

The Cost of Doing Science on the Cloud: The Montage Example

Ewa Deelman, Gurmeet Singh,
Miron Livny, Bruce Berriman, John Good
[SC '08]

数理・計算科学専攻 松岡研究室 島田 大地

これまでの計算資源

- 自分の所有するクラスタを使う
 - 自分の資源を他の機関と共同で使う
 - 例) OSG
 - 国際的なネットワークインフラ資源を使う
 - 例) TeraGrid, EGEE

 - しかし、これらの解決法は
 - 異なるレベルの支払い義務が必要
 - 異なるレベルのサービスとなる
-

クラウドコンピューティング

- 新しい次元のコンピューティングランドスケープ
 - 規模の経済に基づいている
 - ウェブやネットワーク技術の進化
 - 提供者は資源の利用に対して課金する
 - 仮想化技術を使用している
 - アプリケーションをセットアップし、与えられたアプリケーションに適した仮想環境を配置することができる
-

クラウドの利点

- ローカルコンピューティングインフラだと必要なもの
 - 購入
 - 管理
 - 維持
 - 定期的なアップグレード
 - クラウドベースの外部委託の方が潜在的にコストが低い
 - 科学アプリケーションに魅力的
-

クラウド使用についての疑問

- コストと性能のバランスを保ちながらどのようにアプリケーションを使うことができるのか？
 - 天文学アプリケーションMontageをクラウド上で次の条件で実行したときのコストを調べる
 - mosaic要求のばらばらな負荷を扱う
 - 全ての計算のための資源提供
 - 計算と長期間のデータストレージ提供
 - 空全体の計算を行った場合、どのくらいの金額がかかるのか？
-

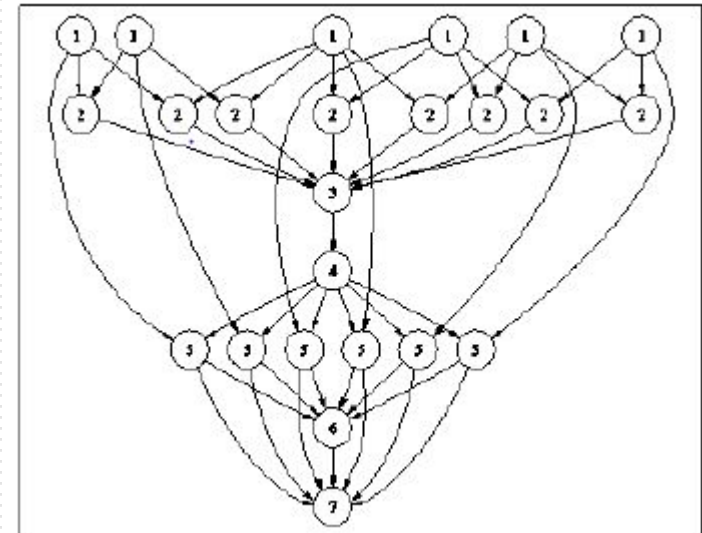
Montageとは

- 入力画像のつなぎ合わせた写真(モザイク)を計算する
 - 入力画像は画像アーカイブから取ってくる
 - 例) Two Micron All Sky(2MASS)

 - 出力モザイクの座標空間に入力画像を投影し、調整、つなぎ合わせる
-

Montage – ワークフロー

- 頂点はワークフローのタスク
 - 数字はタスクのレベル
 - 他のタスクに依存しない
タスクはレベル1
 - 他のタスクは最大親レベル+1
 - 特定のレベルは異なる
入力データごとに同じことを行っている
- 辺はタスク間のデータ依存



Montage – 特徴

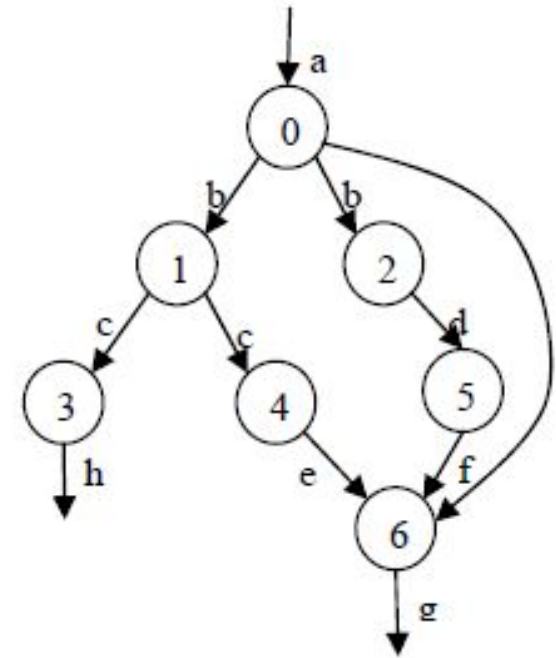
- Montageはデータインテンシブなアプリケーション
 - 入力データ、中間データ、出力データがある
 - かなりのストレージが必要
 - タスクは多くても2~3分の小さな実行時間
 - ストレージ資源が豊富な環境で実行することが望ましい
-

クラウドモデル

- アマゾンサービスを基本モデルとして選択
 - アマゾンは利用回数料金制で計算資源とストレージ資源を提供する
 - 加えて、ストレージへのデータ転送と、ストレージからのデータ転送にも課金する
 - 論文執筆時のレート
 - ストレージ: 0.15ドル/GB-Month
 - ストレージへのデータ転送: 0.1ドル/GB
 - ストレージからのデータ転送: 0.16ドル/GB
 - 計算資源の利用: 0.1ドル/CPU時間
-

提案手法

- クラウドストレージ資源を使用する3つの異なるプランに対応するモデルを実装する
 - 右図を用いて説明
 - 7つのタスクがある
 - それぞれのタスクは1つの入力ファイルを取り、1つの出力ファイルを作成する(6を除く)



Remote I/O

- それぞれのタスクで次を行う
 - 資源へのデータ入力
 - タスクの実行
 - 資源からのデータ出力
 - 資源からの入力と出力ファイルの削除
 - タスクによって使われる計算資源が共有ストレージを持っていないときに使われる
 - タスクがローカルデータアクセスする代わりに Remote I/Oをしていることと等しい
-

Regular

- ワークフロー内で作成されたデータをストアする
 - ワークフロー内のタスクによって使われる
 - 例えば、
 - タスク0が実行を終え、ファイルbを作成する
 - 後でタスク1と2によって使われるためにストレージシステム上に残ることを認める
 - 全てのタスクが実行を終えるまでどのファイルも削除されない
-

Dynamic cleanup

- Regularではデータが長生すると、ストレージ資源を占めるかもしれない
 - 例えば、
 - ファイルaはタスク0の計算後必要とされないが、残ったままになる
 - このモードではストレージから必要でないファイルを削除する
 - ワークフローのレベルでデータ使用の解析をするので、ファイルaはタスク0の後に削除されるが、ファイルbは削除されない
 - ストレージ使用を減らすので節約になる
-

疑問

□ 疑問①

- (金額)コストを最小限にしながら、性能を最適化するためにどのくらいのプロセッサを提供するればいいのか？

□ 疑問②a

- ローカルデータアーカイブを持っていて、計算を任せようとしていたときにかかるコストはいくらか
-

疑問

□ 疑問②b

- クラウドに対して計算とストレージ両方を任せているとき、少ない要求に対して、クラウド上にデータをストレージすることは安くなるのか？

□ 疑問③

- 空全体のモザイクを作成するコストはいくらか？
 - 2Massのデータで4度正方形が3900枚
 - 一度モザイクを計算して、クラウドにストアする
 - 毎回計算するよりもどのくらい効果的か？
-

調査方法

- 前の疑問を解決するためにシミュレーションを行う
 - 人工的にデータセットの大きさを変更し、計算操作割合などの実行コストを見積もることができる
 - 実際の場合は困難
 - アマゾンに料金を払わず、時間的コストを被ることなく性能とコストのトレードオフを調査できる
-

実験対象

- シミュレーションでは3つのMontageを使う
 - M17星雲の1度正方形モザイク作成
 - M17星雲の2度正方形モザイク作成
 - M17星雲の4度正方形モザイク作成

 - これらのワークフローはMontage割り当てのmDAG成分を用いることで作られる

 - ワークフローの解析をしてシミュレータへの入力としてグラフの隣接関係リスト表現を作成するためにプログラムを作成した
-

実験環境

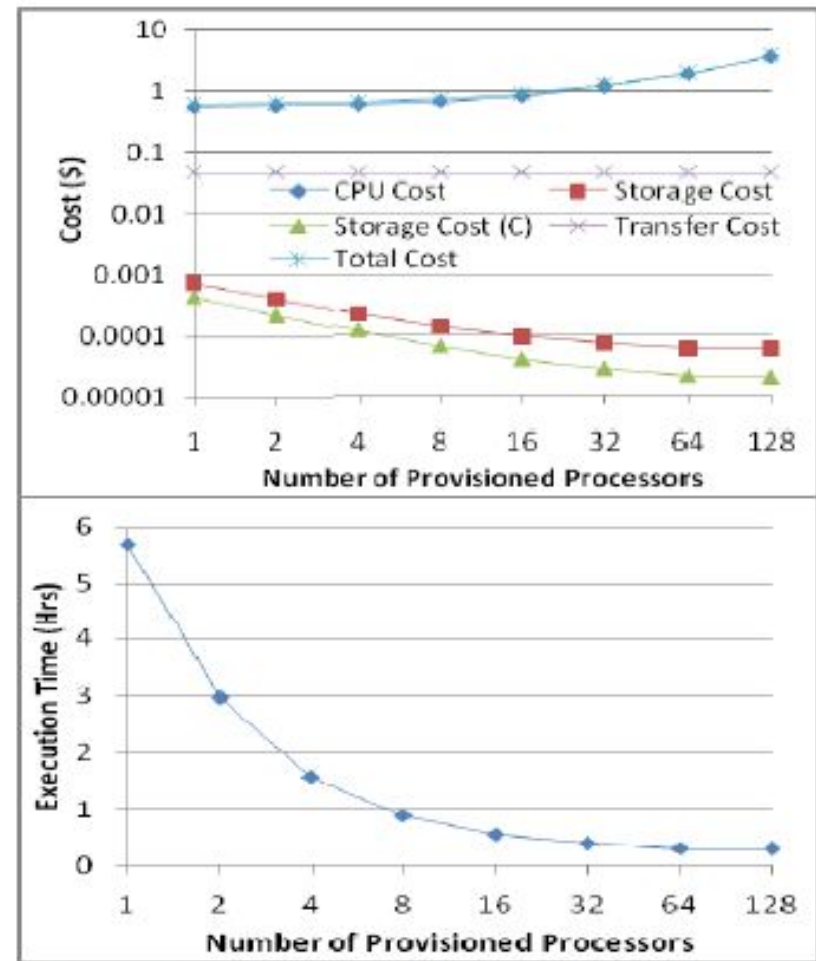
- シミュレートされるワークフローの最大並列以上のプロセッサを持つ単一計算資源をシミュレートする
 - 計算資源は無限の容量のストレージシステムを持っている
 - ユーザとストレージシステム間のバンド幅は10Mbpsで固定
 - ワークフローのための全ての入力データはアプリケーションと同じように配置される
 - 最後に結果として出てくるモザイクは外に出されて、表示されてシミュレーションは完了する
-

実験 - 疑問①

- クラウドを最も良く使う方法を調べる
 - 供給されるプロセッサの数を1~128で変化させる
 - CPUコスト、ストレージコスト、転送コスト、合計コストを比較する
 - 仮想マシンの組み立てや取り壊しは含めない
-

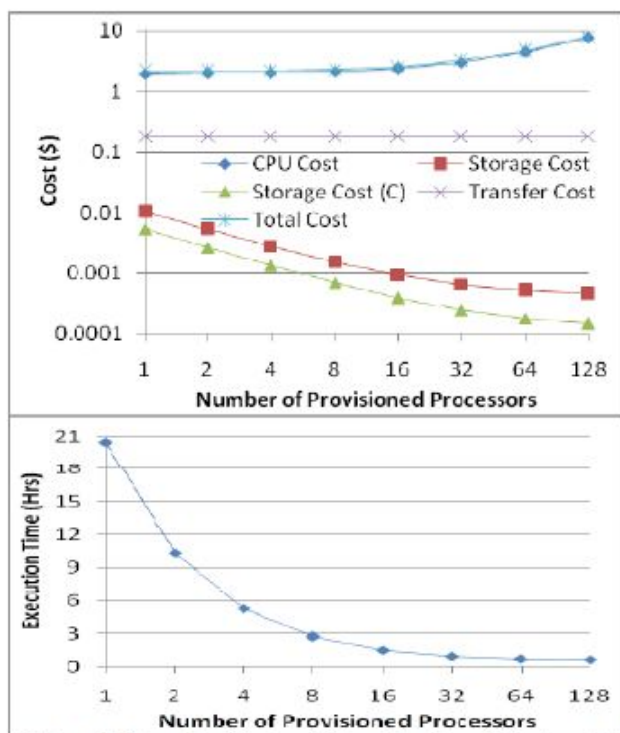
実験結果 - 疑問①

- 1次元正方形ワークフロー
 - 203タスクからなる
 - 合計コストのうちCPUコストが支配的
 - データ転送コストはプロセッサの数によらない
 - プロセッサの数が増えるとストレージコストは減るが、CPUコストは増える
 - プロセッサの数が少なくなると実行時間が増える

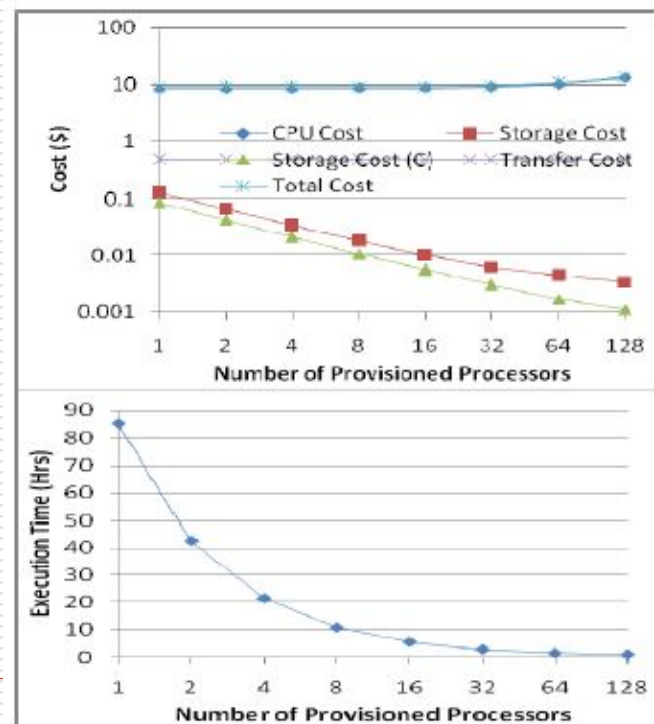


実験結果 - 疑問①

- 2度正方形ワークフロー: 731タスクからなる
- 4度正方形ワークフロー: 3027タスクからなる
- 1次元とほぼ同様の形



2次元



4次元

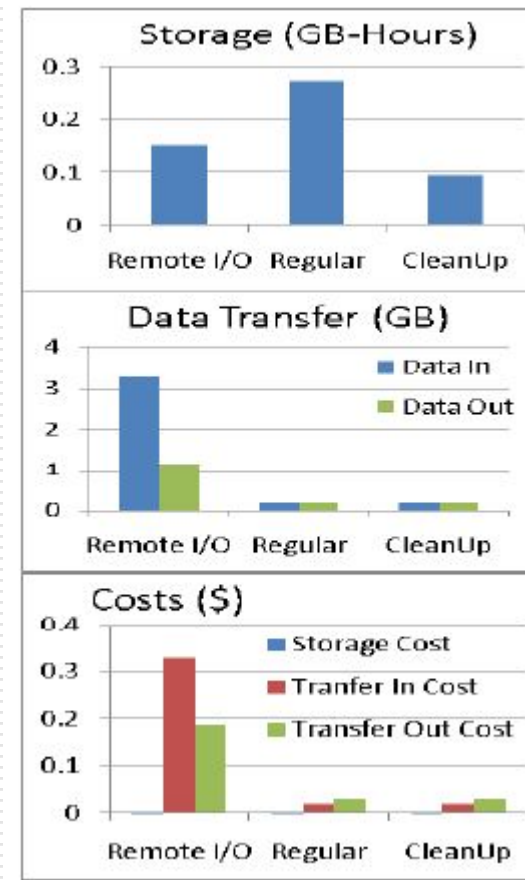
実験 – 疑問②a

- 科学計算のユーザ要求のコストの問題を調べる
 - 多くの資源をクラウドから供給している
 - できるだけ多くの資源を必要な限り使う要求を認める
 - アプリケーションは提供された資源へのユーザ要求をスケジューリングする責任がある
 - プロセッサ時間は必要とされるだけしか使われない
 - データ転送とストレージコストは全体の要求コストのより重要な役割となるかもしれないと予想する
 - 3つの異なるデータ管理解決を使うことでのトレードオフを調べる
-

実験結果 - 疑問②a

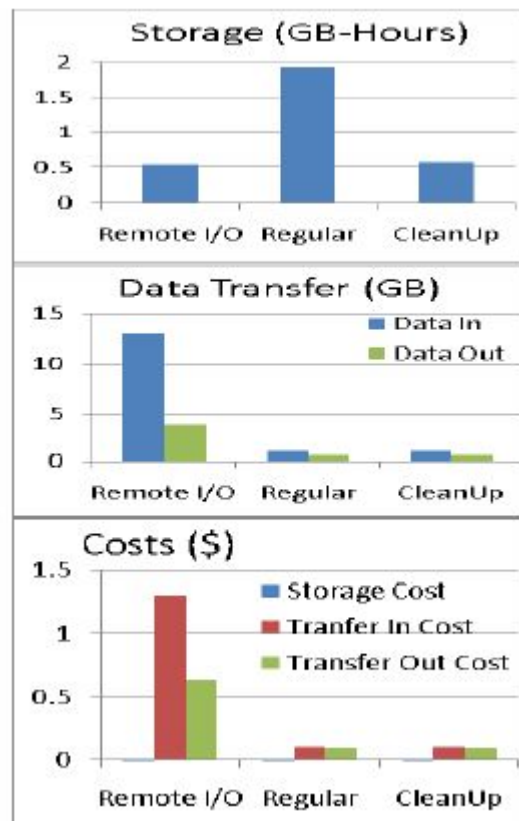
□ 1次元正方形ワークフロー

- 一番上はワークフローで使われたストレージの合計
- 中央はデータ転送の合計
- 一番下はワークフローの実行と関連したコストと合計コスト

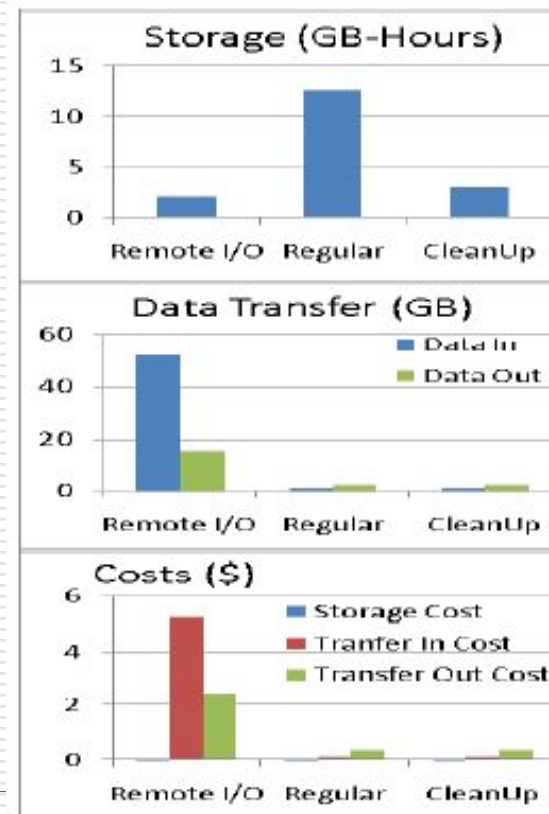


実験結果 - 疑問②a

□ 1次元とほぼ同様の形



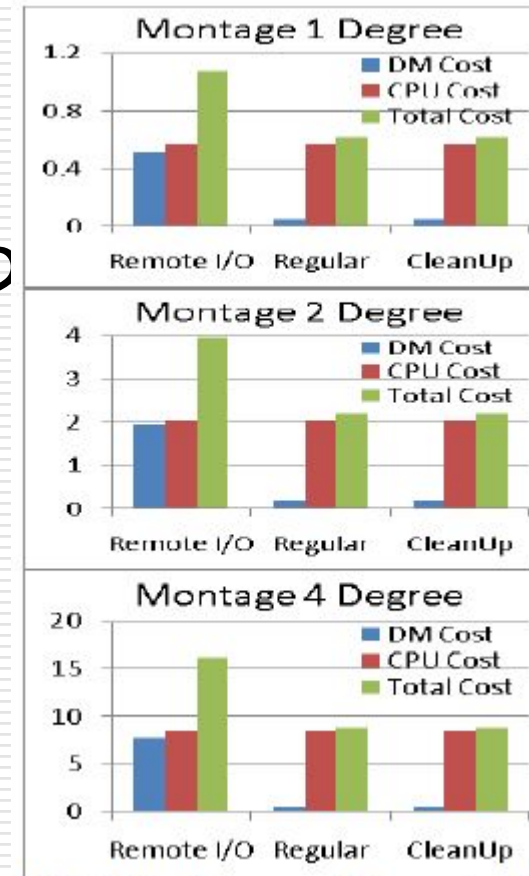
2次元



4次元

CPUコストの比較

- CPUコストはRemote I/Oのデータ管理コストより少し高い
- CPUコストはどの実行でも変わらない



実験結果 – 疑問②b

- クラウド上で計算をし、巨大なデータ集合をストアすることによる利益
 - 全データ集合のサイズは12テラバイト
 - クラウドに全てストアすると、データアクセスの遅延が低くなるが、アクセスのコストが0になる
 - しかし、データストアのコストは1ヶ月1800ドルになる
 - ストレージコストに勝つためには一ヶ月に18000回 mosaic を要求する必要がある
-

実験結果 - 疑問③

- 空全体のモザイクを作成するコストを調べる
 - 3900枚の4度四正方形か、1736枚の6度正方形モザイクになる
 - レギュラーモードで4度正方形は8.88ドルでできた
 - したがって、合計コストは34632ドルと考えられる

 - 入力データがすでにクラウドにアーカイブされていると仮定する
 - 4度は8.75ドルなので、合計コストは34145ドル
-

実験結果 – 疑問③

- 常に入力データからモザイクを生成する代わりに人気のモザイクをアーカイブするとどうなるか
 - 1度MontageのCPUコストは56セント
 - 56セントのコストは、1ヶ月1GBあたりのストレージ料金が0.15ドルとして21.52ヶ月ストアすることができる
 - 同様に2次元では24.25ヶ月、4次元では25.12ヶ月
 - 2年の間に同じ要求が繰り返されるとすると、ストアしておいたほうがいい
-

まとめ

- MontageとAmazonEC2料金体系を用い、粒状の計算のあるデータインテンシブアプリケーションに対して、ストレージコストがCPUコストと比較してわずかであったことを示した
 - したがって、クラウドコンピューティングはデータインテンシブアプリケーションに対してコストエフェクティブ解決を提供している
 - 科学計算にクラウドコンピューティングを用いた一つの側面を調査した
 - 異なるワークフロー実行モードとクラウド資源に対するプラン提供のトレードオフ
-