

ネットワークとクラウドを活用してデータ収集から解析まで

- IoT、ビッグデータ、教育基盤 -



どんな研究？

- クラウド間連携により、複数のクラウドの計算資源を組み合わせた**実験再現環境**や**教育環境**の構築を容易にします
- ゲノムデータ解析の実行にかかった時間や最大メモリ使用量などの**メトリクス**を収集する仕組みを実現します
- メトリクスを活用して、計算資源の選択や、実験に必要な金銭コストの推測などを容易にします

何ができる？

- クラウド選択**の自由度が上がる
- 実行環境の再現**や**流通**が容易になる
- より適切な計算資源環境で実験を行うことが容易になる
- 複数環境での**実行結果を比較・評価**を行うことが容易になる

動機



- どのクラウド基盤を使う？
- どのインスタンスを使う？
- 資源の性能は？
- データに近い計算資源が欲しい

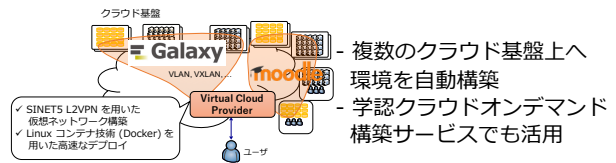


アプリケーション環境の構築は大変

適切な計算資源を使いたい

研究内容

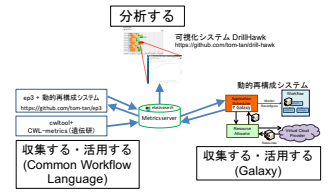
Virtual Cloud Provider (VCP)



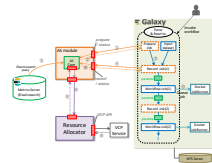
- 複数のクラウド基盤上へ環境を自動構築
- 学認クラウドオンデマンド構築サービスでも活用

- ✓ SINETS LZVPN を用いた仮想ネットワーク構築
- ✓ Linux コンテナ技術 (Docker) を用いた高速なデプロイ

メトリクス収集・活用のエコシステム



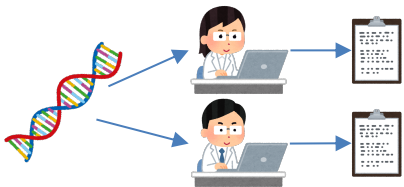
実行環境再構成技術



- 計算資源を動的に確保・削除し、アプリケーション環境を再構成
- 各ステップの実行直前に計算資源を確保して実行

今後の展望

- エコシステムを利用しやすくするためのパッケージの作成・配布



- 実験結果の再現性を確保したい
- 実行時間や金銭コストなどを比較したい

本研究は、JST、CREST の支援 (Grant 番号 JPMJCR1501) を受けたものである