



## VMware® ESXサーバによる ユニファイド・I/Oの重要性

### 1.0 なぜユニファイドI/O なのでしょう？

マルチ・コア CPUテクノロジーとCPUをより効率的に使用するVMWareのような仮想化ソフトウェアの進歩により、物理サーバは、より多くの仮想サーバ(仮想マシン)を扱えるようになります。これによりサーバはより多くのI/Oを必要とし、VMwareのベストプラクティスでは複数のI/Oアダプタを使用して仮想化サーバの要求を満たすことを推奨しています。データセンター基盤におけるサーバあたりのI/Oアダプタとケーブルの標準化を減らす唯一の方法は、高性能なユニファイド・I/Oソリューションを使用することです。このようなソリューションにより、I/O関連の資産と運用コストを減らし、環境に配慮したデータセンターを構築できるようになります。

### 1.1 ユニファイドI/O - 2つの世界

サーバでI/Oを統一するには、10GigE NICまたは、InfiniBand HCAを使う2つの方法があります。VMware ESXサーバ環境では、どのようなものを利用しなければならないのでしょうか？考えかたにもよりますが以下を参考にしてください。どちらの選択も可能です。(a) SAN (storage area network)の選択、(b) すでに利用可能なソリューションを利用する、または、新しいテクノロジーまで待つ (c) コストパフォーマンスの目標SANを使うには、サーバ利用率の増加、高可用性、障害回復へ対応する仮想マシン・マイグレーションを備えたVMware 仮想インフラストラクチャのような機能が必要になります。ファイバーチャンネルSANにしてもiSCSI SANを使うにしても、どのようなユニファイドI/Oテクノロジーが実際に利用できるかによって左右されます。ポートあたりのコスト(アダプタとスイッチのポート)とユニファイドI/Oテクノロジーの成熟度は、2番目に重要な検討材料になります。本書では、この3つの基準について説明します。

### 1.2 Fibre Channel SANに よるユニファイドI/O

ファイバーチャンネル(FC) SANはフォーチュン500社と、VMware仮想化製品を使った仮想化サーバ環境による大規模データセンターで広く採用されています。地理的に離れたサイト全体に広がるVMware 仮想インフラストラクチャ機能は、高可用性、VMotion®、Storage VMotion などを使うためにFC SANを必要とします。VMware ESX Server 3.5が動作しているホスト側では、FCアダプタ(HBA) はI/Oを統一することができません。選択肢は10GigEまたはインフィニバンドになります。FC SANサービスは、重要なストレージデータを扱うために(データやパケットの損失のない)高信頼なファブリックを必要とします。ベストエフォート型のEthernetは、データセンターのミッションクリティカルなアプリケーションのストレージトラフィックに適した高信頼性サービスを提供できません。一方、インフィニバンドは、FCよりも高いバンド幅で、必要とされる高信頼サービスを提供します。つまり、サーバ上のユニファイドI/OやVMware ESX Server 3.5上の複数の仮想マシンのI/O要求を満たせるハイ・パフォーマンスI/O、FC SAN接続が必須であるなら、現在、インフィニバンドが唯一のソリューションとなります。

インフィニバンド上でのエンド・ツーエンドFC SAN接続は、今日、多くのデータセンターで採用されています。

I/O Directorやgatewayソリューションは、多くのインフィニバンド・ソリューションベンダーから入手できます。このソリューションは、サーバへ接続する複数のインフィニバンドポートを使用しています。IT基盤側では、LANやFC SANへ接続するためのEthernetとFCポートを備えたラインカードやI/Oモジュールを含んでいます。これにより、サーバでは、ケーブル配線の複雑度を減らす“1ワイヤー”（最大70%）と、I/Oコストと電力を減らす“1アダプタ”（30-50%）を使用できるようになります。同時に、このようなソリューションでは、VMware ESX Server 3.5のVM上のアプリケーションが（ESX VM上のレガシーNICとHBAインタフェースで）透過的に動作できるフレキシブルな仮想NICと仮想HBAを公開します。VM用のI/OプロビジョニングとVMware VirtualCenterを使用する仮想インフラ機能も同様に透過的になり、使いやすいエンド・ツーエンドEthernetやFC SANを可能にします。

### 1.3 Fibre Channel over Ethernet (FCoE)は？

FCoEを主導する業種では、FC SAN接続を必要とする統一されたサーバI/O用途には、Ethernetでは欠陥があり、インフィニバンドが適していることを確認しています。FCoEは期待される一方、仕様はまだ策定中であり、ソフトウェア、ハードウェア、ソリューションサプライヤーのエコシステムもまだ発達段階にあります。エンドノードでは、FCoEアダプタによるソリューションは、まったく新しい種類のソフトウェアであり、OSベンダーによる採用やスタックの成熟には時間がかかります。インフラでは、Ethernet上で高信頼性サービスを提供するスイッチやゲートウェイは、今日、または、近い将来に利用できる製品がないため、新たに開発する必要があります。新しいテクノロジーが、エンドノードとインフラの同時アップグレードを必要とする場合、歴史的に、実際のデータセンターを運用できるようになるまで数年かかります。

### 1.4 10GigEによる LAN,iSCSI SAN ユニファイドI/O

iSCSI SANを利用している場合には、VMware ESX Server 3.5は、iSCSIとNetQueueを使ったI/Oの高速化による、10GigE NICを使ったユニファイド・I/Oソリューションを提案します。多くのI/Oベンダーは、VMからのLANとiSCSIトラフィック用にラインレートに近いスループットを発表しました。しかし、CPU使用率が高く、スループットを確保するには複数VMと8コアCPUベースのサーバを必要とします。VMware ESX Server 3.5で使われる10GigE NICは、複数のGigE NICを単一のアダプタに置き換え可能なソリューションを提供し、I/Oの性能向上とケーブル接続の複雑さを減らします。このようなソリューションは、PCIスロット数に制限のあるブレード・サーバにも向いています。

### 1.5 最高の製品

すべてのユースケースシナリオをカバーし、将来にわたり、最高の環境を提供するVMware ESX server上のユニファイドI/Oソリューションを使います。つまり、必要に応じてインフィニバンドと10GigEを使い分けます。Mellanox社ConnectX I/Oアダプタは、VMware ESXサーバ3.5環境で、このような柔軟性を提供します。ConnectX™ I/Oアダプタは、同じアダプタで10/20/40Gb/sのインフィニバンドHCAまたは10GigE NICとして構成できるので、プロトコルアジャイルです。プロトコルアジャイルは、VMware ESX Server 3.5環境で、適切なI/Oデバイスドライバとプロトコルを有効にするアダプタのファームウェア・スイッチにより実装されています。VMware ESX Server 3.5用のConnectXインフィニバンドソフトウェアは現在入手可能です。VMware ESX Server 3.5用ConnectX 10GigEソフトウェアは、2008年第2四半期に利用できるようになる予定です。

### 1.6 仮想マシンのスケールリング、サーバ利用率のアップ

VMware ESX Server 3.5で使われるConnectX 10GigE NICは、I/O性能を損なことなく物理サーバのVM数のスケールリングを提供します。例えば、10GigEのスピードは、VM数が（8コアCPUサーバの場合）5から16にスケールするときに維持されます。VMごとにline rateに近づく一方、30-40%のCPUサイクルはVMのアプリケーションのために残すことができます。総合的にみて、Mellanox社の10GigE NICは、I/Oボトルネックを取り除き、サーバあたりの仮想マシン数を増やすことで、サーバ稼働率の改善を手助けします。

## 1.7 コストパフォーマンス

インフィニバンドは市場に出ている今日最も成熟した10GB/s+のI/O技術で、2007年12月現在、およそ3百万ポートを出荷しています。そして、それはインフィニバンド製品が今日、第4世代になり、重要な機能のアップグレード、最適化と集積化されているという事実に反映しています。実際の使用において、ポートあたりのインフィニバンドコスト(スイッチとアダプタ)は、4 Gb/s FCの50%、10GigEの30%に達しています。

ConnectXインフィニバンドを使ったVMware ESX Server 3.5は、エンド・ツーエンドのFC SAN接続と最高のコストパフォーマンスを提供します。4 Gb/s FCに比べ、ESX VMで4倍のSAN I/Oスループットを実現します。また、GigE NICに比べ、ESX VMで3-4倍のLAN I/Oスループットを実現します。このように、単一のConnectX インフィニバンド アダプタで、複数のGigEとFCアダプタを置き換え、50%のI/Oコストと最大30%のI/O電力を削減できます。

## 1.8 将来への対応

10GigEでのFC SAN コンソリデーションが必要になったときでも、ConnectX 10GigE NICは、少ないCPU利用率で高いパフォーマンスを出せるように、CPUに負担をかけるSCSIとFC関連の機能をアダプタハードウェアへオフロードすることで、FCoEをサポートします。

PCI SIG I/O Virtualization—Single Root (SR)とCPU使用率の低減化、のような先進の標準化技術を使ったダイレクトVMアクセスをサポートするI/Oソリューションに対応できるように、ConnectXアダプタはPCI SIGで定義された仮想関数(VF)を実装しています。VFはI/Oアダプタ リソースリソースを分割し、ネイティブOS環境と同じようにVMが直接アクセスできるようにします。フルアイソレーション、プロテクション、パフォーマンスをVMへ提供するために、ConnectXは、VM毎の送信、受信キュー、高度なメモリ管理、フレキシブルなreceive side scalingとフィルタリングにより、これらの機能を補います。ハードウェアベースのI/O仮想化ソリューションは、仮想化ソフトウェアソリューションでそれらを使用できるようになったときに、サポートされるでしょう。

## 1.9 まとめ

VMware ESXサーバのユニファイドI/Oは、コストと電力の節約に大きく貢献します。10GigEとインフィニバンドは、現在の状況を見直すよい機会になります。ベストなコストパフォーマンスとエンド・ツーエンドFC SAN接続によるサーバI/Oの統合には、インフィニバンドが、現在と近い将来において、唯一使用可能なソリューションになります。10GigEアダプタは、iSCSI SAN環境でサーバI/Oを統合する場合に有用で、高スループットなVMが可能になります。Mellanox社のプロトコルアジャイルConnectXアダプタは、インフィニバンドと10GigEの両方を利用できるようにします。さらに、FCoEやPCI SIG IOV SRのような先進機能を取り込んでいるConnectXアダプタは、長期にわたり安心してお使いいただけます。