

## 異種グリッドミドルウェアに跨る アプリケーションホスティングサービス (AHS) の設計と実装

宇佐見仁英<sup>†</sup> 大西尚樹<sup>††</sup> 水澤有里<sup>††</sup> 金澤宏幸<sup>††</sup>

異種グリッドミドルウェアで構成される複数のグリッド環境に跨るアプリケーションホスティングサービス (AHS) の設計と実装について報告する。AHS は、研究室レベルの軽量級グリッドミドルウェア (GridSAM 等) とスーパーコンピュータセンター等の大規模・高信頼度のグリッドミドルウェア (NAREGI) で構成されるような複数のグリッド環境に跨って、アプリケーションの資源要求や配置 (デプロイ) 等の情報をカタログ化して利用者に提供するサービス機能である。これにより、グリッドにおける仮想組織 (VO) で定義される研究コミュニティのメンバーが、研究室レベルの小規模な計算環境から情報基盤センターレベルの大規模な計算環境に跨って、メンバー間のお互いのアプリケーションを共有することが可能となる。

### Design and Implementation of Application Hosting Services (AHS) Utilizing Resources on the Multiple Grid Environments

Hitohide Usami<sup>†</sup> Naoki Onishi<sup>††</sup> Yuri Mizusawa<sup>††</sup> and  
Hiroyuki Kanazawa<sup>††</sup>

This paper describes the application hosting services (AHS) for the multiple grid environments based on heterogeneous grid middleware. AHS provides application catalog information for users that the applications had deployed in the heavy duty grid environments (for example, large scale supercomputer center running on NAREGI grid middleware) and the lightweight laboratory level grid environments (for example, small or middle scale laboratory system running on GridSAM grid middleware). Community members can easy use and co-share the latest version of applications in the communities realized by grid Virtual Organization (VO).

### 1. はじめに

21 世紀の IT 基盤であるネットワークインフラの発展は極めて目覚しく、産業界での生産性向上、実社会生活での利便性の面で世の中を大きく変えてきている。特に、インターネットにグリッドと言う新たな概念を組み込むことにより、サイエンス分野でも新たな変革が始まってきている。サイエンスグリッドは、ナノテクノロジー、バイオテクノロジーなど、これからの日本を支える最先端科学技術開発や、「ものづくり」における新製品開発のプロセスを大きく変える可能性を持っている。我が国のサイエンスグリッドを推進するプロジェクトとして「サイエンスグリッド NAREGI (National Research Grid Initiative) プログラム」が 2003 年度よりスタートし、2008 年 3 月末に当初の目標であるサイエンスグリッド向けミドルウェアを完成させて終了した。研究開発成果は、サイエンスグリッド向けミドルウェアの RPM パッケージング版「NAREGI ミドルウェア Ver.1.0」として、2008 年 5 月 9 日より NAGERI のポータルサイトから提供が開始されている。現在、九大、名大、東工大、筑波大などの 9 大学情報基盤センター、或いは分子研、高エネ研、天文台などの公的研究機関の情報処理センターに導入され大規模なグリッド環境の整備が始められている。しかしながら、グリッドによるスーパーコンピュータセンターレベルでの利用が出来たとしても、実際の日々の殆どの利用は自部門の計算機での利用範囲に限られており、研究室レベルの研究活動にまでグリッド利用が行渡っていないのが実状である。グリッド利用が日々の研究室での研究活動にも使われるようになることが、情報基盤センターの利用者拡大のためにも重要である。そのためにも研究室レベルで手軽に使える軽量級のグリッド環境の整備が NAREGI の普及とともに必要となってきている。

グリッド先進国の米国では、グリッド上に特定領域の研究コミュニティを一つの仮想組織として形成し、その領域で有効な様々なサービスの提供やアプリケーション、ノウハウの共有を可能とすることによって研究者の利便性を高めると共に、グリッドの煩わしさを解消させようとするサイエンスゲートウェイという考え方がテラグリッドプロジェクトを中心に推し進められている。

本論では、GridSAM 等の軽量級のグリッドミドルウェアを使った研究室レベルの計算資源を中心とした小規模なグリッド環境や NAREGI 等の heavy-duty なグリッドミドルウェアを使った情報基盤センター等の大規模なグリッド環境など、異種のグリッド環境を跨って形成された研究コミュニティにおけるアプリケーションのホスティングサービスの設計と実装について報告する。

<sup>†</sup> 玉川大学 学術研究所 / 情報・システム研究機構 国立情報学研究所  
Tamagawa University, Research Institute /  
Reserch Organization of Information and Systems, National Institute of Informatics  
<sup>††</sup> 富士通株式会社 テクニカルコンピューティング・ソリューション事業本部  
Fujitsu Limited, Technical Computing Solution Unit

## 2. 現在のグリッドの問題点

英国 サザンプトン大学 (University of Southampton) の David De Roure (dder@ecs.soton.ac.uk)教授は、現在のグリッドの問題点を以下のように指摘 (IEEE e-Science 2008 conference, 12 December 2007, Prof. David De Roure の Keynote Speech “The New e-Science” より引用) している。

- Everyday researchers doing everyday research  
 BUT heroic Grid infrastructure not being adopted
- A data-centric perspective, like researchers  
 BUT Grid gives APIs to computation not data
- Collaborative and participatory  
 BUT Grid has deeply rooted service provider mindset
- Better not Perfect  
 BUT Grid aims to provide well-engineered perfect solution
- Giving autonomy to researchers  
 BUT Grid imposes institutional control (at this time)
- About pervasive computing  
 BUT Grid is about portals, not the next generation of users

特に最初の2つの指摘は重要で、今のグリッドシステムがスーパーコンピュータセンターでの計算が中心であり、日々の研究室レベルでの利用に至っていないとの指摘である。グリッドの普及、さらに基盤センター利用者の拡大には、いかに現場の研究者を巻き込むか、現場のデータをシームレスに基盤センター等の計算資源と結びつけるかが重要となる。

## 3. 次世代 IT 基盤構築のための研究開発

2006 年度からペタフロップス級の超高性能な次世代のスーパーコンピュータの開発を目的とした「最先端・高性能汎用スーパーコンピュータの開発利用プロジェクト」がスタートした。この次世代のスパコンを頂点に大学・研究機関のスーパーコンピュータ (情報基盤センター等) から研究室レベルのシステムまでグリッドの世界を広げることによって、我が国の科学技術・学術研究の基盤となる次世代 IT 基盤の整備 (図 1) が計画されている。

図 1 は、ペタフロップス級の大規模データを扱う次世代スパコン (NLS: National Leadership System) を頂点に、7センターの情報基盤センター、国研のセンター等の中規模のデータを扱う大学・研究機関のスパコン (NIS: National Infrastructure System)、日々の研究業務での小規模データを扱う研究室レベルのクラスタ (LLS: Laboratory

Level System) を高速ネットワークで接続し、利用者からは研究室のサーバから次世代スパコンまで、計算資源を意識せずにシームレスに利活用できる環境を次世代 IT 基盤として新たに整備する事業が本年度より始まった。このような階層的な計算資源の整備は、既に米国、欧州では始まっており、特に米国では既に実行速度 (LINPACK) で 1 ペタを超えるマシンが設置されており、サイエンスゲートウェイとしての実験室レベルの小規模環境の取り込みなど日本の目指す次世代 IT 基盤環境が着実に構築されてきている。

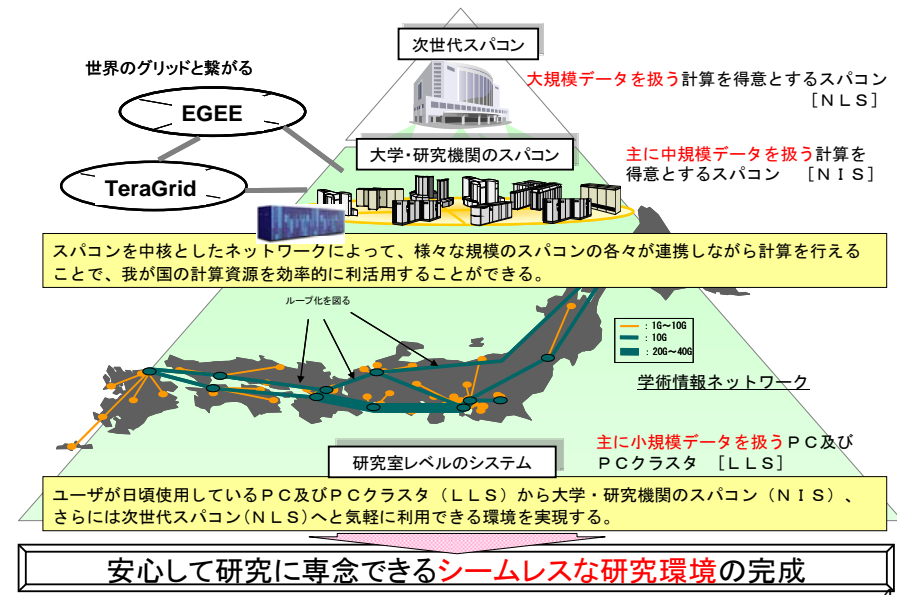


図 1 シームレスな研究環境

### 3.1 日本版 e-サイエンス環境の構築に向けて

文部科学省の次世代 IT 基盤構築のための研究開発「e-サイエンス実現のためのシステム統合・連携ソフトウェアの研究開発」～研究コミュニティ形成のための資源連携技術に関する研究～委託研究が昨年度から RENKEI (REsources liNKage for E-science) プロジェクトとしてスタートしている。情報システム研究機構・国立情報学研究所 三浦プロジェクトリーダーの下、NAREGI 関係者を中心に研究室レベルの日々の研究業務で発生するさまざまなデータを中心に、研究室のクラスタでの小規模な解析、或いは情報基盤センターのスーパーコンピュータを利用するような大規模な解析

を利用者が計算機資源を意識することなく、スムーズに利活用できるようなグリッド環境を実現する新たなグリッドのソフトウェア体系を計画している。今回の計画(図2のRENKEI部)は、次世代スパコン(NLS; Linpack性能で10ペタFlops超えが目標)の2010年度末の供用開始が予定されており、それに対応できるようにNISとLLSレベルの計算資源をシームレスにグリッド環境で統合的に利活用できる新たなミドルウェアを開発して次世代IT基盤環境を構築して行く予定である。具体的には、NAREGIの下方展開として新たにLLSに対応する軽量級のグリッドミドルウェア(Light Weight Grid Middleware)を展開し、NAREGIとの一体的な運営により、利用者から見たらNIS,LLSの区別を意識せずにシームレスにジョブ実行できるとともに、実験データ等をLLS間、或いはLLSとNIS間で共有できる仕組みとするものである。

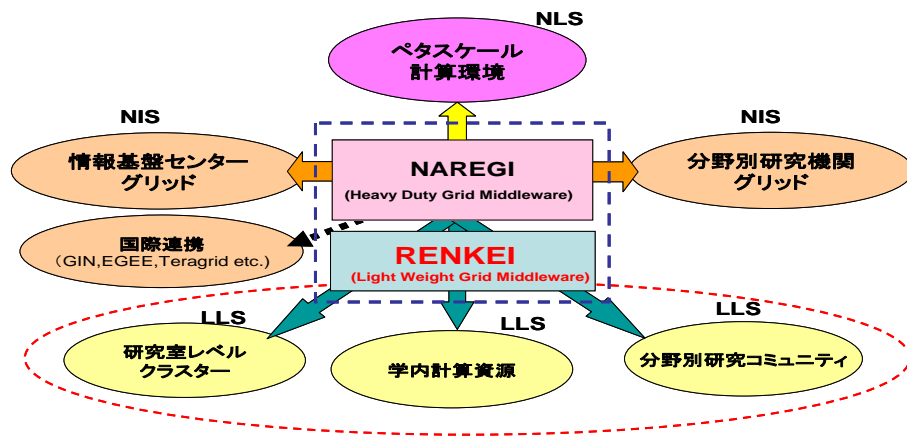


図2 次世代計算研究環境としてのグリッドの展開

### 3.2 e-サイエンス計算連携技術の概要

RENKEIプロジェクトにおいて、我々はアプリケーション実行環境を担当している。その部分をe-サイエンス計算連携技術と呼ぶことにする。e-サイエンス計算連携技術の全体概念図を図3に示す。e-サイエンス計算連携技術は、NAREGIのコンポーネントの内、基本的にはPSEとWFT(Workflow Tool)を複数のグリッド環境に対応できるように拡張するものである。特に複数グリッド環境でアプリケーションを共有するためにNAREGI-PSEを包含する形で、NIS及びLLSのそれぞれのグリッド環境で、NAREGI-PSEと同様なサービスを提供するアプリケーションホスティングサービス(AHS; Application Hosting Service)を開発することにした。WFTは、NIS対応アイコン、LLS対応アイコンを定義し、利用者はそれぞれのアイコンを選択することで、投

入先の環境を指定することになる。ワークフロー内部に解釈エンジンを搭載することで、ワークフローがNIS或いはLLSにジョブを投入するかの判断をして、それぞれの環境にジョブを投入する。また、資源管理に関しては検討段階であるが、英国のOMIIプロジェクトで開発されたGridSAMをベースに全体の仕様を検討した。また、NIS及びLLSでの全体のジョブの投入処理に関しては、国際標準として検討されているBES(Basic execution system)に準拠した仕様とした。

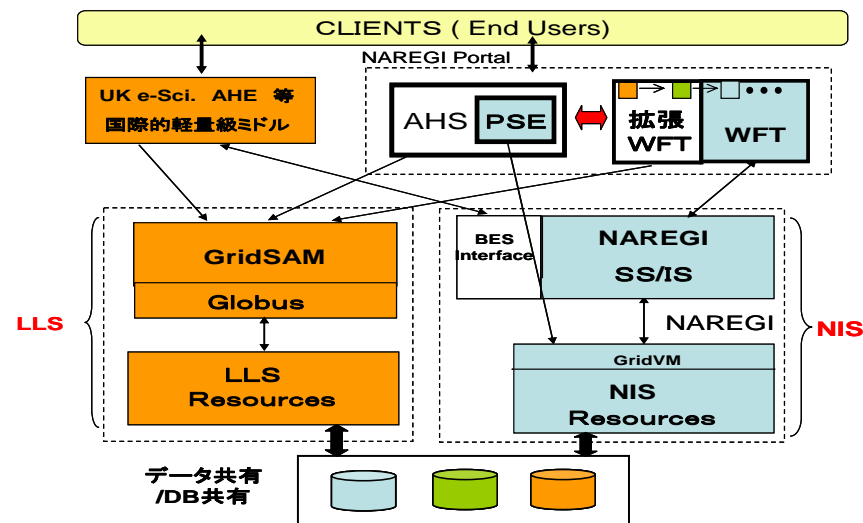


図3 e-サイエンス連携計算技術の全体概念図

### 3.3 アプリケーションホスティングサービス (AHS)

#### (1) AHSの概要

AHSの開発に当たっては、NAREGIの設計コンセプトに合わせながら、LLS環境をシームレスにNAREGI環境と合わせて行くことに留意した。特に、NAREGIで実現した仮想組織(VO; Virtual Organization)、及びグループ間でのアプリケーション共有の枠組みをLLSレベルにも展開する必要がある。NIS及びLLSを含んだ仮想組織、及び仮想組織内でのアプリケーションの実行環境へのDeployment、及びアプリケーション共有のイメージを図4に示す。

アプリ開発者は、研究コミュニティで利用可能なNIS、LLS資源にアプリケーションを配置しておく。アプリ利用者は、例えば、小規模解析ならLLSで、大規模な場合はNIS資源でどのように適時使い分けて同じアプリケーションを同じ操作環境で実行

することができる。解析のための実験データ等は、データ共有機構によりコミュニティの中で仮想的に統合管理されているので、特に意識する必要が無く解析実験が出来る。

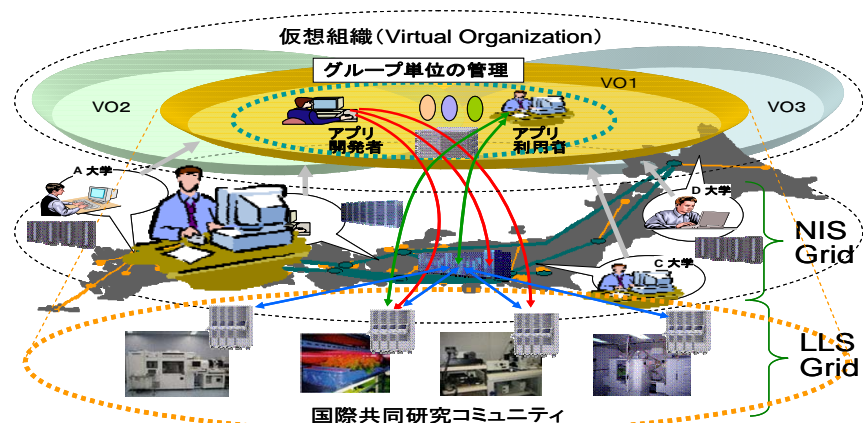


図 4 VO でのアプリケーション Deployment, 及び共有

(2) AHS の基本機能

e-サイエンス計算連携での LLS,NIS に跨るジョブ実行処理イメージを図 5 に示す。

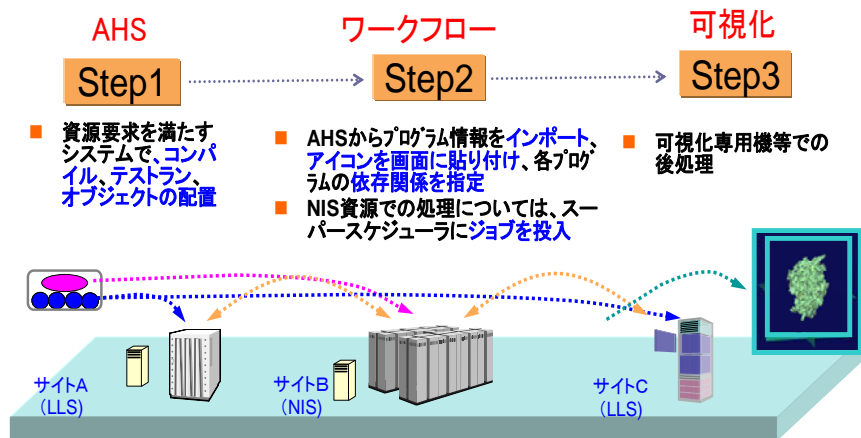


図 5 e-サイエンス計算連携での典型的なアプリケーション実行シナリオ

AHS は、図 5 に示した複数グリッド環境に跨ってアプリケーションを実行するようなシナリオにおいて、利用者がグリッド環境の違いを意識せずにアプリケーションを実行できるような環境を整えるための機能がある。

① 環境定義

アプリケーション開発者は、アプリケーションのソースコードと共に、make コマンド等の環境定義ファイルを AHS サーバにアップロードする。アップロード後、アプリケーション情報登録画面よりアプリケーション情報を登録する。特に、アプリケーションが必要とする資源要求を記述する部分は、配置可能計算機資源を選択するためのマッチングのためのキーワードとなる重要な項目である。この資源要求をベースに NAREGI ミドルウェアの資源管理のプロセッシングのための JSDL の資源要求のドラフトもこの情報から生成する。これらの仕様は、OGF の JSDL 関連 WG で継続的に検討中である。資源要求(JSDL)と分散情報サービスに登録されている計算機資源情報とをマッチングさせ、配置先サーバの候補リストを利用者に提示する。利用者は、これらの配置可能リストから配置させたい計算機資源にマーキングして配置計算機資源を決定する。全ての資源に配置する場合は、All ボタンで一括処理をすればよい。

② 転送

配置先サーバにコンパイル、テストランのためにアプリ環境一式を転送する。コンパイルだけを専用サーバで実行する場合は、コンパイルサーバに転送する。

③ コンパイル

コンパイルを実行し、コンパイルが成功すればバイナリを AHS サーバへ転送する。

④ テストラン

実行サーバへバイナリを転送し、テストランを実行する。

⑤ 登録

テストランが成功すれば、アプリケーションの配置情報を分散情報サービスに登録する。

(3) AHS の構成と実装

アプリケーションホスティングサービスの基本構成を図 6 に示す。NAREGI-PSE と大きな変更を実施していないが、LLS 環境の e-サイエンス計算連携対応のための拡張、及び BES 等への対応のためのスキーマの見直しを行った。特に、登録及びデプロイ機能等の NAREGI-PSE で提供してきた主要な機能については、LLS 環境の e-サイエンス計算連携でも同様なサービスとしてシームレスに利用できるように実装した。グリッドサービスのためのミドルウェアには GT4 を使用している。本来ならばデプロイされたアプリケーション自体をグリッドサービスとして実行サーバ上に立ち上げるべきとの意見もあるが、研究者の個人持ちのアプリケーションを基本とした本システムのような場合での適用等での検討の余地が多いため、実行環境の整備をグリッドサービスとして実現する実装とした。また、ライフサイクルマネージメントに関して検討した



結果、Web サービスとは根本的な思想が異なるので、永続的な指定とした。利用者は意識的にサーバから配置済みアプリケーションを削除することにし、そのためのサービス (Un-deploy) を提供することとした。WSRFの一つの基本的な考え方である「状態を持つリソース」として各コンポーネントを記述し、それらのサービスを連携させることで複雑な処理が可能な枠組みで実装することができた。特に、アプリケーション管理機能として新たにアプリケーションコンテンツ管理機能を作成してアプリケーション情報管理機能と分離した。また、アプリケーションコンテンツ管理機能については、将来的にPSE以外のコンポーネントから利用されることを想定し、OGFにおいて標準化仕様を共同で提案してきたACS(Application Contents Service)の仕様に基づいたWSRFベースのサービスとした。これに伴い、アプリケーション管理機能に関連するコンパイル支援機能、アプリケーション配置機能の各機能については、アプリケーションコンテンツ管理機能に対応させた実装となっている。利用者インタフェースについては、各機能に独立したGUIとすることで、各機能別の再利用性を高めている。

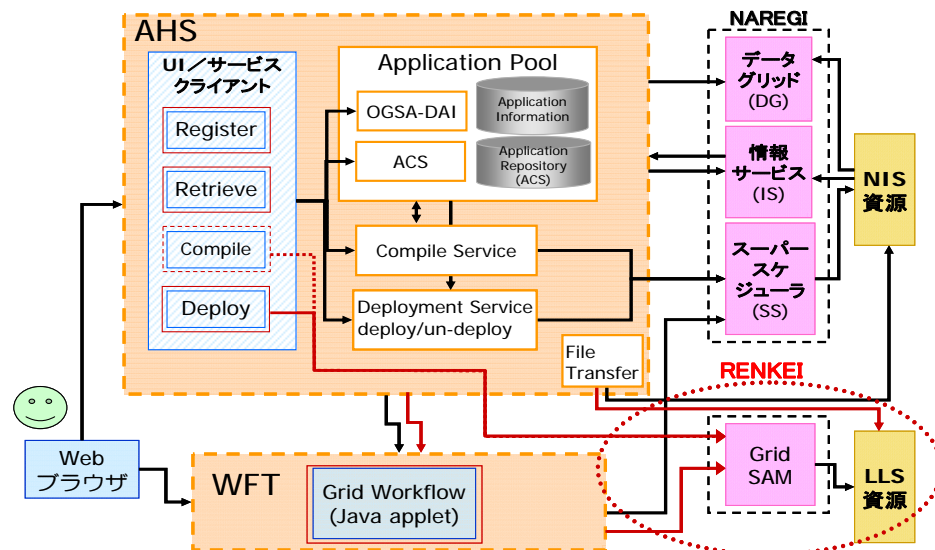


図6 アプリケーションホスティングサービスの基本構成図

#### (4) 基本性能

AHSの基本機能についての性能測定を実施した。基本性能を計測するための前提条件として、研究コミュニティー (VO) での同時利用者数、およびコミュニティーでの共

有するアプリケーション数を仮定して測定を実施した。今回のRENKEIプロジェクトでは、大学等の研究室が中心となって運用する研究コミュニティを想定しており、比較的規模の大きなコミュニティでも利用者200名、同時利用者はその10%の20名程度として性能評価を実施した。

#### ① 測定環境(AHSサーバ)

- ・ハードウェア構成
  - －NAREGIテストベット
- ・ハードウェア (PRIMERGY RX200)
  - －Xeon 3.06GHz
  - －メモリ 1 GB
- ・ソフトウェア
  - －ReadHat9
  - －PostgreSQL 8.2.4

#### ② 性能目標

同時20ユーザがSearch機能で20アプリの検索を実施し、各ユーザが30秒以内にレスポンス(検索画面)を参照できること。

#### ③ 測定結果

表1 AHSの基本性能

| ユーザ数    | 10  |     |     |     |     | 20 | 40  |
|---------|-----|-----|-----|-----|-----|----|-----|
| アプリ数    | 1   | 20  | 40  | 80  | 160 | 20 | 160 |
| 登録(src) | 0.2 | 0.7 | 1.2 | 3.1 | 5.1 | 11 | 60  |
| 登録(bin) | 0.3 | 0.3 | 0.7 | 0.7 | 0.8 | 8  | 15  |
| コンパイル   | 0.5 | 0.6 | 0.3 | 0.4 | 1.2 | 8  | 19  |
| デプロイ    | 0.2 | 0.4 | 0.3 | 0.6 | 0.8 | 8  | 19  |

SEC

同時20ユーザによる20件の登録されているアプリケーションの検索が30秒以内に実行できることを確認(表1)した。しかしながら、srcファイルへの登録において20人規模でも10秒を超えており、利用者にとっては決して早いとはいえない。大学等の研究室を中心とした研究コミュニティにおいては、同時利用者数20名程度と仮定しても妥当な範囲ではあるが、学会、あるいは情報基盤センターを一つの研究コミュニティとするようなVOでは、同時利用者が百人規模となる可能性もあるので、大規模なマルチコアマシン等、並列処理可能な高性能なサーバを利用する必要がある。さらに性能が必要な場合には、複数のサーバによる負荷分散を考える必要がある。

#### 4. おわりに

ブロードバンドネットワークの進展により、地球上に分散配置されているコンピュータを有機的に結合して、あたかも一つのコンピュータのように活用できるグリッド技術が注目を浴びている。一方、インターネット上での様々なサービスが展開され、Web サービス基盤が確立されてきている。次世代の情報基盤として Web とグリッドが融合したグリッドサービスの重要性が増してきている。特に、グリッドはヘテロジニアスなコンピュータ利用環境でのアプリケーション実行を前提としており、利用者にとって必ずしも使い勝手のよいシステムとはならない。グリッドのもたらす巨大なコンピュータパワーを、特殊な IT 知識を必要としなくても研究者が日々の研究活動において簡単に使いこなせるようにするには、操作性が良く簡単に使えるアプリケーションの実行環境が重要となる。今回、NAREGI の下方展開ということで、研究室レベルの計算資源 LLS にもグリッド環境を拡張し、研究者の日々の研究にダイレクトに接続できる環境の検討案を提示することができた。特に、NAREGI-PSE の拡張版として複数のグリッド環境に跨って仮想組織内でのアプリケーションを共有するアプリケーションホスティングサービス (AHS) のコンセプトを纏めることができた。

今後は、プロトタイプングでの検証を実施しながら、2011 年度の提供開始に向けた開発を実施していく予定である。さらに、システムが巨大化し複雑になればなるほど関連知識やノウハウの集大成が重要となってきており、従来の人工知能技術を超える新たな枠組みが必要となってきている。今後 3 年間の RENKEI プロジェクトを通して、より実用的なシステムの構築、及び知識ベースを具備したグリッド PSE の世界標準化などを図って行きたい。

**謝辞** 本論文執筆にあたってご指導・ご議論を頂いた国立情報学研究所三浦教授、合田教授、大阪大学松田教授、筑波大学建部准教授、産業技術総合研究所関口部門長、東京工業大学松岡教授、高エネルギー加速器研究機構佐々木教授、ならびに NAREGI 関係者に感謝します。

尚、本研究の一部は、文部科学省の平成 21 年度科学技術試験研究委託事業「研究コミュニティ形成のための資源連携技術に関する研究 (アプリケーション共有方式のユースケースと実証に関する研究)」の成果である。

#### 参考文献

- 1) 宇佐見, 宮原, 藤崎, 松本, 早勢, 川田, : Web intelligence から Knowledge grid へ, PSE を通して, PSE Workshop2003, 2003
- 2) 宇佐見, 宮原, 藤崎, 早勢, 仙田, 川田, : サイエンスグリッドにおける問題解決環境, 計算工学講演会論文集 Vol.9 (2004 年 5 月)

- 3) 藤生, 稲葉, 北向, 早勢, 宇佐見, 菊池, 川田, : 分散コンピュータ環境におけるジョブ実行支援 PSE の構築, 計算工学講演会論文集, Vol.9 (2004 年 5 月)
- 4) 宇佐見仁英, 黄淳郁, 宮原豊, 山田基弘, 藤崎正英, 早勢欣和, 川田重夫: グリッドサービスを基盤とした問題解決環境「NAREGI/PSE」, PSE Workshop2004, 2004
- 5) S.Matsuoka, S.Shimojo, M.Aoyagi, S.Sekiguchi, H.Usami, K.Miura, : Japanese computational grid research project: NAREGI, Proceedings of the IEEE, VOL. 93, NO. 3, pp.522-533, March 2005.
- 6) 宇佐見, 山田, 宮原, 藤崎, 早勢, 川田; サイエンスグリッドにおける NAREGI-PSE の Provisioning 方式, 計算工学講演会論文集 Vo2.0 (2005 年 6 月)
- 7) 宇佐見, 金澤, 山田, 宮原, 早勢, 川田: WSRF を基盤とした問題解決環境「NAREGI-PSE」, PSE Workshop2005, pp82-57, 2005.9.9
- 8) 金澤, 山田, 宮原, 早勢, 川田, 宇佐見: グリッドサービスを基盤とした問題解決環境「NAREGI-PSE」, 計算工学講演会論文集 D-3-4 (2006 年 6 月)
- 9) 宇佐見, 金澤, 伊藤, 山田, 宮原, 川田: サイエンスゲートウェイにおけるアプリケーション共有機構としての「NAREGI-PSE」, PSE Workshop2006, pp13-18, 2006.9
- 10) 宇佐見, 金澤, 宮原, 藤崎, 川田: 研究コミュニティにおける NAREGI-PSE のアプリ共有方式, 計算工学講演会論文集 D-10-1 (2007 年 5 月)
- 11) 宇佐見, 金澤, 宮原, 川田: NAREGI-PSE によるリアルタイムコラボレーション, PSE Workshop2007, pp19-25, 2007/9/25
- 12) 宇佐見, 金澤: 複数グリッド環境におけるアプリケーションホスティングサービス (AHS) の一考察, PSE Workshop2008, pp18-21, 2008/9/1
- 13) H.Usami, H.Kanazawa, M.Yamada, Y.Miyahara, Y.Hayase, Kawata, "Problem Solving Environment (PSE) for scientific Grid computing, SC2005 CD-ROM Conference Proceedings, Seattle, USA, Nov., 2005
- 14) H.Kanazawa, H.Usami, M.Yamada, Y.Miyahara, Y.Hayase, S.Kawata, "Problem Solving Environment based on Grid Services : NAREGI-PSE ", Workshop on Innovative and Collaborative Problem Solving Environment in Distributed Resources, Melbourne, Australia, Dec., 2005
- 15) H.Kanazawa, Y.Itou, M.Yamada, Y.Miyahara, Y.Hayase, S.Kawata, H.Usami, Design and Implementation of NAREGI Problem Solving Environment for Large-Scale Science Grid, 2nd IEEE International Conference on eScience and Grid Computing, Amsterdam, Netherlands, Dec., 2006.
- 16) H.Usami, H.Kanazawa, Y.Miyahara, S.Kawata, Problem Solving Environment for Scientific Grid Computing, 3rd IEEE International Conference on eScience and Grid Computing, Poster session, Bangalore, India, Dec., 2007.
- 17) H.Usami, H.Kanazawa, Application Hosting Services in a virtual organization that support multiple grid environments., 4th IEEE International Conference on eScience and Grid Computing, Indianapolis, USA, Dec., 2008.
- 18) 田中,合田: 異種ミドルウェア間に跨るワークフロージョブの実行方式と実装. 情報処理学会 HPC 研究会, SwoPP 仙台 2009,
- 19) NAREGI プログラム <http://www.naregi.org/>
- 20) GLOBUS プロジェクト <http://www.globus.org/>
- 21) TeraGrid プロジェクト [dhttp://www.teragrid.org/](http://www.teragrid.org/)