

次世代の通信インフラストラクチャーに対応する 効率的な仮想ネットワーク機能(VNF) アーキテクチャー

Hewlett Packard
Enterprise

中国电信
CHINA TELECOM
Beijing Research Institute

帯域幅の急増にネットワークが対応するためには、ネットワーク機能の仮想化(NFV)を
技術的・経済的に実現する、拡張性の高いアーキテクチャーが必要です。

著者 はじめに

Rajesh Narayanan

Hewlett Packard* Enterprise (HPE)
ディスティングイッシュト・テクノロジスト

Vinay Saxena

Hewlett Packard Enterprise
HPE フェロー & チーフ・アーキテクト CSB

Chuck Tato

インテル コーポレーション
ネットワーキング & NFV ビジネス部門ディレクター

Herman Nakamura

インテル コーポレーション
戦略的マーケティング

Haining Wang

China Telecom* Beijing Research Institute
シニアエンジニア & プロジェクト・マネージャー

Bo Lei

China Telecom Beijing Research Institute
IP & Future Network Research Center ディレクター

Shaoyang Rao, Ph.D.

China Telecom Beijing Research Institute
ネットワーク・テクノロジー & プランニング部門
副ディレクター

(謝辞を参照)

目次

はじめに.....	1
ネットワークとデータセンターの統合.....	2
効率的な VNF の実装が不可欠.....	2
CTNet2025: 2025 年までの China Telecom ネットワークの変革.....	3
拡張性と効率性の高い vBRAS アーキテクチャーへ.....	4
vBRAS のパフォーマンス結果.....	6
効果的な vBRAS で相対コストを削減.....	7
他の CUPS 実装との比較.....	8
まとめ.....	9
謝辞.....	9
参考資料.....	9
詳細情報.....	9

通信サービス・プロバイダーは、利用ニーズの急増に直面しています。例えば、世界の IP トラフィックは、今後の 5 年間で 3 倍に増えると予想されています。¹ モバイル加入者、IoT デバイス、新しいサービスの急増により、ネットワークの構築(設備投資)および運用(運用コスト)の複雑さとコストも増えています。またより多くの帯域幅へのニーズも高まり続けています。そのため世界的な通信プロバイダー各社は、需要に応えるとともに、より低コストの IT インフラストラクチャーを活用しようと、自社のインフラストラクチャーを仮想化する方法を模索しています。現行のシステムに使われている独自規格のハードウェアは高額であり、また簡単に拡張することができません。IT プロフェッショナルは、ハードウェアを効率的に利用するための有効な手段として、一般的に仮想化を用います。IT やクラウドの導入環境に仮想化を用いることにより、複数のアプリケーションでコンピューティング・インフラストラクチャーを共有し、その結果ハードウェア・コストを減らすことができます。サービス・プロバイダーは、標準サーバーと高速化、仮想化されたネットワーク機能を用いてパフォーマンスと拡張性のニーズにバランス良く応えられる、新しいインフラストラクチャーを必要としています。

ネットワーク機能仮想化(NFV)は、サービス・プロバイダーが真の仮想化インフラストラクチャーに移行するための革新的なテクノロジーとして登場しました。AT&T や China Telecom など、仮想化の実現方法を模索するプロバイダーによって、NFV への取り組みは勢いづいています。こうした多くの企業の多大な尽力により、NFV が目指す未来像は着実に実現しつつあります。

China Telecom では、トラフィックやサービスの増加に効率的に対応する新しい統一要件の必要性を認め、自社のネットワーク・アーキテクチャーを再構築するためのユースケースを開発し、目標を定めました。この青写真をもとに、Hewlett Packard Enterprise (HPE)、インテル、China Telecom の Beijing Research Institute は、次世代ネットワークのニーズへの対応が可能な効率性と拡張性の高いアーキテクチャーの概念実証を開発しました。

このホワイトペーパーでは、サービス・プロバイダーが直面している主な問題を明示するとともに、相対コストの問題に効果的に対処する仮想ネットワーク機能(VNF)フレームワークの拡張アーキテクチャーについて説明します。このアーキテクチャーは、ベンダーに依存しないアプローチを促進し、多様性に富む充実したマルチベンダー・エコシステムによってサポートされる、クラウドネイティブかつ実用的な VNF の先駆けとなることを目的としています。

「グローバルな通信産業は、閉鎖的で業種固有のアーキテクチャーから業種を問わないオープンなアーキテクチャーへ移行する時期にきています。これは、多大なチャンスをもたらす一方で、かつてない課題も伴います」 — China Telecom

ネットワークとデータセンターの統合

サービス・プロバイダーは現在、いくつかの大きな課題に直面しています。

- 加入者の増加は横這い状態に達したものの、ネットワーク帯域幅の急拡大に対処しなければなりません。例えば、China Telecomでは需要の増加から、バックボーンのネットワーク帯域幅消費が過去3年で42%増加し、また大都市圏のネットワーク・トラフィックはバックボーン・トラフィックの1.5～2倍の割合で増加しています。²
- 加入者の需要も急速に変化しており、4K TV、8K TV、仮想現実、拡張現実、ますます大規模化するオンライン・バックアップ・ストレージといった新しいユースケースへの対応を迫られています。
- 大規模なモビリティは、データや接続の送信元と宛先が常に移動状態にあるため、ビジネスの複雑性が大幅に増えています。
- IoTテクノロジーや先進運転支援システム(ADAS)テクノロジーにより、接続される「モノ」の数と、連続的に生成されるデータの量は増大し続けています。

「サービス・プロバイダーの最大のコアリソースはネットワーク自体です」— China Telecom

クラウド・コンピューティングや仮想化が誕生した目的は、IT機能の急拡大に効率良く対応することです。しかし現在のところ、クラウドとネットワークは柔軟に連携できていません。コンピューティングやストレージを仮想化するためのクラウド・テクノロジーが高度に洗練され動的になると同時に、ソフトウェア・デファインド・ネットワーク(SDN)を用いたネットワーク構成の自動化は大幅に進歩しています。その一方で、ネットワークの仮想化は遅れています。ネットワークのワークロードは、エンタープライズ・アプリケーションのワークロードとは全く異なります。新しいアプリケーションが高いパケットレートを必要とするのに伴い、QoS/トラフィック・シェーピング、仮想プライベート・ネットワーク(VPN)、ファイアウォール、ネットワーク・アドレス変換、カプセル化/カプセル化解除、ディープ・パケット・インスペクション(DPI)、リアルタイム・モニタリングといった無数のパケット処理アプリケーションの回線速度を維持することが非現実的になり、結果として設備投資と運用コストが大幅に増えていきます。これは、高度に動的で仮想化されたインフラストラクチャーの導入にも直接影響するため、既存の通信ネットワークの変革に遅れが生じる要因となっています。

効率的なVNFの実装が不可欠

VNFとは、独自規格に基づく専用ハードウェア向けに構築されたソフトウェアから、標準サーバー上で動作する仮想ソフトウェアへと機能を移行したものとします。そのためVNFは、独自規格で値段も高額なハードウェアのニーズを減らします。しかしながら、現在のVNF実装の多くは、新しい仮想化環境向けに十分に最適化されておらず、独自規格に基づく従来環境よりもパフォーマンスが低くなっています。NFV、SDN、オーケストレーションを利用すれば、通信クラウド導入後の運用上の複雑性がかなり軽減されますが、パフォーマンス不足を克服するために必要なインフラストラクチャーのコストの高さは変わりません。NFV導入の主な課題は、全体的な設備投資と運用コストを減らすことです。しかし、クラウドベースのNFVモデルはまだ新しいため、コスト削減を達成するために最適化しなければならない変動要因はいくつもあります。このホワイトペーパーでは、効率的な設計によって競争力の高いNFVを実現する方法について説明します。

効率的な(つまりクラウドネイティブの)VNFの必要性は広く認識されていますが、このようなVNFの開発方法や、アーキテクチャーに関する明確なガイドライン/ベンチマークの不足についてはほとんど理解されていません。これまで標準化団体のガイドラインに頼ってきたベンダー各社は岐路に立たされています。適合性が高く、採用しやすく、シンプルに大規模導入できるVNFを、ベンダー・コミュニティが構築しなければなりません。効率的なVNFに必要なものは、単にオートメーションとオーケストレーションだけではありません。VNFは、少なくとも独自規格のハードウェア機能を上回るもの、できればはるかに優れたものでなければなりません。

VNFの効率性の問題は、全体的な視点でとらえる必要があります。例えば、独自規格に基づく物理実装から仮想機能へとネットワーク機能を行きわたる第1段階では、同じシステムがまだコントロール・プレーンとデータプレーンを処理しています。仮想化は運用上の問題を改善するように思えますが、多くの仮想機能が依然としてラインレートのトラフィックを処理する必要があります。分析の結果、3つの異なる重点領域があることが分かりました。

コントロール・プレーンとデータプレーンの個別のスケールアップ

コントロールとユーザープレーンの分離(CUPS)により、コントロール・プレーンとデータプレーンは個別の拡張が可能になります。これは、VNF設計における次の論理段階です。CUPSは第3世代パートナーシップ・プロジェクト(3GPP)アーキテクチャーで定義されていますが、新しい概念ではありません。信号方式No.7(SS7)(1980年)などの通信プロトコルは、分離されたコントロール・プレーンと「ベアラー」プレーンを用いて設計されたものです。セッション開始プロトコル(SIP)(2002年)などの比較的新しいプロトコルは、この概念をVoice over IP (VoIP)対応のIPネットワークに採り入れています。

新しいアプリケーションの帯域幅要件は急速に増えています。したがって、需要の成長を妨げないよう、データプレーンのリソース効率と拡張性を高めることが必要です。CUPSは、コントロール・プレーンとデータプレーンの個別拡張を可能にします。そのためCUPSは、拡張性と効率性が高く最適コストのNFVインフラストラクチャー実現に向けた、次の段階となります。

最適化された標準ベースの再プログラム可能なデータプレーン

NFVの核となるテーマの1つが、標準的な市販のハードウェア・プラットフォームの使用です。コントロール・プレーンは、標準的な市販のコンピューティング/メモリーリソースを用いて容易に拡張できますが、複雑なパケット処理を高速で実行できるようデータプレーンを拡張することは困難です。それを実現するには、データプレーンが再構成可能なパケット・パイプラインと複雑なパケット処理に非常に高いスループットで対応できるようなシステムを設計する必要があります。

さらにデータプレーンは、QoS/トラフィック・シェーピング、暗号化/復号処理、エンコード/デコード、その他の複雑な機能など、演算負荷の高いタスク(L4～L7サービス)をラインレートで実行する必要もあります。これによりCPUは、L4～L7サービスの一部のタスク処理を支援する場合に、高スループットで決まった動作を維持できるようになります。

「データセンターが、単純に別のアクセスノードとしてネットワークのエッジに存在するならば、データセンターはネットワーク・トラフィック・フローの動的な変化に適応できません。プログラマブル・インフラストラクチャーと標準化されたオープン・インフラストラクチャーがなければ、キャパシティを柔軟に増やすことは困難です」— China Telecom

VNFの隔離を最大限に高める

VNFの効率性が低いと、シングルノードのコンテキストではパフォーマンス目標を達成できなくなります。例えば、仮想ブロードバンド・リモート・アクセス・サーバー(vBRAS)にはさまざまな機能があり、容易に拡張可能な機能もあれば、アクセラレーションを利用する機能もあります(vBRASは、仮想ブロードバンド・ネットワーク・ゲートウェイ[vBNG]とも呼ばれます)。さまざまなVNFインスタンスを柔軟に拡張・管理する必要がある場合、サービスのチェイニングはより一層複雑になります。クラウド・データセンターなどの動的環境では、配置上の問題から、このようなVNFが分散型のイントラ/インター・データセンターになり、サービス機能チェイン(SFC)が無秩序に拡大する可能性があります。次のセクションで説明するように、SFCが拡大すると、NFVインフラストラクチャーの管理、拡張、運用がさらに難しくなります。

VNFの拡大と隔離の必要性を理解する

例えばvBRASは、(トラフィック・シェーピングやルーティングに加えて)リモート認証ダイヤルイン・ユーザー・サービス(RADIUS)、動的ホスト構成プロトコル(DHCP)、ファイアウォール(FW)、DPI、VPN機能などの複数のVNFで構成されます。RADIUSやDHCPは演算上のオーバーヘッドが最小限で済みますが、加入者のトラフィックを処理、転送する仮想ルーティングやトラフィック管理などの機能はラインレート(40 Gbps、100 Gbps、およびそれ以上)で動作するため、演算負荷が大きくなります。

仮想マシン(VM)やコンテナでは、1つの機能が1つ以上のCPUで動作することができます。チェイニングが最適でなく、VNFが無秩序に拡大すると、インフラストラクチャーにかかる負荷が増大します。vBRAS → vFW → vDPI → vVPNというSFCを考えてみましょう。この場合は、すべてのVNFがvBRASで、CPU負荷の高いパケット処理(階層型QoSなど)を実装する必要があります。図1は、Closベース³のデータセンターでこれらのVNFがどのように配備、チェイニングされるかを示しています。この例では、アクセスサービスの需要に応えるために、vBRASにさまざまな機能を異なる配列で組み込むことができます。

- 100 Gbpsの合計帯域幅をサポートするには、ファブリックのスループットを6倍にする必要があります。
- 控えめに見て、各VNFがCPU使用率100%で25 Gbpsしかサポートしないと想定すると、各VNFが100 Gbpsをサポートするにはサーバーノードが4つ必要になります。

つまり、VNFの実装が最適でないと、インフラストラクチャーの需要増大により必然的に設備投資と運用コストが増え、それにより全体的なコストも増えてしまいます。さらにこのアーキテクチャーには、ラインレートのトラフィックを複数のvBRASインスタンスに分配するためのロードバランサーも必要です。

図1の例は、クラウドネイティブの設計においてネットワーク・アプリケーションを設計、最適化しなければならない理由を表しています。これを行わなければ、たとえデータセンターのインフラストラクチャーが拡張性の高い導入を可能にし、管理を簡素化するとしても、相対コストの削減は困難です。

「目標は、閉鎖的なプライベート・ネットワークをなくし、ビジネスロジック、ソフトウェア、オープン・アーキテクチャーを主体とする統一ハードウェア・プラットフォームを構築することです」— China Telecom

CTNet2025 : 2025年までのChina Telecomネットワークの変革

2016年7月に発表されたChina TelecomのCTNet2025イニシアチブ⁴は、シンプルかつオープンで俊敏性の高い、クラウドベースのネットワークの構築を目標としています。ネットワークの制御と転送を分け、ネットワークのソフトウェアとハードウェアを切り離し、ネットワーク・クラウドとITを仮想化して、業種固有の閉鎖的なアーキテクチャーから業種を問わないオープンなアーキテクチャーへと効率良く移行するというのが、同社のビジョンです。China Telecomは、次のテクノロジーが新しいネットワーク・アーキテクチャーの開発に重要と判断しました。

- SDNの実装
- 2025年までにインフラストラクチャーの80%を仮想化
- クラウド・コンピューティング・テクノロジーの活用
- オープンソース・ソフトウェアの使用

この変革に不可欠なVNFは、vBRAS、仮想ユーザー宅内機器(vCPE)、仮想進化型パケットコア(vEPC)、仮想IPマルチメディア・サブシステム(vIMS)、仮想オンライン端末(vOLT)です。

チェイニングの影響:

- トラフィックの増大によりファブリックが大規模化
- 同じ量のアクセス帯域幅を処理するために、より多くのサーバーノードが必要
- 設備投資と運用コストが増加

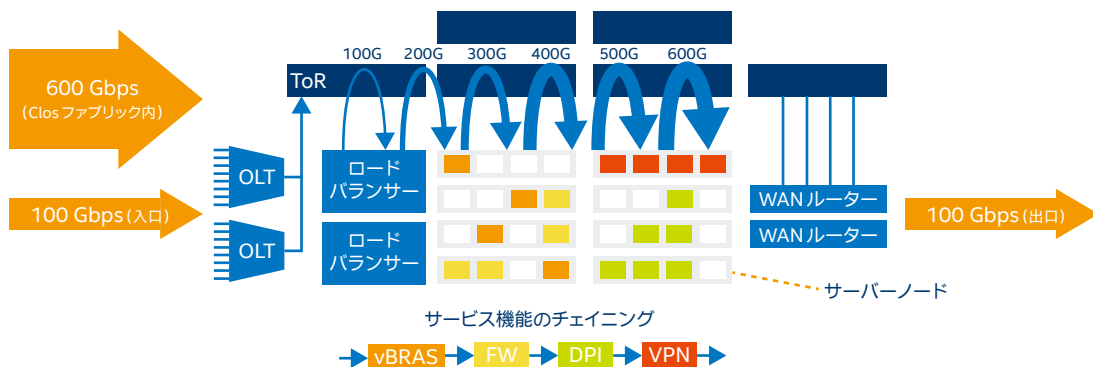


図1. 非効率なvBRAS導入の例

拡張性と効率性の高いvBRASアーキテクチャーへ

設計目標

China Telecom、HPE、インテルは、CTNet2025以降のChina Telecomのネットワーク要件を満たすため、次の設計目標を定めました。

- **アーキテクチャー** — CUPSを第1段階として、市販のサーバー・プラットフォームをベースとする未来のクラウドネイティブ設計原則をサポートしなければなりません。
- **消費電力とパフォーマンス** — 電力消費量は、1テラビット毎秒当たり4 Kワット(4 KW/Tbps)を達成しなければなりません。各ノードのシステムは、数100 Gbpsまたはそれ以上でなければなりません。
- **相対コスト** — NFVの目標の1つは、システム全体の相対コストを減らすことです。帯域幅が増加すると、一定のコスト削減を維持することが難しくなります。このプロジェクトの設計目標は、標準NICではなくSmartNICを用いた50%以上のコスト削減と、将来の世代にも対応できる拡張性を実現することです。

「私たちのビジョンには、高速かつ柔軟な情報フローをサポートし、産業ニーズの進化に応じて新しいアプリケーションに適合する、NFV主導型でソフトウェア主導型のアーキテクチャーも含まれています」— China Telecom

BRASからvBRASへの進化

ブロードバンド・リモート・アクセス・サーバー(BRAS)は、データ量の多い他のサービスを網羅する共通VNFです。図2は、ClosファブリックにVNFを導入するコンテキストで、vBRASの進化を表しています。

- 物理アプライアンス** — 物理的なBNGアプライアンス(シャーシのフォームファクター)からなる従来製品。この実装では、コントロール・プレーンとデータの両方が物理アプライアンスに実装されます。
- 仮想BRAS** — 第1世代のVNF(vBRASを含む)により、ソフトウェアが物理アプライアンスから仮想マシンに移植されました。コントロール・プレーンのワークロードの演算負荷はそれほど高くありませんが、このアーキテクチャーには基本的に、データプレーンのパフォーマンス上の問題があります。なぜなら、データプレーンのオペレーションは並列性、反復性、演算負荷が高いためです。さらに、仮想化によるリアルタイム・ソフトウェアの課題も加わります。
- コントロールとデータの分離** — コントロール・プレーンとデータプレーンを個別に拡張する必要性から、このホワイトペーパーで説明しているvBRASアーキテクチャーによって、これらのプレーンを別々のサーバーに分離します。ただし、コントロールとデータを分離するだけでは不十分です。効率的なvBRAS設計によって、前述したデータプレーン拡張の課題に対処する必要もあります。

効率的なvBRAS設計

China Telecomのユースケースと、HPEおよびインテルとのコラボレーションから生まれた新しいソリューションでは、vBRAS-c(コントロール・プレーン)とvBRAS-d(データ)プレーンを物理的に分離した市販の標準サーバー・プラットフォームに実装することで、コントロール・プレーンとデータプレーンの個別の拡張が可能になります。データプレーンはパケット処理用に最適化され、演算負荷の高いトラフィック・シェーピング機能とQoS機能も高速化されます。図3を参照してください。

vBRAS-c ノード

vBRAS-c ノードは、標準NIC(アクセラレーションなし)を使用した、コントロール・プレーン処理専用の市販のサーバーです。ユーザーは導入ニーズに応じて、コントロール・サーバー内でvBRAS-c インスタンスの数を増減できます。vBRAS-cの設計は非常に軽量で、コンテナ実装に適しています。コントロール・プロセッサを分離することには、必要に応じて新規ユーザーを追加できたり、vBRAS-c インスタンスがロケーションに依存しなくなったりするなどの利点があります。軽量で仮想型(コンテナまたはVM内)のため、vBRAS-c インスタンスを分散型にしてデータ・プレーン・ノードと同じロケーションに置くことも、サービス・プロバイダーのクラウド内に集中させることも可能です。どちらの場合も、vBRAS-c インスタンスは、分散されたデータプレーンをサービス・プロバイダーの中央オフィス・ロケーションで管理します。各vBRAS-cは、1つ以上のデータ・プレーン・サーバーを制御できます。

vBRAS-d ノード

vBRAS-dは、パケット処理専用の別個のサーバーです(コントロール・セッションは扱いません)。この構成ならば、一部のコアにパケット転送のみを実行させ、その他のコアに複雑なパケット処理を実行させることができます。サーバーのパケット処理機能をさらに向上させるため、この新しいアーキテクチャーにはインテル® Arria® 10 FPGAをベースとしたSmartNICが使われています。FPGAは、QoS/トラフィック・シェーピング、VPN、ネットワーク・アドレス変換(NAT)、DPIなどの複雑なパケット処理機能を実行することで、さらにCPUコアを支援できます。このホワイトペーパーで説明する効率性の高いvBRASは、FPGAを用いてQoS/トラフィック・シェーピング機能を高速化しており、次世代の導入ニーズに対応するレベルのパフォーマンスと機能性を実現しています。

FPGAベースの設計には、いくつかの利点があります。例えばこの実装では(パケット転送の支援に加えて)、FPGAがラインレートのQoSと階層型トラフィック・シェーピングをサポートします。インテル® Arria® 10 FPGAは、非常にパワフルで再プログラム可能な転送プレーンを提供します。40 GbpsのQoSとトラフィック・シェーピングにFPGAリソース(インテル® Arria® 10 GT 1150デバイス)を40%しか消費しないため、FPGAの約60%をラインレートの処理が必要な他のネットワーク機能に使用できます。[†] ホスト・プロセッサは、加入者固有でレイテンシーに厳しい他のオンデマンド・サービスを実行できます。FPGAは再プログラム可能なため、デプロイメントの違いに応じて機能を変えるこ

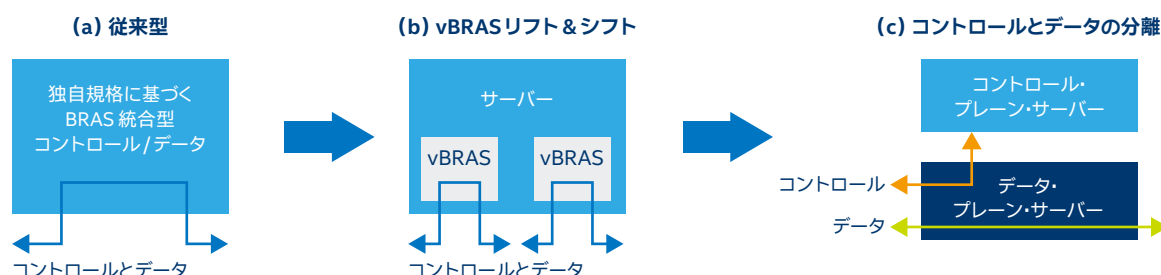


図2. Clos ファブリックでのvBRASの進化

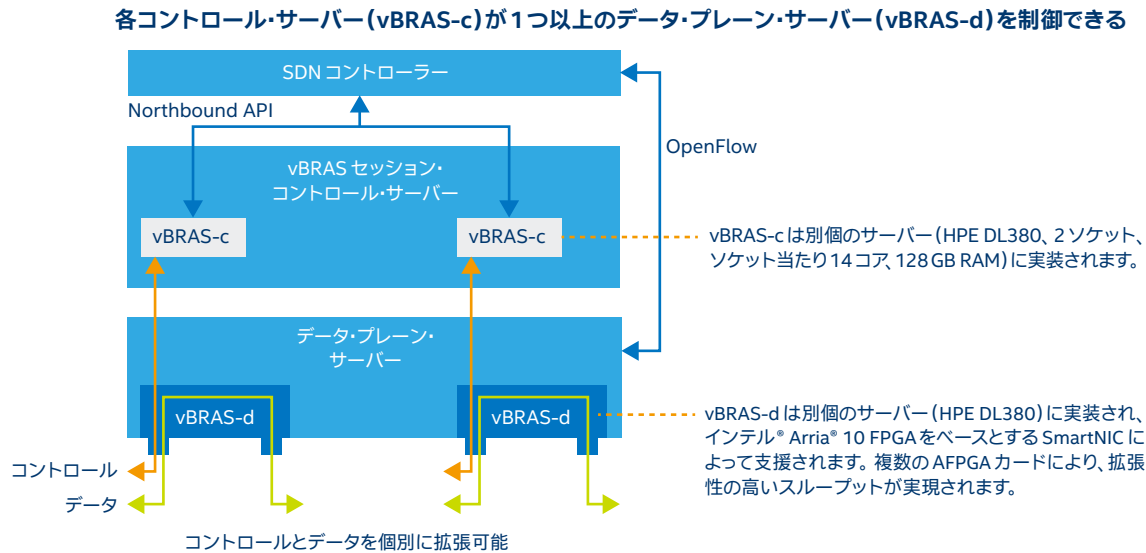


図3. SDN から着想を得た分割 vBRAS

とができます。例えば、独自規格に基づく、または非標準の DPI アルゴリズムを FPGA に実装することが可能になります。さらに、ストリーミング・コンテンツの動的エンコード/デコードなど、アプリケーション固有のユースケースに同じサーバー・ハードウェア構成を利用することもできます。

SDN コントローラー

SDN コントローラーは非常に軽量であり、OpenFlow^{*5} プロトコルを使用して、FPGA ベースの SmartNIC によって実装される転送/QoS プレーンを構成します。このコントローラーは、ネットワーク・ファブリック全体の状態ではなく、vBRAS-d 転送プレーンのみを制御します。主な利点は、既存のアプライアンスと同様に、VNF を個別に制御・管理できることです。ただし、このソリューションは OpenFlow プロトコル・ベースの SDN コントローラーなので、より中央集中型の手法で vBRAS-c インスタンスを制御することもできます。最終的な目標は、サービス・プロバイダーがソリューションを柔軟に導入できるようにすることです。

インテル® Arria® 10 FPGAベースのSmartNIC

インテル® Arria® 10 FPGAのパケット転送設計(図4)は、複数の並列パケット処理パイプライン(それぞれ40G/100G以上のラインレートをサポート)に対応するように構成できます。パイプラインのレイテンシーはハイパフォーマンスのカスタムASIC/ASSPで発生するレイテンシーと同様で、柔軟性と拡張性の高い実装は内部バッファと外部バッファのどちらもサポートします。パケット処理機能は、新しいハードウェアを導入せずに拡張/更新できるため、製品のアップグレード・サイクルを速めることができます。

ベンダーに依存しないオープン・エコシステム

コントロール機能とデータ機能を別々のサーバーに分けることが、このアーキテクチャーの肝であり、これによって、複数のベンダーが寄与するエコシステムが促進されます。このホワイトペーパーで説明するシステム例では、HPEとインテルの製品を使用していますが、さまざまなベンダー（前述のアーキテクチャー原則に従っているベンダー）からvBRASコントローラーやSmartNICを調達できます。HPE、インテル、China Telecomは、vBRAS VNFを足掛かりに、他のSmartNICベンダーやVNFベンダーも標準ベースのアプリケーション・プログラミング・インターフェイス(API)の定義の議論に加わることを期待しています。

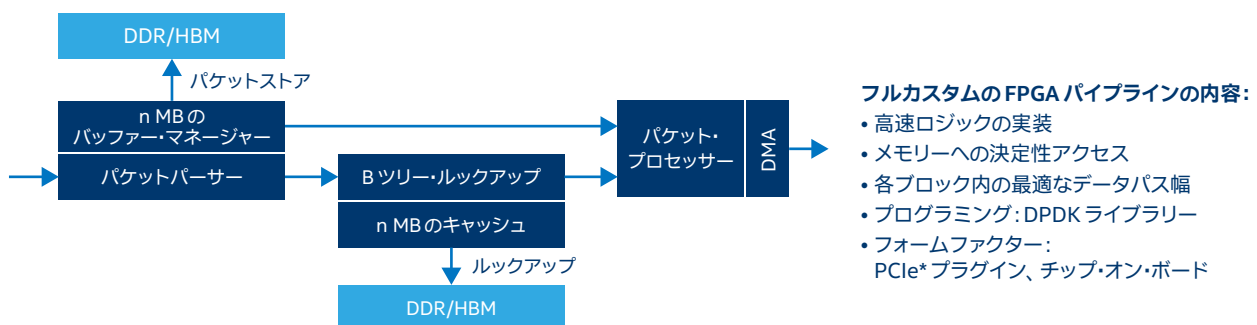


図4. パケット転送のFPGA設計

vBRASのパフォーマンス結果

このアーキテクチャーではvBRAS-cとvBRAS-dを個別に拡張できるため、最初に検討すべきセットアップは、vBRAS コントローラー・プレーンとvBRAS データプレーンの分離です。ソフトウェア・アーキテクチャーは、図3に示す設計に基づいています。表1に、このテストで使用したハードウェアを示します。

コントローラーのパフォーマンス

図5に示すセットアップでは、vBRAS-cのパフォーマンスを個別に測定するために、ソフトウェア・ベースのvBRAS-dを使用しています。

サーバーの分離 — Spirent* PPPoE テスト・ジェネレーターがPPPoEセッションを1万回実行します。これらのセッションはvBRAS-dを介してvBRAS-cサーバーにリダイレクトされます。この構成では、OpenFlowのPacket-Inなどのトンネリング・メカニズムを必要とせず、セッション・コントロール・プロトコルのネイティブ機能を保持しながらコントロール・プレーンとデータプレーンを分離することができます。

コントロール・サーバーの拡張 — vBRASソフトウェアはセッション・コントロール情報のみを扱うため非常に軽量で、コンテナや仮想マシン(VM)にパッケージ化することができます。HPEの予備パフォーマンス・テストでは、(2ソケットシステム内の14コアのうち)1コアのCPU使用率はわずか11%、RAMの使用量はわずか1.5 GBで、各vBRAS-cノードが1万回のセッションを処理できることが実証されました。† vBRAS-cコンテナ・インスタンスの数を最大にすると、各サーバーが384 GBのRAMを使用して、250万回以上のセッションを処理できる計算になります。†

vBRAS データプレーンのパフォーマンス

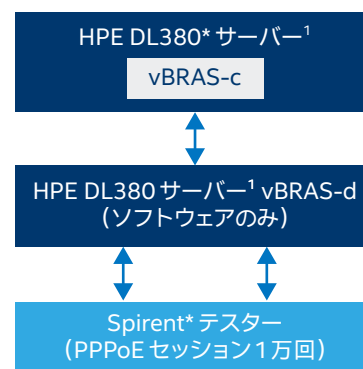
このテストでは、FPGA ベースのSmartNICを使用して、高データ・レート・トラフィックのQoSとトラフィック・シェーピングを高速化できる可能性を検証しました。

データプレーンの転送機能 — 測定は、インテル® Arria® 10 FPGA ベースのSmartNICを3枚搭載したHPE DL380*6サーバー上で行われました(図6を参照)。DL380サーバーは、各ソケットにインテル® Xeon® プロセッサー E5-2660 v3 製品ファミリーを搭載したデュアルソケット・サーバーで、RAMは256 GBです。

このテストの目標は、QoSで120 Gbpsのライン・レート・パフォーマンスを達成することでした。このテストは12,000回の加入者セッションと、それぞれ2つのトラフィック・クラス(高優先度と低優先度)で構成され、1加入者当たり8 Mbpsをシェーピングします。1加入者当たり、5 Mbpsの高優先度トラフィックと5 Mbpsの低優先度トラフィックが生成されます(合計で60 Gbpsの高優先度トラフィックと60 Gbpsの低優先度トラフィックが生成されます)。例えば、この構成では、パケットサイズに関係なくシェーピングされたトラフィックが8 Mbps × 12,000、つまり96 Gbpsで送信されると同時に、シェーピングされないトラフィックが120 Gbps(損失なし)で返信されます。シェーピング・パラメーターの値を変えて(8 Mbps、8.5 Mbps、10 Mbps)テストしたところ、(4,000回のセッションの)合計スループットが全セッションの最大値(それぞれ96 Gbps、102 Gbps、120 Gbps)内に収まることが分かりました。

図7は、3つのインテル® Arria® 10 FPGA SmartNICカードを搭載したデータ・プレーン・ノードの予備パフォーマンスを示しています。トラフィックをシェーピングすると、高優先度トラフィックに帯域幅が完全に割り当てられ、低優先度トラフィックにはシェーピングが適用されます。

データプレーンはPCIe* SmartNICをベースにしているため、このソリューションは、多様なベンダー・エコシステムが定めるロードマップを追従することができます(新しいSmartNICが登場したら、それをシステムに組み込むことが可能です)。例えばHPEでは、SmartNICの進化に伴い、2018年にはDL380の性能が240 Gbps(トラフィック・シェーピングとQoSあり)にまで倍増すると見込んでいます。



1. デュアルソケット、ソケット当たり14コア

図5. パケット転送のFPGA設計

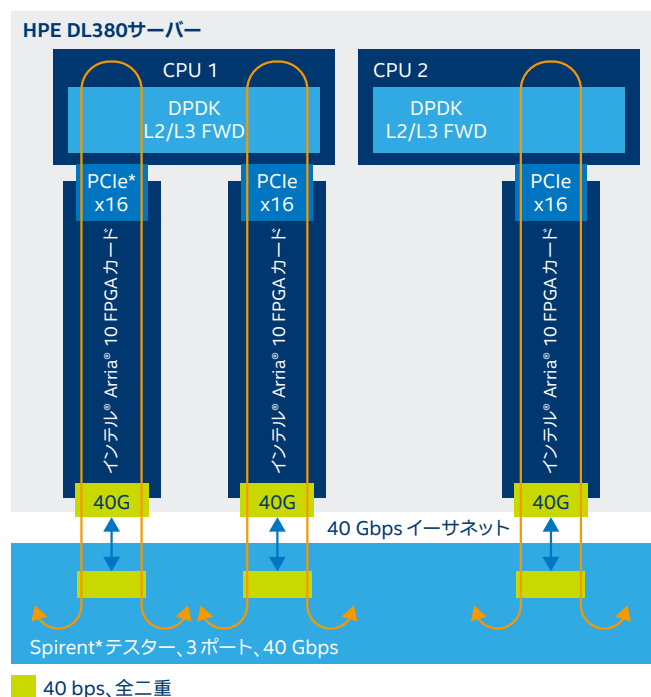


図6. 120 Gbps対応のFPGA SmartNICを3つ搭載したDL380サーバー・セットアップ(QoSとトラフィック・シェーピングあり)

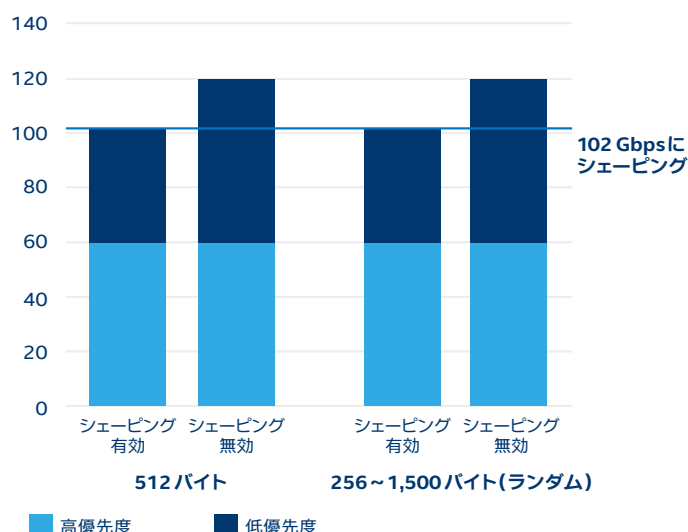


図7. データ・プレーン・ノードのパフォーマンス

これらの初期結果は、コントロール・プレーンとデータプレーンを物理的に分ける分割アーキテクチャーによってコントローラーの拡張性を向上できることを示しています。FPGA ベースの SmartNIC は、QoS の高速化手段としても有望視されています。今後は、ソフトウェア・ベースの vBRAS-d を FPGA と統合されて、QoS で完全 vBRAS のトラフィック・スループットを実現する取り組みが中心になるでしょう。この初期結果は希望をもたらすものであり、インテル、HPE、China Telecom はオープン・エコシステムの推進に向けて、この情報と統合アーキテクチャーを共有したいと考えています。

効果的な vBRAS で相対コストを削減

非効率な VNF は、インフラストラクチャーの需要を増大させる可能性があります。このホワイトペーパーでは、FPGA ベースの SmartNIC を用いて複雑なパケット処理機能を高速化する、経済性の高いソリューションを紹介しています。このセクションでは、価格、パフォーマンス、消費電力など、システム上のメリットについて説明します。

明確な基準やベンチマークなしに、コスト削減率を計算することは困難です。したがって、コスト削減率をモデル化するため、このホワイトペーパーでは明確な目標を定め、各種コンポーネントで公開されているコストモデルに基づいたアプローチを構築します。

モデルの目標とアプローチ

この分析では、同一ワークロードに対する一定のシステム・パフォーマンスを用いて、標準 NIC と SmartNIC を比較します。ここでは、コストに影響を与える 2 つの主要因、消費電力とサーバーコストを比較します。比較の条件を一定にするため、同じ合計パケット・スループットを達成する同等のシステムを構築する際の全体的なコストと消費電力を調べます。

方法

vBRAS サーバーのキャパシティを変え (50G ~ 200G)、サーバー・パフォーマンスは一定を保つよう試み、SmartNIC を使用する場合と使用しない場合でコストをモデル化します。分析の結果、パフォーマンスの向上により CPU サイクルが節約され、コア数が減ることが分かりました。図 8、9、10、11 のチャートは、サンプルシステムで標準 NIC と SmartNIC の消費電力、サーバーコスト、システムコスト、総コストを比較したものです。結果は、50 Gbps の基準帯域幅に対する割合で示されています。

要約:

- SmartNIC ソリューションは、標準 NIC サーバーと比べて総消費電力が最大 50% 低下します。[†] 絶対値では、このソリューションは 4 KW/Tbps の設計目標内に十分収まります。
- インテル® Arria® 10 FPGA ベースの SmartNIC と市販のサーバー CPU を組み合わせると、データプレーンのパフォーマンスが最適化され、高い柔軟性を維持しながらコストを削減できます。
- SmartNIC を使用すると、各サーバーのパフォーマンス向上が 3 倍を超えます。[†] スループットを向上させるには、シングルサーバー・アーキテクチャーのシステム能力では無理です。SmartNIC を使用することで、パフォーマンスが向上し、消費電力とコストの増加分を極力抑えつつ、高いスループットがサポートされます。

ハードウェア	数量
HP DL380 Gen9 8SFF CTO サーバー	1
HP DL380 Gen9 インテル® Xeon® プロセッサー E5-2660 v3 製品ファミリー FIO キット	1
HP DL380 Gen9 インテル® Xeon® プロセッサー E5-2660 v3 製品ファミリー・キット	1
HP 16 GB 2Rx4 PC4-2133P-R キット	16
HP 600 GB 6G SAS 10K SC 2.5 型 ENT HDD	2
HP DL380 Gen9 Secondary 3 Slot Riser Kit	1
HP Smartアレイ P440ar/2G FIO コントローラー	1
HP Ethernet 1 Gb 4ポート 331FLR アダプター	1
HP Ethernet 10 GB 2ポート 560SFP+ アダプター	6
HP 2U SFF Easy Install Rail Kit	1
HP 500 W FS Plat ホットプラグ対応 パワー・サプライ・キット	2
10 GB SFP+ モジュール	12

表 1. ハードウェア構成 - データ・プレーン・サーバー

「NFV を利用することで、デバイスの機能は今後、高額な専用ハードウェアに頼る必要はなくなります。安価な x86 ベースのシステムで、機器投資コストを削減し、新しいアプリケーションを短期間で開発・導入できるようになります」
— China Telecom

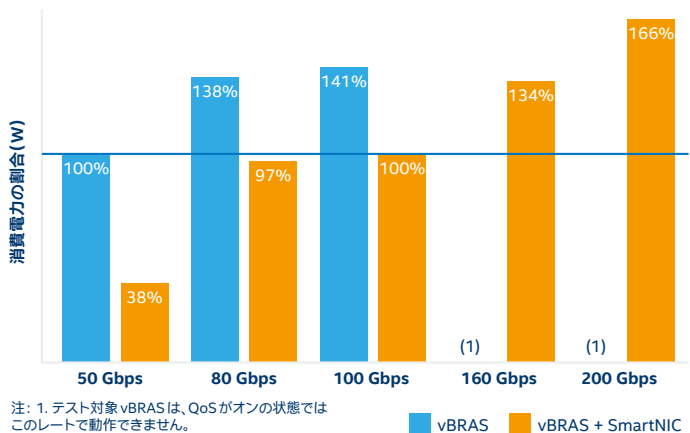


図 8. システムの総消費電力 (W) - 基準値 (50 Gbps) に対する割合

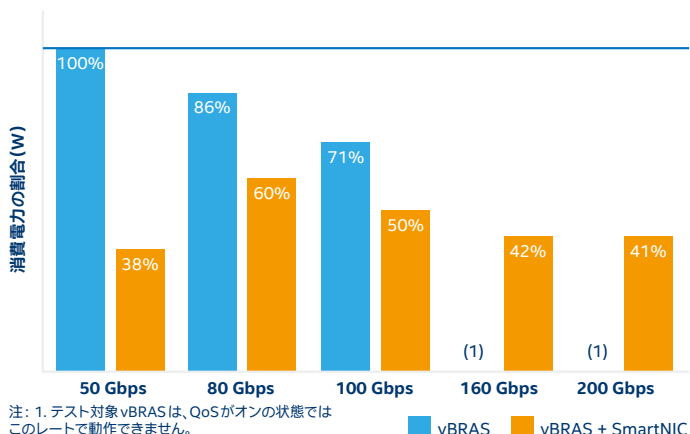


図 9. 総消費電力 (KW/Tbps) - 基準値 (50 Gbps) に対する割合

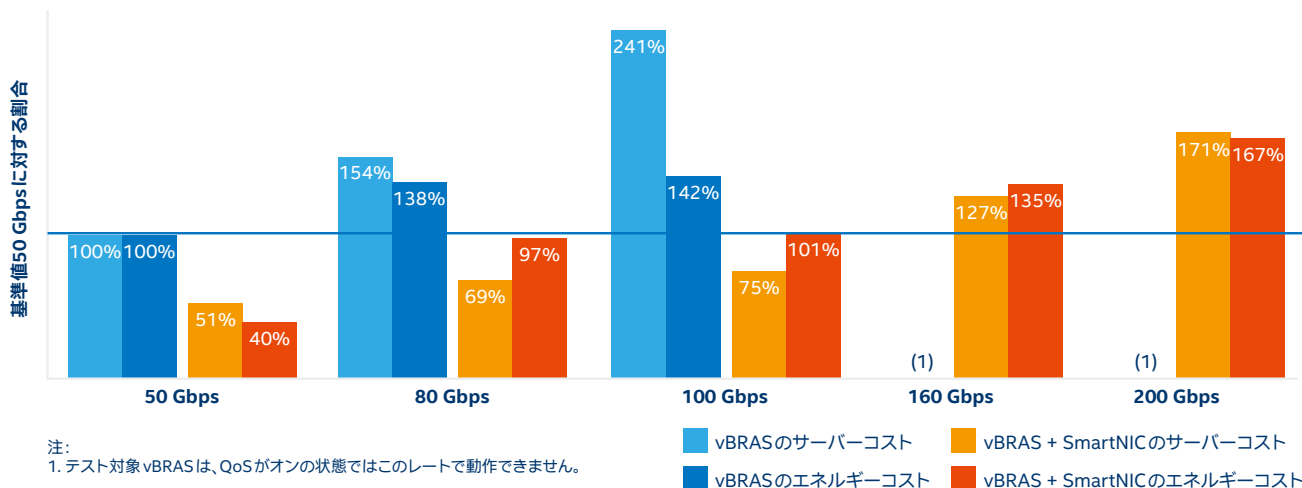


図 10. 総サーバーコスト - 基準値(50 Gbps)に対する割合
3年後の総エネルギーコスト - 基準値(50 Gbps)に対する割合

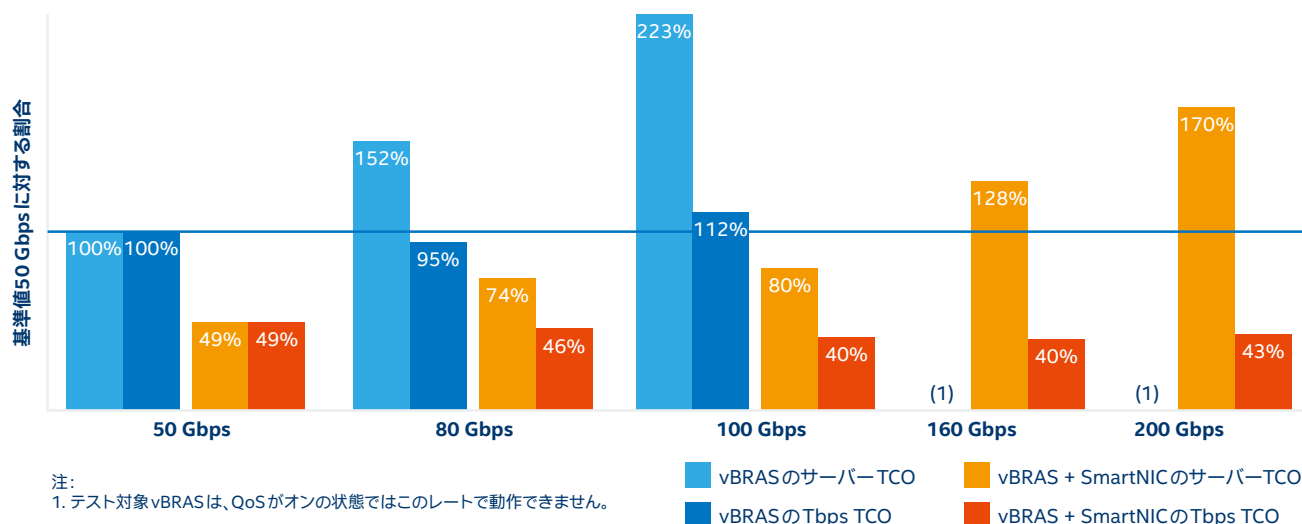


図 11. 3年後のサーバー当たりの総コスト - 基準値(50 Gbps)に対する割合
3年後の1 Tbps 当たりの総コスト - 基準値(50 Gbps)に対する割合

他のCUPS実装との比較

図 12 は、vBRAS ソリューションの代替となる別の CUPS 実装を示しています。vBRAS コントロールは市販の標準サーバーですが、既存の製品やビジネスへの投資を維持するために、データプレーンは独自規格のシャーシを使用します。このソリューションはオープンでもプログラム可能でもありません。サードパーティーのベンダーはデータプレーンを利用できず、プロバイダーは標準ハードウェアを使用して帯域幅を拡張することができません。さらに、独自規格のシャーシ/ラインカードを使用しているため、ベンダーが固定されるリスクもあります。

一方、このホワイトペーパーで紹介している純粋なサーバーベースのプラットフォームならば、vBRAS-c プレーンと vBRAS-d プレーンを個別に拡張できます。この設計では、ベンダーに依存しないコントロール・ノードとデータノードの両方が標準ハードウェアをベースにしているため、図 12 に示すような制約がありません。演算負荷の高いパケット処理用にデータプレーンのスループットをさらに高めるため、この設計では新しいインテル® Arria® 10 FPGA ベースの SmartNIC を採用しています。

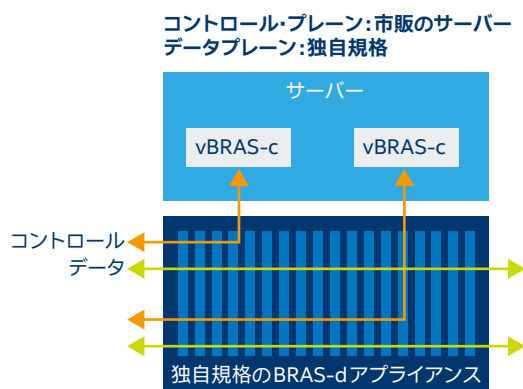


図 12. 代替のCUPS実装

まとめ

サービスの集中に伴い、ネットワーク機器はより複雑になりました。その結果、機器が高度に特殊化され、ハードウェア/ソフトウェア/機器の結び付きが強まり、拡張が難しくなり、アップグレード・サイクルが長くなり、新規ビジネスの導入サイクルはかつてないほど延長しています。

仮想ネットワークの再設計によって、ビジネスルール、ネットワーク機能、インフラストラクチャーをリソースプールから利用することが可能となり、サービス導入コストは削減され、効率性は向上します。ここで紹介した新しいアーキテクチャーは、SDNとNFVを融合したものです。クラウドとネットワークが結合され、未来のデータセンターがネットワークの核となります。完全に仮想化されたインフラストラクチャーに移行すると、発展する市場の需要に容易に対応できる一方で、設備投資と運用コストを減らして全体的なコストを抑えることができます。

「帯域幅の需要増加に伴い、データセンターはネットワーク・アーキテクチャーの核になる必要があります。これは革新的な変化であり、あらゆるサービス導入の構造を決めることになるでしょう」— China Telecom

謝辞

Prodip Sen

Hewlett Packard Enterprise
ネットワーク機能仮想化担当 CTO

Siddhartha Karkare

インテル コーポレーション
SDN/NFV エコシステム & アクセラレーション担当
シニア・マーケティング・マネージャー

Alexander Kugel

インテル コーポレーション
ネットワーク仮想化 & SDN/NFV 担当シニア・アーキテクト
(システム & ソリューション・エンジニアリング)

Mark Lewis

インテル コーポレーション
ネットワーク・システム・ソリューション担当シニア・マネージャー

Frederic Richard

インテル コーポレーション
エンジニアリング・マネージャー

参考資料

- <http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/vni-hyperconnectivity-wp.html> (英語)
- http://www.sohu.com/a/110764570_294469?_f=v2-index-feeds (中国語)
- <http://www.networkworld.com/article/2226122/cisco-subnet/clos-networks--what-s-old-is-new-again.html> (英語)
- <http://www.c-fol.net/news/content/1/201609/20160926081545.html> (中国語)
- <https://www.opennetworking.org/> (英語)
- <https://www.hpe.com/us/en/product-catalog/servers/proliant-servers/pip.hpe-proliant-dl380-gen9-server.7271241.html> (英語)

詳細情報

インテルおよびインテル® Arria® 10 FPGAの詳細については、<https://www.altera.co.jp/products/fpga/arria-series/arria-10/overview.html>を参照してください。



本資料に記載されているコスト削減のシナリオは、特定の状況および構成において、所定のインテル搭載製品が将来的なコストにどのような影響を与え、どの程度コストを削減できるかを例示することを目的としています。状況は変わる可能性があります。インテルは、いかなるコストまたはコスト削減も保証しません。

† テストは、特定システムでの特定テストにおけるコンポーネントのパフォーマンスを測定しています。ハードウェア、ソフトウェア、システム構成などの違いにより、実際の性能は掲載された性能テストや評価とは異なる場合があります。購入を検討される場合は、ほかの情報も参考にして、パフォーマンスを総合的に評価することをお勧めします。性能やベンチマーク結果について、さらに詳しい情報をお知りになりたい場合は、<http://www.intel.com/benchmarks/> (英語)を参照してください。

©2017 Intel Corporation. 無断での引用、転載を禁じます。Intel、インテル、Intel ロゴ、Intel Inside、Intel Inside ロゴ、Intel. Experience What's Inside、Intel. Experience What's Inside ロゴ、Altera、Arria、Cyclone、Enpirion、Intel Atom、Intel Core、MAX、Nios、Quartus、Stratix および Xeon の名称およびロゴは、アメリカ合衆国および/またはその他の国における Intel Corporation またはその子会社の商標です。インテルは、製品やサービスの内容を、いつでも予告なく変更できるものとします。本資料に記載した情報、製品、サービスの適用や使用により生じた損害について、インテルは、書面で明示的に合意した場合を除き、一切の責任や義務を負いません。公開された情報を利用する場合や、製品やサービスをご注文の場合は、事前に最新バージョンのデバイス仕様を入手されることをお勧めします。

* 第三者の社名、製品名などは、一般に各所有者の表示、商標または登録商標です。