

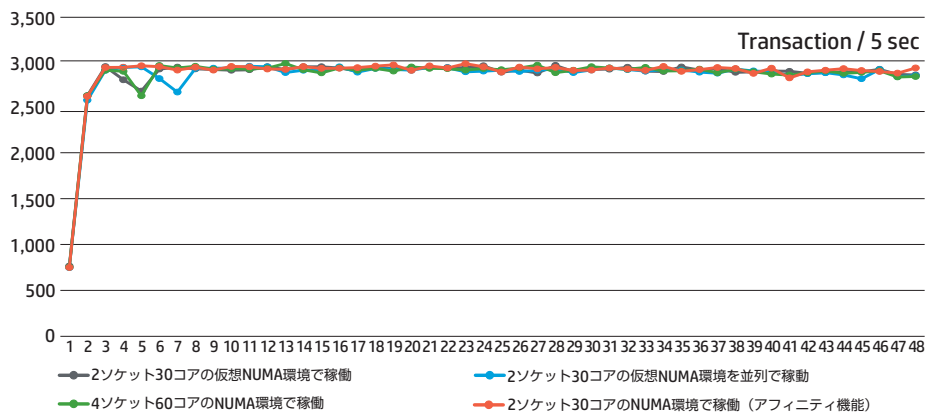
仮想環境における SQL Server 2014 のスケーラビリティを実証



インテル® Xeon® プロセッサ E7 v2 ファミリー搭載 HP ProLiant DL580 Gen8 での仮想 NUMA

HP ProLiant DL580 Generation8(Gen8)は、4ソケットで60コアのインテル® Xeon® プロセッサ E7 v2 ファミリーを搭載し、最大6TBのメモリーをサポートする、ミッションクリティカル用途での利用に最適なマルチプロセッサ・サーバーです。今回、インテルとマイクロソフト、HPは3社共同でHP ProLiant DL580 Gen8 上に構築したHyper-V仮想環境におけるデータベース処理のベンチマークテストをおこないました。その結果、仮想環境のオーバーヘッドはほとんど存在せず、Microsoft® SQL Server® 2014を利用するアプリケーションの仮想環境上のスケーラビリティを実証しました。ビッグデータ分野、大規模な仮想化・サーバー統合における大きな可能性を示唆しています。

■ 図1 NUMA環境と仮想NUMA環境における負荷テストの結果



■ 図2 NUMAノード上のHyper-V

SQL Server				SQL Server			
仮想 Windows Server & Hyper-V Guest				仮想 Windows Server & Hyper-V Guest			
Windows Server & Hyper-V Host							
物理コア 1	物理コア 2	物理コア 3	物理コア 4	物理コア 5	物理コア 6	物理コア 7	物理コア 8
CPU ソケット 1	CPU ソケット 2	CPU ソケット 3	CPU ソケット 4	NUMA ノード 0	NUMA ノード 1	NUMA ノード 0	NUMA ノード 1
ローカルメモリー	ローカルメモリー	ローカルメモリー	ローカルメモリー	ローカルメモリー	ローカルメモリー	ローカルメモリー	ローカルメモリー

仮想NUMAでもオーバーヘッドゼロに近いパフォーマンスを実証

インテル、マイクロソフト、HPは3社共同で高性能のインテル® Xeon® プロセッサ E7 v2 ファミリー搭載のHP ProLiant DL580 Gen8での4ソケット60コアのデータベース処理性能を検証するベンチマークテストをおこないました。その検証の結果、Microsoft® SQL Server® 2014 を利用するアプリケーションにおいてNUMA環境と仮想NUMA(vNUMA)環境の間に大きな差異はなく、仮想マシンのスケールアウトにも性能劣化が発生しないことを実証しました。ベンチマークテストの設計は、日本マイクロソフト株式会社サーバープラットフォームビジネス本部アプリケーションプラットフォーム製品部 技術顧問の熊澤幸生氏がおこなっています。熊澤氏はMicrosoft® SQL Server®の性能検証に長年携わっている第一人者であり、今回の検証においては仮想化時代におけるMicrosoft® SQL Server®の今後の可能性を示唆する新たなデータの取得を実現しています。HP ProLiant DL580 Gen8で、NUMAノード上のHyper-V環境を構築(図2)。4、2ソケットNUMA環境と、2ソケット仮想NUMA環境、4ソケットの仮想NUMA環境における性能評価をおこないました。熊澤氏が作成した負荷テストツールで240秒間の負荷テストをおこなったところ、上限900ユーザーの同時接続と5秒間に3,000トランザク

ションを処理しています(図1)。しかも、4つのいずれの環境においても差のないフラットな数値を示しており、同一マシン上における仮想マシンのスケールアウトによってMicrosoft® SQL Server®アプリケーションの冗長化、スケールアウトの可能性を示すものとなりました。今回の検証をおこなった理由について熊澤氏は次のように述べています。

仮想NUMAにおけるデータベース処理の実現

今回の検証に使用したハードウェア環境は、ソケットあたり15個のコアを集積したマルチプロセッサ構成対応のインテル® Xeon® プロセッサ E7-4890 v2(動作周波数 2.8GHz)を、4ソケット(60コア)搭載することができる、HP ProLiant DL580 Gen8を使用しました。ソフトウェア環境は、Microsoft® Windows Server® 2012 Datacenter R2と、Microsoft® SQL Server® 2014 Enterprise Edition CU1の組み合わせによる、ミッションクリティカル向けソフトウェア環境を使用しました。今回の検証は、4ソケット60コアのマシンを、(1)4ソケット60コアで稼働させたケース、(2)2ソケット30コアで稼働させたケース、(3)2ソケット30コアの仮想環境で稼働させたケース、(4)(3)の2ソケット30コアの仮想環境を2つ並列で稼働させたケースのベンチマークをとることで、

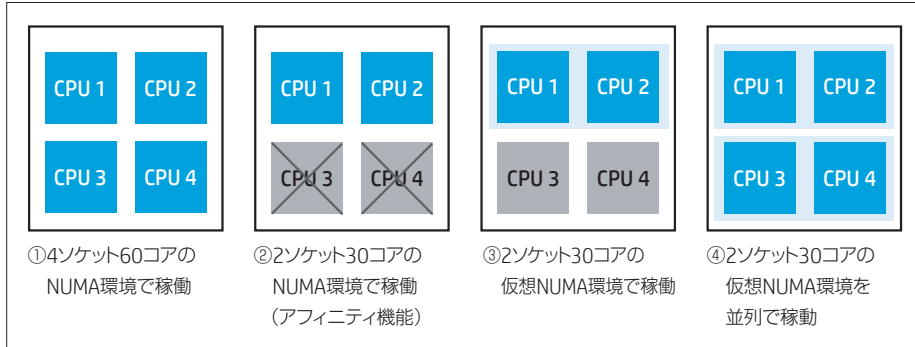


インテル® Xeon® プロセッサ E7 v2 ファミリー

「仮想化によって同一サーバー上にWebサーバーとアプリケーションサーバー、データベースサーバーを運用するお客様が増えている中、NUMA環境に対応したサーバーを導入したのに、NUMAを意識した構成をとっていないために、性能を発揮できていないお客様が多くいます。データベースの仮想NUMAにおける正しい運用を示すためにも、ぜひ今回は検証したいと考えていました」(熊澤氏)。

■ 図3 NUMA環境と仮想NUMA環境での負荷テストシナリオ

2ソケット30コアの物理環境(②)と2ソケットの30コアの仮想環境(③)で負荷テストを行い、仮想環境のオーバーヘッドが非常に小さいことが確認できました。さらに、2ソケット30コアの仮想環境を並列に動作させた場合(④)でも、③の結果と比較したところ、2つの仮想環境が影響し合いパフォーマンスに悪影響を及ぼさないこともわかりました。①の4ソケット60コアの物理環境の測定は、④の結果との比較の参考値として実施。2者の間にパフォーマンスの差異がないことを確認できています。



仮想NUMAにおけるSQL Server 運用の実用性を実証するのが目的です(図3)。

今回の検証においてポイントとなったと熊澤氏が語るのは、Hyper-V仮想環境上においてSQL ServerにNUMA環境を認識させ、パフォーマンスを測定することでした。

「私は今まで仮想環境を構築されているさまざまなお客様を見てきましたが、SQL ServerがNUMAに対応したマシンをSMPマシンとしてしか認識できていないケースがほとんどでした。これだと、NUMAを無視したコアの割当てがされず、NUMAのメリットであるメモリアクセスの効率化が行えず、パフォーマンスは落ちてしまいます」(熊澤氏)。

熊澤氏は、DL 580 Gen8であればHyper-VにホストされたSQL Server アプリケーションと仮想NUMAの組み合わせでも、物理環境にひけをとらないパフォーマンスが出る確信があったといいます。2ソケット30コアの2つの仮想環境を、SQL ServerにNUMAとして認識させる情報は少なく、熊澤氏は米国のSQL Serverチームに問い合わせることで、仮想環境のパラメータ設定をおこない、環境の構築を実現しました。

図2のようにNUMA環境を認識したHyper-Vを実現するには、2つの設定をおこないます。Hyper-VのGuest環境を作成してすぐに、設定ダイアログの「NUMA構成」を開くと、「NUMAトポロジ」欄に「ハードウェアトポロジを使用」をクリックして、仮想NUMAトポロジを物理ハードウェアのトポロジにリセットします。という記述があります。この下にある「ハードウェアトポロジを使用」をクリックしてオンにします(図4)。これがNUMAを使う場合の1番目の設定です。もう1つの設定は、Hyper-V設定のダイアログでおこないます。「サーバー」欄に「NUMAノードにまたがるメモリの割り当て」メニューがあります。初期設定ではチェックボックスがオンになっていますが、これをオフすることでSQL ServerがNUMA単位でメモリーを使用します(図5)。

「SQL Serverはチューニングレスで性能を発揮する思想を持っていますが、チューニングを意識させないことで、すなわちNUMAを意識させない初期設定になっています。そこでHyper-V上で仮想NUMA環境をつくるためには、こちら2つの設定をおこなう必要があるのです。今回2ノードの仮想NUMAを作成したところ、きれいに2ソケットずつのパーティショニングができ、それぞれにNUMAノードが2つずつ、30個の分離CPUが認識できました」(熊澤氏)。

限りなくCPUを使い切るテストツールのチューニング

テストツールには、米国SQL Serverの開発者チームが作成したSQLストレスツールを使用しています。スレッド数とユーザー数、ストアドプロシージャの呼び出し回数など

を制御できるパラメータが用意されていて、連続的にSQL Serverに対して負荷をかけることが可能です。

ベンチマークで使用したデータベースは、ベーシックなオーダーエントリーです。受注ヘッダーが5万件程度の比較的コンパクトなデータベースを用意しました。また、インサート系、更新系、検索系のトランザクションミックスとなっていて、ロックの競合、ラッチの競合、メモリーの圧迫、CPUの高負荷が発生しやすくなっています。

SQL Serverは原則チューニングを行うことなく性能を発揮するように機能設計されていますが、それでも4ソケット60コアをすべて利用し、パフォーマンスを引き出すのは容易ではありません。CPUの限界値まで使い切り、どの程度までユーザーが連続的にデータを投げられるかを調査できるように、CPUをフルに使い切るためのチューニングをおこないました。

Microsoft® SQL Server®の1コネクションあたりの最大使用CPU数の上限を設定する「MAX DOP」は、CPUの1ソケットあたりの物理コア数を超える並列処理が発生(すなわち処理がほかのソケットにまたがる)しないよう、並列処理の発生限度を「1」に設定しました。2つ目のチューニングは、tempdbのデータファイルの分割です。tempdbはMicrosoft® SQL Server®が利用する一時表領域であり、さまざまなクエリの間接処理に利用されます。CPUのコア数分だけtempdbを作成することにより、一時表領域のアロケーションを並列で処理することが可能になります。また、ログファイルもI/O負荷が高いため、ドライブレターを2つに分けて、I/Oの負荷分散もおこなっています。

さらに極力余分なI/Oを抑えるためにデータをオンメモリーに載せました。ディスクI/Oはトランザクションログの書き込みと、いわゆるtempdbと呼ばれる一時表領域のI/Oのみに限定しています。

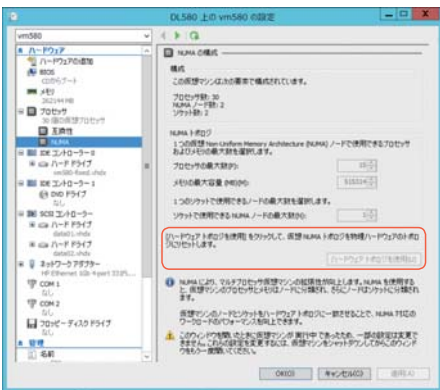
仮想環境の性能を裏付ける新しいアーキテクチャー

今回の検証結果が示している、NUMAと仮想NUMAでオーバーヘッドが変わらない理由の1つに、SQL Serverは自身のアーキテクチャー変更による性能の向上があります。

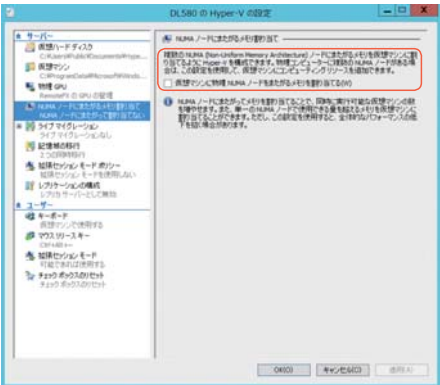
SQL Server 2008 R2まではノード1にSQL OSを配置、残ったノードにSQL OSユーザーを置いていましたが、SQL Server 2012 R2からはWindows OSと同じノード0にSQL OSを置き、残りのソケットをすべてSQL OSユーザーとして使うアーキテクチャーに変更されています。この一番の利点は、1番目のノードにOSとSQL Serverのエンジンを置くことでノードをまたがないディスクI/OとメモリーI/Oが可能になったことです。

もう1点の大きな変更はメモリーの使用方法です。SQL Server 2008 R2までは、OSからSQL Serverがメモリーを獲得する際に8Kバイトのシングルページと2Mバイトのラージページの2つがありました。2012以降はラージページに統一されました。例えば、2012から登場した列ストアインデックスは、大規模なデータセットに対するクエリ

■ 図4 Hyper-V 仮想NUMA設定(1)



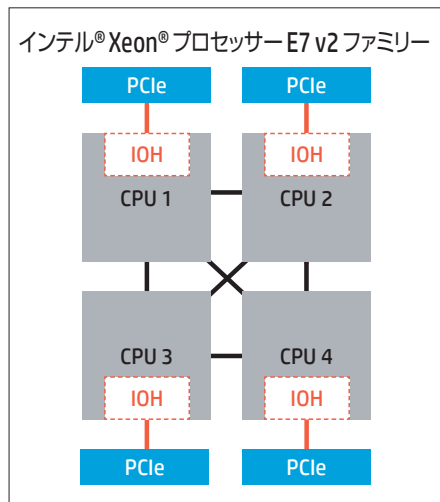
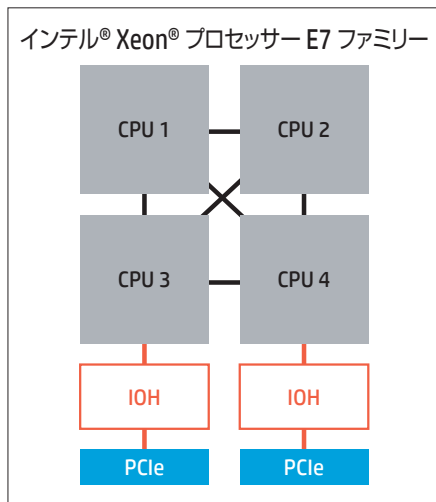
■ 図5 Hyper-V 仮想NUMA設定(2)





日本マイクロソフト株式会社
 サーバプラットフォームビジネス本部
 アプリケーションプラットフォーム製品部
 技術顧問
 熊澤 幸生 氏

■ 図6 各ソケットのCPUにI/Oが内蔵されデバイスが直結
 インテル® Xeon® プロセッサ E7 v2 ファミリー



で効果を発揮します。列ストアインデックスはBLOBを使ってデータ領域を管理します。この高性能化への対応の1つがラージページへの統一です。

さらにSQL Serverの2014からは、インメモリOLTPが登場しました。ファイルストリームを使ってデータベースの領域を用意します。このように、メモリの使い方にさまざまなシナリオが増えてきました。それが相互干渉しないようにするのがリソースガバナーという機能です。これを使うと、SQL Serverのインスタンスの中をあたかもパーティショニングで分けてリソース管理ができます。

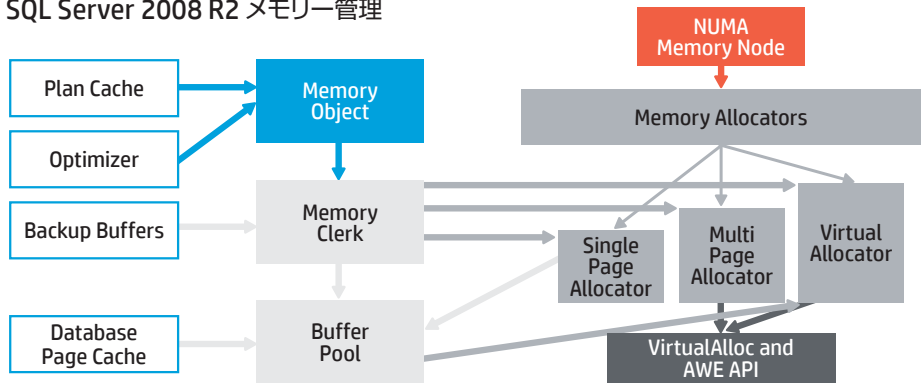
仮想環境でもう1つポイントとなるのはディスクI/Oです。DL580 Gen8が搭載するインテル® Xeon® プロセッサ E7 v2 ファミリーが採用したインテグレートドI/Oでは、DDIO(Data Direct I/O)という仕組みを取り入れたことにより、PCI Expressからのデータを直接プロセッサの

キャッシュに読み込み、メインメモリへ転送せずにプロセッサで処理をおこないます。従来外部にあったI/Oハブが内蔵され、各ソケットに直結となったことにより、レイテンシは30%削減され、今まで以上の高速なデータ転送がおこなえるようになりました。

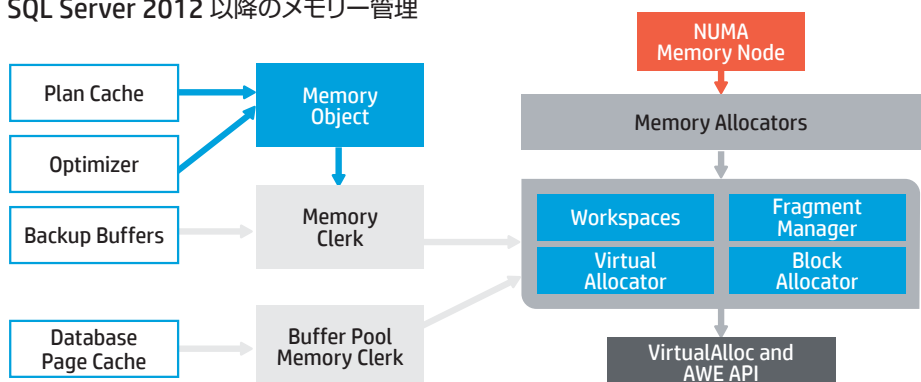
「CPUとPCIバスが直結となりソケット単位でのI/Oが可能になったことは、パフォーマンスチューニングの観点からは、どのCPUにどのデバイスを接続するかを意識しなくてはならないということです。また、仮想環境ではVHD (Virtual Hard Disc) にアクセスする際のオーバーヘッドがあるため、ディスクがボトルネックとならないようSSDなどの高速なディスクを使用する必要が出てくるでしょう」(熊澤氏)。

■ 図7 SQL Serverにおけるメモリー管理の変更

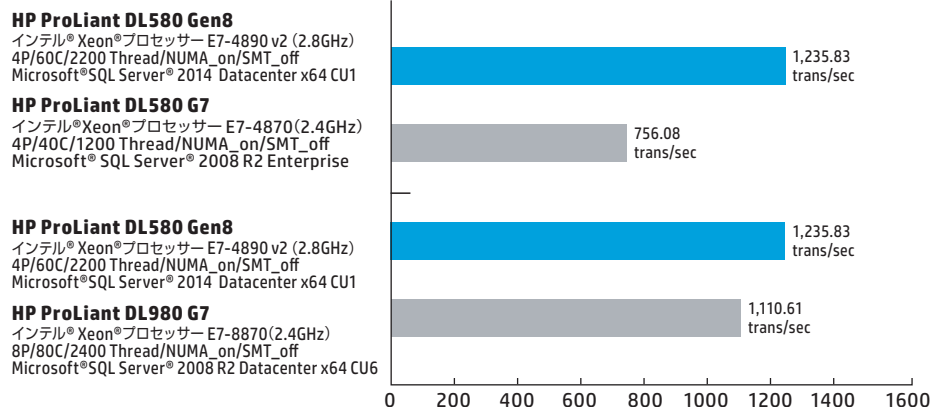
SQL Server 2008 R2 メモリー管理



SQL Server 2012 以降のメモリー管理



■ 図8 OLTP負荷テストにおけるピーク時のスループット



スケールアップ係数値に近い リニアなパフォーマンス向上も確認

OLTP負荷テストにおけるピーク時のスループットについても測定を行いました。4ソケット60コアのインテル® Xeon® プロセッサ E7 v2 ファミリーを搭載したDL580 Gen8と、旧世代の4ソケット40コアのインテル® Xeon® プロセッサ E7 ファミリーを搭載したDL580 G7との比較では、ハイパースレッドオフの状態では1.64倍とリニアに性能が向上することを実証しました。

旧世代の8ソケット80コアのインテル® Xeon® プロセッサ E7 ファミリーを搭載するDL 980 G7と4ソケット60コアのインテル® Xeon® プロセッサ E7 v2 ファミリーを搭載するDL580 Gen8との比較では、ソケットとコア数

ともに少ないDL580 Gen8がハイパースレッドオフの状態でも1.11倍の性能を発揮することを実証しました。これはハードウェアのコストダウンと電力の省力化を図りながら従来と同性能を得られることを証明しており、ユーザーにとっての大きなベネフィットです。

HP ProLiant DL580 Gen8とインテル® Xeon® プロセッサ E7 v2 ファミリー、Microsoft® SQL Server® 2014の組み合わせが発揮するパフォーマンスは、データベースも仮想環境で運用する本格的な時代を迎えることを示唆するといえるでしょう。これらのプラットフォームが、ビッグデータを処理する冗長化データベースや大規模な仮想化サーバー統合の時代にも、大きく貢献しようとしています。絶え間なく進化を続けるこれらのプラットフォーム技術に、これからもぜひご注目ください。

■ 今回の検証環境

HP ProLiant DL580 Gen8	
データベース・プラットフォーム	Microsoft® SQL Server® 2014 Enterprise Edition CU1
サーバー OS	Microsoft® Windows Server® 2012 Datacenter R2
メモリー	2TB
ソケット数	4 ソケット60 コア
プロセッサ(動作周波数)	インテル® Xeon® プロセッサ E7-4890 v2(動作周波数2.8GHz)
HP ProLiant DL360 Gen8	
データベース・プラットフォーム	Microsoft® SQL Server® 2014 Enterprise Edition CU1
サーバー OS	Microsoft® Windows Server® 2012 Datacenter R2
メモリー	128GB
ソケット数	2 ソケット20 コア
プロセッサ(動作周波数)	インテル® Xeon® プロセッサ E5-2690 v2(動作周波数3.0GHz)



安全に関するご注意 ご使用の際は、商品に添付の取扱説明書をよく読みの上、正しくお使いください。水、湿気、油煙等の多い場所に設置しないでください。火災、故障、感電などの原因となることがあります。

お問い合わせはカスタマー・インフォメーションセンターへ

03-5749-8330 月～金 9:00～19:00 土 10:00～17:00(日、祝祭日、年末年始および5/1を除く)

HP ProLiant製品に関する詳細情報は <http://www.hp.com/jp/proliant>

本ページに記載されている情報は取材時におけるものであり、閲覧される時点で変更されている可能性があります。予めご了承ください。

本書に含まれる技術情報は、予告なく変更されることがあります。

Intel、インテル、Intel ロゴ、Intel Inside、Intel Inside ロゴ、Intel Atom、Intel Atom Inside、Intel Core、Core Inside、Intel vPro、vPro Inside、Celeron、Celeron Inside、Itanium、Itanium Inside、Pentium、Pentium Inside、Xeon、Xeon Phi、Xeon Inside、Ultrabook 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 の商標です。

記載されている会社名および商品名は、各社の商標または登録商標です。

記載事項は2014年7月現在のものです。

© Copyright 2014 Hewlett-Packard Development Company, L.P.

日本ヒューレット・パッカード株式会社

〒136-8711 東京都江東区大島2-2-1

