

エグゼクティブサマリー

米国 DOE ASCR (Department of Energy Office of Advanced Scientific Computing Research、エネルギー省先端科学コンピュータ研究局) が資金を提供した Magellan プロジェクトの目的は、特に中規模コンピューティングおよび将来のデータ集約型コンピューティングワークロードに関連して、DOE SC (Department of Energy Office of Science、エネルギー省科学局) のコンピューティングのニーズに応えるクラウドコンピューティングの潜在的役割を調査することであった。性能、使いやすさ、およびコストにおけるクラウドコンピューティングのさまざまな側面を調べるために、一連の調査課題が設定された。これらの課題に取り組むために、分散テストベッド基盤が ALCF (Argonne Leadership Computing Facility) および NERSC (National Energy Research Scientific Computing Center) に導入された。このテストベッドは、さまざまなコンピューティングモデルおよびハードウェア設計ポイントを調べ、さまざまな科学アプリケーションに与える影響を理解するために十分に柔軟かつ強力であるように設計されている。プロジェクト実施中、テストベッドは、応用科学者に価値あるリソースとしても提供された。MG-RAST (メタゲノミクス解析サーバー)、JGI (Joint Genome Institute、共同ゲノム研究所)、RHIC (Relativistic Heavy Ion Collider、相対論的重イオン衝突型加速器) の STAR 実験、LIGO (Laser Interferometer Gravitational Wave Observatory、レーザー干渉計重力波天文台) など、さまざまなプロジェクトのアプリケーションが Magellan プロジェクトでクラウド内のベンチマークのために使用されたが、プロジェクトチームは Magellan クラウドリソースを使用して重要な生産科学を達成することもできた。

クラウドコンピューティングは、特にカスタムソフトウェア環境および利用増加といった幅広いコンピューティングニーズや要件に対応できるモデルとして、産業および研究科学者の双方から大きな注目を集めている。クラウドサービスは、パブリックでもプライベートでも容易にコスト効率良く使用でき、さまざまな企業および Web のワークロードに対応する、拡張性に富んだサービスのセットを提供できることが実証されている。これらの利点は、ネットワーク経由でアクセス可能でユーザによって弾力的に調整可能なプールされたオンデマンドセルフサービスリソースという、クラウドコンピューティングの定義から導かれる直接的な結果である。大規模なユーザベースに対してリソースをプールすることは、スケールメリットを高めるとともに、容易に提供可能で、弾力的に拡張できることで、柔軟な能力を提供できる。

エグゼクティブサマリーに続き、プロジェクトの主要な知見および提案をまとめる。詳細は報告書の本文で説明する。ここでは、研究から得られた知見の大枠を簡単にまとめる。

- クラウドアプローチでは、ユーザが自分のソフトウェアスタックを持ち込んで、管理の手間をあまりかけずに新たなコンピューティング環境で試すことができるカスタマイズされた環境、大きな課題に対応するためにリソースを急速に拡大する機能、およびスケールメリットの増加により得られる経済的な利点など、数多くの利点が提供される。仮想化は、これらの能力を提供する主な戦略である。クラウドを使用する応用科学者と協力した筆者らの経験では、仮想化の機能によって、完全にカスタマイズされた環境と柔軟なリソース管理、およびそれらが科学者にもたらす潜在的価値が実証された。
- クラウドコンピューティングでは、アプリケーションをこのような新しいモデルに移植するために、最初に多大な労力と技術が必要になる場合がある。これはクラウドコンピューティングで使用される

いくつかの新しいプログラミングモデルにも当てはまる。科学者は、クラウドへの移行を決定する際に、あらゆる経済的な分析においてこの初期投資を考慮しなければならない。

- 仮想環境、ワークフロー、データ、サイバーセキュリティ、およびその他の管理の分野には、大きな不足や課題がある。こうした新しいコンピューティングモデルで顕在化した機能を科学者が容易かつ効果的に利用できるようにするには、さらなる研究開発が必要である。これには、クラウド環境を簡単に利用できるようにするツール、オープンソースクラウドソフトウェアスタックの改良、これらをカスタマイズする柔軟性を許容しながらユーザのブートストラップを助ける基本イメージの提供、新しいセキュリティ技術とアプローチの調査、および科学的データおよびワークフローにより適合する MapReduce モデルの強化などが含まれる。さらに、これらの機能を従来の HPC プラットフォームで可能にする方法を調査する機会もある。これにより、クラウドモデルの柔軟性を HPC システムの性能に結び付けることができる。
- クラウドの重要な経済的な利点は、広範なコミュニティにわたるリソースの集中化から得られる。その結果として、利用率の向上、スケールメリット、および運用効率の向上が導かれる。既存の DOE のコンピューティングセンターはすでにクラウドコンピューティングから多くの恩恵を受けている。これらのコンピューティングセンターは、複数のプログラム局にわたるコンピューティングを集中化し、大規模に展開し、運用効率の継続的な改良改善に努めているからである。コスト分析によれば、DOE のコンピューティングセンターはコスト競争力を持っている。商用クラウドプロバイダと比較すると、コストはおおむね 1/3~1/7 である。商用セクターは革新を続けているので、DOE の研究所およびコンピューティングセンターは、そのコンピューティングコストを継続的にベンチマークして、パブリッククラウドと対比し、競争力のあるサービスを提供できていることを確認する必要がある。

クラウドコンピューティングは、最終的にビジネスモデルであるが、クラウドモデルは多くの場合、特定のワークロードに役立つ追加機能と柔軟性を提供する。DOE の研究所およびコンピューティングセンターは、何十年にもわたる開発と精緻化を経て最適化されてきた生産性やコンピューティングプラットフォームの効果を損なうことなく、より多様なワークロードをサポートし、より多くの科学的発見を可能にするために、その運用にクラウドコンピューティングのこうした機能の採用と統合を考慮する必要がある。このアプローチでは、科学者のニーズに応えることができない場合は、まずプライベートクラウドコンピューティング戦略を考慮すべきである。プライベートクラウドは、パブリッククラウドで広く知られているセキュリティ、データ管理、および性能に関する課題を避けながら、商用クラウドの多くの利点を提供できるからである。

提案

Magellan プロジェクトでは、DOE SC (DOE Office of Science、エネルギー省科学局) の資金提供を受けて、クラウドコンピューティングモデルおよびさまざまな関連技術を評価し、科学者のコンピューティングニーズへの対応におけるその潜在的役割を調査した。これにより得られた知見では、DOE SC のアプリケーションにクラウドコンピューティングを利用することの潜在的な価値および課題の両方が強調される。ここでは、DOE SC、DOE リソースプロバイダ、応用科学者、およびツール開発者に対する筆者らの提案をまとめる。提案のいくつかは、現在の適用範囲と予算に収まらず、現在の DOE プロジェクトに提供されている資金を超える追加予算が必要になる場合がある。

DOE SC。 Magellan プロジェクトによる知見では、クラウドコンピューティングには、DOE SC 科学コミュニティのコンピューティングニーズを満たす上で、いくつかの潜在的な利点があることが示されている。カスタマイズされた環境や MapReduce プログラミングモデルなどのクラウド機能は、DOE リソースの一部の現行ユーザのニーズおよび現在の環境を大規模化する必要のある新しい科学アプリケーションのニーズに対応するのに役立つ。また、既存の DOE HPC センターは、商用プロバイダと比較して、コスト競争力があることも判明した。したがって、DOE は DOE HPC センターおよび DOE リソースプロバイダと協力して、科学コミュニティの拡大するニーズを確実に満たす必要がある。こうしたニーズに対応するためにクラウドモデルが必要と思われる場合は、DOE がまずプライベートクラウドコンピューティング戦略を検討することを推奨する。このアプローチにより対費用効果が高まり、より最適化されたオフリングが可能になり、セキュリティ、性能、およびデータ管理の課題への取り組みを向上させながら、クラウド環境の多くの利点が提供される。

DOE リソースプロバイダ。 スーパーコンピューティングセンターで稼働している大規模密結合アプリケーションという現行の従来型モデルに適合しないアプリケーションのクラスが増大しているが、コンピューティングおよびストレージリソースのニーズは増え続けている。さらに、さまざまな科学コミュニティは、固有のコラボレーションまたはワークロード要件を満たすために、カスタムソフトウェア環境または共有仮想クラスタを必要としている。さらに、クラウドにより、大規模リソース要件を持つ高スループットおよびデータ集約型アプリケーションのための機構およびツールが提供される。クラウドモデルによりこれらの要件の一部に対応するソリューションが提供されるが、これらのソリューションは DOE リソースプロバイダが実現することもできる。筆者らはこうした機能を提供するために DOE リソースプロバイダに特定の推奨事項を示す。リソースプロバイダは、Leadership Computing Facilities や NERSC などの大規模な ASCR の施設、多数の研究所によって運用される組織のリソース、および他の科学プログラム局の資金提供による多数の大規模ユーザ施設を伴うコンピューティングリソースを含んでいる。

1. DOE リソースプロバイダは、ますます多くのコンピューティングおよびストレージリソースが必要になる高スループットかつデータ集約型のワークロードなど、より多様なワークロードをサポートするための機構を調査する必要がある。特に、リソースプロバイダは、こうした粗結合コンピューティングワークロードのために、より柔軟なキューイングポリシーを検討する必要がある。

2. いくつかのユーザコミュニティでは、コンピューティング負荷の増加のために、一定の期間にわたり、あらかじめ定められた量のリソースへの排他的アクセスをする必要がある異なるリソースプロビジョニングモデルが必要である。Material Genomes プロジェクト、Joint Genome Institute、およびドイツの大腸菌株解析に関する筆者らの経験では、オンデマンドリソースプロビジョニングは、科学ユーザグループのコンピューティング要件に対応する上で重要な役割を果たすことができることが実証されている。筆者らは、DOE リソースプロバイダが、ユーザグループにオンデマンドリソースを提供するための機構を調査することを推奨する。
3. いくつかの DOE 共同研究は、固有のライブラリを持ち、バージョンの依存関係がある、複雑な科学用ソフトウェアパイプラインに依存している。仮想マシンは、カスタマイズ可能な環境を必要とするエンドユーザにとって便利である。しかし、多くの密結合科学アプリケーションでは仮想化のオーバーヘッドが大きい。こうしたアプリケーションは、カスタム環境を提供するベアメタルプロビジョニングやその他のアプローチの恩恵を受けることができる。DOE リソースプロバイダは、共有リソース上に適切に調整された環境を科学者に提供する機構を検討する必要がある。
4. DOE リソースプロバイダはコスト効率と電力効率が優れている。しかし、商用部門は絶えず革新を続けているので、DOE リソースプロバイダは商用クラウド部門と比較してコスト効率と電力効率を把握することが重要である。
5. プライベートクラウド仮想化ソフトウェアスタックはプロジェクトを通じて成熟している。しかし、性能、安定性、および信頼性の点では大きな課題がある。DOE リソースプロバイダは、こうしたソフトウェアスタックを実際使用するために幅広く導入する前に、プライベートクラウドソフトウェアスタックの開発コミュニティと協力して、こうした欠点に対処する必要がある。
6. DOE 固有の課金、割り当て、およびセキュリティポリシーを現行のクラウドソフトウェアスタックに実装するためには不足がある。クラウドソフトウェアソリューションは、リソース割り当て、セキュリティ、課金、および監視に関連するサイト固有のポリシーに対応するためのカスタマイズが必要になる。DOE リソースプロバイダは、サイト固有のカスタマイズをサポートするために開発するか、プライベートクラウドソフトウェアスタックの開発コミュニティと協力する必要がある。
7. ユーザ作成の仮想イメージは強力である。しかし、科学者の参入障壁を軽減するために、基本イメージおよび簡単なツールの標準セットも必要である。さらに、科学者はしばしば、スーパーコンピューティングセンタースタッフの専門知識が必要な調整済みライブラリを必要とする。DOE リソースプロバイダは、仮想化環境を簡単に使用できるようにする基準イメージおよびツールの提供を検討する必要がある。
8. クラウドにより、そのリソースおよびツールを使用するエンドユーザのトレーニングに追加投資を必要とする新しい利用モデルが顕在化する。また、新しいモデルでは、トレーニングを受けた要員が新しい技術によって生じる新たなプログラミングおよびシステム管理の負担を軽減するためにエンドユーザを支援できる新しいユーザサポートおよび協力モデルが必要になる。DOE リソースプロバイダは、こうした新しいモデルを幅広くサポートする前に、ユーザサポートの課題を注意深く検討する必要がある。

科学グループ。クラウドコンピューティングは、リソースへのオンデマンドアクセスやユーザ環境の制御などの利点があるので、科学アプリケーションに役立つことが期待されている。しかし、クラウドコンピューティングには、性能と信頼性、プログラミングモデル、イメージの設計と管理、コンピューティングリソース全体への作業の分散、およびデータ管理に関して課題があるため、アプリケーションの設計および開発に大きな影響がある。筆者らは、アプリケーションにクラウド技術またはモデルを検討しようとしている科学グループに対して以下のことを提案する。

1. IaaS（サービスとしてのインフラストラクチャ）は、既存のアプリケーションインフラストラクチャの多くを活用しながら、クラウドリソースを活用するための容易な経路を科学グループに提供する。しかし、仮想化クラウドシステムは、インスタンスタイプおよびストレージクラス（ローカル、ブロックストア、オブジェクトストア）に関して、性能および価格帯の異なるさまざまなオプションを提供する。科学グループは、さまざまなオプションについて、アプリケーションを慎重にベンチマークして、最良の価格性能比を見いだす必要がある。
2. クラウドシステムでは、アプリケーション開発者がソフトウェア環境を完全に制御できる。しかし現状では、そうした環境ではワークフローおよびデータ管理に使用できるツールの選択肢が限られている。科学グループは、これらの環境を管理するために、カスタムスクリプトを開発するのではなく、標準ツールを使用することを検討する必要がある。科学者は、ツール開発者と協力して、求めている要件とワークフローが十分に把握され、理解されるようにする必要がある。
3. アプリケーション開発者は、設計と実装の変動と障害の可能性について検討する必要がある。これは一般に良い慣行であるが、クラウドの中で実行されるアプリケーションには、重大な障害や性能の変動が生じるため、さらに重要である。
4. クラウド技術により、ユーザグループは自らのマシンイメージを管理できるようになり、複雑なソフトウェア依存関係があるグループがポータビリティを獲得できるようになる。この柔軟性には、これらのグループがセキュリティの懸念に確実に対応するという課題が生じる。科学グループは、そのイメージに関するセキュリティやその他の構成上の問題を防止するために、標準化されたセキュアイメージの使用を試みる必要がある。また、科学グループはイメージを安全にし、セキュリティパッチで最新の状態に保つための行動計画を持つことも必要になる。
5. メッセージキュー、テーブルストレージ、およびプロブやオブジェクトストアなどのクラウド技術は、必要でない場合には同期化を必要とせずにアプリケーションの規模を調整することができる主要な技術をいくつか提供する。これらの技術により、アプリケーション実行モデルは根本的に変わる。科学ユーザはメッセージキュー、テーブルストレージ、およびオブジェクトストレージといった技術をアプリケーションの設計段階で評価する必要がある。

ツールの開発と研究。 最近の数年間に、HPC およびクラウド環境の科学環境を管理するためのいくつかのサイトレベルおよびユーザサイドツールが進化した。しかし、この分野には大きな不足と課題がある。筆者らは、科学向けにクラウドモデルと技術の採用に関連したツールを開発および研究するためのいくつかの重要な領域を特定する。

1. 仮想マシンまたはプロビジョンされたベアメタルハードウェアは、多くのアプリケーショングループに役立つ。しかし、科学者は、ソフトウェアスタック、ジョブおよびデータコーディネーションを含む、これらのプロビジョンされたリソースの管理を行う必要がある。ツール開発者は、プロビジョンされたリソース上のワークフローおよびデータ管理を単純化および円滑化するツールのサポートを検討する必要がある。
2. 適切なパッチが利用できるようになったときにイメージを更新するための自動機構を実現するツール、およびユーザグループおよびコミュニティ全体でこれらのイメージを組織、共有、および検索するための単純な方法への投資が必要である。ツール開発者はイメージの組織および共有を有効にする開発サービスを検討する必要がある。
3. 仮想クラウド環境は、仮想マシンで利用可能なネットワークおよび I/O オプションの制約を受ける。GPFS や Lustre などの高性能並列ファイルシステム、および仮想マシン内の InfiniBand などの低レイテンシ広帯域相互接続へのアクセスにより、より多くの科学アプリケーションが、性能や使

いやすさを犠牲にすることなく仮想環境の恩恵を受ける。システムソフトウェア開発者は、仮想環境でHPC機能を提供する方法を検討する必要がある。

4. プライベートクラウドを監視し、セキュアに保つための新しい方法が必要であり、この分野における研究がさらに必要である。通常、サイトは多くのセキュリティポリシーを実装するために、OSレベルの制御に依存している。これらの制御の大半をハイパーバイザに移行するか、代替アプローチを開発する必要がある。セキュリティ開発者は、こうした環境を安全に保ち、理想的には仮想化によって提供される先進的な機構を活用する新しい方法を探さなければならない。
5. **MapReduce** は、データ集約型科学アプリケーションに対応するために役立つ場合があるが、科学データと解析方法の特性を明らかにする **MapReduce** の実装が必要である。計算機科学の研究者と開発者は、科学領域で一般的に使用されるデータフォーマットのデータモデルを理解し利用するためのフレームワークを可能にする、**Hadoop** などのフレームワークの変更と拡張を検討する必要がある。
6. いくつかのクラウドコンピューティングの概念と技術 (**MapReduce**、スキーマレスデータベースなど) が「ビッグデータ」および関連メタデータの管理の概念に基づいて発展した。クラウド技術は、自動スケーリング、フォールトトレランス、およびデータ局所性の課題に対処するもので、これらはすべて大規模システムの成功の鍵である。科学データとメタデータを管理するためのクラウド技術および概念の使用法を研究する必要がある。

第 1 章

概要

この数年間に、クラウドコンピューティングは、企業の Web アプリケーションのニーズに対応している。「クラウドコンピューティング」という用語は、多数の異なる概念（例：MapReduce、パブリッククラウド、プライベートクラウドなど）、技術（例：仮想化、Apache Hadoop）、およびサービスモデル（例：サービスとしてのインフラストラクチャ [IaaS]、サービスとしてのプラットフォーム [PaaS]、サービスとしてのソフトウェア [SaaS]）を指す用語として使用されている。クラウドは、コスト削減、迅速な弾力性、使いやすさ、および信頼性を含む、多くの重要な利点を提供することが実証されている。クラウドコンピューティングは特に、重要な IT インフラストラクチャが不足している顧客や既存の処理能力では対応できなくなった顧客において成功を収めている。

科学調査に際限がないこと、そして、科学における計算の役割が増大し続けていることから、計算リソースのニーズは増え続けている。この数年間、こうした需要に対処するためにクラウドコンピューティングの使用の評価に対する関心が高まっている。さらに、クラウド環境には、一部の科学アプリケーションにとって魅力的である多くの重要な機能がある。たとえば、多くの科学アプリケーションには、OS の特定のバージョンへの依存、コンパイラおよびライブラリを含む、固有のソフトウェア要件があり、ユーザは仮想環境が提供できるカスタムソフトウェア環境の柔軟性を求めている。その一例が Supernova Factory である。これは超新星探索のための大規模データに依存しており、多数のカスタムモジュールで構成されるコードベースがある。そのパイプラインは複雑で、特定のライブラリおよび OS のバージョンに依存する。仮想環境によって、科学者が協力者と環境を共有するためのポータブルコンテナの提供が約束される。たとえば、CERN の大型ハドロン衝突型加速器を用いた素粒子物理学実験である ATLAS 実験では、必要なすべてのソフトウェアを分散するための仮想マシンイメージの使用法を追求している [10]。同様に、MapReduce モデルは、データ集約型アプリケーション向けに有望である。このように、クラウドコンピューティングモデルは、データ集約型科学アプリケーション、オンデマンド/サージコンピューティング、およびカスタマイズされたソフトウェア環境を必要とするアプリケーションを含む、新しい分野の科学アプリケーションにとって有望な手段となる。科学コミュニティに提供されるサービスに対してクラウドのソフトウェアとビジネスモデルがもたらす影響については、科学コミュニティの数多くのグループが調査と追跡を実施している。しかし、科学アプリケーションのためのクラウドの運用法と使用法、科学ワークフローの移植法、およびクラウドの費用便益のトレードオフの決定法などについては、あまり分かっていない。

Magellan プロジェクトには、アメリカ復興・再投資法により、DOE SC（エネルギー省科学局）のアプリケーションに対するクラウドコンピューティングの適用可能性を調査するための資金が供与された。Magellan は ALCF（Argonne Leadership Computing Facility）と NERSC（National Energy Research Scientific Computing Center）の合同プロジェクトである。過去 2 年にわたり、IaaS（サービスとしてのインフラストラクチャ）、PaaS（サービスとしてのプラットフォーム）、仮想ソフトウェアスタック、MapReduce とそのオープンソース実装（Hadoop）などのクラウドのさまざまな特質を評価してきた。これらを、リソースプ

ロバイダの観点からは安定性、管理容易性、セキュリティ、そしてエンドユーザの観点からは性能と使いやすさといった、さまざまな基準で評価した。

クラウドコンピューティングには、グリッドコンピューティングやユーティリティコンピューティングなどの、他の分散コンピューティングモデルとの類似点がある。しかし、仮想化技術、MapReduce プログラミングモデル、および Eucalyptus や Hadoop などのツールを使用するために、科学環境に対するクラウドコンピューティングの影響を研究する必要がある。Magellan プロジェクトは、DOE の科学アプリケーションの固有の要件と、クラウドコンピューティングが科学コミュニティで果たすことができる役割を理解することに焦点を当てている。しかし、特定された不足と課題は、より広くクラウド環境を使用する科学アプリケーションに当てはまる。

1.1 Magellan の目標

Magellan プロジェクトの目標は、クラウドコンピューティングビジネスモデルが、DOE SC アプリケーションのニーズにどのように応えるかを調査することである。Magellan プロジェクトでは、特に以下の研究項目について答えを出すことが求められた。

- オープンソースのクラウドソフトウェアスタックはすぐに DOE HPC の科学アプリケーションに使用できるか
- クラウド内で DOE サイバーセキュリティ要件を満たすことができるか
- 新しいクラウドプログラミングモデルは科学コンピューティングに役立つか
- DOE HPC アプリケーションをクラウド上で効率良く実行できるか。クラウドにはどのようなアプリケーションが適しているか
- クラウド環境を科学アプリケーションにどの程度使用できるか
- どのようなときに、クラウドで DOE HPC の科学アプリケーションをコスト効率良く実行できるか

この報告書では、プロジェクトで上記の研究課題に取り組む際に得られた知見と提案をまとめる。

1.2 NIST のクラウド定義

「クラウドコンピューティング」という用語は、ここ数年間、さまざまな概念、モデル、およびサービスを指す用語として使用されている。この報告書では国立標準技術研究所 (NIST) が行ったクラウドコンピューティングの定義を使用する。NIST は、クラウドコンピューティングを、*管理の手間やサービスプロバイダの介入を最小限に抑えながら迅速に提供しリリースする構成可能な計算リソース (ネットワーク、サーバー、ストレージ、アプリケーション、およびサービスなど) の共有プールに対する利便性の高いオンデマンドネットワークアクセスモデル* [63] と定義している。

エネルギー省の資金提供による HPC (ハイパフォーマンスコンピューティング) センターなどの HPC センターは、リソースプール、ブロードネットワークアクセス、ユーザ割り当てに基づく従量サービスなど、このような数多くの重要な機能を提供している。しかし、今日の HPC センターでは、迅速な弾力性やオンデマンドセルフサービスのサポートは限られている。

1.3 影響

Magellan プロジェクトは、多くの領域に影響を与えた。Magellan のリソースは、エンドユーザが、仮想マシン、Hadoop、および従来のバッチキューなど、いくつかの構成で使用することができた。システムは、

すべての DOE SC オフィスのユーザが利用可能である。以下に、Magellan が過去 2 年間に影響を及ぼした重要な領域を示す。

- Magellan プロジェクトは、科学用のクラウドコンピューティングの使用について行われた初めての徹底的な調査である。その成果は、主要な計算機科学の学会およびワークショップで、多数の刊行物として発表された。
- JGI（共同ゲノム研究所）における施設の問題により、ゲノム解読の本運用を支えるバックアップコンピュータハードウェアに対する切迫したニーズが生じた。NERSC は ESnet と協力して、Magellan のハードウェアを HaaS（サービスとしてのハードウェア）モデルで提供して、研究所のニーズを満たすことができた。
- LCRC（Argonne Laboratory Computing Resource Center）は、Magellan プロジェクトのスタッフと協力して、本番用 HPC 計算クラスタを ALCF の Magellan テストベッドに拡大するセキュアな方法を開発することができた。これにより、LCRC のストレージにアクセスし、同じ計算環境内で実行しながら、ジョブで追加リソースを容易に使用できるようになった。
- RHIC（重イオン衝突型加速器）の STAR 実験では、リアルタイムのデータ処理に Magellan を使用した。この解析の目標は、「ミッシングスピン」を探すために衝突結果を精査することである。
- 生物学者は ALCF の Magellan クラウドを使用して、2011 年夏に欧州で発生した大腸菌感染で疑いのある大腸菌株を迅速に解析した。
- Magellan は、2010 年と 2011 年に「クラウドにおける HPC の最適な使用」で HPCwire 読者賞を受賞した。
- Magellan のベンチマーク研究は、CloudCom 2010 と DataCloud 2011 の両方で「Best Paper Award」を受賞した。
- Magellan は、ANI（Advanced Networking Initiative）に導入された 100 Gbps ネットワークの性能を示す上で、重要な役割を果たした。Magellan はストレージと計算リソースを提供し、SC11 で行われた多数のデモで使用された。デモでは通常 80~95 Gbps の速度を達成した。

1.4 アプローチ

このプロジェクトでは、クラウドコンピューティングおよび関連技術のさまざまな側面を評価するアプローチを使用した。また、プロジェクトで主要な決定を行いながら、科学コミュニティに固有のニーズを評価する、アプリケーション主導のアプローチを採用した。

分散テストベッドインフラストラクチャが ALCF（Argonne Leadership Computing Facility）と NERSC（National Energy Research Scientific Computing Center）に導入された。テストベッドは、IBM iDataPlex サーバーおよびアクティブストレージ、ビッグメモリー、および GPU サーバーを含む Infiniband ファブリックで接続された各種の専用サーバーで構成される。また、テストベッドには、分散型およびグローバル型双方のディスクストレージ、アーカイブストレージ、および 2 つのクラスのフラッシュストレージを含む、各種のストレージオプションもある。システムは広帯域、低レイテンシ、4 倍速 Infiniband ネットワークと一般的な Gigabit Ethernet ネットワークを提供する。この構成は典型的なクラウドインフラストラクチャとは異なっているが、科学アプリケーションにはより適している。たとえば、Infiniband は通常の商用クラウドオフリングでは一般的ではないが、これにより Magellan のスタッフは性能ポイントの範囲を調べ、アプリケーション性能への影響を測定することができた。

Magellan テストベッド上のソフトウェアスタックは、多様かつ柔軟であり、Magellan のスタッフおよびユーザはテストベッドでさまざまなモデルを検討することができた。テストベッド上のソフトウェアには、複数のプライベートクラウドソフトウェアソリューション、ワークロードの監視機能を備えた従来型

のバッチキュー、および **Hadoop** が含まれていた。ソフトウェアスタックは、使いやすさ、管理しやすさ、DOEセキュリティポリシーの適用能力、安定性、性能、および信頼性について評価した。

プロジェクトでは、マイクロベンチマークとアプリケーションベンチマークを使用して通信プロトコルと I/O を含むクラウドコンピューティングのさまざまな側面を評価する徹底的なベンチマーキングとワークロード分析を実施した。また、クラウドのコスト効率を理解するためにコスト分析も提供した。プロジェクトの主な目的は、プライベートクラウドソフトウェアがどのように運用されるかについての理解を深め、その不足と限界を特定することであった。しかし、完全を期すために、商用サービスとの性能とコストの比較も実施した。

さまざまな性能のトレードオフを理解し、科学パイプライン用のモデルの適用可能性を理解するために、**MapReduce** と **Hadoop** をさまざまなワークロードで使用して評価した。科学グループと密接に協力して、クラウドリソースを使用しながらアプリケーションの設計および関連する課題の複雑さを特定した。ユーザグループとのこの密接な作業は、現在の科学向けクラウド技術におけるいくつかの主要な不足を特定するために役立った。

1.5 報告書の概要

この報告書の以降の章は、次のように構成されている。第 2 章では、クラウドサービスモデルについて説明し、アプリケーションモデルを検討する。第 3 章では、**Magellan** プロジェクト活動のタイムラインと概要を示す。第 4 章では、DOE のユーザグループとの議論から、クラウドのための特定された要件やユースケースの要約を示す。第 5 章では、**Magellan** のテストベッドについて述べる。第 6 章では、普及しているさまざまな仮想ソフトウェアスタックオフリングの機能と使用経験を比較し、対比する。第 7 章では、ユーザサポートの問題について検討し、第 8 章では、セキュリティ分析について詳述する。第 9 章では、ベンチマーク調査について述べる。第 10 章では、特にオープンソースの **Apache Hadoop** の実装について、**MapReduce** プログラミングの経験の詳細を示す。第 11 章では、**Magellan** 上のいくつかの主要なアプリケーションのケーススタディを示し、クラウド環境を使用する際の課題をまとめる。第 12 章では、コスト分析を示す。第 13 章では、結論を示す。

第2章

背景

「クラウドコンピューティング」という用語は、ある範囲の提供およびサービスモデルを対象としている。これらのサービスモデルに共通する特徴は、従量料金制と弾力性、および使用するサービスを必要に応じて素早く拡大・縮小できることである。したがって、クラウドコンピューティングの成長につれて、分散コンピューティングおよびデータ分析の新しいアプローチも登場してきた。これには、MapReduce や拡張可能な鍵値ストア (Big Table [11]) などのモデルが含まれている。

クラウドコンピューティングの技術とサービスモデルは、既存のシステムを置換または補完するためにリソースにオンデマンドでアクセスできる上、ソフトウェア環境を制御できるため、科学計算ユーザにとって魅力的である。科学計算ユーザおよび科学計算ユーザにサービスを提供するリソースプロバイダは、これらの新しいモデルと技術の影響を検討している。この章では、クラウドサービスのモデルと技術について簡単に説明し、議論のための基盤を提供する。

2.1 サービスモデル

クラウド製品は通常、IaaS (サービスとしてのインフラストラクチャ)、PaaS (サービスとしてのプラットフォーム)、および SaaS (サービスとしてのソフトウェア) に分類される。各モデルは、科学計算においてそれぞれの役割を担っている。

サービスモデルの区別は、サービスがエンドユーザに抽象化される階層に基づいている (たとえば、ハードウェア、システムソフトウェアなど)。そして、エンドユーザは、抽象化されたレベルの上位のソフトウェアスタックを完全に制御する。つまり、IaaS の場合、仮想マシンまたはハードウェアがエンドユーザに提供され、ユーザはオペレーティングシステムとソフトウェアスタック全体を制御する。以下、各サービスモデルについて説明し、それぞれの特徴を理解するために、商用クラウド空間の既存の例を示す。

2.1.1 サービスとしてのインフラストラクチャ

サービスとしてのインフラストラクチャ提供モデルでは、組織はストレージ、ハードウェア、サーバー、およびネットワークコンポーネントを含む機器をアウトソーシングする。サービスプロバイダが機器を保有し、収納、稼働、および保守に責任を負う。商業分野では、顧客は通常、機器の使用料を従量制で支払う。

Amazon Web Services は現在最も広く使用されている IaaS クラウドコンピューティングプラットフォームである。Amazon は料金に応じて、何通りもの異なる計算パワーを提供する。Amazon EC2 におけるデータストレージのプライマリメソッドは S3 と EBS (Elastic Block Storage) である。S3 は拡張性が高いキーベースのストレージシステムで、フォールトトレラントとデータ完全性をトランスペアレントに処理する。EBS は弾力的な計算インスタンスに関連付けることのできる仮想ストレージデバイスを提供する。S3 の課金

は、月当たりの使用スペース、データ転送量、およびメタデータ処理回数（1000 回単位）に従って行われる。EBS の課金は月当たりの格納データに対して行われる。S3 と EBS のいずれも、同じ領域（たとえば、米国または欧州）内の EC2 とのデータ転送には課金されない。

Eucalyptus、OpenStack、および Nimbus は、プライベートクラウド IaaS サービスの作成に使用できるオープンソースソフトウェアスタックである。これらのソフトウェアスタックは、Amazon の EC2 によって提供される多数のサービス（イメージ管理、持続的なブロックストレージ、仮想マシン制御など）を模したサービス群を提供する。これらのサービスのインターフェースは多くの場合、Amazon EC2 互換で、同じツールとメソッドのセットを使用できる。

Magellan では、他の相乗的な活動と組み合わせて、既存のクラウドプラットフォームを理解し、比較するための商用クラウドプラットフォームとして Amazon EC2 を使用している。また、科学のワークロード向けにクラウド環境を提供する詳細な経験を積むために、Eucalyptus と OpenStack を使用して Magellan のハードウェア上にプライベートクラウド IaaS プラットフォームを設定している。IaaS モデルの場合、ユーザは、複雑なソフトウェアスタックを持つ科学者に便利である自分のソフトウェアスタックを制御できる。

2.1.2 サービスとしてのプラットフォーム

PaaS（サービスとしてのプラットフォーム）は、アプリケーションの構築と提供の完全なライフサイクルをサポートするサービスとしての計算プラットフォームを提供する。PaaS は、多くの場合、アプリケーションの設計、開発、導入、テストのための機能、およびセキュリティ、拡張性、ストレージ、状態などを管理するためのインターフェースを含んでいる。Windows Azure、Hadoop、および Google App Engine は商業分野でよく使われている PaaS 製品である。

Windows Azure は Microsoft のクラウドサービスオペレーティングシステム製品である。Azure は開発、サービスホスティング、およびサービス管理環境を提供する。Windows Azure は、アプリケーションをホスティングしてコストをスケールリングするために、オンデマンドの計算およびストレージリソースを提供する。Windows Azure プラットフォームは、Web ロールインスタンスと Worker ロールインスタンスの 2 つのプライマリ仮想マシンインスタンスタイプをサポートしている。また、仮想マシンインスタンスからデータを格納し、それにアクセスするための簡単な手段としてブロッブも提供している。キューは、Worker ロールインスタンスに、Web ロールインスタンスからの作業割り当て（work quantum）にアクセスする手段を提供する。Azure プラットフォームは主に Web アプリケーション向けに設計されているが、科学アプリケーションに使用することも試されている [58、69]。

Hadoop は、MapReduce [13] モデルにより大規模データセットの分散処理にコモディティクラスタを利用する機能を提供するオープンソースソフトウェアである。Hadoop ストリーミングモデルにより、任意の実行ファイルまたはスクリプトをマッパーまたはリデューサとして、マップ/リデューサジョブを作成できる。これは、何年も前からコードのある複雑な科学プロセスをキャプチャする科学アプリケーションに最適のモデルである。

HDFS（Hadoop ファイルシステム）は、Hadoop で使用されているプライマリストレージモデルである。HDFS は Google ファイルシステムを模したモデルとなっており、Hadoop/MapReduce に特に適したいくつかの機能を備えている。これらの機能にはデータ局所性やデータレプリケーションのエクスポーズなどがある。Hadoop は計算をデータの近くで行おうとするので、データ局所性は Hadoop が良好なスケールリングおよび性能を達成する上で重要な機構である。これは最も I/O 集約的となることが多いマッピングフェーズで特に当てはまる。

Hadoop は粗結合データ集約型アプリケーションを管理するプラットフォームを提供する。Magellan の相乗的活動では、Yahoo! M45 Hadoop プラットフォームを使用して BLAST をベンチマークした。さらに、

Hadoop は、ユーザがそのプラットフォームを体験できるように Magellan ハードウェア上に展開されている。PaaS はユーザに、アプリケーションで拡張性、フォールトトレランスなどを扱うための構成要素とセマンティクスを提供する。

2.1.3 サービスとしてのソフトウェア

サービスとしてのソフトウェアは、特定の機能を持つアプリケーションまたはソフトウェアへのアクセスをエンドユーザに提供する。商業分野における例としては、SalesForce や Gmail などのサービスがある。Magellan プロジェクトの活動では、Windows Azure BLAST サービスを使用して、Windows Azure プラットフォーム上で BLAST ジョブを実行している。科学ポータルは、リモートユーザが Web インタフェースを通じてデータセットを分析および閲覧できるようにしていることが多いので、サービスとしてのソフトウェアを提供しているとみなすこともできる。このモデルは、ユーザがアプリケーションの導入、構成、および保守の責任を移転でき、背後にあるソフトウェアの複雑さがエンドユーザから隠されるので、魅力的な場合がある。

2.1.4 サービスとしてのハードウェア

HaaS (サービスとしてのハードウェア) は、「ベアメタルプロビジョニング」としても知られる。このモデルと IaaS との主な違いは、ユーザが提供するオペレーティングシステムソフトウェアスタックが、生のハードウェア上に置かれ、ユーザが独自のカスタムハイパーバイザを提供したり、InfiniBand などの高性能ハードウェアの仮想化の性能への影響とともに、仮想化を完全に避けたりすることができることである。HaaS と他のサービスモデルとのもう 1 つの違いは、ユーザがリソース全体を「借り切る」ことであり、HaaS は同じ仮想空間で他のユーザと共用されないことである。HaaS の場合、サービスプロバイダが機器を所有し、その収容、稼働、および保守の責任を負う。HaaS は IaaS の利点の多くを提供し、ハードウェア構成の制御レベルを高めることができる。

2.2 展開モデル

NIST の定義によれば、クラウドは、クラウドインフラストラクチャの運用形態に応じて、(a) パブリック、(b) プライベート、(c) コミュニティ、(d) ハイブリッドのいずれかの展開モデルを持つことができる。

パブリッククラウド。パブリッククラウドは、クラウドサービスを販売する大規模な産業によって一般社会に提供されるインフラストラクチャを指す。Amazon のクラウド製品は、この分類に該当する。これらのサービスは従量制で、通常はクレジットカードを使用して購入することができる。

プライベートクラウド。プライベートクラウドインフラストラクチャは、特定の組織のためだけに運用され、特定のポリシーをサポートする固有の機能を持つ。ユーザに仮想マシンを提供するために、Eucalyptus、OpenStack、Nimbus などのクラウドソフトウェアスタックが使用される。この観点からすると、Magellan は、DOE SC のユーザにサービスを提供するプライベートクラウドと考えられる。

コミュニティクラウド。コミュニティクラウドインフラストラクチャは、複数の組織によって共有され、共通の目的を持つ固有のコミュニティのニーズに応じて提供される。FutureGrid [32] は、コミュニティクラウドと考えられる。

ハイブリッドクラウド。ハイブリッドクラウドは、独立して運用されるが、アプリケーションのポータビリティを実現する技術整合性によって結び付けられている 2 つ以上のクラウドインフラストラクチャを指す。

2.3 その他の関連作業

Magellan プロジェクトでは、現行のプライベートクラウドソフトウェアの評価、不足と限界の理解、アプリケーションソフトウェアのセットアップなどを含む、さまざまなトピックを検討した。筆者らの知る限り、科学アプリケーション向けのクラウドコンピューティングのさまざまな側面に関するこのような包括的研究はこれまでなされていない。

FutureGrid プロジェクト [32] は、地理的に分散している異種コンピューティングシステムなどのテストベッドを提供しており、これにはクラウドリソースが含まれている。同プロジェクトの目的は、研究者が計算機科学の複雑な研究課題に取り組むことを可能にする機能を提供することにある。**Magellan** の場合は、科学のニーズに対応することに重点を置いている。

さまざまなグループが、**Amazon** クラウドで科学アプリケーションを実行する実現可能性とベンチマーク研究を行ってきた [67、40、15、52、53、51]。標準ベンチマークも **Amazon EC2** 上で評価されている [62、23、66、74、82]。これらの研究は、ハイエンドの密結合アプリケーションが現行のクラウド環境の性能特性によって影響を受けることを示す筆者らの実験を補完するものである。

第 4 章

アプリケーションの特性

クラウドコンピューティングを科学に利用する上での課題は、科学コミュニティのニーズが通常の企業顧客とは大きく異なる場合があることである。アプリケーションは、大半の企業アプリケーションが必要とするよりも大きな規模で、密結合で実行されることが多い。このため、帯域幅とレイテンシの要件は、大半のクラウド顧客よりも厳しくなることが多い。また、科学アプリケーションは、大量のレガシーデータを含む大量のデータへのアクセスを必要とするのが普通である。このため、開始コストとデータストレージコストが嵩むことになる。Magellan の目標は、先進的なテストベッド、柔軟なソフトウェアスタックを使用し、さまざまなデータ収集作業を実施して、こうした要件を分析することであった。この章では、科学アプリケーション向けのクラウドの実現可能性を理解するために必要な、いくつかの重要なアプリケーションの特性を要約する。セクション 4.1 では計算モデルをまとめ、セクション 4.2 ではさまざまな使用シナリオについて論じる。セクション 4.3 ではユーザ調査の結果をまとめ、セクション 4.4 ではいくつかの科学向けの使用事例について検討する。

4.1 計算モデル

科学のワークロードは、そのリソース要件に基づき、大きく 3 つのカテゴリに分類される。それは大規模密結合計算、中規模計算、および高スループット計算である。このセクションでは、リソース要件に基づいて科学分野におけるワークロードを大まかに分類し、このようなアプリケーション分野でクラウドコンピューティングが魅力的である理由について詳しく検討する。

大規模密結合。これらは通常、国内の大規模スーパーコンピューティングセンターで実行される複雑な科学用コードである。通常、多数のプロセッサ（しばしば数千個単位）を使用する MPI コードである。規模により、単一のジョブに数千～数百万コア時間が費やされる。これらのジョブは通常、バッチキューシステムを通じてスーパーコンピューティングセンターで実行される。ユーザは要求したリソースへアクセスするために管理されたキューの中で待ち、ジョブは、必要なリソースが空いて、優先順位リストの上位に他のジョブがなくなったときに実行される。大半のスーパーコンピューティングセンターは、このようなアプリケーションのストレージと I/O のニーズに対して、アーカイブストレージと並列ファイルシステムアクセスを提供している。このクラスのアプリケーションを仮想クラウド環境で実行する場合、性能が大きく低下することが予想される [46]。

中規模密結合。これらのアプリケーションは、大規模ジョブよりも小さな規模で実行される。これらは数十～数百のプロセッサを必要とする多数のコードである。これらのアプリケーションの一部は、スーパーコンピューティングセンターで実行され、キューを埋めている。より一般的には、これらのニーズを満たすためにユーザは科学グループ自身で管理している小規模な計算クラスターに依存している。これらの中規模アプリケーションは、いくらかの性能低下が生じるとしても、クラウドコンピューティング向けの有望な候補である。

高スループット。科学的探究の一部は、デスクトップまたはローカルクラスタで行われ、非同期の独立した計算を実行する。大規模な科学的問題の場合でも、可視化など、多数のデータの前処理と後処理の段階は多くの場合、研究者のデスクトップ上で実行されている。低コストセンサーやその他の技術によるデジタルデータの規模の拡大により、こうしたアプリケーションの規模の拡大が必要になっている [58]。これらのアプリケーションはしばしば、スーパーコンピューティングセンターで使用されているスケジューリングポリシーにより不利益を被っている。こうしたアプリケーションの要件は、現在クラウドコンピューティング空間の大半を占めているインターネットアプリケーションの要件に類似しているが、データストレージとスループットの要件ははるかに高い。これらのワークロードは、MapReduce プログラミングモデルから、このクラスのアプリケーションのプログラミングおよび実行の簡略化という恩恵を受ける場合もある。

4.2 使用シナリオ

科学アプリケーションおよびユーザのニーズを理解し、既存のクラウドコンピューティングプラットフォームおよびソリューションの観点からその要件を分析することが重要である。筆者らは、従来からの DOE HPC センターのユーザの他に、DOE HPC センターにおけるクラウドコンピューティングリソースの恩恵を受ける可能性のある 3つのカテゴリの科学コミュニティユーザを識別した。

4.2.1 オンデマンドのカスタマイズされた環境

商用クラウドコンピューティングによって一般に提供される IaaS（サービスとしてのインフラストラクチャ）施設は、大規模グリッドおよび HPC システムの深刻な短所に対応する。つまり、アプリケーションのポータビリティはあまりない。この問題は、グリッドシステムの大きな課題の 1つと考えられている。というのも、地理的に、そして組織の境界を越えて分散しているシステム間でソフトウェアスタックを展開し、維持するためには、多大な労力が必要だからである [30]。こうした統一的ソフトウェアスタックの主要な設計目標は、最も広い範囲のアプリケーションに、最高のソフトウェアを提供することである。残念ながら、科学アプリケーションでは、特定のバージョンのインフラストラクチャライブラリが必要である場合が多い。そのようなライブラリが使用できない場合、アプリケーションは機能が大きく低下するか、まったく使用できないことがある。たとえば、Supernova Factory プロジェクトでは、宇宙およびダークエネルギーの膨張を測定するためのツールを構築している。このプロジェクトでは多数のカスタムモジュールを使用している [1]。そのパイプラインは複雑なため、特定のライブラリおよび OS のバージョンを使用することが重要であり、衝突や非互換性のため、多数の大規模リソースを利用することが困難になっている。この問題に対処するには、アプリケーショングループが提供する特定のアプリケーションのために調整された、ユーザカスタマイズのオペレーティングシステムイメージが役立つ。

4.2.2 仮想クラスタ

一部の科学ユーザは、いくつもの理由から、自分自身のプライベートクラスタを運用することを好む。このようなユーザはスーパーコンピューティングセンターで達成できるレベルの並行処理は必要としないが、特定の時間にリソースにアクセスできる保証を必要とすることが多い。また、ユーザごとにソフトウェア環境を設定すると、手間がかかり、時間の無駄になるため、共同研究者との共用環境を必要とすることも多い。クラウドはこのようなニーズを満たすことのできるプラットフォームである。

4.2.3 科学ゲートウェイ

明確に定義された計算ワークフローのユーザは、アプリケーションワークフローおよびデータアーカイブに対する単純な Web ベースのインタフェースを好むことが多い。Web インタフェースであれば、専門家で

なくてもリソースに容易にアクセスでき、科学データを一般分野のユーザコミュニティに幅広く提供できる（たとえば、仮想組織）。クラウドコンピューティングは、このような使用シナリオを実現できるようにする多数の技術を提供する。

4.3 Magellan ユーザ調査

クラウドコンピューティングにより、新しい使用またはビジネスモデルがもたらされ、従来は科学アプリケーションには広く適用されていなかった新たな技術が導入された。Web 2.0 アプリケーション向けのクラウドコンピューティングビジネスモデルを成功させるための仮想化技術を使用して、科学ユーザが管理および制御できる私有の仮想クラスタを構成できる。さらに、クラウドコンピューティングにより、サイトのリソース管理、プログラミングモデル、およびユーザレベルのツールに、非常に多くの新しい技術が導入されている。

筆者らは、ユーザコミュニティの要件と期待を把握するために調査を実施した。NERSC ユーザおよびクラウドコンピュータに関心を持っている DOE のその他のコミュニティに、それぞれのアプリケーションの特性とクラウドコンピューティングに対する期待について詳しい説明を求めた。調査は NERSC Magellan Web サイトを通じて実施された。

表 4.1 に、調査回答者の所属する DOE 部局を示す。調査フォームは付録に掲載する。以下にその詳細と調査結果をまとめる。

表 4.1 : DOE 部局ごとの調査回答者の比率

ASCR (先端科学コンピュータ研究)	17%
BER (生物環境科学)	9%
BES (基礎エネルギー科学)	10%
FES (核融合科学)	10%
HEP (高エネルギー物理学)	20%
NP (核物理学)	13%
ANI (Advanced Networking Initiative) プロジェクト	3%
その他	14%

4.3.1 クラウドコンピューティングの魅力的な機能

最も魅力を感じるクラウドコンピューティングの機能をユーザに尋ねた。ユーザには複数の選択肢を提示し、同一カテゴリにおける複数の選択を可能とした。図 4.1 に、ユーザが選択した回答を示す。追加のリソースへのアクセスが調査対象ユーザにとって最も一般的な動機であり、回答者の 79% が選択した。また、多くのユーザが、ソフトウェア環境の制御 (59%)、および現行のスーパーコンピューティング組織では困難な、共同研究者とのソフトウェアや実験設定の共有 (52%) を求めている。さらに、クラウドはローカルクラスタよりも操作が容易 (52%) で、サイエンスゲートウェイのエンドユーザへのアクセス (48%) の点でも魅力を感じている。一部のユーザは、科学の課題に対する Hadoop および MapReduce プログラミングモデルの使用および Hadoop ファイルシステムの使用にも関心を示した。

さらに、回答者の 90% は、所属している科学コミュニティにクラウドコンピューティングを調査しているか、クラウドコンピューティングの調査に関心を持っている人がいると回答した。また、ユーザの大半は、他の共同研究者も自分のクラウド設定を使用できるようにすると答えた。

4.3.2 アプリケーションの特性

クラウドコンピューティングは、現在のスーパーコンピューティングセンターとは根本的に異なるリソースアクセスのモデルを提供する。クラウドコンピューティングの恩恵を受ける可能性のアプリケーションを調べ、設計についての判断を検討するために、アプリケーションの特性をより良く理解する必要がある。ユーザにアプリケーションについての多数の質問をした。クラウドコンピューティングを使用することに関心のあるユーザは、従来からのスーパーコンピューティングの並列モデルから、MapReduce や、これらのプログラミングモデルの 1 つまたは複数を使用したカスタムコードまで、さまざまなプログラミングモデルを使用していた。図 4.2b に、ユーザのアプリケーションの実行時間を示す。大半のユーザは時間単位 (48%) または分単位 (14%) であった。さらに、これらのアプリケーションの大半の占有メモリは、10 GB 未満で、必要な帯域幅は 1 つのアプリケーションでおおむね 1 Gbps であった。

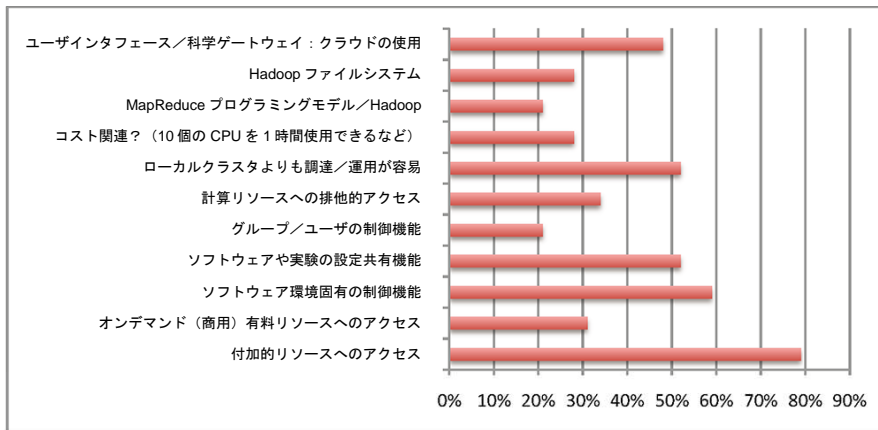
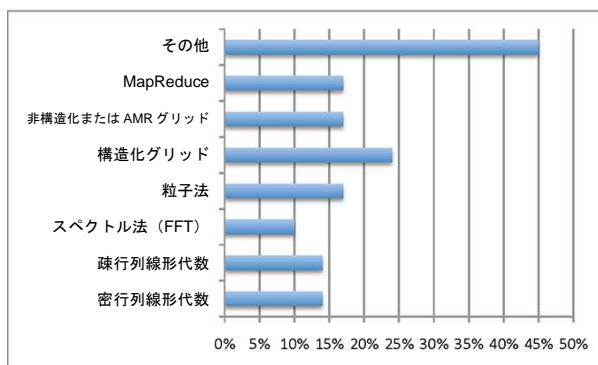
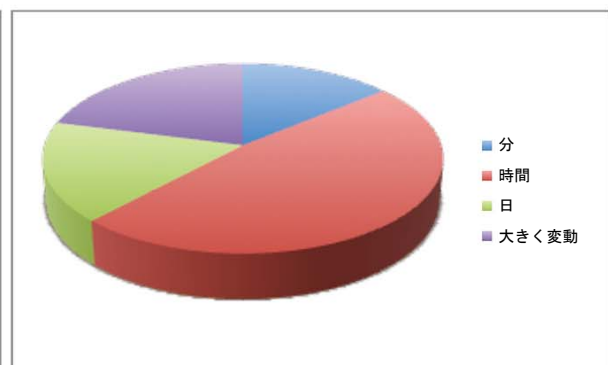


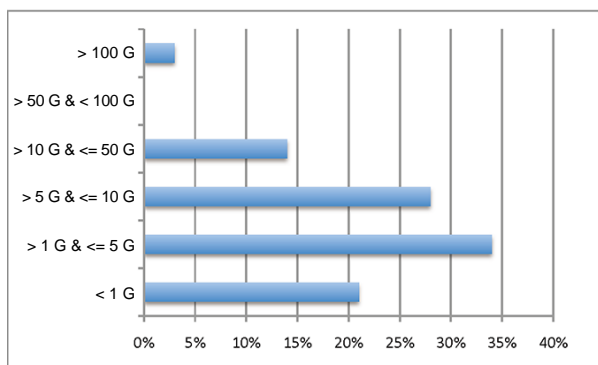
図 4.1: 科学ユーザが興味を持つクラウドコンピューティングの機能



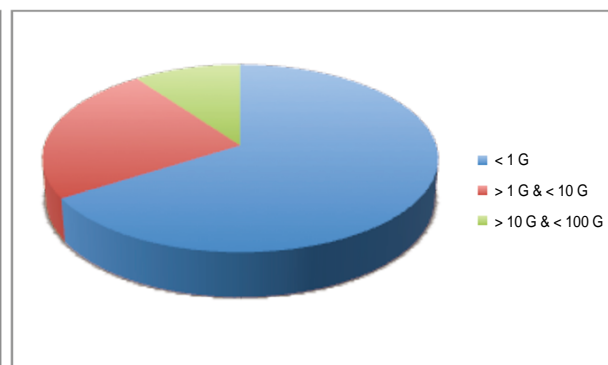
(a) プログラミングモデル



(b) コードの実行時間

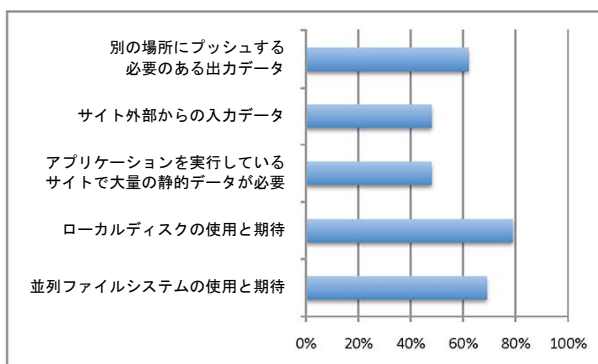


(c) メモリ要件

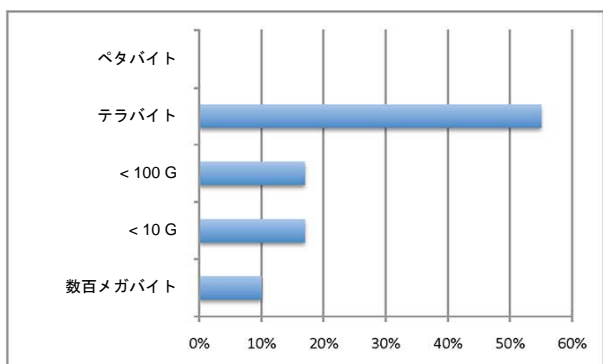


(d) 帯域幅要件

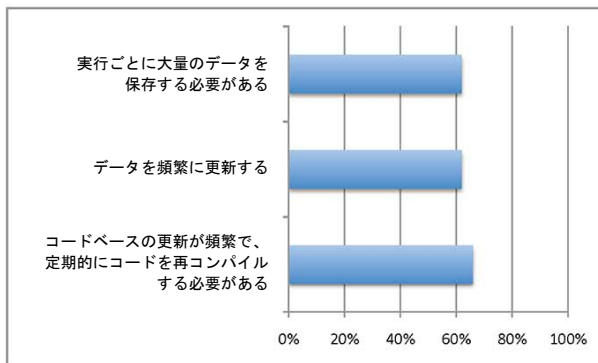
図 4.2: プラットフォームとしてのクラウドコンピューティングに関心を持つユーザのアプリケーションコードの特性の理解



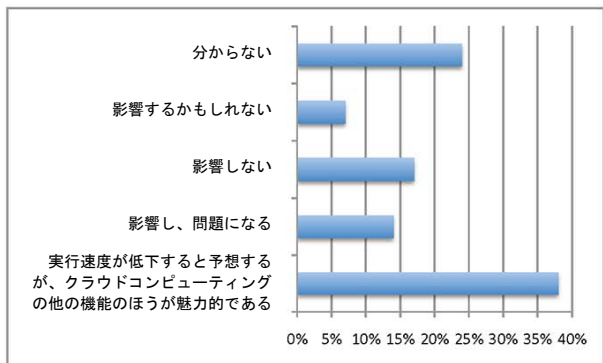
(a) アプリケーションのデータ特性



(b) 固定ディスク要件



(c) イメージ生成に影響する特性



(d) 性能へのクラウド環境の影響

図 4.3 : クラウド環境で設計の決定に影響を与える可能性のあるアプリケーションの特性の理解

同様に、クラウドにおけるデータ要件に関してユーザに調査を実施した。多数の科学アプリケーションは、並列ファイルシステムに依存しており、ノードにローカルディスクがあることを想定している。また、アプリケーションがサイトで使用可能であるとみなす大量のデータ（ギガバイトまたはテラバイトのオーダー）やリモートサイトから送られてくるデータがある。一部のユーザはアプリケーションの性能が仮想化によって悪影響を受けると予想しているが、多くのユーザがクラウドコンピューティングには他の利点があり、自らの科学分野にとって魅力的なプラットフォームだと感じている。科学アプリケーションをクラウド環境で実行する際には、別の課題にも直面することになる。コードベースはしばしば変わるので、それに応じて各コードを定期的に再コンパイルして、新しい仮想イメージを作成する必要がある。また、データが頻繁に更新されるのに加え、仮想マシンは持続した状態を維持しないので、実行ごとに大量のデータを保存する必要がある。

4.4 アプリケーションの使用事例

いくつかのアプリケーショングループについて個別に追跡調査を行った。クラウドを早期に導入したユーザとの検討と、Magellan からの要求事項をまとめる。その後、これらのユーザの多くは両方のサイトにおいて Magellan リソース上で実行している。各ユーザの経験については、第 11 章で詳述する。

4.4.1 Climate 100

気象学者は、気候変動の潜在的影響をより正確に把握し、取り得る緩和策の効果を評価するために、大量のデータの生成と共有を行う能力への依存を強めている。Climate 100 プロジェクトの目標は、ミドルウェアとネットワーク研究者を結び付け、ANI (Advanced Networking Initiative) 100 Gbps ネットワークを使用して前例のない量のデータを移動するために必要なツールと技術を開発することである。Climate 100 プロジェクトのデータは 100 万ファイル単位で構成されている。各ファイルのサイズは平均 100 MB 前後である。Climate 100 では、気象データの大規模データ解析を行うために、仮想マシンおよび Hadoop といったクラウド環境を役立てることができる。そのデータ量により、ネットワーク、計算、およびディスクリソースを調和良く組み合わせる必要がある。

4.4.2 オープン科学グリッド／STAR

STAR 核物理実験では、ブルックヘブン国立研究所の重イオン衝突型加速器で収集したデータから核物質の基本的性質を研究している。STAR 実験は、実験データの定期的な処理に使用される、多数のさまざまなリソースサイトにアクセスしている。

STAR 実験は、極めて並列度の高いアプリケーション（すなわち、非 MPI コード）であり、各ジョブは 1 つのプロセッサまたはコアで扱われる。STAR のアプリケーション群の使用事例は、さまざまな解析およびシミュレーションプログラムからなる。STAR におけるモンテカルロ事例は、データ転送および I/O の要件が非常に小さいため、クラウド環境での実行に適している。仮想マシンのソフトウェアスタック制御機能は、ソフトウェアスタックが複雑なため、コミュニティにとって非常に魅力的である。既存の仮想マシンイメージを使用するバースト時の規模拡大能力により、科学の生産性が大きく向上し、解決時間が大幅に短縮できる。グループはすでに Amazon EC2 リソースを使用して、クォーク物質物理学会に間に合うように、大量のデータ処理を行えることを実証した。STAR コミュニティのさまざまなグループが、クラウドコンピューティングと仮想マシンに興味を示している。コミュニティは、将来の実験に使用するソフトウェアのパッケージ化と配布の手段として仮想マシンイメージを使用することについての調査を継続する。

4.4.3 Supernova Factory

Supernova Factory プロジェクトは、宇宙の膨張とダークエネルギーを測定するためのツールを構築している。実験では、カスタム環境を必要とする多数のシミュレーションを行う。ここでは、他の共同研究者との最終結果の共有も必要であるため、計算およびネットワークリソースの共有およびストレージリソースへのデータの移動も必要になる。**Supernova Factory** は、超新星探索の大量のデータに依存しており、コードベースは多数のカスタムモジュールからなる。パイプラインが複雑なため、固有のライブラリと OS パージョンが必要であり、そのことが他の大規模なリソースを利用する妨げとなっている。**Supernova Factory** プロジェクトでは、ソフトウェア環境を制御でき、ユーザアカウントとグループのソフトウェアへのアクセスの管理と制御ができる点で、クラウドコンピューティングが魅力的であるとみている。**Magellan** プロジェクトの人員と協力して **Amazon EC2** 上で実施した初期の実験により、クラウドがプラットフォームとして適していることが判明した。また、疎結合ジョブの協調と管理に **Hadoop** を使用することにも関心を示している。

4.4.4 ATLAS

ATLAS プロジェクトでは、クラウドプラットフォームを使用した解析ジョブのサポートに関する調査を進めている。**ATLAS** プロジェクトには、数テラバイトのデータを運用する数百のジョブがあり、クラウドリソースへのタイムリーなアクセスが大いに役立つ。クラウド環境は確実に、科学コードをデスクトップでのテストから、大規模クラウドリソースに移行するための効果的なプラットフォームにもなる。グループは、必要なすべてのソフトウェアの配布に仮想マシンイメージを使用することについて調査している [10]。これにより、ソフトウェア管理の作業をほとんどあるいはまったく必要とせずに、別のサイトの仮想マシンをブートできるようになる。

4.4.5 IMG (Integrated Microbial Genomes) パイプライン

DOE JGI (共同ゲノム研究所) の **IMG** (Integrated Microbial Genomes、統合微生物ゲノム) パイプラインは、すべての公開された基準分離株微生物ゲノムを統合したコンテキストでの微生物群集メタゲノム解析を提供する。**IMG** には、コンテンツの保守のために、数週間~数カ月ごとに実行する必要のあるワークロードがある [8]。ワークロード完了の適時性は、コミュニティにとって重要である。また、こうしたデータセットのサイズが恐ろしい勢いで大きくなっているため、数多くのリソースへのアクセスが重要である。計算ステージは、個別の遺伝子の機能的アノテーション、ペアワイズ遺伝子の識別、および染色体クラスタの識別から構成されている。最も計算集約的な部分は、基準データベースに対する **BLAST** 解析である。以降のステップでは、**BLAST** によって報告されたアラインメントに基づいて遺伝子の特徴付ける。**BLAST** の出力のみで通常 1 テラバイトを超える。したがって、一致率の高いものを見つけて、タクソンを特定するための出力の解析には時間がかかるため、並列に処理する必要がある。**Hadoop** などのテクノロジーを使用して、これらのゆるやかに結合されたアプリケーションの実行を管理することに関心が持たれている。

現在、「全データ対全データ」のペアワイズ遺伝子シーケンス比較は、圧縮ファイルに格納されている。しかし、この形式では個々のエントリの変更およびデータの検索が容易ではない。グループとしては、データの管理に **HBase** を使用することに関心がある。これにより、ユーザは個々の行を更新し、簡単に検索できるようになる。

4.5 要約

詳細な調査結果は、科学がクラウド環境に求める要件を理解するのに役立ち、プロジェクトの研究の方向に影響を与えた。ここでは、ユーザ調査で収集された要件と対応するプロジェクト活動をまとめる。

- クラウドコンピューティングに対するユーザの要求は、カスタム環境へのアクセスから、MapReduce プログラミングモデルまで多岐にわたる。これらの多様な要求により、柔軟なソフトウェアスタックが開発された。Magellan のユーザは以下にアクセスした。(a) xCAT を通じてカスタムソフトウェア環境の追加機能を備えた従来型のバッチキューアクセス、(b) Eucalyptus または OpenStack フロントエンドを通じて、ユーザが商用プロバイダとプライベートクラウドの間での移植を可能にするカスタマイズされた仮想マシン、(c) ユーザが MapReduce プログラミングモデル、Hadoop 分散ファイルシステムおよびその他のフォールトトレランスなどジョブ管理機能にアクセスできるようにインストールされた Hadoop。ソフトウェアスタックの詳細は第 5 章で述べる。また、筆者らの Hadoop の評価を第 10 章に示し、ユーザ体験を第 11 章にまとめる。
- Amazon EC2 などの商用クラウドプラットフォームおよび Eucalyptus や Hadoop などのプライベートクラウドソフトウェアが科学のニーズに適合しているかどうかを理解することが重要である。筆者らは既存の不足を特定し、ポジションペーパー『Defining Future Platform Requirements for e-Science Clouds (e 科学クラウドのための将来のプラットフォーム要件定義)』にまとめた。
- 不足を把握するのに加え、商用クラウドプラットフォームの計算拡張性を把握するためにベンチマーク作業に取り組んだ。その結果は第 9 章に示す。
- ユーザは、クラウド環境に加えてスーパーコンピューティングセンターで提供されているその他の機能を利用できるプライベートクラウドにも関心を示していた。たとえば、HPC システムは通常、共用ファイルシステムへの広帯域大容量の並列書き込みが可能な高性能並列ファイルシステムを提供している。また、HPC センターは通常、重要な出力および結果をアーカイブするためのアーカイブストレージを提供している。
- クラウドコンピューティングは、現行のグリッドシステムにおける既知の問題である、ソフトウェアスタックの可搬性に対処する。現在のスーパーコンピューティングセンターでは、各サイトが複数のグループに必要なオペレーティングシステムと共通ミドルウェアを管理している。ユーザは、特定のシステムに自分のアプリケーションをコンパイルしてインストールする（その際、特定のハードウェアに必要な最適化のためにサイト担当者の支援を受けることが多い）。クラウドコンピューティングモデルに移行する際に、各サイトは特定のグループによって必要とされるカーネルおよびオペレーティングシステムの多様なセットを管理するためのツールとサポートを提供できる必要がある。ソフトウェアのアップグレードとオペレーティングシステムのパッチ適用に関しては、もはや明確な責任の分離は存在しない。サイトはユーザによってサポートされるイメージをサポートすることと、サイトのセキュリティポリシーとのギャップを埋める仕組みを必要とする。ユーザが自分のイメージを完全に制御するクラウドシステムは、サイトのセキュリティポリシーに一定の関連を持つ。筆者らのセキュリティへの取り組みは、これらのカスタム環境を指針としており、第 8 章で論じる。