7050X3 シリーズの中核には Arista EOS®（Extensible Operating System）があります。アリスタが 2004 年の創業以来、コア・テクノロジーのイノベーションを利用して一から構築してきた EOS には、1,000 万行以上のコードと、1,000 人年を超える高度な分散システム・ソフトウェア・エンジニアリングの成果が詰まっています。EOS はオープンな標準ベースの手法で構築され、その最新アーキテクチャは高い信頼性を実現し、すべてのシステム・レベルで独自にプログラム可能になっています。

EOS は、クラウド・ネットワークに存在する 2 つの基本的な課題を解決するために開発されました。その課題とは、ノンストップの可用性を実現することと、質の高いソフトウェアを短期間で開発することです。アリスタは自社のエンジニアが 30 年以上におよぶネットワーキング製品の開発で培ってきた経験と、最先端のオープン・システム・テクノロジーおよび分散システムを利用して、クラウド時代に適したオペレーティング・システムの開発に一から着手しました。

その基盤になっているのは、状態情報をパケット転送から切り離し、プロトコル処理やアプリケーションのロジックから独立した形で管理する、マルチプロセス型の独自の状態共有アーキテクチャです。EOS では、システム状態とデータは効率性の高い一元化されたシステム・データベース（SysDB）に保存されます。保存されたデータには、自動化されたパブリッシュ/サブスクライブ/通知のモデルを使用してアクセスします。この独特のアーキテクチャ設計原則は、当社ソフトウェアの自己回復型の耐障害性、ソフトウェア・メンテナンスの容易化やモジュールの独立性、全体的なソフトウェア品質の向上、お客様の求める新機能の迅速な市場投入に貢献しています。

Arista EOS の開発アプローチは、1980 年代に開発されたネットワーク・オペレーティング・システムのものとは大きく異なります。かつてのネットワーク・オペレーティング・システムは、システム状態を個々の独立プロセス内に埋め込み、システム全体での状態管理のためにプロセス間通信（IPC）メカニズムを多用し、サブシステムの統合は自動化された構造化コアを使わず手動で行うというものでした。こうした旧式のネットワーク・オペレーティング・システムでは、大規模ネットワークでの処理や、システム・プロセスの障害や再起動時に動的なイベントが発生するため、リカバリが不可能ではないにしても、困難になる場合があります。

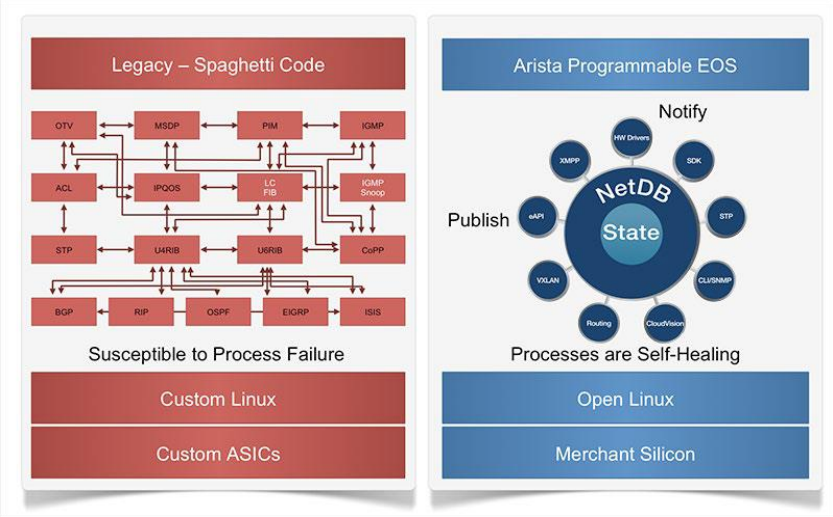

図 27：ネットワーク・オペレーティング・システムに対する旧来のアプローチ（左）、Arista　EOS（右）

Arista EOS は、オープンソース分野の教訓に従い、改変されていない Linux カーネル上に構築されています。また、Linux シェルおよびユーティリティに対しては、完全なセキュア・アクセスが維持されています。そのため EOS では、Linux コミュニティのセキュリティ、機能開発、ツールを継続的に利用できます。これは、オリジナルの OS カーネルを改変したり、十分なメンテナンスが行われていない古いバージョンの UNIX をベースにしたりする旧来のアプローチとは大きく異なる点です。EOS は Docker コンテナなどをネイティブにサポートするため、Arista スイッチ上のアプリケーションの開発や展開はシンプルです。　Arista EOS は、単純ながら強力なアーキテクチャ的アプローチをとることで、新しい機能のスピーディな提供を可能にする、高品質のプラットフォームを実現します。

EOS はすべてのレベルで拡張性があり、すべてのレベルでオープン API を使用できます。管理プレーン、コントロール・プレーン、データ・プレーン、サービス・レベルの拡張性、アプリケーション・レベルの拡張性のほか、シェルレベルのアクセスを含むすべての Linux オペレーティング・システム機能へのアクセスが提供されます。Arista EOS は、ネットワークのエンジニアリングおよび運用のニーズを満たすために、変更されていない Linux アプリケーションや増え続けているオープンソース管理ツールを用いて拡張することができます。

EOS API（eAPI）、OpenConfig、EOS SDK などのオープン API は、詳細なドキュメントと、構成、管理、監視に対する広範なプログラムによるアクセスを提供することで、リアルタイムのネットワーク・テレメトリのストリーミングを可能にします。これは、従来のポーリング・メカニズムに代わる優れた手段になります。

まとめ

7050X3 は、Smart System Upgrade、LANZ、ネットワーク・テレメトリ、パケット・タイムスタンプなど、他の Arista X シリーズ・システムで既にサポートされている、一貫性のある転送アーキテクチャと高度な EOS 機能セットをサポートします。運用と機能の一貫性を維持しながらパフォーマンスとスケールを強化するため、新製品導入の際に通常行われる評価の時間を短縮できます。さらに、7050X3 システムは既存のネットワークにシームレスに接続できます。

7050X3 プラットフォームは、高いパフォーマンスとスケール、低遅延、電力効率の向上、一貫性のある機能とネットワーク・イノベーションを提供し、大規模なエンタープライズ、ビッグデータおよび機械学習環境や、サービス・プロバイダのエッジ・ネットワーキングの役割の進化に最適です。

アリスタネットワークスジャパン合同会社

〒100-0004 東京都千代田区大手町 1-7-2 東京サンケイビル 27F
Tel:03-3242-6401

西日本営業本部
〒530-0001 大阪市北区梅田 2-2 ヒルトンプラザウエストオフィスタワー 19F
Tel: 06-6133-5681

お問い合わせ先

Japan-sales@arista.com

www.arista.com/jp

Copyright © 2023 Arista Networks, Inc.
Arista のロゴ、および EOS は、Arista Networks の商標です。その他の製品名またはサービス名は、他社の商標またはサービス商標である可能性があります。

2022 年 10 月