

1 ビジネスグリッドの狙い

吉野松樹 >> (株)日立製作所 : matsuki.yoshino.pw@hitachi.com

阿部欣成 >> 富士通(株) : yoshinari.abe@jp.fujitsu.com

中誠一郎 >> 日本電気(株) : s-naka@ab.jp.nec.com

ビジネスグリッドコンピューティングプロジェクトは、2003年度より3年間、経済産業省の主導による官民連携プロジェクトとして実施された。本プロジェクトは、科学技術計算の分野で発展しすでに広く利用されているネットワーク接続された複数のコンピューティングリソースを協調動作させるグリッドコンピューティングの概念を、ビジネスで利用する情報システムに応用することで、情報システムの信頼性向上をリーズナブルなコストで実現する技術を確立することを狙ったものである。

本稿では、グリッドの概念の概説を行い、それを踏まえてこのプロジェクトの狙いを解説する。

グリッドの概念

❖ グリッドという言葉の由来

「グリッドコンピューティング」とは、「パワー・グリッド(送電線網)」という言葉に由来する概念である。電力は、原子力、火力、水力など発電方式や、発電所の場所など一切気にせず、コンセントにプラグを差し込めば利用できる。これと同じようにネットワーク経由でコンピューティングリソースをその場所や機種などを気にせず利用できるようにしようというのが「グリッドコンピューティング」である。

❖ 科学技術計算分野におけるグリッド

科学技術計算分野においては、コンピュータを利用したシミュレーションが広く行われているが、その計算には膨大なCPU資源が必要であり、膨大な量のデータが使用される。最先端の研究を行うためには、単独のスーパーコンピュータでは不十分な状況になっている。研究に必要な複数のスーパーコンピュータをつなげて仮想的に巨大なスーパーコンピュータとして使いたいというのが、科学技術計算分野においてグリッド技術が発展した理由である。

科学技術計算分野の最大規模のグリッドの例として、2004年1月から本格稼働し始めたTeraGridがある。これは、北米の9つのスーパーコンピューティングセンタの合計100TFLOPSを超える計算パワーを40Gbit/secの高速ネットワークで接続し、合計15ペタバイトの情報を管理するというものである。それぞれのセンタは、計算パワーやストレージを提供するだけでなく、特色ある機能を提供している。たとえば、アルゴンヌ国立研究所は、計算結果をビジュアルにみせるための計算を高速に行う専用ハードウェアを持っており、別のセンタのスーパーコンピュータで計算した結果をアルゴンヌ国立

研究所に送って、ビジュアル化しリモートでその結果を見ることができる。日本国内の同様の例として、2003年に文部科学省がスタートさせた「超高速コンピュータ網形成プロジェクト(National Research Grid Initiative: 通称NAREGI)」で構築したNAREGIテストベッドがある。NAREGIテストベッドは、国内各地の大学、研究機関の計算資源(プロセッサ数で合計約3,000)を超高速学術情報ネットワーク(スーパーSINET)で結び17TFLOPSの性能を達成している。NAREGIテストベッドは、2006年から開始した文部科学省プロジェクト「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」に引き継がれ、グリッド基盤ソフトウェアの研究開発、ナノサイエンスによるグリッド実証研究のための研究基盤として利用されている。

複数のスーパーコンピュータシステムを接続して仮想的に1つのスーパーコンピュータのように利用できるようにするためには、連携するシステム間の情報のやりとりを保証し、さまざまな連携先の違いを意識せずに連携できるように、必要な標準仕様が整備されることが必要である。グリッドに関する標準化の議論を行う場として、GGF(Global Grid Forum)が2001年から活動しており、科学技術計算分野だけでなくビジネスへの応用を含め幅広い議論が行われている。さらに、2004年にビジネス分野でのグリッド技術の普及を目指して設立されたEGA(Enterprise Grid Alliance)とGGFは、2006年6月に合併してOGF(Open Grid Forum)として活動することを発表、さらにビジネス分野でのグリッド技術の適用の普及・発展のための活動を推進していくことが期待されている。

❖ PCグリッド

「グリッド」という言葉を聞くと、科学技術計算向けのグリッドよりも、多くのPCの空き時間を利用して

膨大な計算を行う「PC グリッド」を思い浮かべる人も多いであろう。「PC グリッド」で最も有名なプロジェクトは、プエルトリコにあるアレシボ天文台の電波望遠鏡で捉えた電波を解析して地球外知的生命体から発信された電波を探そうという SETI@Home (The Search for Extraterrestrial Intelligence with a Home Computer Screen-Saver) プロジェクトである。

日本国内の同様のプロジェクトとして、2002年12月から2003年4月にかけて、(株)NTTデータが実施した cell computing プロジェクトがある。プロジェクトの結果は、以下の URL で報告されている (<http://www.nttdata.co.jp/release/2003/062600.html>)。結果の要点をまとめると、12,206 台の PC の参加があり、ピーク時性能で 3TFLOPS を記録している。全 PC の合計 CPU 時間は、611 年 162 日 13 時間 24 分 23 秒であり、1 台の PC で実行すると 600 年以上かかる計算を約 4 カ月で終了したことになる。なお、この実験では、「ヒトの遺伝子情報からの周期性の発見」と「光学的に新たな特徴を持つ材質の設計図の作成」が問題として選択され、それぞれ有意義な結果が導かれたとのことである。

SETI@Home、cell computing いずれの場合も、解析対象データは小さな断片に分割され、インターネット経由で参加している PC に送付される。解析プログラムはスクリーンセーバになっており、PC の空き時間に解析を行う。配布されたデータ断片の解析は、他の PC の解析結果とは独立に行えるため、このような方法で、大規模な解析をきわめて効率的に行うことが可能である。このような種類の問題 (embarrassingly parallel と呼ばれることもある) を効率的に解くためには PC グリッドはきわめて有効である。

シミュレーションの技法に乱数データを用いてシミュレーションを繰り返すモンテカルロシミュレーションという技法がある。シミュレーション回数を増やせばその分だけシミュレーション精度が向上する。それぞれのシミュレーションの結果は他のシミュレーションの結果とは独立なので、複数の PC で乱数データが重複しないよう工夫して同時に並行してシミュレーションを実行すれば、短時間に精度のよい結果が求められる。創薬、遺伝子情報の解析、設計シミュレーション、金融のリスク計算などさまざまな分野でこのような形態での PC グリッドの利用が進んでいる。

ビジネスグリッドコンピューティングプロジェクト

❖ プロジェクトの目的

ビジネスにおいても、スーパーコンピュータを駆使して行う計算が必要な場合もあり、またモンテカルロシ

ミュレーションが有効なケースが多々あることは事実であるが、ビジネスグリッドコンピューティングプロジェクトの狙いは科学技術計算向けのグリッドや PC グリッドのビジネス分野への適用拡大ではない。科学技術計算向けのグリッドを実現するために創り出された概念、技術、標準仕様などを情報システムの信頼性向上と運用効率向上に利用することが目的である。

❖ リソース共有によるコスト削減

ビジネスグリッドコンピューティングプロジェクトで対象としている情報システムは、主にサーバを使った分散型の情報システムである。通常、情報システムで取り扱う業務量には日時による変動がある。当然業務量のピーク時に安定稼働可能なリソースを用意する必要があるため、結果として情報システム全体としては過剰な設備を抱え、コスト増大を招くことになりかねない。図-1 に現状の情報システムの状況と、ビジネスグリッドによる解決案を示す。

特に、近年増加しているインターネットベースの WWW システムでは、情報システムに不特定多数のユーザがアクセスすることから、従来の企業内等に対象が限定される情報システムと比較して必要な処理能力の正確な見積りが困難であり、さらにピーク時とオフピーク時のトランザクション量に大きな差があることから、この傾向が顕著となっている。ユビキタス情報社会が進展し、RFID、センサネット、自動車などさまざまなものがデータの発信源になるとますますこの傾向が強くなると予想される。このような状況においてもあまりコストをかけずに信頼性の高い情報システムを提供するための基盤を確立することが、ビジネスグリッドコンピューティングプロジェクトの狙いの1つである。

❖ リーズナブルなコストでの信頼性の確保

もう1つの大きな狙いは、リーズナブルなコストで信頼性の高い情報システム基盤を提供することである。図-2 に、情報システムの信頼性と構築・運用コストの関係を示す。

図-2 の右上の領域は、高信頼性を必要とする情報システムである。ここでは、システムの二重化、自然災害等に備えた地理的に離れた複数のセンタでのバックアップ設備の確保、緻密な高信頼性設計などにより、相当なコストをかけて信頼性を確保している。

それほどコストをかけられない一般的なシステムでは、図-2 の左下の領域になってしまっているのが現実だと考えられる。業種・業態、事業規模を問わず、情報システムの信頼性は事業継続性に直結するケースが増えてきており、リーズナブルな範囲のコスト増加で、情報

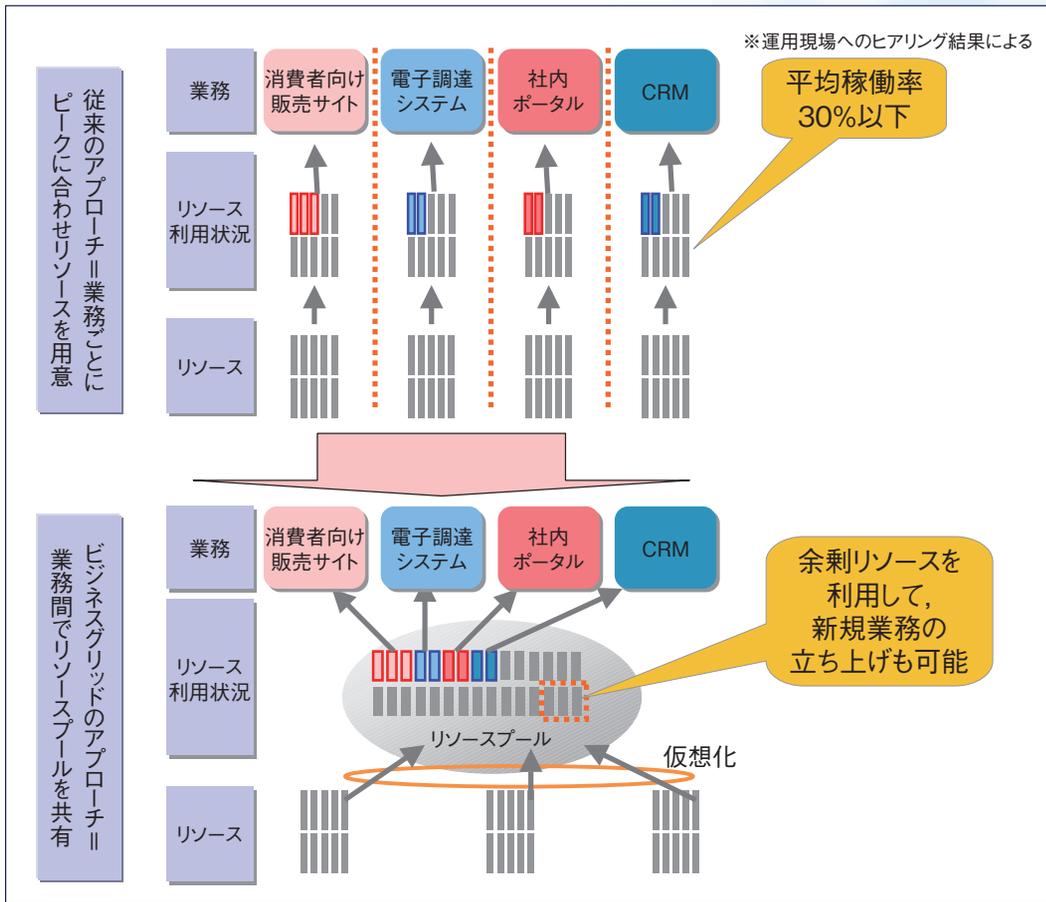


図-1 ビジネスグリッドでのITリソース共有のアプローチ
業務ごとに準備してITリソースをプール化して必要なだけ割り当てることで平均稼働率を向上させる

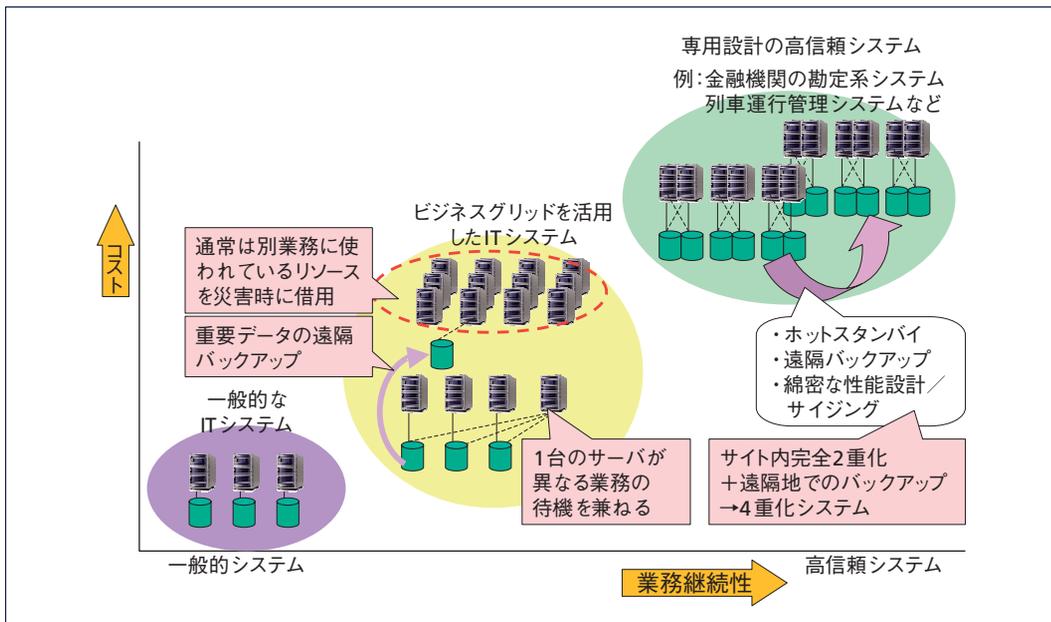


図-2 信頼性とコスト

システムの信頼性を向上させるソリューションの提供が急務である。ビジネスグリッドでは、1台のサーバで複数のヘテロジニアスな（異機種混在している）サーバのバックアップを可能とする技術や、専用の待機リソースを用意せずにディザスタリカバリ（災害によるシステム障害からの業務復旧）を可能とする技術を開発することで、図-2の真中の領域を埋めることを狙いとしている。

❖ 運用・保守コストの低減

これまで述べた2つの狙いを実現することで、情報システムの運用・保守コストを低減することがビジネスグリッドの目的である。

情報システムは、初期開発が終了し実際に業務で利用され始めると、運用フェーズに入り一定のコストが継続的に必要となる。また、業務環境の変化等に合わせた改

変その他の保守の作業も必要である。この運用・保守のコストを削減する手段を講じないと、IT投資の枠が一定だと仮定すると、運用・保守のコストが累積し新規のIT投資を行う余裕がなくなってしまう。情報システムの運用にグリッド技術を適用することで、運用・保守のコストを下げ、新たな戦略的なIT投資を行う余裕を生み出し、事業における競争力を強化することが重要である。この状況を図-3に示す。

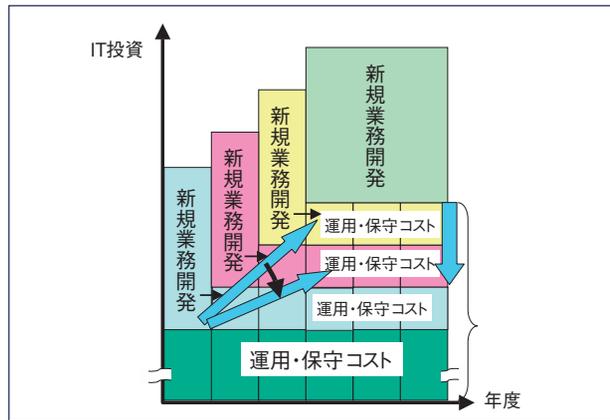


図-3 運用・保守コストの低減

❖ プロジェクトの推進体制

このような目的を達成すべく、経済産業省主導の下、日本電気、日立製作所、富士通の3社が共同でプロジェクトを推進した。プロジェクトは経済産業省から参加企業に対する委託事業として行われたが、官民連携プロジェクトとして参加企業側も委託事業費と同額の開発投資を行うマッチングファンド方式で行われた。

プロジェクト全体は、推進委員会が推進し、推進委員会幹事会で実務レベルの推進を行った。プロジェクト全体の事務局を、(独)情報処理推進機構が務め、技術アドバイザーとして(独)産業技術総合研究所グリッド研究センターが参画した。産業技術総合研究所グリッド研究センターは、本プロジェクトの一部として、科学技術計算分野向けグリッドを幅広く利用可能とするためのGridASPの開発を行っており、その概要を本特集でも紹介している。また、国際標準化に関する情報交換・戦略に関する議論の場として標準化WGを組織し活動した。プロジェクトの体制を図-4に示す。プロジェクトには、ビジネスグリッドコンソーシアムとストレージシステムコンソーシアムの2つのコンソーシアムがある。本特集はビジネスグリッドコンソーシアムの活動について記述している。

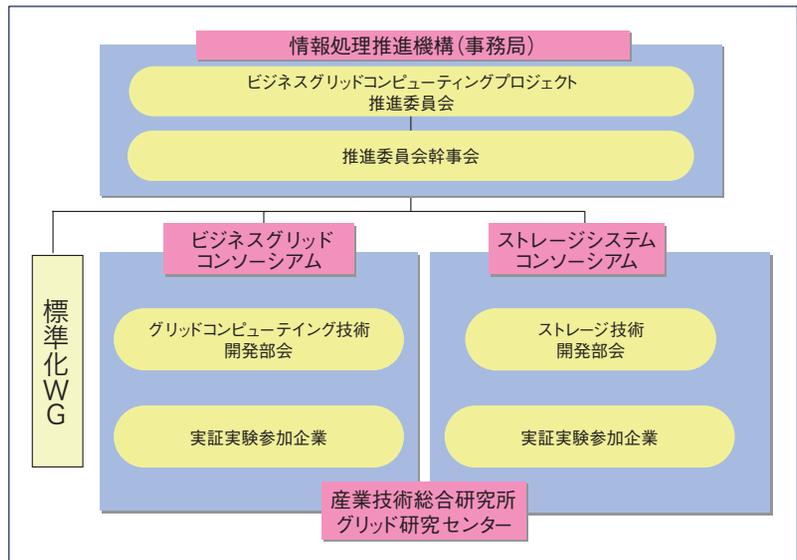


図-4 プロジェクト推進体制

とした実証実験を行う。

- 国際標準を適用するだけでなく、本プロジェクトで開発した技術を国際標準の場に積極的に提案する。本特集では、本稿に続き、上記の項目に関してそれぞれの成果を別稿として報告している。

❖ 技術開発のフェーズ分け

本プロジェクトは、3年間にわたるプロジェクトである。長期間にわたるプロジェクトでは、最終目標だけでなく、中間目標を明確に定めて遂行することがプロジェクトを成功に導くために重要である。本プロジェクトにおいても、年度ごとの目標を定めてプロジェクトを推進した。

年度ごとの目標を設定する上で考慮した大きなポイントは下記の3点である。

- ビジネスグリッドミドルウェアを適用する範囲が同一のセンタ内であるか、複数のセンタをまたがるのか
- リソース割当て戦略
- 自律制御のレベル

❖ 目的達成に向けた推進方法

具体的なプロジェクトの目標としては次のようなものを掲げてプロジェクトを推進した。

- グリッド技術のビジネス適用に向けた先進的な技術開発を行う。また、単なる技術開発だけにとどまらず、その技術を適用し実際に動作するミドルウェアを開発する。
- 開発したミドルウェアを実際のユーザ企業の情報システムで適用しその効果の検証とフィードバックを目的

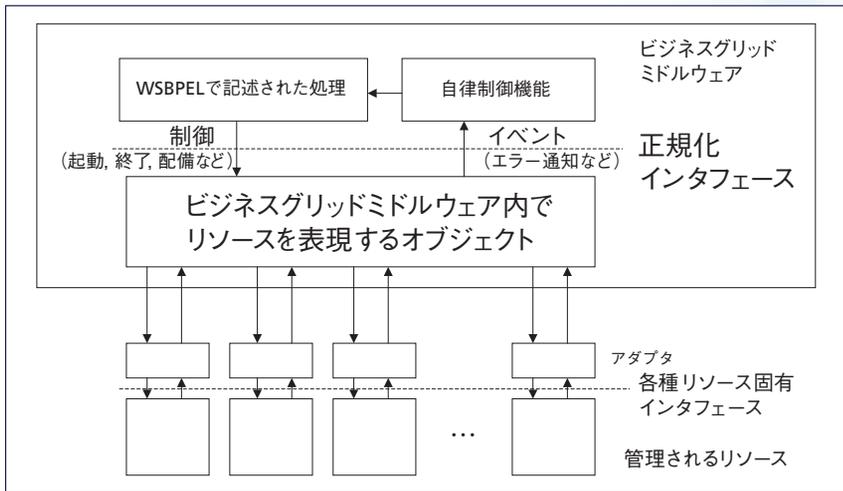


図-5 リソースの仮想化のメカニズム

以下、それぞれについて説明を行う。

❖ ビジネスグリッドミドルウェアの適用範囲

情報システムにおけるリソースの有効活用ということを考える場合に、まず考えるのは、同じセンタの中にあるリソースをうまく有効活用することである。しかし、ビジネスグリッドの目的であるリーズナブルなコストで信頼性の高い情報システムを実現するためには、広域負荷分散、ディザスタリカバリなど複数のセンタが協調して動作するための技術が不可欠である。

ここで、同一センタ内、複数センタ間を問わずビジネスグリッドミドルウェアが提供する基本的機能について若干説明しておく。

ビジネスグリッドミドルウェアは、リソースの有効活用のために、業務で使うリソースは事前に決めずビジネスグリッドミドルウェアが割り当てる。事前に業務が使うリソースが決まらないため、リソースの環境設定や、アプリケーションの配備等の作業を従来のように人手で行うことは難しく、ビジネスグリッドミドルウェアが自動的に行う必要がある。このような処理をプロビジョニングと呼ぶ。ビジネスグリッドミドルウェアでは、プロビジョニングの処理手順をあらかじめ WSBPEL (Web Services Business Process Execution Language) で記述しておくことで自動実行を可能としている。WSBPEL は、業務を複数のサービスの組合せで実現する SOA (Service Oriented Architecture) でサービス呼び出しの処理フローを記述するために一般的に使われる XML の記法であり、ここではプロビジョニング手順の記述に使っている。

先にも述べたように、プロビジョニングの処理を記述する時点では、業務で利用するリソースが決まっていない。プロビジョニングする対象が決まっておらず、そのためプロビジョニングに必要な処理のインタフェース

も分からない状態で、プロビジョニングの処理が記述できるのかという疑問が湧くかもしれない。これを解決するのがビジネスグリッドミドルウェアで提供するリソースの仮想化機能である。

図-5 にビジネスグリッドミドルウェアでのリソース仮想化のメカニズムの概要を示す。管理対象となるリソースとビジネスグリッドミドルウェアの間にあるアダプタが実際のリソースのインタフェースを正規化インタフェースに変換する。ビジネスグリッドミドルウェアの世界では、正規化インタフェースだけを意

識すればよい。WSBPEL で記述するプロビジョニングの処理もこの正規化インタフェースで記述するので、業務で利用するリソースが決まる前に記述することができる。

複数のセンタを協調して動作させるためには、単一のセンタ内のプロビジョニングだけではなく、複数のセンタをまたがった広域プロビジョニング技術の開発が必要となる。

広域プロビジョニング技術を含め、複数のセンタが協調して動作するための広域連携の技術開発は難易度が高いと判断し、初年度で同一センタ内でのリソース有効活用、信頼性向上のための基本技術を開発し、2年目以降広域連携機能の開発および、初年度開発した基本技術のブラッシュアップを図る計画とした。

❖ リソース割当て戦略

リソースの有効活用を図るためには、ビジネスグリッドミドルウェアが管理しているリソース群から要求に応じてリソースを割り当てる際の戦略が重要である。大きなセンタでは、サーバ台数が数千台規模に上るところもあり、規模が大きくなっても性能的に問題が発生しないような戦略が必要である。また、将来考案されるものを含め複数のリソース割当て戦略の中からセンタごとにあるいはその時の状況に応じて最適な戦略を選択したいというニーズが想定される。

初年度は、あらかじめ定められたスケジュールに基づく静的なリソース割当てと、比較的単純なアルゴリズムに従う動的リソース割当て戦略に基づいた実装を行った。動的リソース割当てとは、業務実行中に障害が発生したリソースに対して代替リソースを割り当てたり、業務負荷が増大した時にリソースを追加したりする場合のリソースの割当てのことである。2年度以降、下記のような高度なリソース割当て戦略の技術開発と実装を行った。

- 複数のリソース割当て戦略から最適なものを選択できるようにするためのプラグイン機構
- 業務に対する具体的なリソースの割当てを実行直前に行うことでよりリソースの柔軟な割当てを可能とするレイトバインディング
- 業務の優先度を考慮したリソース割当て戦略
- 広域連携におけるリソース割当て戦略

❖ 自律制御のレベル

情報システムの運用コストを抑えながら信頼性を向上させるためには、システム運用中に起こるさまざまな事象に対する対処を極力自動化し、オペレータの介入を最小限にすることが一番効果的である。

ビジネスグリッドミドルウェアでは、各々のリソースから上がってくるイベントも仮想化しているので、イベントに対応するアクションをプロビジョニングと同様にWSBPTELで記述しておくことにより、イベントに対するアクションを自動的に行うことができる。

対処を必要とするイベントは、障害の発生および障害発生の予兆となる監視項目の閾値超過等である。監視項目としては、CPU 負荷、メモリ使用量等の業務とは独立したシステムレベルの監視項目と、業務ごとの応答性能、スループット等の業務レベルのSLO (Service Level Objective) となる監視項目の2種類がある。

初年度は、システムレベルの監視項目を対象とした単純なイベント-アクションによる自律制御の機構を実装した。

システムレベルの監視項目を対象にイベント-アクションを定義するよりも、業務レベルのSLOを対象にイベント-アクションを定義する方が、より利用者視点に即したサービスレベルを維持できる。さらにWeb3階層アプリケーションのように複雑な構成の業務アプリケーションでは単純なイベント-アクションの関係だけでは不十分である。たとえば、応答性能が低下したからといって、常にWebサーバを追加すればよいというものでもない。応答性能の低下がDBサーバでの処理が滞留していることにあってDBサーバの追加がアプリケーション構成上不可能であるという場合には、新しいトランザクションの流入を抑止して(たとえば「ただいま込み合っております。後ほどアクセスしてください」などと表示するsorryサーバに誘導するなど)滞留している

	2003年度	2004年度	2005年度
基本線表	基本機能の設計 ／開発	マルチベンダ対応／複数サイト対応等の機能拡張および実用化を目指した実証実験	
技術研究 ／開発内容	サイト内で業務配備	複数サイトにまたがる業務配備 ・ディザスタリカバリ 機能強化	
	サイト内でのリソース割当て	複数サイトにまたがるリソース割当て 機能強化	
	自律フレームワーク	負荷・障害時、自律的にサービス品質を維持 複数サイト対応	
	マルチベンダリソースの仮想化	リソースを仮想化するためのSDK	OASIS WSDM 対応
実証実験		実証実験	

図-6 各年度の開発内容

処理が完了するのを待つというのが、正しい対処となる。

このような状況に対応できるよう、イベントの発生に対して、複数の状態を確認し、その上で複数の対処候補から適切な対処を選択する。また、さらに対処の効果を確認しSLOがまだ満足されていないければ、さらに別の対処を実施するといった、人間の運用オペレータの思考をシミュレートする高度な自律制御および、広域連携における自律制御機能の技術開発は2年目以降に実施した。

❖ 各年度の開発内容まとめ

図-6に各年度の開発内容のまとめを示す。マルチベンダリソースの仮想化は、当初本プロジェクト固有のインターフェースで行っていたが、国際標準の制定に合わせて、最終年度はOASIS WSDM (Web Services Distributed Management) 仕様に準拠したインターフェースも開発した。

実証実験の詳細については、本特集の「ビジネスグリッドの実証実験」に譲る。

今後に向けて

本プロジェクトの成果を踏まえ、ビジネスグリッドの普及・発展を目的としてビジネスグリッド推進コンソーシアム (<http://businessgrid.ipa.go.jp/consortium>) を2006年5月11日に設立し、本プロジェクト参画メンバーに加えユーザ、SI事業者、IDC事業者など幅広い会員企業・団体で活動を行っている。

(平成18年7月28日受付)