

分散クラウドコンピューティング、 およびそれがデータセンター内の ケーブルインフラストラクチャに及ぼす影響



目次

高頻度取引.....	4
クラウドゲーム	4
高性能計算.....	4
より多くの処理能力をよりコンパクトなスペースに収容.....	5 — 6
プラグ接続かケーブルか	7 — 8
Cisco 部品番号.....	9 — 10
パンドウイト部品番号	10
まとめ	11

高頻度取引、高性能計算、AI、ゲームなどのアプリケーションでは、高度な数値計算が実行され、低遅延が要求されます。こうしたワークロードの処理を最適化するには、計算処理の負荷の大部分を、データが発生するネットワークのエッジ方向に移動させる必要があります。このクラウドの分散化によって、データセンター運用者側には、データセンターの処理能力を増強し、上記のアプリケーションの要求を満たす方法を見つけるという課題が発生します。本書では、エッジコンピューティングへの移行を生み出すさまざまなアプリケーション、およびそれによって空間的な制約条件の厳しい環境のケーブルインフラストラクチャがどのような影響を受けるかについて説明します。

これまでのクラウドコンピューティングのアーキテクチャでは、すべてのデータを集中的に計算および保存します。このタイプのクラウドは、膨大な計算および保存能力を活かして、大規模なデータ分析や保存に対応できます。こうしてインフラストラクチャのインテリジェンスを引き続きクラウド内に保持する一方、厳しい遅延要件のあるデータはエッジコンピューティングモデルを使用して処理します。エッジコンピューティングは、データの発生源に物理的に近接した場所でデータを処理する方法で、遅延と帯域幅消費の低減に有効です。

コンピューティングリソースの設置スペースを拡大すれば、エッジでの処理能力を最大化できます。つまり、データセンターに計算サーバー、ストレージ、ネットワーク機器が収納されたラックを追加するというアプローチですが、コロケーションまたは自社所有の小規模データセンター環境では、物理的にも経済的にも実行可能とは思われません。データセンター運用者にとって代替の選択肢となるのが、密度を高めたラックを構築して、現在の設置スペースを活用する方法です。ラック密度の最大化は、計算サーバー、ストレージ、とりわけネットワークケーブルインフラストラクチャに影響を及ぼします。

図 1 に示すエッジコンピューティングモデルでは、計算、保存、ネットワーキングなどの機能がデータの発生源に近い場所に移されています。

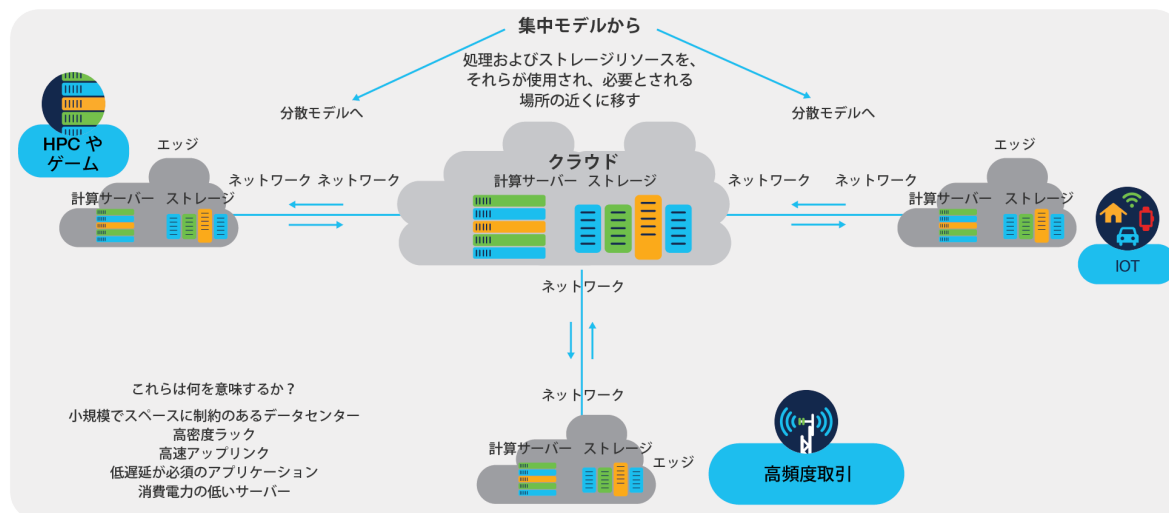


図 1. エッジコンピューティングによる集中モデルから分散モデルへの移行

図 1 のように分散モデルで複数の小規模なデータセンターを運用するメリットは、データソースまたはユーザーとの距離が近いゆえに、クラウドプロバイダーからエンドユーザーへのコンテンツ配信を大幅に高速化できることです。

エッジコンピューティングモデルのメリットを最大限に享受できるアプリケーションは、高性能計算 (HPC)、クラウドゲーム、高頻度取引 (HFT) など、どれも高度な数値計算と低遅延が要求されるアプリケーションです。

高頻度取引

高頻度取引では、秒単位で何百万という注文を実行する必要があります。金融機関は、トレーディングオフィス街の取引所にできるだけ近い不動産を購入するか、さもなければサービスプロバイダーのデータセンター内のスペースを借りるコロケーションモデルを採用する必要があると考えています。金融機関が提供するコロケーションサービスで、ネットワークその他の遅延が最小化されます。

コロケーションモデルや小規模なデータセンターを構築する手法ではスペースが限られるため、電力を大量に消費するラック間の密度がどうしても高まります。高頻度取引には、コアごとの性能、高帯域幅/スループット、低遅延が要求されます。高頻度取引に最適化されたデータセンターを設計するには、高密度ラックを構築する必要があります。今後、データセンターの設計手法は、I/O 接続数の増加、100 ギガビットイーサネット以上のダウンストリームポートを備える高密度 Top of Rack スイッチ、およびプラグ接続可能トランシーバまたはアセンブリ済みのアクティブ光ファイバーケーブル (AOC) を利用した接続インフラストラクチャに基づいた、サーバー密度の最大化に向かうでしょう。

クラウドゲーム

エンドユーザーへのサービス配信で遅延やタイムラグをどれだけなくせるかが成功の鍵を握っているクラウドゲームは、エッジコンピューティングのもう 1 つのユースケースとして注目されています。最速のインフラストラクチャを提供するゲーム会社が最大の成功を収めるでしょう。20 ~ 30 μ s の短縮が他社との決定的な差別化をもたらす可能性があります。ゲーム会社が低遅延、低タイムラグの配信を実現するには、自社でスペースを所有するか、コロケーションモデルで既存のデータセンターからスペースを借りる必要があります。自社でデータセンターを所有する方法には高いコストがかかり、管理も難しくなる可能性があるため、ほとんどのゲーム会社がコロケーションモデルを選ぶと見られます。

スペースを借りるこれらのゲーム会社が、密度を最大化したラックを構築して、強力な GPU (Graphical Processor Unit) を搭載したゲーミングサーバーをサポートできれば有利です。

高性能計算

高性能計算は、生命科学、地球科学、天気予報、エンジニアリング、シミュレーション、設計、防衛、セキュリティなどの科学的アプリケーションに適しています。高性能計算には、低遅延の維持と高いネットワークスループットの両方が必要です。

データの処理をエッジに移動する場合に 3 つのユースケースすべてに共通するのが、スペースの制約がある中で I/O 接続数を増やし、I/O 速度を上げる必要があるという点です。比較的小規模なデータセンターの場合、運用者に大量のラックを配備する余裕はなく、1 ラックあたりの処理能力を最大化するしかありません。

より多くの処理能力をよりコンパクトなスペースに収容:

今日の標準 19 インチラックの一般的な設計では、1 ラックあたり最大 32 台のサーバーを収容でき、図 2 に示すように、ネットワークインターフェイスカードから送信速度 1 G または 10 G の I/O 接続で Top of Rack スイッチへのアップストリームを実現できます。

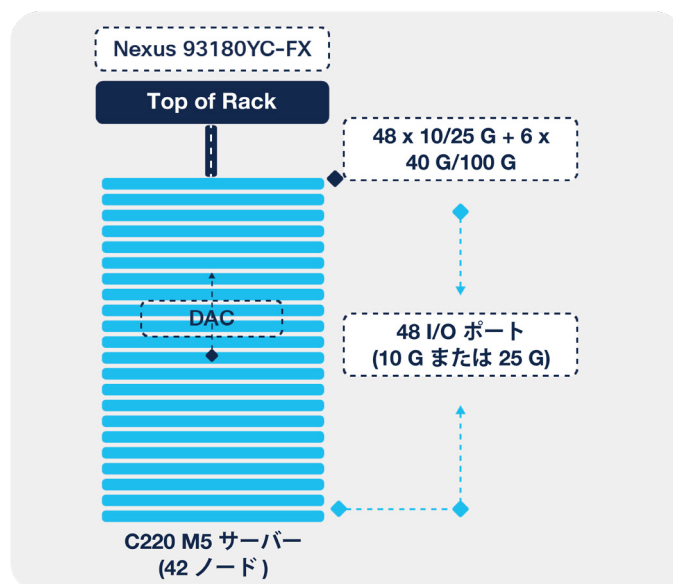


図 2. 1 RU ラックサーバーを使用したサーバー接続

図 2 の例で示す既存ラックでは、Cisco 社の C220 M5 サーバーから速度 10 G または 25 G で Nexus 93180YC-FX Top of Rack スイッチにアップストリーム接続しています。この例では、48 x 25 G ポートの 2:1 オーバーサブスクリプションモデル、または Top of Rack スイッチから送信される 48 x 10 G ポートのノンブロッキングモデルを想定しています。

もっと多くのリソースを消費し、低遅延要件が厳しいワークロードや計算タスクを処理する場合、さらに多くの処理能力をよりコンパクトなスペースに収容するアプローチが考えられます。この手法でデータセンターの密度を最大化すると、物理インフラストラクチャの面でコストを削減でき、スペースが限られているデータセンターでも状況に対応できるようになります。密度を最大化するこうした設計によって、管理するラックの数が減るため、運用も簡素化されます。

処理能力を増強する最初のステップは、1 ラックあたりのサーバーの数を増やすことです。これは、高密度ラックサーバーを使用すれば実現できます。

今日、Cisco 社をはじめとする各ベンダーは、高密度 2 RU のラックサーバーシャーシを提供しており、1 RU 内の計算ノードの数を倍にすると同時に、Top of Rack スイッチへの高速 I/O 接続の数も倍にできます。

Cisco UCS C4200 シリーズラックサーバーシャーシは、モジュラー式の高密ラックサーバーシャーシで、最小 2 個、最大 4 個の UCS C125 M5 ラックサーバーノードをサポートし、高密度計算フォームファクタが必要な環境での使用に最適化されています。

高密度ラックサーバーを使用すると、計算ノードの数が増え、表 1 に示すように 1 ラックあたりの I/O 接続数が 2 倍になります。

表 1. サーバー接続比較表

	UCS C4200 (C125)	UCS C220
ラックあたり最大サーバー数	84	42
ラックユニットあたりサーバー数	2	1
ラックあたり最大 I/O 数	168	84

ラック内の I/O 接続の数が増えると、Top of Rack スイッチの能力も増強する必要があります。Top of Rack スイッチをアップグレードしてサーバーへのダウンストリームで 100 G ポートをサポートすれば、100 G ポートを 4 x 25 G に分岐させて I/O 接続数を 128 まで増やせるようになります。

図 3 の例では、速度 10 G でアップストリーム接続する Cisco UCS C220 M5 サーバーから、Nexus 9336C-FX2 Top of Rack スイッチにアップストリーム接続する高密度 UCS C4200 モジュール式ラックサーバーにアップグレードして、I/O 接続数を 2 倍にしています。

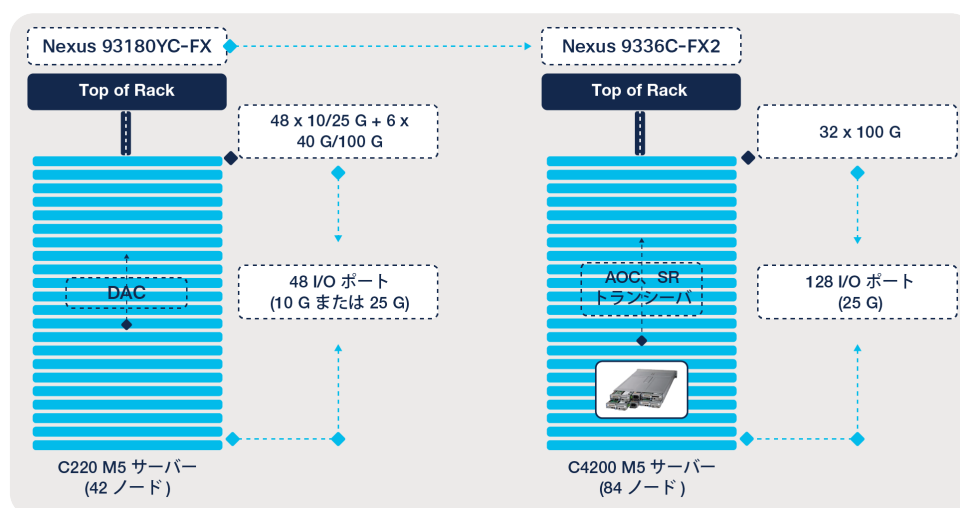


図 3. マルチモードシャーシを使用したラック密度の最大化

接続要件の高度化に取り組む 1 つの選択肢として、Top of Rack スイッチの 100 G ポートから 4 x 25 G ブレークアウトケーブルを使って速度 25 Gbps のダウンストリーム構成にするとする方法があります。現在、Cisco 社は、非常に短いリンク用に 100 ギガビットイーサネットから 4 x 25 ギガビットイーサネットへのダイレクトアタッチ銅線 (DAC) ブレークアウトケーブルを提供し、ラック内部および隣接するラック間を接続するコスト効率の高い方法を提案しています。これらのブレークアウトケーブルの一方の端を Cisco スイッチの 100 G QSFP28 ポートに接続し、他方の端にある Cisco サーバーの 4 つの 25 G SFP28 ポートに接続します。

別の選択肢は、プラグ接続可能な QSFP-100G-SR4-S 光ファイバーを Top of Rack スイッチと組み合わせ、ブレークアウトケーブルを使用して速度 25 G でサーバーにダウンストリームする方法で、25 G 光ファイバーと Cisco SFP-25G-SR-S、または到達距離要件が 100 m を超える場合は Cisco SFP-10/25G-CSR-S を使用します。

プラグ接続可能モジュールは、スイッチ、ルーター、サーバーのポートに差し込んで接続できるトランシーバです。データセンターで使用するこの種のプラグ接続可能な光ファイバーの一般的なフォームファクタは、以下のとおりです。

- SFP (Small Form-factor Pluggable): 1 ギガビットイーサネット速度、通常は LC ファイバーコネクタ
- SFP+ (Small Form-factor Pluggable Plus): 10 ギガビットイーサネット速度、通常は LC ファイバーコネクタ
- QSFP (Quad Small Form-factor Pluggable): 40 および 100 ギガビットイーサネット、MPO
- QSFP-DD (Quad Small Form-factor Pluggable – Double Density): 200 および 400 ギガビットイーサネット、MPO または LC

プラグ接続かケーブルか

銅線/AOC アセンブリ済みケーブルとプラグ接続可能光ファイバーのどちらを配備するかは、コストと管理性の要素に左右されます。将来のアップグレードの際に投資が無駄にならないかも考慮されます。まったくの新規案件でコストだけが考慮要素だとすると、ほとんどのデータセンター運用者は銅線 (DAC) を選ぶでしょう。とりわけ、Top of Rack スイッチとサーバー間の接続がすべてラック内で行われる場合はそうです。

ラック内の密度要件が高まると、プラグ接続可能トランシーバと比べて、銅線ケーブルが発展性のあるソリューションかどうか問題になります。密度が高まるにつれ、銅線ケーブルの重量が無視できない要素になるからです。特にケーブルが長い場合は問題で、増加した重量のためにスイッチのポートに応力が発生し、それが誤動作の増加につながる可能性があります。銅線ケーブルが多くのスペースを占有すると空気の流れが妨げられ、それが機器の過熱を引き起こして、ハードウェアの故障の原因になることもあります。アクティブ光ファイバーケーブル (AOC)、プラグ接続可能トランシーバのどちらでも、銅線による応力とかさばりを減らせます。

ケーブルとトランシーバのどちらにするかを決定する際には、トラブルシューティングやケーブルの交換、特に到達距離が長い場合の運用面の課題も考慮に入れてください。複数のラックを横断しているケーブルの交換は、コストも時間もかかります。トランシーバを使用している場合は、エンドポイントの交換作業で済む可能性があります。プラグ接続可能トランシーバのほうが初期費用は高いですが、万一接続トラブルが発生しても、一方または両方のエンドポイントを交換するだけで、接続ケーブルまで交換する必要はありません。これで運用コストを大幅に削減できます。

最後にもう 1 つ重要な要素となるのが、インフラストラクチャのアップグレード時の投資の保護です。ファイバーケーブルプラントを敷設すると高い費用がかかるうえに、データセンター運用に問題が起こってダウンタイムが発生する可能性もあります。プラグ接続可能トランシーバを使用すれば、たとえば、データセンター運用者が 10 Gbps トランシーバを取り外して 25 Gbps にアップグレードするだけで済むので、既存のインフラストラクチャを活用できます。銅線または AOC を使用していると、速度のアップグレード時にケーブル全体の交換が必要になります。つまり、初期投資では、銅線または AOC のほうがプラグ接続可能トランシーバより有利ですが、帯域幅のアップグレードが必要になるとその利点は失われます。

Cisco[™] 社は、100 G から 25 G へのブレイクアウト要件用に、以下のプラグ接続可能トランシーバを提供しています。

表 2. 10/25/100 Gbps 対応 Cisco 社製プラグ接続可能トランシーバ

SKU	詳細
QSFP-100G-SR4-S	Cisco 100GBase SR4 QSFP トランシーバ、MPO、OM4 MM5 100 m
SFP-25G-SR-S	Cisco 25GBase-SR SFP トランシーバ、OM4 MM5 100m
SFP-10/25G-CSR-S	Cisco 10/25GBase-CSR SFP28 トランシーバ、OM4 MM5 400m

本書ではここまで、高性能計算、クラウドゲーム、高頻度取引などのアプリケーションに適用できる、サーバー、スイッチ、接続の論理レイアウトに注目してきました。ここからは、物理的なインフラストラクチャレイアウト、および上記アプリケーションの実現に必要な製品について説明します。パンドウイットは、Cisco 社のスイッチ、サーバー、その他のデバイスをサポートするインフラストラクチャのための完全なソリューションを提供しています。

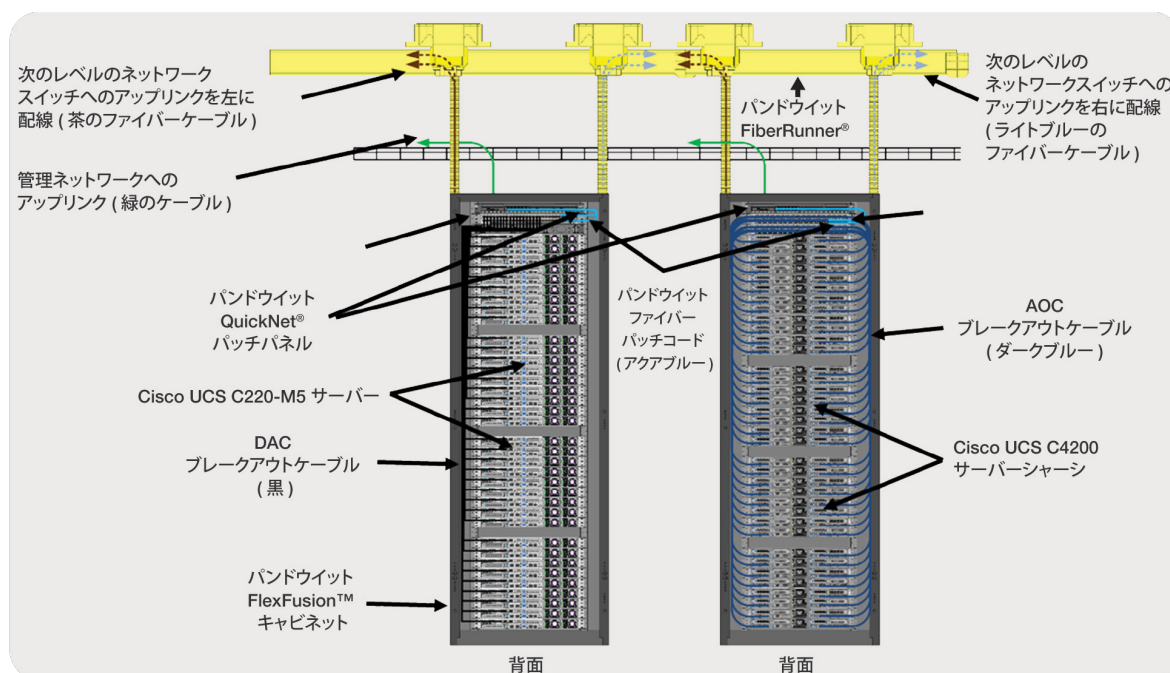


図 4. スペースに制約のあるラック

上の図 (図 4) には、サーバーとスイッチ間の 2 つの異なる接続方法と、アクセス/リーフスイッチから次のレベルのネットワークスイッチへのトランシーバおよびファイバー接続が示されています。

図の左側のキャビネットでは、Cisco UCS C220-M5 サーバーが Cisco Nexus 93180YC-FX スイッチに DAC ブレークアウトケーブルで接続されています。ここでは、スイッチの 100 G ポートから、二軸銅線ケーブルを使用するサーバーの 4 x 25 G ポートに分歧接続しています。DAC ブレークアウトケーブルは、Cisco 社とパンドウイトのいずれからも提供されています。部品番号については下掲の表をご覧ください。

図の右側のキャビネットでは、Cisco UCS C4200 シャーシサーバーが Cisco Nexus 9336-FX スイッチに AOC ブレークアウトケーブルで接続されています。ここでは、スイッチの 100 G ポートから、ケーブルの両端にアセンブリ済みのトランシーバとのファイバー接続を使用するサーバーの 4 x 25 G ポートに分歧接続しています。AOC ブレークアウトケーブルは、パンドウイトから最大 30 メートルのさまざまな距離で提供されています。部品番号については下掲の表をご覧ください。

どちらのキャビネットでも、QSFP-100G-SR4-S トランシーバと PanMPO OM4 12 芯パッチコード、およびスイッチから次のレベルのネットワークスイッチへのアップリンクのトランクを使用しています。

ネットワークインフラストラクチャの接続と管理に使用される主要コンポーネントは以下のとおりです。

- QuickNet™ PanMPO™ パッチコードおよびトランク: 細径できわめて柔軟性に優れ、短いケーブル施工距離に適しています。これらのケーブルで使用する PanMPO ファイバーコネクタは、現場でピンあり/ピンなしおよび極性の変換が可能な特許取得済みの MPO 設計です。
- QuickNet™ パッチパネル: RU あたり 96 芯 (48 デュプレックスポート) の高密度パッチパネルでラックスペースを節約できます。ストレート型またはアングル型のパッチパネルとして提供され、曲げ半径を適切に調整し、必要な水平ケーブル管理パネルを最小限に抑制できます。
- FiberRunner™ ケーブルルーティングシステム: チャンネル、フィッティング、およびブラケットから成るシンプルな敷設と構成が特長のシステムで、光ファイバーケーブルと高性能ツイストペアケーブルを分離、配線、保護するように設計されています。

Cisco 部品番号

表 3. Cisco Nexus スイッチ部品番号

スイッチ	詳細
Cisco Nexus 93180YC-FX	48 x 1/10/25-Gbps SFP/SFP28 ポート、および 6 x 40/100-Gbps QSFP28 ポート
Cisco Nexus 9336C-FX2	36 x 40/100-Gbps QSFP28 ポート

表 4. Cisco UCS サーバー部品番号

サーバー	詳細
UCSC-C220-M5SX	UCS C220 M5 SFF 10 HD (CPU、メモリ、HD、PCIe、PSU なし)
UCSC-C220-M5SN	UCS C220 M5 SFF 10 NVMe (CPU、メモリ、HD、PCIe、PSU なし)
UCSC-C4200-SFF	UCS C4200 ラックサーバーシャーシ、ワンレールキット、およびベゼル
UCSC-C125	UCS C125 M5 ラックサーバーノード、AMD EPYC CPU ベース

表 5. Cisco ネットワークインターフェイスカード部品番号

ネットワークインターフェイスカード	詳細
UCSC-PCIE-C25Q-04	Cisco VIC 1455 VIC PCIE – Quad-Port 10/25G SFP28
UCSC-MLOM-C25Q-04	Cisco UCS VIC 1457 Quad-Port 25G SFP28 mLOM
UCSC-PCIE-ID10GF	Intel X710-DA2 Dual-Port 10Gb SFP+ NIC
Cisco UCSC-PCIE-QD25GF	Intel X710 Quad-Port 10Gb SFP+ NIC
UCSC-PCIE-ID25GF	Qlogic QL41212H Dual-Port 25Gb NIC
UCSC-P-M4D25GF	Intel XXV710 Dual-Port 25Gb SFP28 NIC
UCSC-P-M5D25GF	Mellanox MCX4121A-ACAT Dual-Port 10/25G SFP28 NIC
UCSC-P-M5D25GF	Mellanox CX-5 EN MCX512A-ACAT 2x25/10GbE SFP PCIe NIC

表 6. Cisco 100 G 銅線ブレイクアウトケーブル部品番号

100 G 銅線ブレイクアウトケーブル	詳細
QSFP-4SFP25G-CU1M	Cisco 100GBase QSFP to 4xSFP25G パッシブスプリッターケーブル、1 m
QSFP-4SFP25G-CU2M	Cisco 100GBase QSFP to 4xSFP25G パッシブスプリッターケーブル、2m
QSFP-4SFP25G-CU3M	Cisco 100GBase QSFP to 4xSFP25G パッシブスプリッターケーブル、3m
QSFP-4SFP25G-CU5M	Cisco 100GBase QSFP to 4xSFP25G パッシブスプリッターケーブル、5m

表 7. Cisco プラグ接続可能トランシーバ部品番号

プラグ接続可能トランシーバ	詳細
QSFP-100G-SR4-S	Cisco 100GBase SR4 QSFP トランシーバ、MPO、OM4 MM5 100 m
SFP-25G-SR-S	Cisco 25GBase-SR SFP トランシーバ、OM4 MM5 100m
SFP-10/25G-CSR-S	Cisco 10/25GBase-CSR SFP28 トランシーバ、OM4 MM5 400m

パンドウイト部品番号

表 8. パンドウイトのケーブル、キャビネット、およびファイバー接続部品番号

製品	詳細
DAC ブレークアウトケーブル 100G to 4 x 25G	
PHQ4SFP2A1MBL	QSFP28 to SFP28 DAC ブレークアウトケーブル、30 AWG、黒、1 m
PHQ4SFP2A2MBL	QSFP28 to SFP28 DAC ブレークアウトケーブル、30 AWG、黒、2 m
PHQ4SFP2A3MBL	QSFP28 to SFP28 DAC ブレークアウトケーブル、30 AWG、黒、3 m
PHQ4SFP2C4MBL	QSFP28 to SFP28 DAC ブレークアウトケーブル、26 AWG、黒、4 m
PHQ4SFP2C5MBL	QSFP28 to SFP28 DAC ブレークアウトケーブル、26 AWG、黒、5 m
AOC ブレークアウトケーブル 100G to 4 x 25G	
AZ83NQ2S2AQM002	アクティブ光ファイバーケーブルアセンブリ、QSFP28 to SFP28、100-4x25G、OM4、アクア、1 m
AZ83NQ2S2AQM003	アクティブ光ファイバーケーブルアセンブリ、QSFP28 to SFP28、100-4x25G、OM4、アクア、3 m
AZ83NQ2S2AQM005	アクティブ光ファイバーケーブルアセンブリ、QSFP28 to SFP28、100-4x25G、OM4、アクア、5 m
AZ83NQ2S2AQM007	アクティブ光ファイバーケーブルアセンブリ、QSFP28 to SFP28、100-4x25G、OM4、アクア、7 m
キャビネットおよびファイバー接続	
QPP24BL	QuickNet™ 24 ポートパッチパネル
FZTRP7N7NYNF003	OM4 12 芯 PanMPO™ インターコネクト、ピンなし - ピンなし、メソッド B、3 フィート
FQMAP66CG	SFQ シリーズ光ファイバーアダプタパネル - 6 MPO アダプタ - メソッド B 接続用
QPPBBL	QuickNet™ パッチパネル用ブランクパネル
XG74222BS0001	FlexFusion™ キャビネット 700 mm 42 RU、奥行き 1200 mm
CMPH1	ケーブル管理パネル 1 RU
TLBP1S-V	ツールレスブランクパネル 1 RU



まとめ

エッジコンピューティングによって、データセンターインフラストラクチャの構築および管理方法が変わりつつあります。密度を最大化したデータセンターをリモートのエッジに近い場所で運用して、遅延やタイムラグを減らす必要性に拍車をかけているのが、高頻度取引、クラウドゲームなどの高度な数値計算が実行されるアプリケーションです。集中モデルから分散モデルに移行するこの流れに伴い、計算サーバー、ストレージ、ネットワークインフラストラクチャをラックに収容する際にさまざまな新しい要件が発生します。データセンター運用者は、管理性と拡張性の高い、高密度ラックサーバー、高速 Top of Rack スイッチ、およびスペースを最適化したケーブルインフラストラクチャへのアップグレードによって、安全で信頼でき、あらゆる面で改善されたエンドユーザーエクスペリエンスを提供することができます。

スイッチやサーバーネットワークインターフェイスカードの接続に必要な光ファイバーやケーブルについて、詳しくは、以下のサイトをご覧ください。

Cisco 社の光ファイバーとデバイスの
互換性マトリックス:
<https://tmgmatrix.cisco.com/>

Cisco 社の光ファイバーと光ファイバーの
相互運用性マトリックス:
<https://tmgmatrix.cisco.com/iop>

パンドウイトと Cisco 社の提携:
www.panduit.com/panduit-cisco-alliance

お客様とのつながりこそ、 当社の最大の財産です。

当社には、お客様のインフラストラクチャ投資の効果を最大化するための知識と経験があります。

www.panduit.co.jp

ぜひご相談ください

www.panduit.com/jpn/ja/support/contact-us.html