

名大情報連携基盤センターにおけるグリッドの取組みについて（その3） － NAREGI グリッドミドルウェアの構築－

津 田 知 子 石 井 克 哉

I. はじめに

名大情報連携基盤センターでは、グリッドコンピューティングに対して、グリッドコンピューティングのミドルウェアである Globus や UNICORE で動作するアプリケーションの開発やその性能評価に取り組んできました [1]。昨年度からは、国立情報学研究所の CSI (Cyber Science Infrastructure) の委託事業の一環として、わが国で開発された NAREGI グリッドミドルウェアの導入と検証を行っています。ここでは、NAREGI グリッドミドルウェアがどのようなものであるかを紹介すると共に、本センターでの取組みをシステムの構築を中心に紹介することとします。

II. NAREGI グリッドミドルウェアとは

NAREGI (National Research Grid Initiative) は、サイエンスグリッド環境の構築の実現を目的として、平成 15 年 4 月にスタートしたプロジェクトです。NAREGI のホームページ [2] によると、NAREGI では、つぎの目的を達成するために NAREGI グリッドミドルウェアの開発を行っています。

- ▶ 異なったメーカーのさまざまな種類のスーパーコンピュータやハイエンドサーバなどの計算資源を、利用者から見て仮想的な一つの大きな計算資源にみせる。
- ▶ 利用者にとってユーザフレンドリーな環境を提供する。
- ▶ 実組織とは異なる組織横断的な研究コミュニティ (仮想組織, VO : Virtual Organization) を必要に応じて動的に形成する。

NAREGI グリッドミドルウェアのソフトウェアの構成を図 1 に示します。

この NAREGI ミドルウェアは、平成 16 年度に UNICORE をベースにした α 版が開発され、平成 17 年に公開に至り、全国共同利用の情報基盤センターで導入作業が行われました。平成 18 年 5 月には OGSA (Open Grid Service Architecture: <http://www.globus.org/ogsa/> 参照) をベースにした β 1 版が、オープンソースソフトウェアとして公開されました。

また、NAREGI ミドルウェアは、平成 18 年度に始まった「最先端・高性能汎用スーパーコンピュータの開発利用」プロジェクトの次世代スーパーコンピュータの基盤ソフトウェアとして認められ、ベタスケール時代の計算研究環境にとって中心となるミドルウェアとなるものと考えられます。次期のスーパーコンピュータシステムを考える上でも、センターとして NAREGI ミドルウェアに対する取組みが必要になってきています。

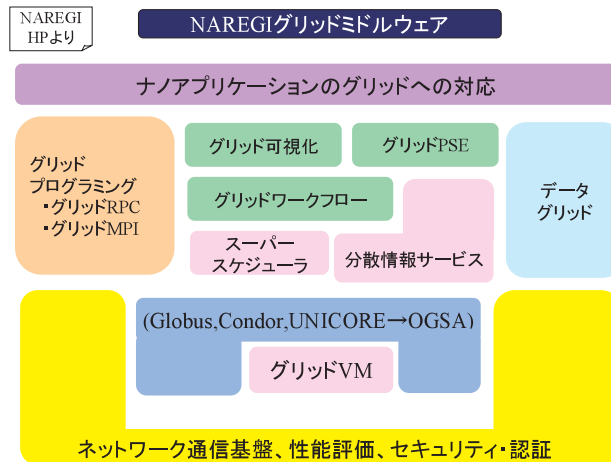


図1 NAREGI グリッドミドルウェアのソフトウェア構成

Ⅲ. NAREGI ミドルウェアβ1版について

1. NAREGI ミドルウェアのβ1版の実験環境

NAREGI ミドルウェアの図1に示すソフトウェア環境を実現するために、本センターでは、表1に示すサーバ構成をとっています。

表1 NAREGI ミドルウェアのサーバ構成

| ノード種類 | 機能 | 主な NAREGI ソフトウェア |
|------------------------------|--------------------|---|
| IS ノード | 資源情報サービス | Information Service |
| SS ノード | 資源管理, ジョブ管理 | Super Scheduler |
| Portal ノード | Authority / Portal | NAREGI-CA, Portal, PSE, My Proxy, Work Flow Tools, GVS Service Provider, Super Scheduler client |
| VOMS | ユーザ管理 | gLite VOMS |
| GridVM Server ノード | GridVM Server | GridVM Scheduler, GridVM Engine, GVS parallel Visualizer, InformationService LRPS, GridMPI |
| GridVM Client ノード (計算サーバ) | GridVM Client | GridVM Engine, GVS parallel Visualizer, GridMPI |

本センターでは、図2に示すように複数の GridVM ノード（計算ノード）からなるβ1版の実験環境を構築しています。ここで、金田研究室と太陽地球環境研究所の GridVM ノードは、CSI 委託事業の e-サイエンス研究分野として採択され、本センターで構築された NAREGI ミドルウェアシステムと連携しているものです。

NAREGIミドルウェアβ1版実験環境

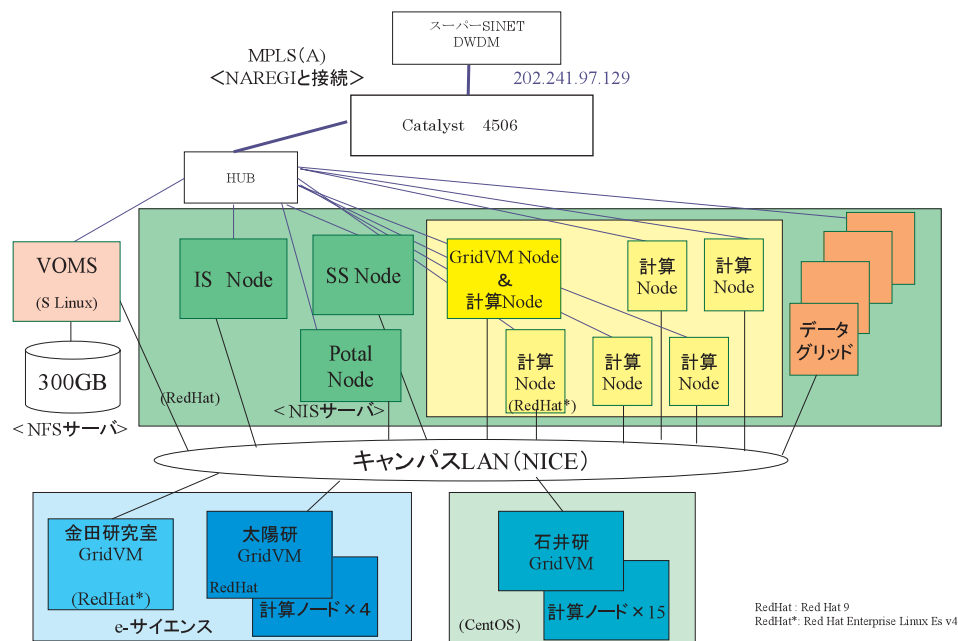


図2 NAREGI ミドルウェアβ1版実験環境

2. β1版のインストールと問題点

β1版のインストールに関しては、以下のような問題点がありました。

- ・インストールに際し、本センターの運用に合わせ、多数の設定ファイルを変更する必要があった。設定ファイルの入力ミスによりエラーが誘発され、インストール作業をより困難なものにした。
- ・ユーザのジョブがエラーで終了した場合、NAREGI システムから出力されるログファイルが膨大で、エラー原因追求を困難にしている。
- ・NAREGI ミドルウェアでは、複数のノードが連携して動作している。エラー原因を追究するためには、ノードにまたがってログの解析をする必要がある。それは非常に大変な作業である。
- ・インストールを優先するため、OS の update を後回しとした。このことにより、殊更、セキュリティに注意を払う必要があった。

インストール時のトラブルを纏めると表2に示すようになります。このことから、インストール方法の改善を行うことにより、多くのトラブルを回避できることが期待できます。

また、実際にβ1版を利用しての問題点を以下に纏めます。

- Work Flow Tool を利用して、ジョブを実行させ、“Exception” で終了した場合、利用者の使用上の誤りによるものなのか、NAREGI ミドルウェアの問題によるものなのかを判断する情報が利用者に与えられない。

表2 β1版のインストールトラブル

| インストールトラブル | % |
|-------------|----|
| 設定ファイルの入力ミス | 53 |
| マニュアルの不備 | 26 |
| その他 | 21 |

- Portal ノードの Portal ユーザの証明書の有効期限が、当初、12 時間であったため有効期限切れによりエラーとなった。しかし、証明書の有効期限切れが適切に利用者に伝わらなかった。これに対処するため、セキュリティ上は問題を残すかもしれないが、利便性を優先し、証明書の有効期限の 12 時間を 200 日とした。
- NAREGI ミドルウェアでは、多くのノードで種々のサービスが動作している。これらのサービスのすべてが正常に動作していないと、利用者のジョブは、正しく処理されない。複数ノードにまたがるサービスの状況をすべて把握することは、困難である。

3. 利用例

NAREGI ミドルウェアは、Web ブラウザ（Java スクリプトの実行が許可されている必要があります）から利用できます。NAREGI の Portal ノードに接続すると、図 3 に示す Naregi Portal の画面が表示されます。Proxy 証明書を作成して、“Sign On” に成功すると、図 4 に示す“Grid Tools”画面が表示されます。

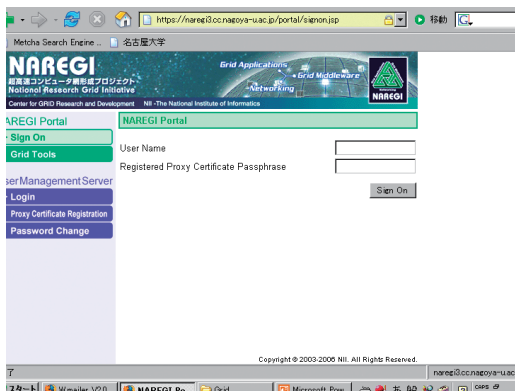


図 3 NAREGI Portal 画面

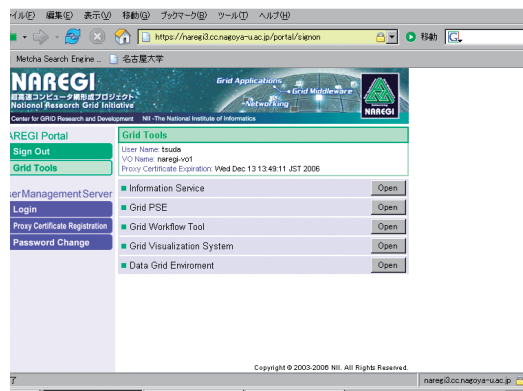


図 4 Grid Tools 画面

ここで、NAREGI ミドルウェアで提供されている以下のサービスが利用できます。

[Information Service]

実行ホストの CPU 情報等の表示

[PSE] (Problem Solving Environment)

ユーザ作成のアプリケーションのグリッド環境への配置（登録・展開）

[Workflow Tool] (WFT)

GUIによるグリッド環境でのプログラムの実行・制御

WFTでは、GridMPIによりMPIのプログラムも実行可能

[Grid Visualization System] (GVS)

グリッド上に分散された大規模データのリモート可視化

[Data Grid Environment]

グリッド上に分散された大規模データを共有ファイルシステムに登録・配置

図5にWFTの利用例を、図6にその実行例を、図7にGVSの利用例を示します。

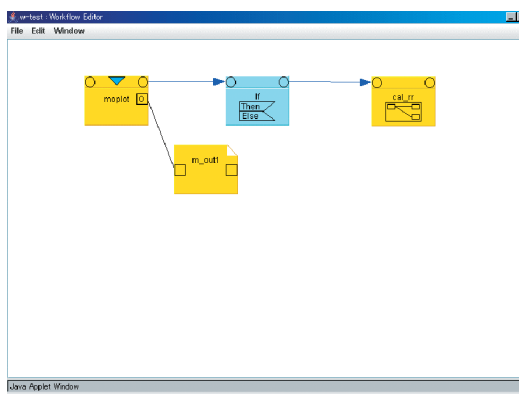


図5 WFTの利用例

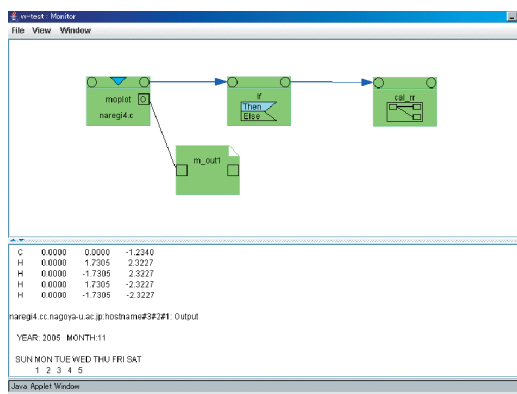


図6 WFTの実行例

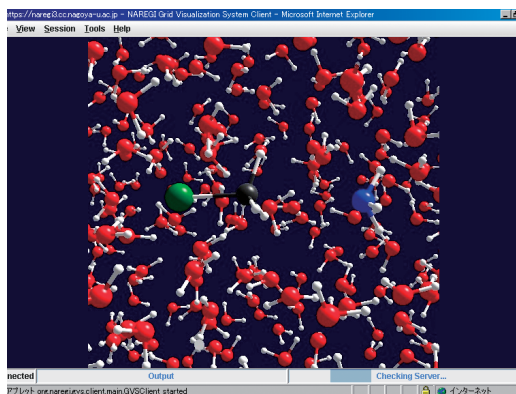


図7 GVSの利用例

IV. おわりに

NAREGI ミドルウェアは、 β 1版で洗いだされた問題点の改善と更なる改良が行われ、 β 2版の基盤センター等への公開は、今年の秋頃が予定されています。インストールの困難さを改善するためのLinux APT-RPMパッケージングも制限つきでサポート予定とのことです。これで、導入のハードルがグンと低くなることが期待されます。

現在、NAREGI グリッドミドルウェアは、本センターのスーパーコンピュータなどの運用マシンではなく、実験マシン環境で動作しています。運用マシン上でのサービスが可能となるまでに、ペタスケール時代のスーパーコンピュータの利用者にとって便利な道具の一つとなるように、いろいろな知見を蓄えていきたいと思っています。手持ちのソフトウェアを NAREGI グリッドミドルウェア上で動かしてみたいとか、NAREGI グリッドミドルウェアに興味をもたれた方のこの実験への参加を歓迎します。筆者までご一報くだされば幸いです。

最後に、NAREGI ミドルウェアシステムの構築にあたっては、富士通（株）SE の加藤氏及び NAREGI のサポート部隊の方々に多大なご協力を頂きました。ここに記して感謝します。

【参考文献】

- [1] 津田知子, 永井亨, 長谷川明生, 石井克哉: 名大センターにおけるグリッドの取り組みについて (その2), 名古屋大学情報連携基盤センターニュース Vol.3, No.2, pp.89-94, 2004.5
- [2] <http://www.naregi.org/>

(つだ ともこ: 名古屋大学情報連携基盤センター)
(いしい かつや: 名古屋大学情報連携基盤センター)