

NAREGI グリッドミドルウェアによる大規模連携接続実証実験

東田 学[†] 友石 正彦^{††} 坂根 栄作[†] 佐藤 仁^{††}
山梨 毅^{††} 大庭 淳一^{†††} 小林 泰三^{†††} 水谷 文保^{††††}
山田 清志^{††††} 津田 知子^{†††††} 河野 貴久^{†††††}
合田 憲人^{††††††} 松岡 聡^{††,††††††} 青柳 睦^{†††} 下條 真司^{†††††††}

大阪大学と東京工業大学、九州大学、名古屋大学の各情報基盤センターは、NAREGI 研究開発拠点である国立情報学研究所と分子科学研究所と連携し、実運用中の大規模計算機システムに NAREGI ミドルウェアを導入し、短期間ではあるが、全国規模でのこれまでにない大規模なシステム構成による連携接続を実証的に行った。本実証実験では、NAREGI ミドルウェアによって、複数の認証局が発行する証明書を用いて仮想組織を構成し利用者情報や計算機資源情報などのグリッド型資源管理が可能であること、および、複数のメタスケジューラが連携して資源予約を行った上で実アプリケーションを認可された計算資源において実行可能であることを検証した。

Interoperability Testing of NAREGI Grid Middleware for Large-Scale Cooperation

Manabu Higashida[†] Masahiko Tomoishi^{††} Eisaku Sakane[†] Hitoshi Sato^{††}
Takeshi Yamanashi^{††} Junichi Ooba^{†††} Taizo Kobayashi^{†††} Fumiyasu Mizutani^{††††}
Kiyoshi Yamada^{††††} Tomoko Tsuda^{†††††} Takahisa Kono^{†††††}
Kento Aida^{††††††} Satoshi Matsuoka^{††,††††††} Mutsumi Aoyagi^{†††} Shinji Shimojo^{†††††††}

Four Information Technology Centers in Osaka University, Tokyo Institute of Technology, Kyushu University and Nagoya University deployed NAREGI Grid Middleware to their Supercomputing Resources under real operational scenarios, and tested its interoperability for nation-wide large-scale cooperation with two NAREGI R&D centers: National Institute of Informatics and Institute for Molecular Science. We successfully demonstrated that we're able to formulate virtual organizations with certificates from multiple authorities and manage their grid resources, and also able to submit real grid applications to authorized computing resources with resource reservations, with coordinated execution across multiple meta-schedulers that issue reservation requests independently to potentially a same resource.

1. はじめに

本実証実験では、実際に運用中のノードも含めた基盤センターの実運用システムに NAREGI ミドル

ウェアを導入し、NAREGI ミドルウェアの研究開発拠点である国立情報学研究所 (NII)、分子科学研究所 (IMS) との大規模な連携接続を行った。まず、2007 年 8 月から、阪大と東工大が九大と NII の協力を得て連携実験を開始し、2007 年 11 月に米国 Reno にて開催された SC07 に 2 拠点の連携環境をデモンストレーションする出展を行った。その成功を受けて、NAREGI ミドルウェアの最終評価として 100TFLOPS 級の計算資源管理が可能であることを実証するため、2008 年 1 月より IMS が参加、さらに 2008 年 3 月から名大が参加し、2008 年 3 月末に 6 拠点での連携実証を行うことになった。実証に際しては、特に以下の 3 項目の検証を行った:

1. 複数の認証局が発行した証明書を参照できる NAREGI 計算資源環境を構築し、動作させるこ

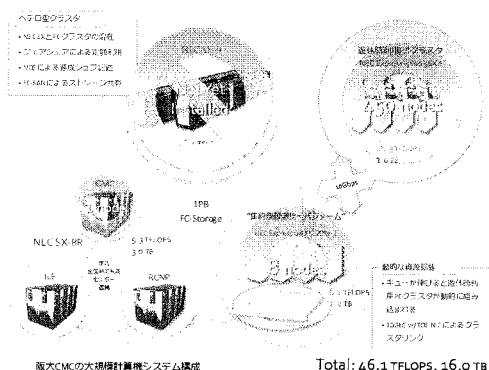
[†] 大阪大学
Osaka University
^{††} 東京工業大学
Tokyo Institute of Technology
^{†††} 九州大学
Kyushu University
^{††††} 分子科学研究所
Institute for Molecular Science
^{†††††} 名古屋大学
Nagoya University
^{††††††} 国立情報学研究所
National Institute of Informatics
^{†††††††} 情報通信研究機構
National Institute of Information and Communications Technology

2. 実際に運用中の基盤センターの大規模計算機資源に対して NAREGI ミドルウェアからジョブ投入できる環境を構築し、動作させることができること

6 拠点は、NAREGIミドルウェアが導入された管理

2. 各拠点の概略とこれまでの取り組み

同センターは、下條がサブリーダーを務めた NAREGI WP5 を中心に NAREGI ミドルウェアの開発に貢献した。2006年には、NAREGI CA によるグリッド認証局を立ち上げると同時に、既存の全国共同利用業務をベースとした CP/CPS を策定し、すべて



の利用登録者に対して 1-click でグリッド証明書を発行可能な業務システムと利用者インターフェイスを構築した。全国共同利用登録に際して、NAREGI CA の証明書払出に必用となるライセンス ID を仮に払出し、そのライセンス ID を利用者に通知するのではなく、同センター側の LDAP データベースにストックする。利用者には全国共同利用の ID と仮パスワードを書面で通知し、グリッド証明書発行に先だって、Kerberos クレデンシャルの取得を要請している。最終的に、SSO による権限委譲を伴って RA から証明書を UMS に格納するという業務フローを構築した [1][2]。この業務フローは、証明書発行に先立つ本人性確認の要請以外では、APGrid PMA Minimum CA Requirements を満たすことを確認している。さらに、同センターは、2004 年以降、 α 版および β 版の導入評価に貢献し、特に $\beta 2$ 版からは、ベクトル型スーパーコンピュータを含む実運用システム全体を NAREGI ミドルウェアによって資源提供できるように対応を進めてきた。

– 134 –

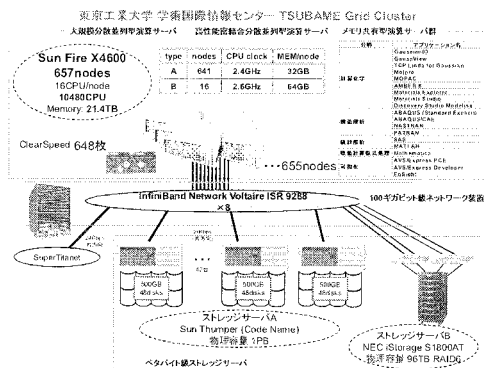


図 2: 東工大 Tsubame システム構成

供しており (図 2)、学内へは 10GbE、SINET3 による CSI-Grid VPN 網へも 1GbE で接続している (実証時は 4GbE に増強)。

同センターは、松岡がリーダーを務めた WP1 を中心に開発に貢献を行った。特に、同センター運用環境への導入をにらみ、Sun GridEngine への対応は、同センターにおいて行ってきた。さらに、初期開発版においてインストールの複雑さ・難解さが問題になったことに際し、九州大学において rpm 開発が始まったのと歩調を合わせ、その rpm を利用した「Xen を使った NAREGI VO サービスのデプロイメント」機能の開発を行った。この機能により、UMS-Portal-ISNAS-SS という VO サービスを 1 時間程度で導入できるようになった。本実証実験においてもこの機能により、東工大拠点での VO サービスの迅速な開始が実現された。

またセンター外からも次の 2 クラスタの参加があった: グローバル COE「計算世界観の深化と展開」で 2008 年 3 月から稼働させている Xeon ベースのクラスタ「TSUBASA」(90 ノード、720CPU コア、8.1TFlops, 0.9TB)、松岡研究室 Opteron ベースのクラスタ「PrestolIII」(252 ノード、512CPU コア)。

九州大学情報基盤研究開発センターは、1970 年に全国共同利用施設として設置された大型計算機センターを起源とし、主に九州地区の研究者を中心にスーパーコンピューティング・サービスを提供している。全国共同利用研究用計算機システムとしては、富士通 PRIMEQUEST 580 (13.1TFLOPS, 4TB)、富士通 PRIMERGY RX200S3 (18.4TFLOPS, 3TB)、日立 SR11000 (3TFLOPS, 2.9TB)、IBM eServer p5 モデル 595/570 (3.3TFLOPS, 2TB) によるサービスを提供している。

同センターは、青柳がサブリーダーを務めた NAREGI WP6 を中心に開発に貢献した。NAREGI WP6 では株式会社日立製作所と共同で異なるアプリケーション間での高度意味変換を伴うデータ交換を実現するグリッド連成ミドルウェア Mediator の開発、ファイルによる同期型データ交換を実現する SBC (Storage Based Communication)、NAREGI グリッド環境構築を簡単に実現する apt-rpm パッケージとインストールツールを開発した。Mediator は本実証研究でも検証した RISM-FMO 連成計算プログラムに適用し、SBC は IMS にて研究開発が進められた GIANT ミドルウェアに適用された。apt-rpm パッケージは既に国立情報学研究所より国内の大学及び研究所に公開されている。

国立情報学研究所 (NII) は、2003 年度から開始された NAREGI ミドルウェアの研究開発拠点である。開発および実証評価を目的として、Xeon 2.8GHz/3.0GHz をベースとした PC クラスタ (780CPU、メモリ 455GB、ストレージ 37.6TB) のほか、SPARC64V、Itanium2、Power4 などから構成される SMP サーバ群が運用されており、全システムの合計ピーク性能は 5TFLOPS である。また、NII では、ミドルウェアのメンテナンスやユーザサポートを目的とした利用支援組織 (GOC) を運用している。

NII では、NAREGI ミドルウェアの研究開発を進めるとともに、大学や研究所と連携してグリッド環境の構築を行っている。具体的には、2004 年に NAREGI ミドルウェア α 版、2005 年に β 版、2006 年に β 2 版のミドルウェアのリリースを行い、2008 年にはバージョン 1.0 をリリースした。ミドルウェア運用に先がけて開発した認証局ソフトウェア NAREGI CA は、海外も含めた多くの組織・プロジェクトで利用されている。NII でも、このソフトウェアを用いた APGrid PMA (IGTF) から Production Level CA の認定を受けた認証局を運営しており、主に NAREGI ミドルウェアの研究開発や利用を行う機関に対して、グリッド用のユーザ証明書およびホスト証明書を発行している。また今後、証明書の発行対象を拡大する予定である。

分子科学研究所 (IMS) は、1977 年 5 月に設置された電子計算機センターを起源とした自然科学研究機構 岡崎共通研究施設 計算科学研究センターを運営しており、現在は国内唯一の分子科学計算のための共同利用基盤センターとして、超高速分子シミュレーションシステム Altix4700 (4TFLOPS, 8TB)、PRIMEQUEST (4TFLOPS, 2.5TB)、及び汎用高速

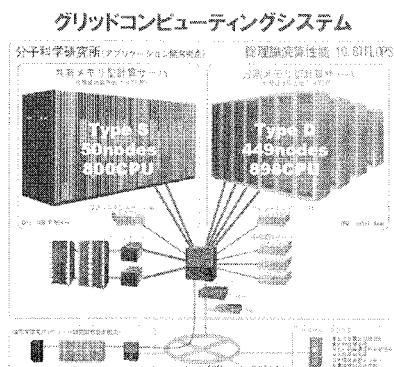


図 3: IMS NAREGI システム構成

演算システム SR16000 (5.4TFLOPS, 2.3TB) により全国共同利用サービスを行っている。これらに加え、IMS が 2003 年 4 月より NAREGI プロジェクトに参画したことを受け、プロジェクト推進のためにグリッドコンピューティングシステムとして、異なるアーキテクチャである共有メモリ型計算サーバ SR11000 (5.4TFLOPS, 3.2TB) 及び分散メモリ型計算サーバ HA8000 クラスタ (5.4TFLOPS, 1.7TB) を導入し、産官学の各研究者に研究開発環境を提供してきた (図 3)。

IMS は、NAREGI プロジェクトのナノサイエンス実証研究における中核拠点として、ミドルウェア開発拠点である NII と連携して、グリッド環境で実行可能なアプリケーション開発及び実証計算を行ってきた。この実証研究のポイントとして、1) メタコンピューティング、2) ハイスループット、3) リアルタイムコラボレーション、の 3 つを設定した。「メタコンピューティング」では、分散またはヘテロな計算機を 1 つの巨大な仮想計算機とする実証を行った。そのために産学官それぞれで開発した 15 本のアプリケーション (ヘテロ環境ではその中の 3 本) により GridMPI を用いて大規模計算を行った。また MPI ベースで開発されたアプリケーションを次々と GridMPI で実行させることにより、グリッド・アプリケーションとしての移植性の検証も行った。「ハイスループット」では、パラメータサーバタイプの多数計算の同時実行を実施した。そのために GridRPC を使った 2 本のアプリケーションの開発し、さらに商用アプリを用いた WFT からの大量ジョブ投入を行った。「リアルタイムコラボレーション」では、PSE を使用して物理的に離れたサイトへのアプリ自動配信や、連成計算を実施した。特に後者においては、開発アプリと商用アプリでの連成、IMS と NII のサイト間にまたがった大規模連成を行った。ま

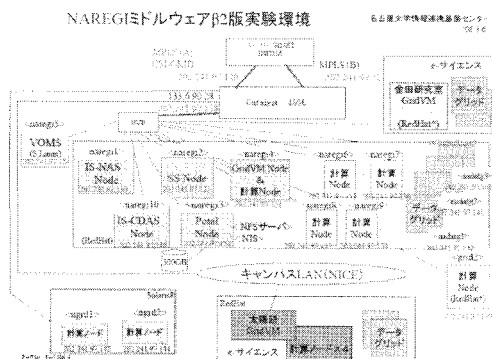


図 4: 名大β2システム構成

たコラボレーション支援のために、2 つのツール (IGNITION、GIANT) を開発してきた。これら実証を行うために、IMS、NII に加えて 5 拠点 (東北大金研、KEK 物構研、東大物性研、京大化研、九大) でナノサイエンス VO を構築した。

名古屋大学情報連携基盤センターは、1971 年に全国共同利用施設として設置された大型計算機センターを出発点とし、以来、トップレベルのスーパーコンピュータや汎用コンピュータ等を導入して、研究者の多様な科学技術計算の要求に応じてきた。現在は、スーパーコンピュータとして分散並列スカラ型のマシン (システム全体で総合理論性能 13.8TFLOPS、総メモリ容量 12TB) によるサービスを行っている。なお、2009 年の春から秋にかけて、現有の数倍のシステムへの更新を予定している。

同センターでは、2002 年度よりグリッドコンピューティングに対する取り組みを開始し、ここ数年は、NAREGI ミドルウェアによる環境を構築し、その実証実験を中心に行っている。2006 年度は、β1 版により同センターと合わせて 4ヶ所の GridVM サーバからなる実験環境を構築した。OS は、いずれも Linux でスケジューラは PBSPro であった。また、データグリッドの構築も試みた。2007 年度には β2 版による構築を行い、前記に加え、Solaris 上にも GridVM サーバを構築した。スケジューラは Parallelnavi の配下で NQS を用いた。β2 版では、GridMPI に加え GridRPC を用いたアプリケーションの利用実験も行った (図 4)。

3. 連携構成

連携に際しての管理ノードおよび計算ノードの構成は、組織の実運用を優先したため、結果として非常に多様な資源提供が行われることとなった (表 1、

Fact Sheet 1: 管理ノード構成

Phase1 / Phase2	SS/NETの接続	ブロードバンド回線	ポータル	SS	IS-NAS	IS-CDAS
大阪大学	10Gbps	○	○	○	○	○
東京工業大学	1Gbps	—	○	○	○	○
九州大学	—	—	—	—	—	—
NAREGI	1Gbps	○	○	○	○	○
分子科学研究所	2000E10月	1Gbps	—	○	—	○
名古屋大学	2008年3月	1Gbps	—	△	△	△

○: NAREGI/TSUBASA (PrestoIII)の管理ノードに接続
△: 名古屋大学が管理ノードは300台以上、1000台程度を予定。NAREGI/TSUBASAと連携して運用する

表 1: 管理ノード構成

表 2、図 5)。本章では、各拠点のノード構成を紹介する。

阪大: Xen VM 上に構築した管理ノード (ポータル、SS、IS-NAS、IS-CDAS) に対して連携設定を投入した。ベクトル型スーパーコンピュータ SX-8Rを含む3つの計算ノードを資源提供した。

東工大: 前述のデプロイメント機構を用いて VO サービスを立ち上げ、VO、スーパースケジューラの連携実証への参加を行った。TSUBAME クラスタへの各参加組織 SSからのジョブ投入は、新 SSの連携機能により、非競合での資源予約の上、実行確認された。また、VO ポータルは、ホスト証明書は大阪大学認証局、ユーザ証明書を NAREGI 認証局からの発行のものでの実験となり、証明書混合環境での動作確認となった。

当初は、TSUBAME 予備ノードの一部による連携動作確認を行っていたが、3 月末のメンテナンスの時期と揃えることで学内の理解を得て、運用ノードから120 ノード、1920CPUを計算ノードとして追加参加させた。この規模は NAREGI ミドルウェアの動作実績としては最大 CPU 数となる。

さらに、より多くの計算資源による実証をめざし、稼働直後の TSUBASA についても連携参加を検討した。TSUBAME と同一ネットワーク・パッチシステムによって管理されているので、NAREGI ミドルウェアのうち GridVM scheduler/engine のみをインストールし、IS-CDAS、VO 環境は上記 TSUBAME 上のものを利用する構成で、基本的なジョブの起動を確認したが、MPI による大規模ジョブが動作せず、早急な問題究明ができなかった。このため、スケジュールの短さを鑑み、連携システムからは切り離すこととした。

また、今後想定されるミドルウェア利用の下方展

Fact Sheet 2: 計算ノード構成

	アーキテクチャ	OS	スケジューラ	ノード数	CPU
	gridvm scheduler/engine	SX-8R	SUPER-LIX	NEC/NOS-1	1 0.3
大阪大学	gridvm scheduler/engine	x86	Linux	NEC/NOS-1	8 0.4
	gridvm scheduler/engine	x86	Linux	NEC/NOS-1	4,50 15.8
東京工業大学	gridvm scheduler/engine	x86/CleanSpeed	Linux	SunGridEngine	120 18.3
九州大学	gridvm scheduler/engine	x86	Linux	PrestoIII	0.3
	gridvm scheduler/engine	SX-8R	SUPER-LIX	NEC/NOS-1	2 0.2
WPIアカデミー	gridvm scheduler/engine	x86	Linux	TSUBAME	14 0.3
	gridvm scheduler/engine	x86	Linux	PrestoIII	278 3.4
分子科学研究所	gridvm scheduler/engine	POWER7	Linux	TSUBAME	12 3.5
	gridvm scheduler/engine	x86	Linux	PrestoIII	6 0.2
名古屋大学	gridvm scheduler/engine	SUPER-LIX	SunGridEngine	Parallel	7 0.3

表 2: 計算ノード構成

開の第一歩として、研究室でミドルウェアを導入し、連携に参加した。具体的には、「PrestoIII」に GridVM scheduler/engine をインストールし、IS-CDAS ノードも用意した上で、独立の計算クラスタとして VO 環境への接続を行った。当初、クラスタに導入されている OS が NAREGI ミドルウェアのサポート対象外であるなどの困難があったが、仮想マシン (Xen) を利用して環境の差異を吸収でき、ミドルウェアが柔軟な実装であったことにより対応することができた。最終的には、構築した環境に、各組織のポータル経由での基本的なジョブの投入・動作し動作確認を行った。

九大: グリッド実験用に導入したユーザサイドクラスタ (Intel Xeon 2.8GHz×2CPU 4GB) の4ノードを提供し、IMS にて構築されている管理ノードよりジョブ実行が可能となるように構築を行った。

NII: PC クラスタ Xeon 2.8GHz/3.0GHz を搭載するノードから11 台を管理ノードとして、19 台を計算ノードとして運用した (3 台は兼用)。管理ノードにおける SS には、連携の中核となるリソース連携用のモジュールが NII に配置された。小規模なクラスタ中心ではあるが、これらの参加、不参加を、頻繁に、また、様々なパターンにおいて行ったときの手順について、確認を行った。

IMS: NAREGI β 2 により7 拠点からなるナノサイエンス VO を構築して実証を行ってきたが、今回の連携実証を行うにあたって、複数 VO 対応版 SS への移行及び IS-NAS のグローバル IP 化が必要であった。SS の移行作業は問題無く行われたが、既存の IS-NAS のグローバル IP 化は困難であったため、NII で運用していた IS-NAS を利用することによって解決した。これによりナノサイエンス VO は、他機関との連

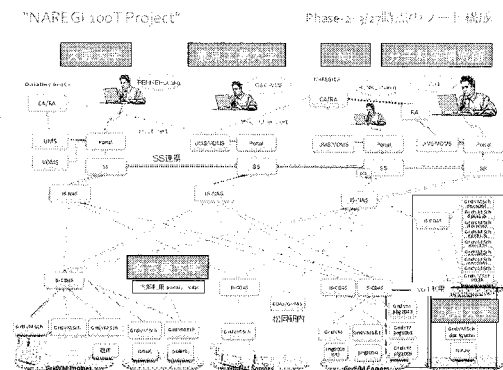


図 5: 管理ノードおよび計算ノード構成

携が行えるようになり、ナノサイエンス VO ポータルサーバより NII、阪大、東工大、名大へのジョブ投入が可能になった。しかしナノサイエンス VO 内に存在する各 VM サイトを他 VO へ提供するには、VM サーバのグローバル IP 化が必要であり大きな構成変更を伴うため断念した。

名大: 今回の連携実証実験では、準備期間が殆どなかったこともあり、同センターで構築した β2 版の計算ノードによる資源提供サイトとしての参加となった。具体的には同センターの GridVM サーバ (OS は Linux と Solaris) を阪大の IS-CDAS ノードに接続し、阪大のポータルから利用する形とした。実験に際して行うべき作業は、阪大のホスト証明書を取得し、その証明書の置き換えと “RENKEI-Osaka” と “vol” のそれぞれの VO に属するユーザ ID の登録であった。これらを行うことで簡単に接続完了となるはずであったが、実際には幾つかのネットワーク的なトラブルと DN 情報の登録及び証明書関係のエラーを解決するのに数日を要した。なお、阪大のホスト証明書は置き換えではなく追加する形をとったので、同センターで構築した GridVM サーバは名大のポータルからもそのまま利用可能であった。

4. グリッドオペレーション

今回の大規模実証実験では、NII が中心となって NAREGI ミドルウェアの導入支援、計算機資源の運用支援、認証局運用を行った。

4.1 ミドルウェア導入支援

大阪大学および東京工業大学の研究者や技術者と連携し、これらの大学の計算機への NAREGI ミドルウェア β2 版の環境構築を行った。また、ここで得られたミドルウェアの改善点は、次期バージョンの開発へフィードバックされている。本実証実験実施

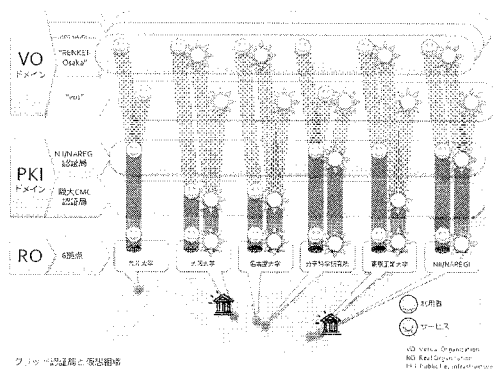


図 6: PKI ドメインと VO ドメイン

時期は、NAREGI ミドルウェアバージョン 1 の開発中であつたため、実験は β2 版を用いて行われたが、スーパースケジューラについてのみ、スケジューラ間の資源予約情報を共有する機能の付いたバージョン 1 のモジュールを先出し、広域での連携という機会を利用してその機能の確認を行った。

4.2 認証局運用とアカウント管理

NII と阪大が実験参加者に対する証明書発行を行った。同時に、グリッドの連携運用には不可欠であるアカウントの相互発行の業務についての実証を NII にて行った。

情報基盤センターのユーザ管理方法は各センターによって異なり、ユーザアカウント申請手続きも異なることが一般的である。本実証実験では、このような情報基盤センター間のユーザ管理手続きの違いを隠ぺいするため、NII においてユーザアカウント申請の代理窓口を実験的に運用した。これにより、実験に参加するユーザは、代理窓口に必要な情報を届け出て申請することにより、実験に参加する組織の計算機上にアカウントを作成することができる体制を整えた。なお、阪大では、グリッド認証局としてユーザ証明書 (50 通) およびサーバ証明書 (59 通) を発行した。

VO は、基本的には、各ポータルを運用する拠点ごとに構成したが、IMS に関しては、提供される計算機資源の AUP から、IMS 所有の計算資源を他の VO に見せない認可設定となる構成を行った (図 6)。

4.3 運用監視

阪大にて開発した SS の動作検証システムによって連携開始時点から逐次連携状況をモニタし、ミドルウェアの品質改善、特に安定性の向上に貢献し

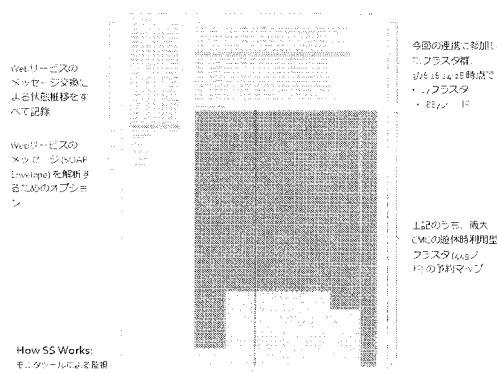


図 7: SS の状態監視システム

た。

同システムは、ソースコードのクロスリファレンスに基づく静的解析システムとコアイメージのデバッグ情報に基づく動的解析システムを Web インターフェイスによってマッシュアップしている。ログ情報やコアダンプ時のデバッグ情報を元に、該当するソースコードを追跡表示し迅速な問題特定を可能とすると同時に、定期的なコアイメージのスナップショットを取得し、過去に遡ってサーバプロセスの状態遷移を検証することを可能としている。これまでの Web サービス型システムにおいては、SOAP メッセージ交換のトラッキングによる受動的な解析が主であったが、本システムによって、サーバプロセスそのものの状態遷移を検証可能とした。これによって、複数の GridVM サーバとの WS-Notification の交換が適切に行われることを検証し、全体の安定運用を促進した (図 7)。

5. アプリケーション実行による評価

阪大: レーザーエネルギー学研究センターの協力を受け、「レーザー生成高エネルギー密度金属の破壊ダイナミックスの分子動力学シミュレーション」を行った。GridMPI による並列化に際し、各ノードのロードバランスを行うために、動的に計算分割領域を最適化している。3 月の連携実証期間後も同センター内でサービスを継続し、220 並列のシミュレーションを 2 ヶ月に及んで安定して実行することができた。

九大: 九州大学情報基盤研究開発センターにて研究開発を進めている OpenFMO 及び RISM-FMO 連成計算プログラムを本実証研究の全拠点 (阪大、東工大、名大、NII、IMS、九大) の計算機システムへの移植と動作確認を行い、拠点内及び拠点間に跨るグリッド上での並列ジョブとして検証を行った。その結果、OpenFMO プログラムに関しては約 16,000 原

子で構成する真空中の分子の電子構造計算を行い、RISM-FMO 連成計算プログラムに関しては約 500 原子で構成する水溶液中の分子の電子構造計算を実現し、大規模なグリッド基盤上で高速に正しい計算結果が得られることを検証できた。

IMS: 1) 商用アプリによる長時間実行およびパラメータサーベイ、2) GridMPI アプリ 5 本を使った移植性確認、3) 特定アプリによる複数サイト実行確認、の 3 パターンで評価した。1) では阪大の SX8 及び東工大の Tsubame へ、2) では阪大の Express-5800 へ、3) は、MD アプリを使用して阪大、東工大、名大へ、それぞれナノサイエンス VO ポータルサイトより WFT を使用してジョブを実行した。3 パターン全て正常に動作させることができた。

6. まとめ

今回の連携実証実験は、ごく短期的とはいえ、開発時の統合環境や NAREGI 拠点の資源と比較するまでもなく、これまで動作実績のない規模での実証ができたことは、ミドルウェアの可能性の確認として非常に有意義であった。

情報サービスやスケジューラの連携機能は、基盤センター間の連携において、非常によく機能を果たしたと判断している。信頼性が心配されていた SS の連携用モジュールを NII に配置し、そこを起点に動作検証を精密に行うことができた。NII 環境への移植作業も含めた、大阪大学のデバッグによる多くの問題発見が大きな効果を上げた。特に、GridVM・SS 間のメッセージの不整合がかなり厳密に解析されたため、開発部隊が具体的解決手段を提供でき、最終の最大規模の実証では、非常に安定した状態で参加・動作を確認することができた。

複数の仮想組織を連携させた環境において、異なる VO 配下マシンで様々な形態のジョブがセキュアに安定性を保ちながら正しく計算結果が得られることを検証できた。今後、各機関間で異なる VO が構築されたとしても、各機関のリソースは相互に共有出来ることが今回示されたことは評価に値する。

また、実際の運用環境への導入と短期ではあるが運用における試行錯誤を通じて、運用環境を対象とした特有の問題について多くの知見を得た。実証実験全般を通じて、NII の支援・運用組織がよく機能し、また、各参加組織の運用担当者レベルでの人的連携が取れるようになったことも大きな成果だと考えられる。また、問題が具体的に見つかったこともあり、

各開発部隊の方々にも非常に迅速にご対応いただいた。終盤には、各拠点の運用人員が、NAREGI ミドルウェア、および、連携動作への理解を深めた結果として、自主的に並行して問題の解決にあたり、短期での問題解決、連携動作実現にこぎつけることができたことは、人材育成の観点でも大きな成果であった。

連携実験を行うに当たっては、連携する拠点の証明書をお互いに信用することが必要であり、それぞれの拠点の認証局がどのレベルの PMA に準拠しているかが重要であることを改めて認識した。今後、グリッドによるサービスを運用レベルで行う場合には、どのような認証局をどのように運用していくかが大きな課題であることを再認識した。

アカウント申請の代理窓口の試行では、アカウントの相互発行の業務の確認のため、仮想的な相互発行のための組織として運用を行い、かなりの短期間でアカウント発行、証明書 DN に関する情報交換を成立させることができた。また、問題となる事例、特に、一部アカウント・証明書を取得済みの場合における注意・考慮すべき点などについての知見を得ることができた。

一方で、運用および現機能では対処しきれない問題も多々残っている。

基盤センター毎に導入システムや運用形態が異なるが、ミドルウェアがその差を吸収しきれていない部分も多い（特にローカススケジューラインターフェイス）。今後は、基盤センター毎に適用のための改変・国立作業を行う必用が生じている。その改変を集約し、より普遍的な基盤センター運用に適用可能とする努力とその維持が必用である。

モジュールの段階的導入において正常に導入・動作できているか確認の術がないため、かなりの規模のソフトウェアであるにもかかわらず、動作の確認が全体導入後のテストジョブまでほとんどできない。これは、連携によってさらにシステムを大きくしていく過程においては、かなりの足かせであった。研究室からの参加による知見として、連携動作を確認するためには、ソフトウェア単体の確認でなく、システム全体の構成や状況を把握すべきことが多い。都度、研究室からシステム全体を把握することは、非常に煩雑・困難であるため、インストール環境の差異を考慮したデプロイツールやミドルウェアのレイヤ間のモニタリングを容易に行うツールの整備などが課題である。

また VO 間連携にあたり、直接的な IP 接続性が要求されるサーバ数が多く、公開が必要なポートも雑多であり、マシン運用環境に多大な影響を与えている。

アプリケーションの実行インターフェイスとしての課題も多い。アプリケーションプログラムが実行できるまでの試行錯誤に多くの時間を要した。各サイトの利用可能資源状況や WFT でのジョブの状態表示がエンドユーザにはわかりにくく、問題となっている原因が利用者側にあるのかシステム側にあるのかの判断材料に乏しい。利用者にとって有用となる情報（エラー原因、予約時間、キューの利用状況など）を明示するなどをし、より一層の利便性向上を目指すべきと考える。また、各サイトでジョブを実行させる準備として、同じプラットフォームでもサイト毎の環境差異への適合作業に時間を費やした。アプリケーションを移植する際に各サイト環境を強く意識しなければならない状況は、グリッドの利点であるジョブの可搬性を阻害する要因になると思われる。今回は実証できなかった PSE がサイト環境を吸収するサポートしてくれることを期待したい。

今後は、この実証を通じて得られた知見を反映すると共に、バージョン 1 においてもより大規模な連携動作を継続し、より安定したシステムを目指していく。

参考文献

- 1) 坂根栄作, 東田学, 岡村真吾, 寺西裕一, 秋山豊和, 馬場健一, 下條真司, “全国共同利用環境へのグリッドミドルウェアの適用”, 情報処理学会 研究報告 (IPSI SIG Technical Reports 2007-DSM-45), pp. 25-30, May 2007.
- 2) Eisaku Sakane, Manabu Higashida, Shingo Okamura, Toyokazu Akiyama, Yuuichi Teranishi, Ken-ichi Baba and Shinji Shimojo, “An Authentication System Development in Application of Grid Middleware to A Nationwide Joint-Use Environment for Computing”, to be appeared in *Proceedings of IEEE SAINT Workshop 2008*, July 2008.

謝辞 実験に関係された皆様、参加組織の皆様、開発組織からご協力いただいた方々にお礼申し上げます。特に、昼夜を問わず参加していただいた方々、どうもありがとうございました。