

平成 21 年 6 月 17 日現在

研究種目：基盤研究 (B)
研究期間：2006 年度～2008 年度
課題番号：18300018
研究課題名 (和文) ユーザ主導型グリッドコンピューティングを実現する高度スケジューリング技術
研究課題名 (英文) Advanced Scheduling Technique Enabling Grid Computing Controlled by Users
研究代表者
合田 憲人 (AIDA KENTO)
国立情報学研究所・リサーチグリッド研究開発センター・教授
研究者番号：80247212

研究成果の概要：これからのグリッドコンピューティングでは、計算の完了時刻や費用、計算精度等、ユーザの様々な要求に柔軟に対応することが求められる。本研究では、ユーザからの多様な計算要求に対応するための高度なスケジューリング技術の確立を目的とし、計算機やネットワークの最適な利用計画を自動生成するアルゴリズムの提案、提案手法を評価するためのシミュレーション技術およびグリッドエミュレーション技術に関する研究開発を行った。

交付額

(金額単位：円)

	直接経費	間接経費	合計
平成 18 年度	3,800,000	1,140,000	4,940,000
平成 19 年度	7,900,000	2,370,000	10,270,000
平成 20 年度	3,500,000	1,050,000	4,550,000
総計	15,200,000	4,560,000	19,760,000

研究分野：総合領域

科研費の分科・細目：情報学・ 計算機システム・ネットワーク

キーワード：ハイパフォーマンスコンピューティング、ネットワーク

1. 研究開始当初の背景

インターネットに代表されるネットワーク技術および PC クラスタに代表される高性能計算機技術の発展に伴い、インターネット等の広域ネットワークに接続された計算機やデータベース等の計算資源を安全に安定してかつ容易に利用することにより高性能計算を実現するグリッドコンピューティング技術が注目されている。グリッドコンピューティングは、インターネット上の複数の計算資源を統合することにより得られる莫大な計算能力を利用することにより大規模問題を高速に解く手法として、情報工学、オペレーションズリサーチ、生命科学、高エネルギー物理学等の様々な分野において期待が大きい。

従来のスーパーコンピュータ等を利用した高性能計算では、ユーザは拠点内に設置された計算機上でアプリケーションプログラムを実行していたため、与えられた計算機の性能や状態によって、アプリケーションプログラ

ムの実行時間、結果の精度、課金額等が決まる、言い換えれば計算資源主導の形で計算が実行されていた。これに対してグリッドコンピューティングでは、ユーザが利用可能な計算資源は 1 拠点内に留まらず、インターネット上の様々な計算資源を利用可能であり、またユーザのアプリケーションプログラムの性質や実行状態に応じて、利用する計算資源の構成を動的に変更することが可能になる。従って、今後のグリッドコンピューティング技術の普及により、ユーザの計算に対する要求が多様化し、計算資源の側がユーザの計算要求に柔軟に対応するユーザ主導の形で計算を実行することが求められる。例えば「○月○日○時までに計算を終了せよ」、「○円以下の予算でできるだけ早く計算を実行せよ」、「○ライブラリを利用して精度△以下の結果を求めよ」等、従来の計算ではなかったユーザからの多様な計算要求が発生することが予想される。

図 1 は、このようなユーザからの計算要求

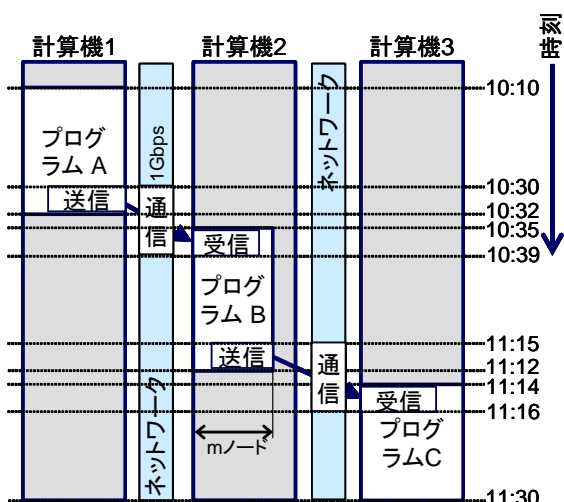


図1. 計算要求の例

の一例を示している。本例では、ユーザのアプリケーションは3つのプログラム(A, B, C)から構成され、これらが連携して3台の計算機上で実行される。また、プログラムBおよびCの実行はそれぞれ、事前に実行されるプログラムAおよびBの出力結果の入力が必要とする。このような例は、例えばローカルサイトの計算機(計算機1)上で入力データの前処理(プログラムA)を実行し、リモートサイトの高性能計算機(計算機2)上で高い計算能力を必要とする計算処理(プログラムB)を実行した後、画像処理用計算機(計算機3)上で結果の可視化(プログラムC)を実行する場合が該当する。ここでユーザからアプリケーションの実行を11時30分までに終了せよとの計算要求が提示された場合、これを保証するためには、各プログラムの計算時間および計算機間の通信時間の見積を行い、例えば10時35分から11時12分まで計算機2上の計算ノードm台をプログラムBのために利用する、計算機1・2間のネットワークの帯域1Gbpsを10時30分から10時39分まで利用する等、計算資源の利用計画を作成し、利用計画通りに各計算資源を確保する必要がある。

このような多様なユーザの要求を満足するためには、インターネット上の計算資源の中からユーザの計算要求に合うものを選択し、計算処理を適切に割り当てるスケジューリング技術が重要となるが、従来のグリッド上のスケジューリング技術では、上述の例のようなユーザの多様な要求を満足することはできない。

2. 研究の目的

本研究では、ユーザからの多様なアプリケ

ーションプログラムの実行要求に対して、柔軟に対応するための高度なスケジューリング技術を開発することを目的とする。

具体的には、ユーザからのアプリケーションプログラムの実行要求に対して、インターネット上の計算資源群(計算機、データベース、ネットワーク等)の最適な利用計画を自動的に作成し、利用計画に基づいて計算資源群の事前予約を行うことにより、ユーザの多様な要求を満足する資源割り当てを行う技術を開発する。

3. 研究の方法

本研究では、グリッド上でのユーザ主導型の高度なスケジューリングを実現するため、計算資源の利用計画を自動的に生成するアルゴリズムを提案するとともに、シミュレーション等によりアルゴリズムの性能評価および改良を行う。以下に本研究の実施方法を3項目に分けて説明する。

(1) 計算資源利用計画自動生成アルゴリズムに関する研究

本研究が目指すユーザ主導型高度スケジューリングでは、ユーザの計算要求を満足するための計算資源をインターネット上の計算資源の中から選択し、その利用計画を作成するとともに、選択した資源を予約することが必要となる。これを実現するため、本研究項目では、グリッド上の資源予約のためのモデルを構築するとともに、本モデルに基づいて計算資源の利用計画を自動生成するためのアルゴリズムを提案する。

(2) グリッド上のユーザ主導型高度スケジューリングの性能評価技術に関する研究

グリッドのみならず分散計算環境上でのスケジューリング問題の解決には、提案するアルゴリズムの性能評価技術が重要となる。本研究項目では、(1)で提案したアルゴリズムの性能評価のためのシミュレーションソフトウェアの開発を行う。また、より現実的な評価モデルに基づく性能評価を実現するため、グリッドのエミュレーション環境の構築も行う。

(3) 経済原理を導入したスケジューリング技術に関する研究

近年のグリッド技術の成熟に伴い、グリッド上のスケジューリングに経済原理を導入し、より高度なスケジューリングの実現を目指す研究が注目されている。本研究項目では、計算資源利用計画の自動生成に経済原理を導入するモデルの検討を行う。

4. 研究成果

本研究では、3 節にあげた 3 つの研究項目について研究開発を行い、スケジューリングアルゴリズムを提案するとともに、性能評価を行った。

並列・分散システム上でのスケジューリング技術は、これまで国内外で活発な研究が行われ、成熟した研究分野であるが、グリッドコンピューティング環境で今後必要とされるよりユーザ主導型のより柔軟なスケジューリングを実現するためには、既存技術では不十分である。これに対して本研究では、計算資源利用計画の自動生成を行う新たなアルゴリズムを提案した。

また本研究では、スケジューリング技術をより現実に近い条件で性能評価するためのシミュレーションおよびエミュレーションモデルを開発し、提案アルゴリズムの性能評価を行った。本研究で開発した性能評価モデルは、今後、グリッドコンピューティングのみならず、並列・分散システム上でのスケジューリング技術の性能評価にも利用可能である。

本研究では最後に、近年注目されている経済原理を導入した資源割り当て技術を提案アルゴリズムに導入するための検討を行った。経済原理をスケジューリングに導入する技術は、現在欧米を中心に研究が始まったところであるが、本研究では、これらの最先端の技術を取り入れたスケジューリングモデルの検討を行った。

本節では以後、各研究項目の成果について、その詳細を説明する。

(1) 計算資源利用計画自動生アルゴリズムに関する研究

本研究項目では、グリッド上の資源予約のためのモデルを構築するとともに、本モデルに基づいて計算資源の利用計画を自動生成するためのアルゴリズムを提案した。

①資源予約モデルの構築

資源予約のためのモデル構築では、ユーザからの計算要求の受付と解釈、インターネット上の計算資源に関する情報収集、計算資源利用計画の作成、複数の計算資源に対する予約手続き等の処理をどのような手順で組み合わせるのかという問題を解決するためのプロトコルの検討を行った。

図 2 は、そのプロトコルの一例を示している。主な処理は、計算資源利用計画を作成する **planning** 処理と複数の計算資源に対して予約手続きを行う **reservation request** 処理からなる。**planning** 処理では、ユーザからの計算要求 (**job request**) およびインターネット上の計算資源の状態 (**resource info**) を入力とし、後述のアルゴリズムを用いて計算資

源の利用計画 (**plan**) を作成する。**reservation request** 処理では、**planning** 処理が作成した

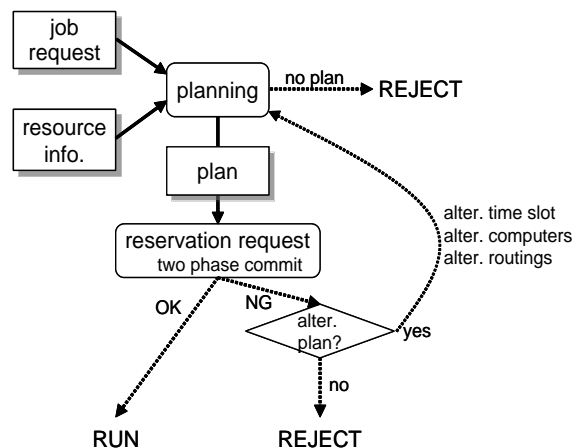


図2. スケジューリングプロトコル

利用計画に記述された計算資源に対して、予約手続きを実施する。また、**reservation request** 処理では、複数の計算資源を同時に予約することは不可能であるため、個々の計算資源に対して仮予約を順次行い、全ての計算資源の仮予約が成功した段階で、本予約に切り替える **two phase commit protocol** を用いている。

②計算資源利用計画自動生成アルゴリズム

図 2 の **planning** 処理では、計算資源の利用計画を自動生成するアルゴリズムが必要となる。本研究項目が目指す計算資源の最適な利用計画の作成は、ユーザからの計算要求(例えば利用する計算資源のアーキテクチャやソフトウェアに関する制約、締切時刻等)を満足することを条件として、図 1 の例のようにインターネット上の計算資源およびネットワークに対してプログラムの処理や通信処理を割り当てる最適な方法を求める最適スケジューリング問題に帰着できる。一般のこの種の最適化問題は利用する計算資源の負荷等の状態が不変の場合でも **NP** 困難な問題であり、ユーザからの計算要求がある度に最適解を求めることは実用的ではない。そこで本研究項目では、計算機やネットワーク上での処理に関する性能予測モデルを構築し、本モデルに基づいてユーザの計算要求を満足するように処理の割り当てを決定する 2 つのヒューリスティックアルゴリズム (予約プラン探索アルゴリズム、**Mixed-Parallelism** スケジューリングアルゴリズム) を提案した。

予約プラン探索アルゴリズムは、ワークフロー形式で表現されたアプリケーションの実行要求に対して、ユーザの条件を満足する計算機およびネットワークの組み合わせをバックトラック法により探索し、これらの資源の予約を行う。また **Mixed-Parallelism** スケジ

ューリングアルゴリズムは、ワークフローを構成するジョブの並列性（1つのジョブに割り当てる CPU 数）とジョブの計算資源への割り当て順序の2つを同時に決定し、計算資源の予約を行う。

以上の研究成果は、雑誌論文(1)(2)および学会発表(3)(7)として発表されている。

(2) グリッド上のユーザ主導型高度スケジューリングの性能評価技術に関する研究

本研究項目では、(1)で提案したアルゴリズムの性能評価のためのシミュレーションソフトウェアの開発およびグリッドエミュレーション環境の構築を行うとともに、アルゴリズムの性能評価を行った。

シミュレーションソフトウェアの開発では、(1)①の資源予約モデルに基づいた離散事象シミュレータを開発した。本シミュレータでは、図2に示すプロトコルに基づく資源予約が可能のほか、インターネット上の計算機の仕様やネットワークトポロジとそれらの動的な振る舞い（計算機の負荷、メモリ容量、ネットワークのバンド幅等の変動）を再現する機能を持つ。

シミュレータの実装では、シミュレーション環境上に、計算機資源の予約管理モジュール、ネットワーク資源の予約管理モジュール、グリッド上の情報サービスモジュール、資源の利用計画生成モジュールを実装し、またこれらのモジュール間のAPIも定義した。これにより、より実環境に近い環境での事前予約スケジューリング技術の性能評価を行うことを可能とした。

本シミュレータを用いた性能評価では、対象とするグリッド環境として人工的に作成したモデルおよび日本国内の学術ネットワークをモデル化したものを作成し、事前予約スケジューリング技術の性能評価に用いた。また、世界各国の計算機センターやグリッド環境上のジョブ実行履歴を解析することにより、本研究が対象とする予約ジョブのモデル化を行い、性能評価に用いた。

本研究では、シミュレーションによる評価のみではなく、より実際のグリッド環境に近い状況での評価も重要である。このような性能評価環境の構築として、グリッドのエミュレーション環境の構築を進め、アルゴリズム評価に必要な機能の検討をシミュレーションと並行して行った。

以上の研究成果は、雑誌論文(3)および学会発表(2)(4)(5)(6)(8)として発表されている。

(3) 経済原理を導入したスケジューリング技術に関する研究

本研究項目では、計算資源利用計画の自動

生成に経済原理を導入するモデルの検討を行った。本モデルでは、計算資源に対する需要と供給の状況によって計算資源が提供する計算能力（例えば CPU×時間）の利用料金（価格）が決定され、計算資源割り当てが行われる。例えば、図3では複数の利用者と資源提供者の間でオークション形式による計算資源割り当てを行う例を示しており、利用者と資源提供者の条件、即ち、購入希望価格(\$11) > 資源提供価格(\$10)が合致する利用者 B と資源提供者 A との間で計算資源割り当てが実行される。

本研究項目では、図3に示すようなオークションメカニズムによる資源割り当てモデルを構築するとともに、本モデルを用いて計算資源利用計画の自動生成を行うヒューリスティックアルゴリズムを提案した。

以上の研究成果は、学会発表(1)として発表されている。

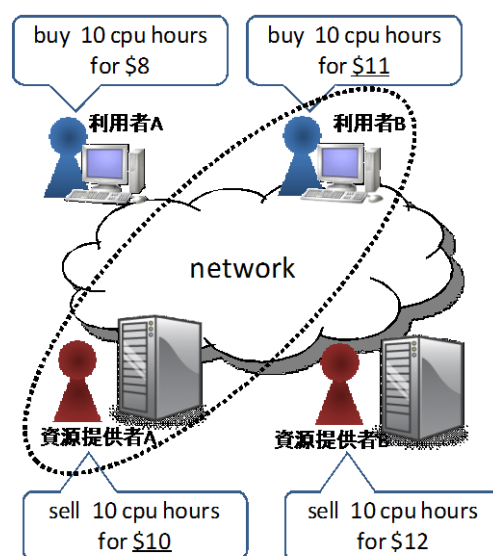


図3 オークションによる計算資源割り当て例

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 3 件)

- (1) Kento Aida, Henri Casanova, "Scheduling Mixed-Parallel Applications with Advance Reservations", Cluster Computing, 2009, 査読有.
- (2) Kento Aida, Henri Casanova, "Scheduling Mixed-Parallel Applications with Advance Reservations", Proceedings of the 17th ACM/IEEE International Symposium on High-Performance Distributed

- Computing (HPDC'08), 2008, 査読有.
- (3) 笠井武史, 西村元一, 前田高宏, 大澤清, 合田憲人, "グリッドアプリケーション評価のためのネットワークエミュレーション", 情報処理学会論文誌, Vol.48, No. SIG13(ACS19), pp.145-155, 2007年, 査読有.

[学会発表] (計 8 件)

- (1) Ikki Fujiwara, Kento Aida, Isao Ono, "Market-based Resource Allocation for Distributed Computing", 情報処理学会先進的計算基盤システムシンポジウム SACSIS2009, pp.149-150, 2009年5月28日, 広島.
- (2) 西村元一, 合田憲人, "グリッド上の計算機エミュレーション手法", 情報処理学会研究報告 2008-HPC-114, 2008年3月6日, 札幌.
- (3) 安藤誠士郎, 合田憲人, "グリッド上の事前予約スケジューリング手法の性能評価", 情報処理学会研究報告 2007-HPC-113, pp.37-42, 2007年12月7日, つくば.
- (4) 笠井武史, 西村元一, 前田高宏, 大澤清, 合田憲人, "グリッドアプリケーション評価のためのネットワークエミュレーション", 情報処理学会先進的計算基盤システムシンポジウム SACSIS2007, pp.293-301, 2007年5月25日, 東京.
- (5) 笠井武史, 西村元一, 前田高宏, 大澤清, 合田憲人, "グリッドネットワークのエミュレーション手法", 情報処理学会研究報告 2007-HPC-109, pp.175-180, 2007年3月, 札幌.
- (6) 笠井武史, 西村元一, 前田高宏, 合田憲人, "グリッドエミュレーション手法に関する研究", インターネットコンファレンス 2006 論文集, pp.134, 2006年10月23日, 東京.
- (7) 安藤誠士郎, 合田憲人, "グリッド上の資源予約スケジューリングアルゴリズムの性能評価", 先進的計算基盤システムシンポジウム SACSIS2006 論文集, pp.225-226, 2006年5月22日, 大阪.
- (8) 笠井武史, 西村元一, 前田高宏, 合田憲人, "擬似グリッド実験環境の構築", 先進的計算基盤システムシンポジウム SACSIS2006 論文集, pp.223-224, 2006年5月22日, 大阪.

6. 研究組織

- (1) 研究代表者
合田 憲人 (AIDA KENTO)
国立情報学研究所・リサーチグリッド研究開発センター・教授

研究者番号 : 80247212

- (2) 研究協力者
藤原 一毅 (FUJIWARA IKKI)
総合研究大学院大学・複合科学研究科情報学専攻・博士課程
Sun Hao
東京工業大学・大学院総合理工学研究科物理情報システム専攻・博士課程