

LC4RIによる情報システムの構築

稲葉 正樹(株式会社デンソー)
木村 佳弘(富士通株式会社)
山本 貴大(株式会社日立製作所)

開発における問題点

手順書の記載漏れ

技術・ノウハウの属人化

運用自動化による弊害

インフラ構築・運用の現場では、手順書の整備不足や属人化等の問題があった。

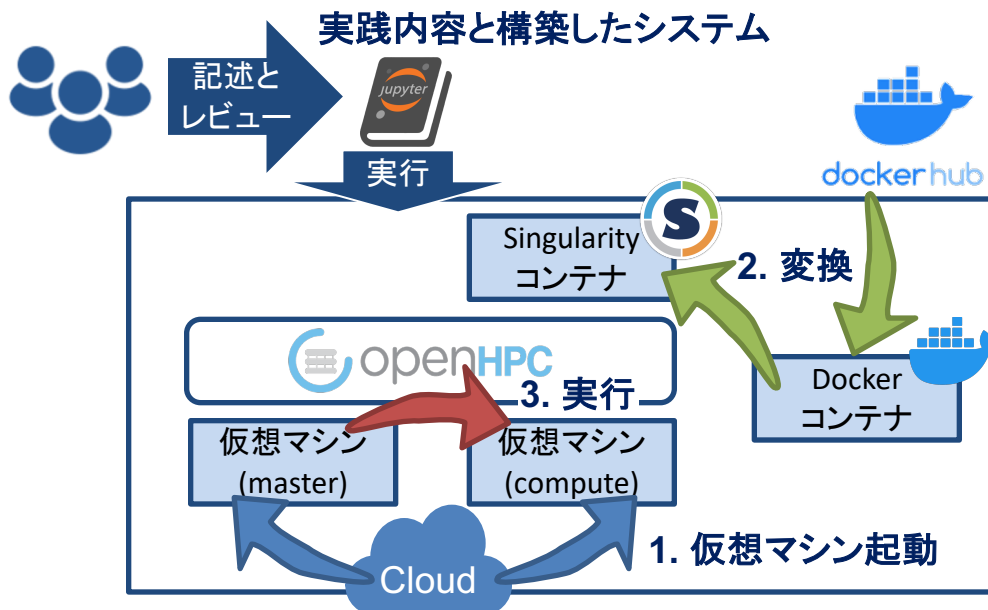
手法・ツールの適用による解決

LC4RIによるインフラ構築・運用

Jupyter Notebookによりドキュメントとコードを一元管理することで解決を目指すLC4RIが提唱されている。本実践では、GPU計算機クラスタの構築手順書をLC4RIに基づき作成することで、LC4RIの有効性を検証し、またLC4RIによる運用プラクティスの拡張の可能性について検討する。

GPU計算機クラスタ構築へのLC4RI適用

- インフラ構築・運用の方法論であるLC4RIに基づきGPU計算機クラスタの構築手順書を作成。
- 手順書作成の過程で既知プラクティスが活用されたか評価し、さらに本実践により得られた知見から新たなプラクティスとなり得る工夫・改善点を検討・抽出。



LC4RI:Literate Computing for Reproducible Infrastructure

説明文と実行コードを一元管理

Singularityイメージに変換

ビルド時にtmpの容量を使うので30GB程度余裕が必要。
[実行時間] 15分程度

tmpの空き容量を確認。

ドキュメント

実行コード

```
in [65]: ansible -i {ansible_hosts} compute -a
```

実行結果

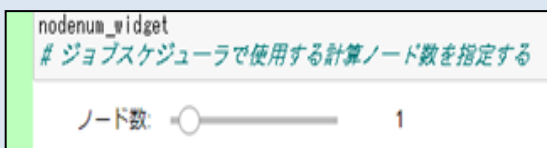
Filesystem	Size	Used	Avail	Use%
devtmpfs	30G	0	30G	0
tmpfs	30G	0	30G	0
tmpfs	30G	17M	30G	1
tmpfs	30G	0	30G	0

プラクティスの評価と工夫・改善点の考察

- 既知プラクティスの有効性を確認(12個中11個のプラクティスを活用)。また本実践を通して 5個の工夫・改善点を抽出。

1. 実行環境の変数設定

LC4RIは利用者の判断による柔軟な運用ができる反面、不明な変数があると誤判断・誤操作につながってしまう。利用者が困らないコメントを残し、かつウィジェットで強調する。



2. 作業手順の非同期化

並列に実行してよいセルは非同期実行して効率的に処理する。
MLはデータダウンロードなど他の処理に依存しないが時間がかかる処理がある

処理3で利用する旨をコメント

処理1 (非同期)

処理2

処理1の結果確認待ち

処理3

#	説明	効果
3	分割したそれぞれのNotebookに後処理を記載する	後処理忘れの防止
4	Notebookを分割した場合、概要を説明しておく	Notebook理解を助ける
5	Notebookに未対応事項があれば明記しておく	Notebook改版の方向性の支援

今後の課題

- 工夫・改善点の整理と検証