## データ収集・蓄積・解析クラウド基盤 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

# 動機 ロザデリーク構築ネットワーク構築 ソフトウエア設定

アプリケーション環境の構築は大変

- 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

適切な計算資源を使いたい



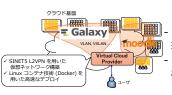
実験結果の再現性を 確保したい 実行時間や金銭コスト などを比較したい

## 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

### 研究内容

Virtual Cloud Provider (VCP)



複数のクラウド基盤トへ 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

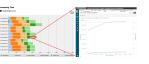
#### 実行環境再構成技術

- 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

#### メトリクス収集機構



#### 可視化システム DrillHawk



https://github.com/tom-tan/drill-hawk

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