

PCクラスタワークショップin 神戸2022「クラウドとHPC」

日立製作所 研究開発グループにおけるオンプレ/クラウドのAI/HPC環境

2022年6月23日

(株)日立製作所 研究開発グループ AI研/知能情報研究部&技術戦略室 清水 正明



Contents

- 1. 日立製作所 研究開発グループにおける研究IT環境
- 2. オンプレミスITとクラウド環境の併用に向けて
- 3. AI・ML用途におけるHPCシステムへの期待と課題
- 4. まとめ

1. 日立製作所 研究開発グループにおける研究IT環境



OT × IT × Productsの総合力を発揮するために研究開発力を強化

■研究開発グループの概要

- ·研究者: 2,700人(*1)
- ・デジタルテクノロジ(AI、メディア、データサイエンス)、制御、エレクトロニクス、エネルギー、ヘルスケア、材料、機械、生産、顧客協創、基礎研究、の広範囲の研究領域をカバー
- ·研究開発費: 2,937億円(対売上比3.4%, *1)

■オンプレIT環境

- ・会社共通のOAインフラ(Office PC, ネットワーク、ファイルサービス等)
- ・研究開発グループ固有のAIクラスタ、HPCクラスタ

■各国・地域の研究拠点

・日本、米国、中国、インド、シンガポール/タイ、ヨーロッパ、オーストラリアに研究拠点

1-1. 研究開発グループ固有のAIクラスタ、HPCクラスタ



協創および最新デバイス利用のために、オンプレに加えクラウド活用加速中

- ■オンプレミスAIクラスタ
- ·AI,データサイエンス研究者向けGPUクラスタ
 - -最新GPU、FPGA、高速ネットワーク、大規模高速ファイルシステム
 - -OSS中心に構成
 - -計算資源利用率向上のためバッチ処理中心
- ■オンプレミスHPCクラスタ
- ・設計、シミュレーション研究者向けHPCクラスタ
 - -ISVソフトウェアを多用
 - -CPU only、高速ネットワーク、高速ファイルシステム、大規模テープアーカイブ
 - →顧客やパートナーとの協創、コンペ参加、最新GPU利用などでクラウド活用加速中

2. オンプレミスITとクラウド環境の併用に向けて



社外に出せるデータの処理、学会コンペは積極的にクラウドへ移行

- ■AI学習はABCI/AWSを活用
 - ・2018年度より産総研ABCI活用中
 - -2019年度~21年度で、多数の国際会議での採録、国際コンペでの入賞を達成
 - -深層学習型ロボット制御技術の開発に活用
- HPCはHPCクラウド/AWSを活用
 - ・2020年度より試行中
 - -オンプレのISVライセンスと共用できないなど課題もあり(契約による)
- →AWS上の環境は、AI用HPC用を併用
 - →本日ご紹介する、AI/HPConAWS

2-1. オンプレミスITと親和性の高いクラウド環境の構築



オンプレミスとの親和性 & クラウドの利点の両立

Issues

■オンプレミス環境との互換性

管理者視点:同じ実行ログファイル形式が欲しい(課金のため)

ユーザ視点:同じ方法でジョブを投入したい、システムによって使い方を変えたくない

■クラウドの利点も享受したい

管理者視点:使った分だけ払いたい、任意の国で簡単に構築したい

ユーザ視点:GPUの有無、ノード間ネットワークなど自由に選びたい

Challenges

- ・スクリプトでクラスタを自動構築(10分程度)
- ・PBSジョブスケジューラ(互換性)
- ・LDAPによるユーザ管理(マルチユーザ環境、互換性)
- ・MPIジョブの実行(for HPC)
- Docker/Singularity
- ・LinuxのPosixの世界と、AWSにおけるユーザ権限のマッチング(S3)
- ・アイドル状態では計算ノードを持たず、ジョブが投入されると最適な計算ノード(EC2)を生成して登録する
- ・ジョブが終了し、一定時間計算ノードが再利用されないとEC2を削除する(定常課金の防止)
- ・ジョブが終了すると利用したファイル群はS3にアーカイブ(LustreよりS3が二桁ほど安価なので)
- ・ログインノードにloginしなくても、PCからS3にファイルを置くだけでジョブの実行が可能(産総研共研成果)

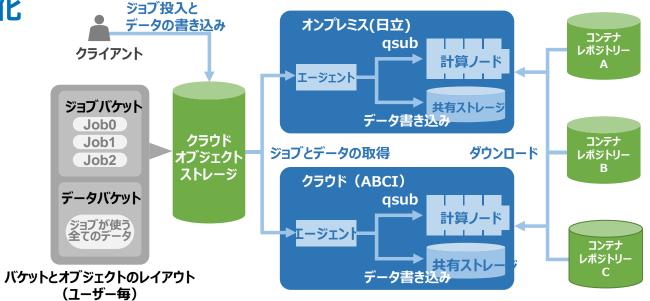
2-2. 産総研との共同研究成果を活用



オンプレミスAIクラスタとABCIのハイブリッドクラウド研究の推進

- ■オンプレミスとクラウド(ABCI)連携で産総研と共同研究中(2019~)
- ・オンプレミス(社内)環境からセキュアにクラウド(ABCI)を利用する
- ・2020/12/21に産総研様から学会発表
- ・2020/12/14に産総研様および日立からニュースリリース(研究トピックス)発信(*2,*3)

・CloudQとして産総研様からOSS化



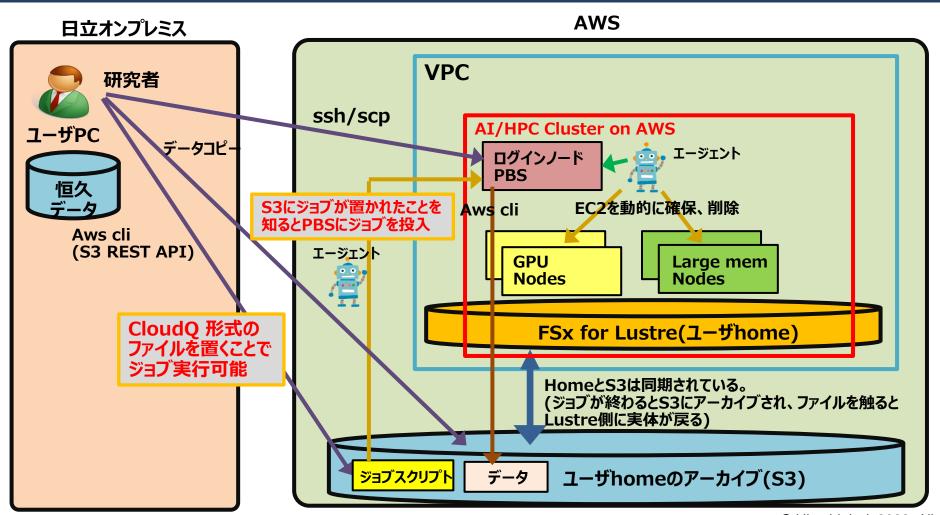
*2: https://www.aist.go.jp/aist_j/new_research/2020/nr20201214/nr20201214.html

*3: https://www.hitachi.co.jp/rd/news/topics/2020/1214.html

2-3. AI/HPC Dynamically Allocate Cluster on AWS



- ・ジョブはログインノードにログインしてsubmitするか、S3にジョブスクリプトを置くか(REST API)で実行可能
- ・ジョブが投入されるエージェントが適切なEC2を選んで計算ノードを起動し、PBSに登録する
- ・ジョブが終了、またはユーザがログアウトすると、関連ファイル、ユーザhomeがS3をアーカイブする







2-4. 社内からでも使いやすいクラウド

89.369

0.000

0.000

0212993838, 533412, 0311390263, 6,

0472860547, 6,



翌日早朝に本人および管理者にAWS利用状況を自動でメール

■PBSのログとAWSのログを対応させて報告

現時点でのS3利用量(89GB)



3. AI・ML用途におけるクラウドHPCシステムへの期待と課題



現在はHPCシステムを流用するも、将来はAI・MLに適した機能を

- ■現在はHPCシステムを流用のため、性能は出るが利用者を選ぶ
- ・HPCシステム + GPU
- ·Docker、Singularity対応
- ・機械学習の小単位IOに耐えるストレージ設定
- ·学習Only
- ·GPU利用効率重視でバッチ強要(一部インタラクティブで、jupyter notebook等を許可)

■将来への期待

- ・Kubernetesのようなサービスデプロイにも使える環境への進化
- ・オンプレとクラウドのハイブリッド機能の浸透(ジョブデプロイの容易化)
- ・企業でも使える暗号化実行環境(物理サーバを共有する仮想サーバの完全隔離など)

4. まとめ



- ・日立製作所 研究開発グループではオンプレの研究IT環境を保持
- ・オンプレミスITとクラウド環境を併用して実用中

AI/HPC on AWSの実装についてはSWoPP2022で産総研様と発表予定

・AI・ML用途におけるHPCシステムへの期待と課題



END

日立製作所 研究開発グループにおけるオンプレ/クラウドのAI/HPC環境

2022年6月23日

(株)日立製作所 研究開発グループ AI研/知能情報研究部 技術戦略室

清水 正明