

## 広域分散環境を提供する HPCI先端ソフトウェア運用基盤の設計

滝澤 真一朗<sup>†1</sup> 棟 朝 雅 晴<sup>†2</sup> 宇 野 篤 也<sup>†3</sup>  
小 林 泰 三<sup>†4</sup> 實 本 英 之<sup>†5</sup>  
松 岡 聡<sup>†1,†6</sup> 石 川 裕<sup>†5</sup>

平成 24 年秋の運用開始が予定されている HPCI では HPC 研究者がスーパーコンピュータ「京」を有効活用することの支援を目的とし、京と基盤センター群が保有するスーパーコンピュータ間の認証基盤統一、データ共有の実現から開始する。しかしながら、スーパーコンピュータはバッチキューでジョブ管理されていることや、計算ノードでの管理者権限がないため、OS や分散システムの研究を行う HPC 研究者向けの利用環境条件を満たさない。そこで我々は、利用者に対してシステムへの管理者権限を付与する広域分散システムのホスティング機能を提供する、先端ソフトウェア運用基盤を設計する。本稿では先端ソフトウェア運用基盤の設計、および、先行システムとして運用されている RENKEI-PoP による事例を紹介する。

### Design of Advanced Software Deployment Infrastructure in HPCI Wide-area Distributed Environment

SHINICHIRO TAKIZAWA,<sup>†1</sup> MASA HARU MUNETOMO,<sup>†2</sup>  
ATSUYA UNO,<sup>†3</sup> TAIZO KOBAYASHI,<sup>†4</sup>  
HIDEYUKI JITSUMOTO,<sup>†5</sup> SATOSHI MATSUOKA<sup>†1,†6</sup>  
and YUTAKA ISHIKAWA<sup>†5</sup>

The purpose of HPCI, which will be operated from autumn 2012, is to support HPC researchers to use K supercomputer, and its initial services are a federated authentication and global file sharing between K and supercomputers provided by computer centers in Japan. However, supercomputers are not suitable for HPC system researchers as their operations do not give users enough privileges. We design the advanced software deployment infrastructure that hosts distributed systems where researchers can have administrator privileges. We introduce the design of the system and a precedent system implemented on

RENKEI-PoPs that use the same software.

#### 1. はじめに

理化学研究所 計算科学研究機構に構築されるスーパーコンピュータ「京」を中核とし、基盤センター群が保有するスーパーコンピュータや他の国内計算機資源を連携させて利用するための環境として「革新的ハイパフォーマンス・コンピューティング・インフラ」(以下、HPCI)の構築が進んでいる。HPCIの基本目標はハイパフォーマンスコンピューティング(以下、HPC)研究者が世界最先端の計算性能を誇るスーパーコンピュータ京を有効活用することの支援であり、研究者が実行環境を他の計算機資源から容易に京に移すための技術として、統一認証基盤、広域データ共有環境の整備を第一の業務として行っており、京の運用開始と合わせ、平成 24 年度秋からの運用が予定されている。

HPCIに所属する基盤センターが所有するスーパーコンピュータは、差異はあるが主に以下のような運用上の特徴がある。

- プライベートネットワーク、あるいはファイアウォールによる外部との通信制限
- 特定の OS・システム・アプリケーションソフトウェアのみの提供
- 一般利用者としての利用権限
- バッチジョブスケジューラによるジョブ管理

このため、各種シミュレーションを行う HPC 応用分野研究者が計算資源としてスーパーコンピュータを使うには問題ないが、一方で OS や分散システムの研究者は、管理者権限が無

<sup>†1</sup> 東京工業大学  
Tokyo Institute of Technology

<sup>†2</sup> 北海道大学  
Hokkaido University

<sup>†3</sup> 理化学研究所 計算科学研究機構  
RIKEN Advanced Institute of Computational Science

<sup>†4</sup> 九州大学  
Kyushu University

<sup>†5</sup> 東京大学  
University of Tokyo

<sup>†6</sup> 国立情報学研究所  
National Institute of Informatics

いこと、外部とのネットワーク到達性が無いこと、が理由に使用できない者もいる。これら研究者は個別に計算資源を調達し、場合によっては他の機関との一定の合意のもと、資源を融通しあって研究に使用していた。あるいは近年では仮想マシンを用いたホスティングサービス、Amazon EC2に代表される IaaS (Infrastructure as a Service) 型のクラウドサービスを用いている者も多い。

HPCI の主要計算資源となるスーパーコンピュータには利用に際し上記の制約があり、一部の研究者には研究に適した環境ではないこと、また、研究プロジェクト毎に情報共有・ポータルサイトが必要となるが多々あるため、我々はこれら需要をまかなうための計算機環境を HPCI の枠組みの中で別途提供する必要があると考えている。そこで我々は「先端ソフトウェア運用基盤」と題し、利用者に対してシステムへの管理者権限を付与する広域分散システムのホスティングを提案する。先端ソフトウェア運用基盤は全国に分散配備された小型クラスター、あるいは仮想マシン (以下 VM) 実行機能を持つ計算資源を集約したシステムであり、利用者の VM 実行要求に基づき、オンデマンドに管理者権限が付与された VM を指定されたサイトで実行するものとして設計を進めている。本システム利用のための認証は HPCI 認証基盤と連携可能であり、本システムにより起動される VM の OS イメージや VM で使用するデータは、HPCI で運用する広域データ共有環境である、HPC ストレージに格納可能である。

本稿では HPCI 先端ソフトウェア運用基盤の設計概要、および、先行システムとして運用されている RENKEI-PoP による事例を紹介する。以降、2 章で HPCI に於ける先端ソフトウェア運用基盤の役割について述べ、その役割を果たすためのシステム要件を 3 章で述べる。4 章でその要件を満たすシステムの提案を行い、5 章で実装案と、先行システムとして導入されている環境での予備評価結果を示す。6 章で関連研究を紹介し、7 章で本論文をまとめる。

## 2. HPCI における先端ソフトウェア運用基盤の役割

表 1 に想定される HPCI 利用者の利用シナリオを載せる。この表は HPCI の利用者区分毎に、想定される利用シナリオと、そのシナリオで要求される計算資源のタイプをまとめている。応用 HPC 研究者は各種シミュレーションや大規模データ解析等の大規模並列計算を求める応用計算分野の研究者であり、CS(コンピュータサイエンス)系研究者は OS やシステム・ライブラリソフトウェア、分散システムなどの計算機システムの基盤となるソフトウェアの研究開発を行う研究者を表す。HPCI を構成するスーパーコンピュータ群はバッチ

表 1 HPCI 利用シナリオ  
Table 1 HPCI Usage Scenario

HPCI 利用者区分	利用シナリオ	要求環境
応用 HPC 研究者	HPC リソースを用いた大規模計算 (データ解析・シミュレーション等) を実行	HPCI 拠点リソース (スパコン) を利用
	特定 OS・ライブラリに依存するアプリケーションを実行	高信頼なサーバでのホスティングが求められる (1)
	HPC リソースにアクセスするためのポータルサービスを実行	
CS 系研究者	成果公開・プロジェクトメンバー内情報共有のための WEB ポータル運用	
	root 権限を必要とする OS 研究開発	管理者権限を持ったサーバからなる、必要に応じて分散配備可能な開発環境が求められる (2)
	開発したライブラリの、複数種 OS 上での検証	
	多拠点にグローバル IP を必要とする分散システムの開発	

キューを介したジョブ実行サービスを提供するシステムであり、表 1 の内、満たせる利用シナリオは 1 つのみである。先端ソフトウェア運用基盤ではそれ以外の利用シナリオを実現するために必要な、表中の (1), (2) に該当する要求環境を提供するシステムである。

これらの環境は、従来は研究者が個別に計算機を用意する、共同研究等の一定の合意の元で他機関の計算資源を利用する、クラウド計算資源や各種ホスティングサービスの商用サービスを用いる、といった方法で実現されていた。しかしながら現在、全国共同利用・共同研究拠点となっている計算機センターからもサーバや Web 等のホスティングサービスが提供されているため、HPCI にそれらサービスを組み込むことにより、研究者はそれらを有効活用できるよう運用できる。また、グリッドにおける Super Scheduler や Resource Broker 等の並列計算機およびストレージを連携させるサービスの実験を行う際に、Amazon EC2 の様な IaaS クラウドサービスでは環境整備、金額面でのコストが大きい。一方で、HPCI の枠組みの中でスーパーコンピュータに隣接したサーバのホスティングを提供することで、既存のスーパーコンピュータ資源を活用した実証・評価が行える。

先端ソフトウェア運用基盤は、各拠点に配備されたホスティングサービスを統一的に管理し、利用者が要求する環境を提供するシステムとしての役割を担う。

### 3. システム要件

先端ソフトウェア運用基盤を実現するための要件を以下にまとめる。

#### 利用者の要求する権限内でのシステム提供

利用者毎にシステムに要求する権限が異なるため、利用者毎の利用シナリオにあう環境を提供できることが望ましい。例えば、計算機システム研究者は管理者権限を要求するが、利用可能 OS に制約のある ISV (Independent Software Vendor) アプリケーションを実行しただけの応用分野研究者には管理者権限が必要ないこともある。特に、Web ホスティングのみを要求している利用者はコンテンツのアップロード権限さえあれば十分で、煩雑なサーバソフトウェアの管理から逃れたいと考えている者も多い。

#### 分散環境の動的配備

分散システム研究者の実験・評価環境提供のために、広域分散環境を配備できる機構が必要である。また、ホスティングサービスを提供する機関が大学の場合、定期的な全学停電が行われる組織もあり、このようなイベントに際して、利用者が動的に環境を構築・解放できる仕組みが必要である。また、利用者によっては複数のソフトウェア、複数のバージョンを切り替えて比較評価する場合もあり、システム側で支援する仕組みを有することが望ましい。

#### リソース利用制限

分散環境を利用者に提供するものの、個別ホスティングサービスへの利用権限は別途管理されなければならない。また、オーバプロビジョニングとならないよう、利用者毎の利用可能資源量を制限する機構も必要である。

#### 多種ネットワーク環境への接続

利用者が使用するサーバ・サービスを接続するネットワークは、利用者毎の用途によって異なる。一般的な Web 公開用のサーバであれば、Internet 接続しているネットワークにサーバを接続するだけで十分であるが、多サイトにまたがる分散システムの研究や、サイト内の特定資源と接続する必要のあるサーバ・サービスの構築のためには、専用のネットワークに接続できる必要がある。このように計算機システムだけでなく、ネットワークに対しても利用者が要求する環境を整備できる必要がある。

#### 操作・管理インターフェースの統一

各拠点に配備されたホスティングサービスを統一的に管理する役割上、サービス間での操作・管理のためのインターフェースの統一が求められる。しかしながら、既存システ

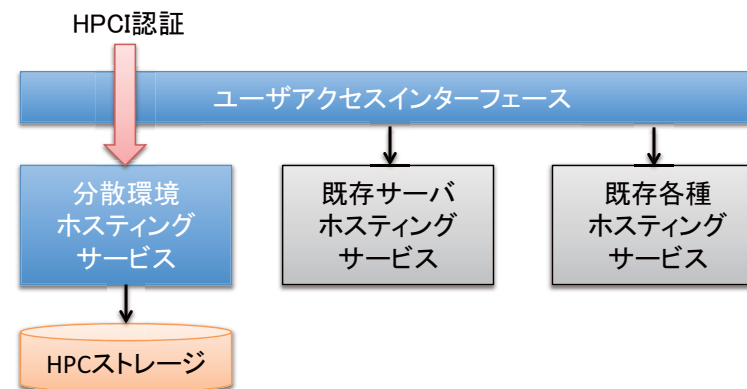


図 1 先端ソフトウェア運用基盤概要図  
Fig. 1 Advanced software development/operation environment

ムではこのような連携を意識して設計されていないものも多く、API レベルでの細粒度での統一は困難である。

#### HPCI 認証基盤・ストレージとの連携

HPCI の運用では課題が承認された研究者には HPCI ID が割り当てられ、その ID による Shibboleth や GSI による各種資源間の統一認証が提供される。また、その統一認証基盤を用いて HPC ストレージも利用可能である。ログイン操作の統一、および、HPC ストレージ利用のためには先端ソフトウェア運用基盤も HPCI 認証基盤と認証を統一する必要があり、また、先端ソフトウェア運用基盤により提供される利用者環境からも利用可能であることが望ましい。

### 4. 設 計

以上の要件を満す、先端ソフトウェア運用基盤のアーキテクチャを図 1 に示すとおり提案する。先端ソフトウェア運用基盤は各種ホスティングサービスをまとめるための「ユーザアクセスインターフェース」と分散環境を利用者に提供する「分散環境ホスティングサービス」を提供する。

ユーザアクセスインターフェースは各地に存在するホスティングサービスへの窓口として提供する。理想としては API レベルで異種ホスティングサービスの操作ができる、HPCI

認証基盤と連携したシングルサインオンで各サービスにログインできる、ことが望ましいが、実現には既存サービスへの大がかりな修正が要求される。HPCI は平成 24 年度秋から運用を開始するため、現実的にはホスティングサービスへのリンク集となるだろうと考えている。

分散環境ホスティングサービスは、利用者に root 権限を付与したサーバからなる広域分散環境を提供するサービスである。図中「既存サーバホスティングサービス」だけでも、複数組み合わせることで利用者が分散広域環境を構築することは可能であるが、どれだけの組織が資源供出可能か現時点で不明であること、操作インターフェースや利用可能資源タイプが異なる可能性が多々あり、サーバを連携させた運用が保証されないため、先端ソフトウェア運用基盤で個別対応することとした。

#### 4.1 分散環境ホスティングサービス

先端ソフトウェア運用基盤の分散環境ホスティングサービスは利用者にサーバを VM として提供する。VM を採用した理由は以下の通りである。

- サーバ運用の安全性. 利用者によるファームウェア等のハードウェア更新を許可しない。
- 資源の有効活用. 利用用途にもよるが、メニーコア・大容量メモリを搭載したサーバの普及により、多数の利用者の要求を少数サーバに集約可能。
- ソフトウェア資源の共有と再利用. VM で使用する OS イメージとして、各種ソフトウェアがプリインストールされたイメージを共有・再利用することで、利用者が求める環境を容易に迅速に構築できる。

分散環境ホスティングサービスのシステム構成を図 2 に示す。本システムは VM を実行するホスティング用リソース、OS イメージ保存用共有ストレージ、利用者からのリクエストを処理する管理サービスの 3 種類のコンポーネントより構成される。

##### ホスティングリソース

分散配置された VM 実行機能を持つサーバ、あるいはクラスタ構成の計算機資源である。利用者の VM を接続するための、各種ネットワーク回線が引き込まれている必要がある。後述する管理サービスからの VM 実行リクエストを受け取った後に、共有ストレージから OS イメージを取得し、VM を実行する。

##### 共有ストレージ

VM が使用する OS イメージが格納されるストレージである。分散配置されている全ホスティングリソースからアクセス可能な場所にある必要がある。OS イメージとしてはシステム管理者が用意した標準的なものだけでなく、利用者による登録も可能である。

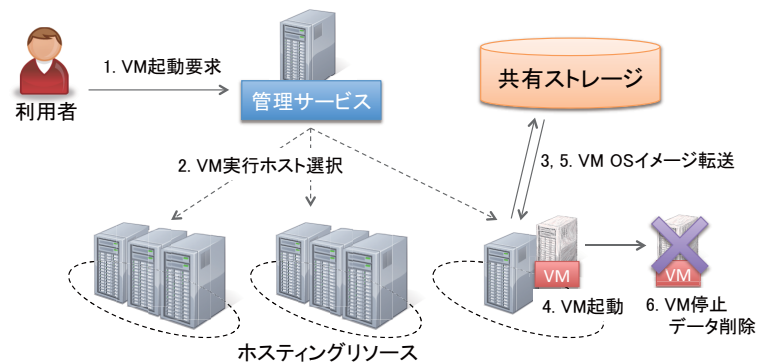


図 2 分散環境ホスティングサービス

Fig. 2 Distributed environment hosting service

HPC ストレージを共有ストレージとして使用することも検討している。

##### 管理サービス

利用者情報管理と、利用者からの VM・OS イメージ操作リクエストの受理、VM 実行・イメージ操作を司る。VM 実行要求に対しては、利用者が指定したサイトのホスティング用サーバを 1 つ選択し、そこで実行する。VM 停止要求が発行された場合には、VM 停止後、ホスティング用サーバから VM 関連データを全て消去する。OS イメージ操作に対しては、操作内容に対して、共有ストレージへのイメージの格納・取得・削除、属性情報の更新などを行う。

分散環境ホスティングサービスにおける、利用者による VM 起動から停止までの流れは以下になる。

- (1) 管理サービスに VM 起動を要求
- (2) 管理サービスは利用者の要求に基づき、VM を実行するサーバを選択
- (3) 選択されたサーバは共有ストレージより、利用者が指定した OS イメージを取得
- (4) サーバ上で、利用者のアクセスキーや取得した OS イメージを用いて、VM を起動し、指定されたネットワークに VM を接続
- (5) 利用者は任意のタイミングで VM の OS イメージを共有ストレージに保存可能
- (6) 利用者が明示的に VM 停止命令をシステムに発効することで、VM は停止され、サーバ上から関連データが消去される



## 5. 実装と予備評価

平成 24 年度後半の HPCI 本格運用に向けて、先端ソフトウェア運用基盤の構築、段階的な機能追加を行っていく。本章では、分散環境ホスティングサービスの実装と予備評価の結果を示す。本サービスは、筆者等が RENKEI プロジェクトで開発・配備を行っている RENKEI-VPE (RENKEI: REsources liNKage for E-science, VPE: Virtual Private Environment) をベースに構築を進める。

### 5.1 RENKEI-VPE

RENKEI-VPE は多拠点に分散配備された計算機・ストレージ資源を相互連携させたグリッドサービスをホスティングするために開発された分散仮想マシン管理システムである。VM をホストするサーバを拠点資源と外部ネットワークの境界に設置し、その上での拠点資源と疎結合するグリッドサービスを VM として容易に実行可能なシステムとして設計されている。RENKEI-VPE は VM ネットワークの設定、起動、アクセスまでを提供し、VM への個別ソフトウェア設定は利用者が行う。

アーキテクチャを図 3 に示す。RENKEI-VPE は 1 台の管理サーバ、1 台以上のホスティング用サーバ、管理サーバとホスティング用サーバからアクセスできる 1 つの OS イメージ用共有ストレージからなる。

管理サーバでは GSI 認証に基づく SSH アクセスを提供しており、管理サーバ上で Unix アカウントを持つ者が RENKEI-VPE を利用できる。また、RENKEI-VPE で管理されるリソース制御用の xmlrpc を提供しているため、外部サーバからの操作も可能である。RENKEI-VPE の内部では、リモートサーバ上の VM 管理にはクラウド管理システム OpenNebula<sup>1)</sup> を用い、OS イメージ保存用ストレージには Gfarm<sup>2)</sup> を用いる。RENKEI-VPE の主な特徴を以下に挙げる。

- 利用者が指定した拠点への指定した OS イメージを用いた VM 起動。VM 接続先ネットワークの指定も可能であり、マルチホーム VM 構築にも対応。
- VM の suspend/resume。VM 状態を OS イメージとして保存し、後の利用、他者との共有。
- 利用者毎の利用可能拠点、リソース量の制限。
- 利用者への IP アドレスの事前割当。

ホスティング用サーバでは、仮想マシン管理ライブラリ libvirt インターフェースを経由して、kvm による VM 実行を提供する。管理サーバからの VM 操作命令、利用者毎の情報

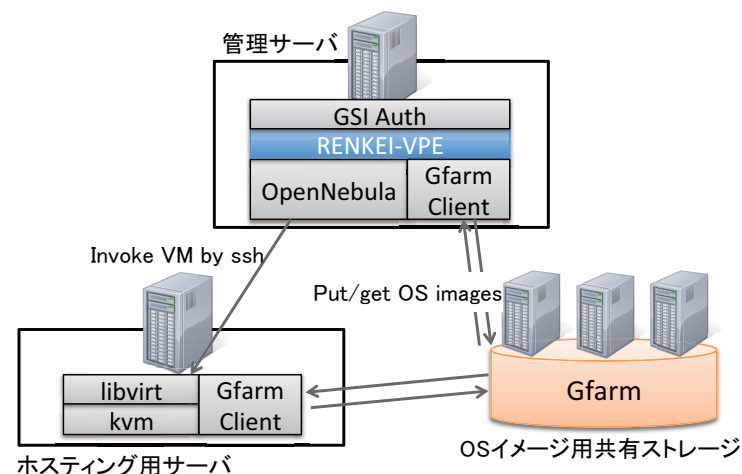


図 3 RENKEI-VPE の構成図  
Fig. 3 Architecture of RENKEI-VPE

(ssh 公開鍵など) は ssh にて転送される。Gfarm client 機能を備えており、VM 起動時には Gfarm から OS イメージをホスティング用サーバのローカルディスクにコピーしてから起動し、VM の状態を保存するときには Gfarm にイメージを転送する。VM 利用終了後、関連データは全て削除される。

OS イメージ用共有ストレージは Gfarm により構成され、管理サーバやホスティング用サーバからの OS イメージ転送要求を処理する。

### 5.2 RENKEI-PoP に置ける予備評価

RENKEI-VPE は RENKEI-PoP (Point of Presence)<sup>3)</sup> に既に導入され、運用されている。RENKEI-PoP とは計算機センターや研究室の 1 ゲートウェイとして働くストレージサーバであり、SINET が提供する 10Gbps VPN で相互接続されている。RENKEI-PoP 間で構築した共有ファイルシステムを用いることで拠点間でのデータ共有が行え、その上で VM を実行することで、TSUBAME にジョブ投入を行う NAREGI を含む、グリッドミドルウェアのホスティングを行っている。RENKEI-PoP は 2011 年 6 月時点で、9 拠点、11 台に導入されている。全てに Gfarm v2.3.2 が導入されており、合計 200TB のストレージを構成している。そのうち、東京工業大学に設置されている 2 台、国立情報学研究所 (以下

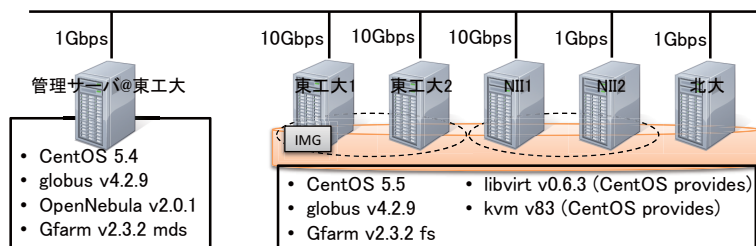


図4 RENKEI-PoP 上の RENKEI-VPE  
Fig.4 RENKEI-VPE on RENKEI-PoPs

NII) 千葉分館に設置されている2台、北海道大学の1台の計5台はRENKEI-VPEによりVM実行制御が管理されている。

図4にRENKEI-PoP上でのRENKEI-VPEの構成図を示す。GfarmのMetadata Server機能を提供するRENKEI-PoP管理サーバをRENKEI-VPE管理サーバとして構築し、Gfarm File Server機能を提供するRENKEI-PoPをVMホスティング用サーバとして構築している。OSイメージはこのGfarm上に保存している。

RENKEI-VPEによる分散環境へのVM配置の性能を、RENKEI-PoPの環境を用いて予備評価を行った。VM用OSイメージとして、QCOW2フォーマットされた物理サイズ653MBの最小構成のCentOS5.5イメージを用意し、図中、東工大1のRENKEI-PoP上に格納した。このイメージを用いて、東工大1、東工大2、NII1、NII2、北大上で1CPU、1024GBメモリを持つVM起動にかかる時間を測定した。表2に結果を示す。表中の単位は秒であり、「Time to Boot」はVM起動要求受信からVM起動完了（VMで全initスクリプト実行完了）までに要した時間である。「Time to Prepare」はRENKEI-PoP上でのVM起動準備に要する時間であり、管理サーバやGfarmからのファイル転送時間、及び、VM用Swapファイルの作成時間を含む。「Time from Prepared to Boot」はRENKEI-PoP上でのVM起動完了までの時間であり、「Time to Prepare」との和が「Time to Boot」となる。「Time for gfexport」はGfarmからRENKEI-PoP上にOSイメージをコピーする際の時間であり、「Time to Prepare」に含まれる。

VM配置の性能はRENKEI-PoP上でのVM起動準備に要する時間に影響して変化することが確認できる。VM起動準備で行う作業の内、大容量のデータ転送を行うgfexportコマンドによるOSイメージの取得は、表よりネットワーク帯域、地理的距離に応じてRENKEI-PoP毎に大きく異なることがわかる。その他、RENKEI-PoP毎のVM起動準備に要する

表2 RENKEI-PoP上でのRENKEI-VPEによるVM起動時間  
Table 2 Time for VM creation on RENKEI-PoPs using RENKEI-VPE

RENKEI-PoP	Time to Boot	Time to Prepare	Time from Prepared to Boot	Time for gfexport
東工大 1	116	19	97	0.549
東工大 2	111	14	97	1.47
NII1	118	21	97	7.58
NII2	141	44	97	23.2
北大	164	68	96	60.3

時間の違いを生じる要素として、RENKEI-PoP管理サーバからの数KByteのデータ転送が考えられるが、管理サーバでは各種監視サービスやWebサービスも提供しているため、これらのサービスによるジッタの影響により、差として現れづらいと思われる。運用時にはアクセス性能向上、および、データ保護を考慮して、OSイメージは複数の複製を作成して管理するため、短距離でのデータ転送が中心となり、2分程度の起動時間となると見込んでいる。

### 5.3 HPCI先端ソフトウェア運用基盤実装計画

先端ソフトウェア運用基盤の実装は、分散環境ホスティングサービスの実装から始める。分散環境ホスティングサービスはRENKEI-PoPと同様なモデルで構築する。具体的には、ホスティング用サーバを東工大・東大等の基盤センター中心にボランティアベースで提供していただき、RENKEI-VPE管理サーバを東工大に用意する。また、RENKEI-VPE専用のGfarmを、RENKEI-VPE管理サーバにMetadata Serverを、ホスティング用サーバにFile Serverを導入し、構築する。VMが接続するグローバルネットワーク環境としてはSINET4によるVPNを1つ用意する。

この環境を用いて、拠点からの資源供出方法、資源の管理方法、課金方針、ヘルプデスク等のシステム運用のための実験・評価・検討を行い、平行して機能使いのために以下の技術検討を行う。

- ホスティングリソースのクラスタ化とホスティング用サーバの管理ネットワークのプライベート化
- クラスタ化されたホスティングリソース内部でのVM配置スケジューリング機能
- 中央管理サーバで管理すべき情報と、拠点ホスティングリソース毎に個別管理すべき情報の分離
- 利用者・課題毎のネットワーク環境の分離

## 6. 関連研究

PlanetLab<sup>4)</sup>は全世界 507 拠点 1091 ノードから構成される分散システム開発・実験のためのテストベッド環境である。PlanetLab ではリソースを slice という仮想単位で利用者に利用権限を与えるが、我々は専用の VM を利用者に割り与える。また、PlanetLab に参加するサーバはファイアウォールのない Internet 接続するグローバルネットワーク 1 つに接続するが、我々のシステムでは各種ネットワーク環境を利用者 VM に引き込み可能である。

InTrigger<sup>5)</sup>は全国 17 拠点に設置されたクラスタより構成されている。利用者が InTrigger を使用する際には、ssh ログインし直接プログラムを実行する方法と、バッチシステムを用いた方法が提供されており、Gfarm による拠点間のファイル共有も提供されている。基本一般利用者権限で利用するシステムであり、自由に管理者権限が得られるわけではなく、また VM の運用はされていない。

先端ソフトウェア運用基盤では OpenNebula を内部で使用する RENKEI-VPE を用いるが、IaaS 型クラウドを構築することを目的に、このような VM 管理を備えたオープンソースシステムとして Eucalyptus<sup>6)</sup> や OpenStack<sup>7)</sup>、Nimbus<sup>8)</sup> 等多数存在する。今回 OpenNebula を内部システムとして利用した理由は、コアとなる VM 管理のコードベースが小さいことや、スクリプト言語により拡張が容易であることに加え、複数のネットワークセグメントに接続された VM 構築がサポートされていることである。上記他システムでは複数のグローバル IP アドレスを VM に割り当てる機能はサポートされていない。

## 7. まとめ

HPCI の計算機資源として、基盤センター群が保有するスーパーコンピュータに加え、第 3 の資源として、先端ソフトウェア運用基盤を提案した。先端ソフトウェア運用基盤は各組織で個別管理されているホスティングサービスを集約し、特に利用者への広域分散環境提供のためのシステムを提供することを目的とする。平成 24 年度秋の HPCI 本格運用に向けての先端ソフトウェア運用基盤の実装は、この分散環境ホスティングサービスの実装から開始する。システムには筆者等が RENKEI プロジェクトで開発・配備を行っている RENKEI-VPE を用いる。RENKEI-VPE は内部で分散 VM 管理のために OpenNebula を、OS イメージ管理に Gfarm を用いる。筆者等が構築・運用を行っている基盤センター間データ共有・資源連携用のサーバである RENKEI-PoP 上で RENKEI-VPE による VM 起動時間を予備評価したところ、VM をホストするサーバと OS イメージが存在するファイルサーバ間のネッ

トワーク性能に応じて起動時間は異なるが、2 分程度で起動することを確認した。今後はこのシステムを用いて、拠点からの資源供出方法、資源の管理方法、課金方針、ヘルプデスク等のシステム運用のための実験・評価・検討を行い、同時に利用者ネットワーク環境の追加機能を中心とした、機能追加のための技術検討を行っていく。

**謝辞** 本稿をまとめるにあたり御議論頂きました「HPCI の基本仕様に関する調査検討」委員の皆様、ならびに、HPCI システム WG 委員の皆様にご感謝致します。本研究の一部は、文部科学省委託「HPCI の基本仕様に関する調査検討」および「HPCI の詳細仕様に関する調査検討」による。

## 参考文献

- 1) Sotomayor, B., Montero, R.S., Llorente, I.M. and Foster, I.: Virtual Infrastructure Management in Private and Hybrid Clouds, *IEEE Internet Computing*, Vol.13, No.5, pp.14–22 (2009).
- 2) Tatebe, O., Hiraga, K. and Soda, N.: Gfarm Grid File System, *New Generation Computing*, Vol.28, No.3, pp.1–6 (2010).
- 3) 滝澤真一郎, 松岡聡, 佐藤仁, 東田学, 友石正彦, 實本英之: e-サイエンス基盤としての計算機センター POP(Point-of-Presence) 連携, 並列/分散/協調処理に関するサマー・ワークショップ (SWoPP2010) (2010).
- 4) Chun, B., Culler, D., Roscoe, T., Bavier, A., Peterson, L., Wawrzoniak, M. and Bowman, M.: PlanetLab: an overlay testbed for broad-coverage services, *ACM SIGCOMM Computer Communication Review*, Vol.33, No.3, pp.3–12 (2003).
- 5) 斎藤秀雄, 鴨志田良和, 澤井省吾, 弘中 健, 高橋 慧, 関谷岳史, 頓 楠, 柴田剛志, 横山大作, 田浦健次朗: InTrigger: 柔軟な構成変更を考慮した多拠点にわたる分散計算機環境, 情報処理学会研究報告 2007-HPC-111, pp.237–242 (2007).
- 6) Nurmi, D., Wolski, R., Grzegorzczak, C., Obertelli, G., Soman, S., Youseff, L. and Zagorodnov, D.: The Eucalyptus Open-source Cloud-computing System, *9th IEEE/ACM International Symposium on Cluster Computing and the Grid*, pp.124–131 (2009).
- 7) OpenStack Project: OpenStack Open Source Cloud Computing Software, <http://www.openstack.org/> (2011).
- 8) Nimbus Project: Globus Nimbus Homepage, <http://www.nimbusproject.org/> (2010).