

社会インフラを支える オープンソースのクラウド基盤ソフトウェア

Open Source Cloud Platform for Public Infrastructure

福田 安宏

Fukuda Yasuhiro

庄子 智誉

Shoji Tomotaka

植田 良一

Ueda Ryoichi

依藤 慈孝

Yorifuji Yasutaka

現在、プライベートクラウドが注目されている。それも、これまでの「どのようにAmazon EC2^{※1)}のようなIaaSクラウドを構築するか」から、ビジネス要件に「IaaSクラウドの持つメリットをどのように活用するか」に関心が移ってきている。北海道大学アカデミッククラウドは、この状況を踏まえた事例の一つである。これは、Amazon EC2 相当のIaaS機能を実現した本格的なプライベートクラウドであり、Amazon^{※2)} 同様に最新HPC/Hadoop^{※3)} 向けのサービスを実際に運用している。

日立グループは、日立クラウドソリューション「Harmonious Cloud」のプライベートクラウドソリューションの適用実績を多数有しており、今後の社会インフラとなるクラウド基盤ソフトウェアを中心としたクラウドエコシステムとともにソリューションを強化していく。

1. はじめに

導入実績の増加とコストパフォーマンスの高さから、オープンソースのクラウド基盤ソフトウェアに対する関心が高まっている。このクラウド基盤ソフトウェアとは、NIST (National Institute of Standards and Technology) が定義する三つのサービスモデルの一つであるIaaS (Infrastructure as a Service) を実現するためのソフトウェアである。

このクラウド基盤ソフトウェアが社会インフラを支える技術となるためには、三つの条件を備える必要があると考える。まず、サステナビリティの観点から信頼できる実績のあるコミュニティが管理するオープンソースであることが挙げられる。次に、ハードウェア資源をサービス化するためのフレームワークを提供していること、最後に、パブリッククラウドとプライベートクラウドの両方に適用できることである。

特に、パブリッククラウド提供事業者と同じクラウド基盤ソフトウェアが導入できれば、ハイブリッドクラウドと

呼ばれるクラウド間連携がより具体的に推進でき、社会インフラとしての発展を期待できる。

ここでは、クラウド基盤ソフトウェアの代表であるApache^{※3)} CloudStack^{※3)}を導入した北海道大学アカデミッククラウドの事例を基に、得られた見識、その後の日立グループのプライベートクラウド導入実績から共有できたユーザーエクスペリエンスで重視すべき視点、クラウド基盤ソフトウェアを取り巻くクラウドエコシステムの状況を考慮したマーケットの動向、および今後の展望について述べる。

2. 北海道大学アカデミッククラウドの事例

「Harmonious Cloud」のプライベートクラウドソリューションの代表的な導入事例が、北海道大学アカデミッククラウドである。このシステムは、北海道大学情報基盤センターの「学際大規模計算機システム」の一部として導入された、国内最大規模の学術クラウドシステムである(図1参照)。

北海道大学を含む全国共同利用施設を持つ大学は、クラウドの登場前から学術研究の発展を目的に、全国の研究者に対して、保有している計算機資源を学術ネットワーク経由でサービスとして提供している。

北海道大学では、大学全体の計算機資源を統合し、消費電力の削減や運用効率の向上を目的とした最適化計画を掲げている。その手段としてはクラウドが適しており、最適化計画に大きく貢献している。

※1) Amazon EC2は、米国Amazon Technologies Inc.の登録商標である。

※2) AmazonおよびAmazonのロゴは、Amazon.com, Inc.またはその関連会社の商標である。

※3) Hadoop, Apache, CloudStackは、Apache Software Foundationの登録商標または商標である。

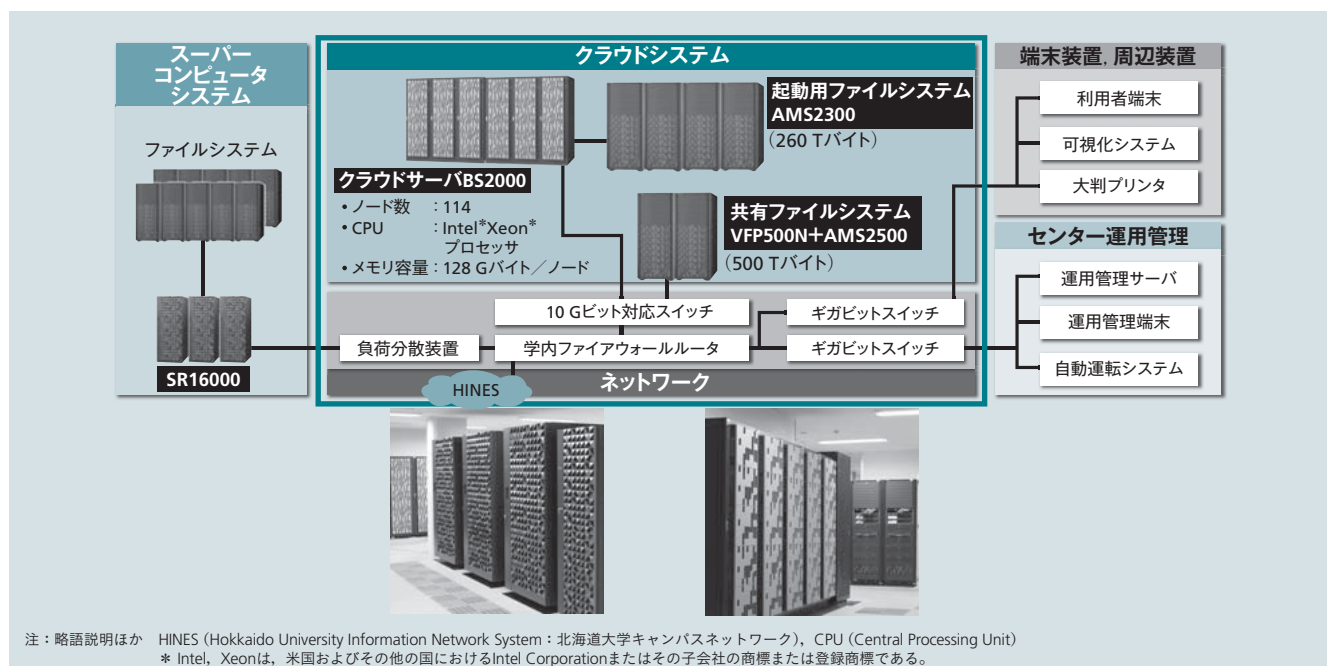


図1 | 北海道大学アカデミッククラウドの概要

北海道大学アカデミッククラウドは、北海道大学情報基盤センターの「学際大規模計算機システム」の一部として構成される。

2.1 北海道大学アカデミッククラウドの特長

北海道大学アカデミッククラウドは、170テラフロップス（TFLOPS：浮動小数点演算を1秒間に1兆回実行する能力）を超える演算性能を有するスーパーコンピュータシステムとともに運用される。HPC（High-performance Computing）用途を意識した性能やオンデマンドで利用できる複数の計算機をまとめたクラスタパッケージによるMPI（Message Passing Interface）環境、Hadoopなどの分散処理環境サービスを提供している。

このHPC用途の性能要件に対応するため、1ノードが高い演算能力を有する統合サービスプラットフォーム「BladeSymphony」のハイエンドモデルBS2000〔CPU（Central Processing Unit）：Intel Xeon E7-8870（10コア）を4プロセッサ、メモリ：128 Gバイト〕が導入された。また、クラスタ内の高速な通信を実現するため、10 Gビットイーサネット^{※4）}をノード間通信に採用した。

クラウド環境で重要なストレージ部分は、ミッドレンジディスクアレイ Hitachi Adaptable Modular Storage 2000 シリーズを8 Gビット/sのファイバチャネルで演算ノードと接続する構成とした。利用者の仮想マシン環境を保存するストレージシステムは、高い信頼性に加え、システム稼働中の性能調整を自動化するダイナミックロードバランスコントローラ機能や、必要に応じてストレージ容量の割り当てを自動的に行うボリューム容量仮想化機能Hitachi Dynamic Provisioningを活用した設計を行い、クラウド

サービスに求められる性能と運用の柔軟性を確保している。

2.2 クラウド基盤ソフトウェアの採用

北海道大学アカデミッククラウドにおいて、Amazon EC2（Amazon Elastic Compute Cloud）相当の使いやすさを備え、高性能で柔軟性を有するIaaSをどのように構築するかを検討した答えが、クラウド基盤ソフトウェアApache CloudStackの採用である。2010年の検討において、日立グループが重視した項目は以下の3点である。

- (1) 安定稼働を見込める信頼性や実績があること
- (2) 顧客の利用者管理や課金処理業務との連携が容易であること
- (3) 実績のあるコミュニティが管理するオープンソースであること

検討方法としては、実績を含めた機能比較などの調査に加え、オープンソースソフトウェアに対するソースコードレベルのレビューと評価環境による動作確認を実施した。選択肢の中でApache CloudStackのみが重視する項目を満たすことができた。主な評価結果は以下のとおりである。

- (1) 複数の海外データセンター事業者の採用実績があり、信頼性が高いこと
- (2) Web API（Application Programming Interface）が十分に整備されており、カスタマイズが容易で保守性も高いと判断できたこと
- (3) 比較した一部のオープンソースソフトウェアでは、完結していないコードが含まれているなど、ソースコードの可読性が低かったが、Apache CloudStackでは、ソースコー

※4）イーサネットは、富士ゼロックス株式会社の登録商標である。

ドの不整合は確認されず、十分に管理されていると判断できたこと

ただし、日立グループとして初めてのIaaS構築経験であったため、米国Cloud.com社（現 Citrix社）とサポート契約を行った。現在、国内および海外データセンター事業者が多数採用している実績が示すとおり、IaaS構築に必要な機能要件を網羅しており、実運用に十分耐えうる信頼性を備えていることが証明されている。

2.3 北海道大学アカデミッククラウドが求める要件

北海道大学アカデミッククラウドに求められた要件は主に二つあり、一つ目は、IT（Information Technology）が専門ではなく、パブリッククラウド利用は難しいと感じる教員や研究者でも、簡単に利用できるような直感的なインタフェースを提供することである。二つ目は、HPC用途でも利用可能なクラスタ環境を研究者に素早く提供することである。

一つ目の要件を達成するために、顧客と議論を重ね、利用者管理・課金処理業務と密接に連携した利用者向けポータル画面を作成した（図2参照）。これにより、Apache CloudStackの標準ポータル画面のみでは実現困難であった、顧客が望む利用者に優しいインタフェースを提供できた。

なお、Apache CloudStackの標準ポータル画面は、利用者が作成したVM（Virtual Machine：仮想マシン）を管理するための機能に加え、利用者ごとの仮想ルータ（ファイアウォール）や新規のOS（Operating System）セットアップに必要なコンソール機能などの優れた機能を有するため、利用者向けポータル画面と連携させて運用している。

二つ目の要件を達成するため、HPC用途にMPIライブラリおよびHadoopが利用可能なテンプレートを作成し

た。Hadoopは、Apache Software Foundationが開発・公開している、大規模データを効率的に分散処理・管理するためのソフトウェアである。このテンプレートは、Apache CloudStackがVM上のOSのrootパスワードを設定する仕組みを改良し、クラスタにおけるマスターノードとスレーブノードの役割を設定する仕様を追加した。これにより、利用者は利用者向けポータル画面からクラスタの台数を指定するだけで、必要なクラスタ環境を作成できる。なお、クラスタ起動テストにおける実績として、2,000 VMのクラスタ生成まで検証を行っている。

2.4 IaaS構築と稼働実績から得られた見識

IaaS構築においては、ハードウェアからハイパーバイザ、クラウド基盤ソフトウェアまで一体となった支援体制が重要である。加えて、利用者に提供するシステムの安定性や品質、性能、消費電力の実体を評価するため試験を実施する必要がある。

北海道大学アカデミッククラウド構築においてはこれらに留意し、以下の見識を得ることができた。

(1) Hadoop用途に必要な性能設計

Hadoop用途に適した性能要件を満たすためには、IaaS環境に特化したVM作成方式を設計したことが有効であった。

(2) 消費電力の評価

実際の消費電力を評価するために、LINPACKベンチマークを用いた消費電力の測定が有効であった。LINPACKベンチマークは、LINPACKに基づいたベンチマークプログラムであり、システムの浮動小数点演算性能を評価する際に用いられるプログラムである。

(3) 品質確保に必要な負荷試験の実施と対策

実際に提供する環境で負荷試験を実施することで、潜在不良を検出できた。この不良対策には、ハードウェア工場と一体の対応が有効であった。

2.4.1 Hadoopに求められるVM作成方式

Hadoopは、ビッグデータの分散処理に優れたソフトウェアである。このHadoopが前提としているハードウェアは、計算ノードごとにディスクを有する構成である。これに対し、IaaS環境では、計算ノードとして作成したVMが利用するディスクを明示的に指定することはできない。

北海道大学アカデミッククラウドの場合は、ハイパーバイザにCitrix^{※5)} XenServer^{※5)}を使用している。Apache CloudStack標準のVMアロケーション仕様では、一つのXenServerクラスタ内に集中させてVMを作成する。この



図2 | 北海道大学アカデミッククラウドの利用申請画面

あらかじめ設定されているサービスレベル、パッケージの組み合わせを選択するだけで、簡単にクラウドサービスを利用することができる。

※5) Citrix, Citrix XenServerは、Citrix Systems, Inc.の商標または登録商標である。

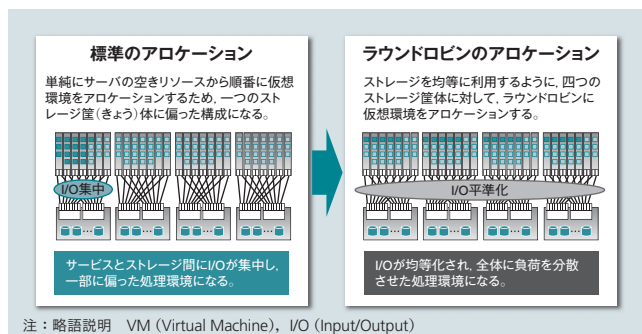


図3 | VMのラウンドロビンアロケーション

VMのラウンドロビンアロケーションは、一つの共用ディスクにI/O（ディスク入出力処理）が集中して、ボトルネックになることを防ぐ。

ため、一つのXenServer クラスタが持つ一つの共有ディスクボトルネックとなり、Hadoop 本来の分散処理による高スループットを実現することができない。

この対策として、すべての共有ディスクに割り当てるVM数を均等にするため、Apache CloudStackが提供するAPIを利用し、VMをXenServer クラスタごとにラウンドロビンでアロケーションする方式を開発した（図3参照）。この方式の追加により、256 VM以下であればApache CloudStack標準のVMアロケーション仕様と比較して、Hadoop 付属のTeraGenとTeraSortの実行時間を最大±程度に短縮することができた。TeraGenとTeraSortは、Hadoopとともに配布されている大規模なデータの生成、ソートに使うサンプルプログラムである。

2.4.2 LINPACKベンチマークを用いた消費電力測定

近年のハードウェアに求められる要件には、性能や機能以外に省電力化や電力消費量の可視化が含まれる。BladeSymphonyは電力監視機能を備えており、運用中のサーバの消費電力測定が可能である。

実際にシステムの最大消費電力を測定するには、サーバのCPU利用率を上げる必要がある。LINPACKベンチマークは、サーバの限界に近い性能を引き出すことを目的とするベンチマークプログラムであり、対象サーバのCPU使用率を100%にする負荷試験を実施することができる。そのため、この負荷試験実施時間に、サーバの最大消費電力

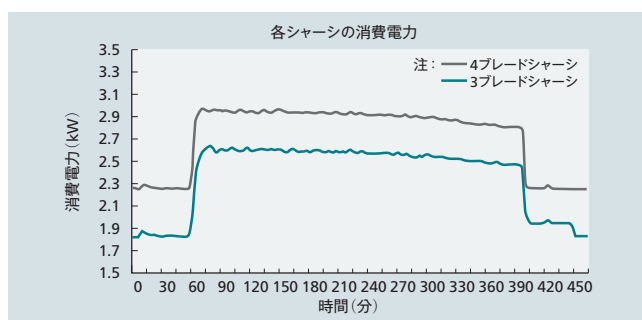


図4 | LINPACKベンチマーク実行時のシャシ当たりの消費電力
1シャシのブレード搭載数ごとに消費電力を測定できる。

を測定できる（図4参照）。

ブレード型のサーバは、ブレード搭載数が少ない場合など、定格で表示される消費電力量では実際の消費電力が分からず、実測しなければ把握できない場合がある。また、複数サーバすべてのCPU使用率を100%にするアプリケーションを用意することも難しいため、今回の負荷試験に用いたLINPACKベンチマークは、実際の消費電力を確認する目的に有効である。

2.4.3 ハイパーバイザの潜在不良対策

北海道大学アカデミッククラウド構築の現地作業期間は、1か月であった。この期間内にIaaS環境の構築とテストを完了し、サービス提供を開始する必要があった。

提供するすべての機能テストに加えて、ハードウェアを中心とする非機能要件を効率よく確認するために、実際に提供する環境に2,000 VMのHadoop用クラスタを作成した。このHadoopクラスタ上でTeraGenとTeraSortを用いて負荷試験を実施し、2件のハイパーバイザの潜在不良検出と対策を実施することができた。

1件目は、ハイパーバイザが予期せず不当にリポートする障害である。この障害は、10 Gのネットワーク機器のソフトウェア更新後に顕在化した不良である。

障害が発生した初期の切り分け結果では、1 Gのネットワークでは発生しないことから10 Gのネットワーク機器が疑われた。ところが、ハードウェア設計技術者と一体で追加調査を行ったところ、ハイパーバイザに原因があると考えられた。そこで、ハードウェアやドライバを追加検証し、速やかにハイパーバイザベンダーに情報提供した。その結果、原因はハイパーバイザの潜在不良であることが分かり、パッチ適用によって短期間で問題を解決することができた。

2件目は、Hadoopを用いた負荷試験において、高負荷の状態になるとハイパーバイザ上のディスクドライバがディスクアクセスのパスを閉塞させる障害である。

この障害では、ディスクアレイシステムのパフォーマンスモニタ結果をハードウェア設計部で評価し、ディスクアレイが高負荷になった場合に発生することが特定できた。対策として、ハイパーバイザ上のディスクアクセス制御プログラムのパラメータを変更し、現象の発生を抑制することにより、障害を回避することができた。

3. プライベートクラウドのユーザーエクスペリエンス

IaaSを実現する「Harmonious Cloud」のプライベートクラウドソリューション提供期間を通して、顧客のユーザーエクスペリエンスを共有することができた。この経験から得た顧客が重視する指標の一つが、レスポンス時間で

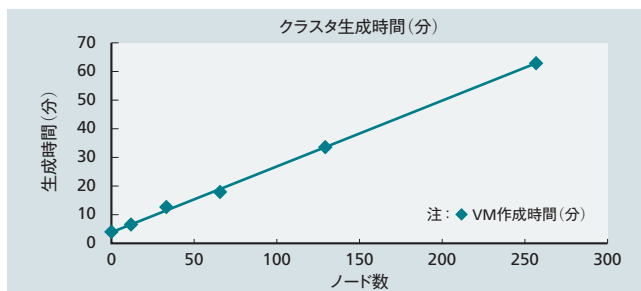


図5 | クラスタ環境作成時間

1ノードの場合4分、250ノードクラスタを生成しても1時間程度で利用可能になる。

あった。

レスポンス時間は、非機能要件であるため機能性や使い勝手に比べて見落とされやすいが、ユーザーエクスペリエンスを左右する重要な指標であり、ハードウェアの選択からIaaS環境の設計に至るまで、日立グループの経験を一貫して提供できる要件である。

3.1 VMを用いたクラスタ環境作成時間

IaaSは、利用者がオンデマンドでVMを作成できる。一つのVM作成にかかる時間の想定は10分とした。この時間であればオペレータに処理を依頼する方法と比べて、十分に短いと考えていた。しかし、実際にプライベートクラウドを導入した顧客は、対象となる利用者を広く想定しているため、利用者が最初に体験するVM作成時間を重視した。

この経験から、利用者が快適と感じるレスポンス時間の重要性に対する認識を新たにした。VM作成時間は、個人の感覚に影響されるため、性能指標として定義することは難しいが、時間短縮に必要なハードウェア設計、VM作成処理設計などの、有用な見識を蓄積できた。

北海道大学アカデミッククラウドの事例では、一つのVM作成時間が4分程度であり、想定よりも短かった。さらに、複数VMのクラスタ環境作成方式をチューニングした結果、クラスタ生成時間が、顧客が想定していた時間の $\frac{1}{2}$ から $\frac{1}{3}$ 程度となり、高速にクラスタ環境が提供されると利用者からの評価も高い(図5参照)。

3.2 ハイパーバイザごとのVM作成時間比較

クラウド基盤ソフトウェアは、複数のハイパーバイザに対応しているため、互換性や予算などの機能要件に合わせてハイパーバイザの選択や組み合わせができる。

非機能要件に着目した場合、VM作成時間がハイパーバイザによって異なる(図6参照)。

同図にあるシステムBとCは、同一ハードウェア上の異なるハイパーバイザで実測した結果である。システムC

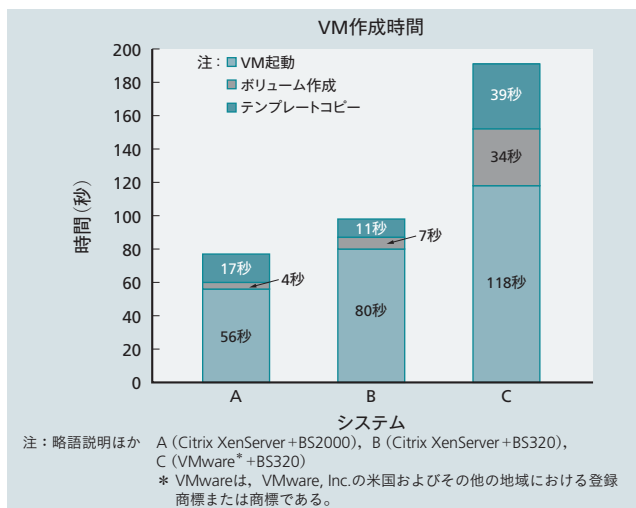


図6 | VM作成時間

Apache CloudStackではさまざまなハイパーバイザを利用することができるが、ハイパーバイザの種類によってVM作成の処理時間が異なる。

のVM作成時間は、システムBに対して2倍近い。参考のシステムAは高性能なハードウェア構成であるが、同一ハイパーバイザのシステムBとの比較ではレスポンス時間の差異は2割程度であり、ハードウェアの構成よりもハイパーバイザの種別に依存することが分かる。もし顧客がVM作成時間を意識しているのであれば、システムBのハイパーバイザの提案を検討したほうがよい。

このほかにも、ハイパーバイザによって一斉に作成できるVM数にも大きな差があるため、非機能要件として把握しておくべきである。

3.3 Hadoop環境における実行時間の比較

Apache CloudStack が提供するIaaS環境において、Hadoop用途に特化した複数VMのクラスタ環境作成方式が性能設計に有効な事例を述べた。このApache CloudStackはベアメタルプロビジョニング機能も有しており、VMのクラスタ環境作成と同じ方式で、オンデマンドで物理サーバを使ったHadoopクラスタを生成できる。

VMのクラスタ(仮想サーバA, B)と、Hadoop用途に適したノードごとにハードディスクを有する高密度実装の物理サーバHA8000-bd/BD10で実行したHadoopのTeraSortのデータ処理速度を図7に示す。

1ノードの実行結果では、VMと物理サーバの違いによる処理速度の差が出ない。

3ノードの実行結果では、仮想サーバAのデータ処理速度が向上しない。この仮想サーバAは、一つの共有ディスクを利用するため、ディスクがボトルネックとなる。

VMノードごとに物理ディスクを占有する仮想サーバBは、物理サーバと同様にデータ処理速度が向上することから、3ノードでのVMと物理サーバの性能差はない。

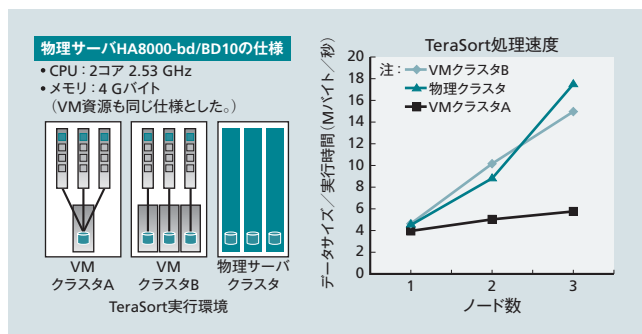


図7 | Hadoop TeraSort実行結果

VMクラスターAは、すべてのノードが一つのディスクを共有する。VMクラスターBと物理サーバはノードごとに独立したディスクを有する。

ただし、ノード数を増やした場合や実行するアプリケーションを変えた場合は、VMは物理サーバに比べてデータ処理速度がノード数の増加に合わせて伸びない結果となる場合がある。

プライベートクラウドの非機能要件を設計するためには、予想されるボトルネックを回避することに加え、これらの実測値や実績から得た見識を生かすことが重要である。

4. 社会インフラとなるクラウド基盤ソフトウェア

クラウド基盤ソフトウェアであるApache CloudStackは、2010年にCloud.com社からマーケットに登場した。当初からコミュニティを重視しながら、すでに存在する製品やサービスと連携することで不足しているソリューションを補いつつ成熟してきた。また、網羅的に整備されたWeb APIやAmazon互換のAPIを提供することで、ハイブリッドクラウド利用に対応し、クラウドエコシステムと共存してきたと言える。

2012年になり、Cloud.com社を買収したCitrix社によってApache Software FoundationにApache CloudStackが寄贈された。これにより、クラウドエコシステムに関わるコミュニティや企業が、Apache CloudStackの周辺に新たな価値を創造し、クラウドエコシステムを拡大させていくと考えられ、デファクト標準をめざす動きにも読み取れる。

一方、社会インフラとして確立されるには、フォーラム標準や、デジュール標準により、互換性・相互運用性・性能・品質・信頼性・安全性などの規格化が必要であり、グローバルな視点でこれらの標準化の動向に意識を向けるべきである。

最近のIaaSマーケットの動向は、クラウド基盤ソフトウェアの普及とともに、パブリッククラウド提供事業者の特色が出しにくくなる状況にあり、より上位のPaaS (Platform as a Service) やハイブリッドクラウド、データセンター間サービスなどが重要視されてきている。

このことは、クラウド基盤ソフトウェアが規格化され、

長期保守に向く社会インフラに必要な要件を備えつつある証拠だと考えられる。

今後は、国、自治体、企業、大学などの組織も、同じ規格のクラウド基盤ソフトウェアを採用し、パブリッククラウド事業者を含む連携した社会インフラとして、実社会全体を支えていくと予測する。

5. おわりに

ここでは、Apache CloudStackを導入した北海道大学アカデミッククラウドの事例を基に、得られた見識、その後の日立グループのプライベートクラウド導入実績から共有できたユーザーエクスペリエンスで重視すべき視点、クラウド基盤ソフトウェアを取り巻くクラウドエコシステムの状況を考慮したマーケットの動向、および今後の展望について述べた。

クラウド基盤ソフトウェアを中心に考えた場合、サーバ、ストレージ、ネットワークといったハードウェアと仮想化機能は、サービス提供を前提に根本的な価値が見直され、サービス指向を強めていくべきである。

企業など、組織内の情報システム部門は、このクラウド基盤ソフトウェアを使って、組織内に競争力を強化するためのクラウドサービスを提供することができる。

このクラウドサービスは、情報システム部門に、利用者との積極的なコミュニケーションや企業戦略に必要な俊敏性を重視する行動を促し、企業の変革を促進させるための重要な役割を期待させることになる。

執筆者紹介



福田 安宏

1992年日立製作所入社、情報・通信システム社 公共システム事業部 学術情報ソリューション本部 学術情報システム部 所属
現在、公共分野のクラウドビジネス推進に従事
技術士（情報工学部門）



植田 良一

1994年日立製作所入社、横浜研究所 情報プラットフォーム研究センター ソフトウェアプラットフォーム研究部 所属
現在、大量データ処理システム、クラウドシステムの研究開発に従事



庄子 智誉

2007年株式会社日立東日本ソリューションズ入社、第三ソリューション事業統括本部 公共ソリューション本部 先端技術ソリューション部 所属
現在、クラウドシステムの設計、構築に従事



依藤 慈孝

1996年日熱エンジニアリング株式会社（現 株式会社日立ソフトテック）入社、ビジネスソリューション本部 ビジネスソリューション部 所属
現在、クラウドシステムの設計、構築に従事