



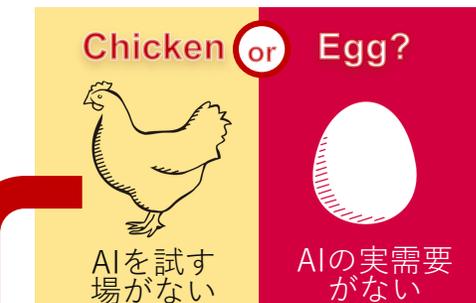
ABCI-Qの挑戦

量子・AIハイブリッド技術のテストベッド構築と 共通ライブラリ整備

国立研究開発法人 産業技術総合研究所
高野 了成



省電力AIスパコン研究



産業界のAI導入が進まない

ABC I AI Bridging Cloud Infrastructure



- 産総研は2000年代初頭からスパコン研究開発を継続
- 特に近年AIの計算需要に応え、同時に運用・管理コストの圧縮を実現する省電力AIスパコンに取り組む
- 一方で、当時AIについては産業界の関心の高さの割に導入が進んでいないという課題があった



- AI橋渡しクラウド（ABC I）：膨大な計算とデータを要するAIをすみやかに試す「場」を提供し、我が国における産学官によるA I 研究開発と社会実装を加速するオープンイノベーションプラットフォーム
- 高い計算能力を活用した人工知能技術の研究開発・実証、社会実装の推進、A I 分野の最重要課題への挑戦が目的

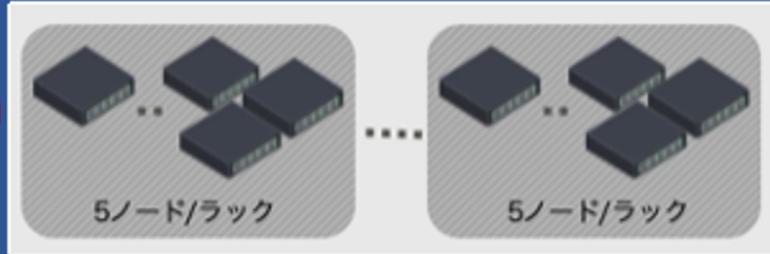


2018年8月1日 運用開始
2021年5月10日 2.0運用開始

- 経産省「人工知能に関するグローバル研究拠点整備事業」（2016年）の一環として整備
- 経産省「人工知能に関する橋渡しインフラ拡張」により2021年にアップグレードを実現

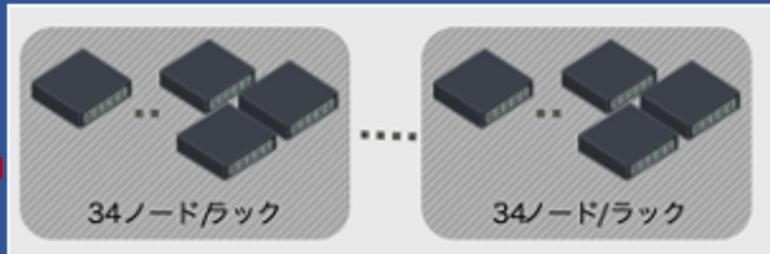
計算ノード (A)

GPU (A100) 計 960 個



計算ノード (V)

GPU (V100) 計 4,352 個



大容量ストレージ (34PB+17PB)

- 多数の計算ノード・GPUを使った**大量データの高速な並列処理に特化したハードウェア構成**
 - 高速なAI計算に不可欠な高性能GPU
 - 広帯域・低遅延な計算ノード間通信
 - 大容量・広帯域・高可用性ストレージ
- AI研究開発を効率化し、**新規ユーザの敷居を下げるソフトウェア構成**と先進的な取り組み
 - 最先端のミドルウェアや並列化コンパイラ、最新のGPU向け開発環境や深層学習のツール類を提供
 - ABCIに最適化された「コンテナ」開発により分散並列学習などの最先端技術をユーザが簡単に利用可能に
 - データと学習済みモデル、実行環境をひとまとめにしてAI開発を容易にするABCI上のサービス「AIハブ」の一部機能がすでに提供 (ABCIデータセット)

産総研がこれまで培ったスパコン構築・運用のノウハウが、日本を代表する「高性能」で「使いやすい」AIスパコンの実現を可能にした

ABCIの今後の方向性

- 2010年代は第3次AIブーム、そして第2次量子コンピュータブームであり、現在進行中
- 生成AIや量子コンピューティングといった研究開発のニーズに応じて2系統に進化
 - 実世界基盤モデル開発向け大規模クラウド基盤（ABCI3.0「シンABCI」）
 - 量子・AI融合処理向け大規模クラウド基盤（ABCI-Q）

2011 D-Wave

2019 量子超越性

2023 48論理量子ビット

2025 ABCI-Q

2018 ABCI

2021 ABCI2.0

2024 ABCI3.0

2012 AlexNet

2015 AlphaGo

2023 ChatGPT

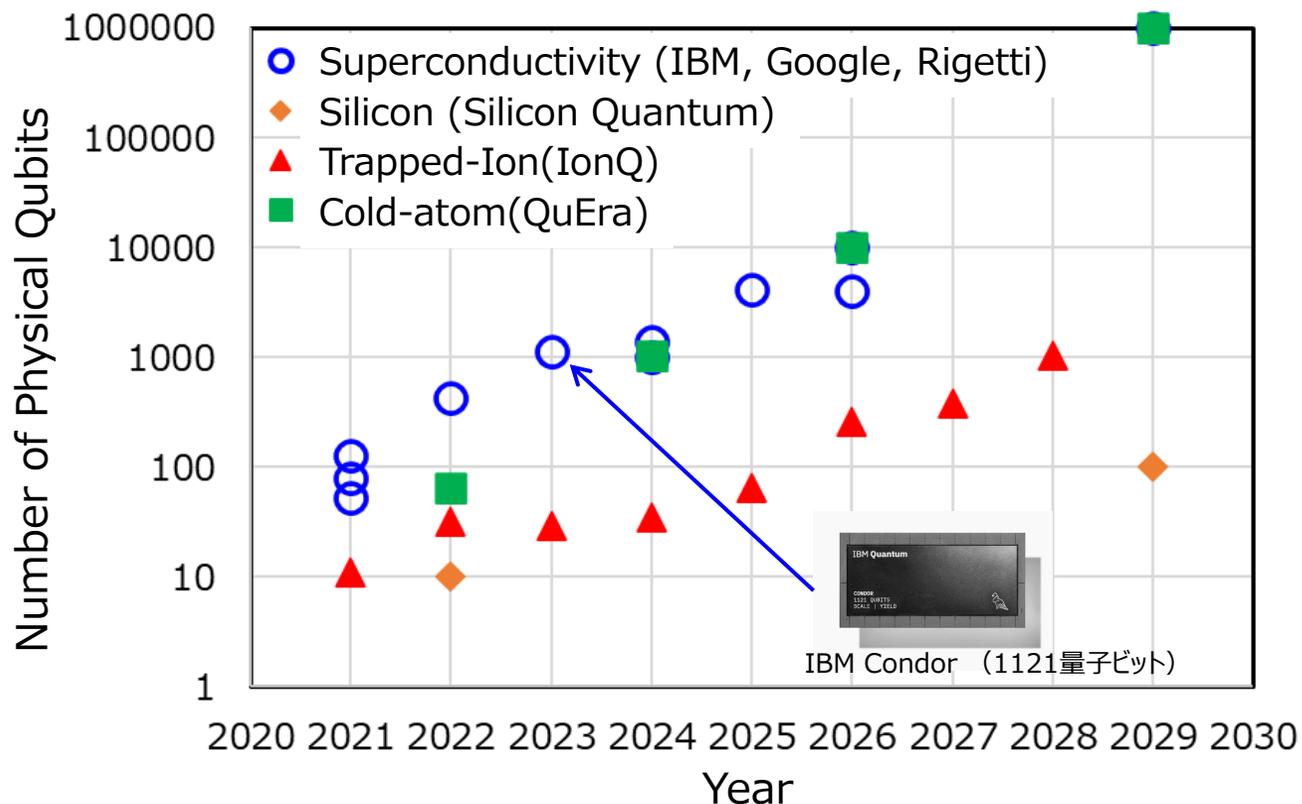
- 深層学習
- 強化学習
- 確率的探索法
- シミュレーション



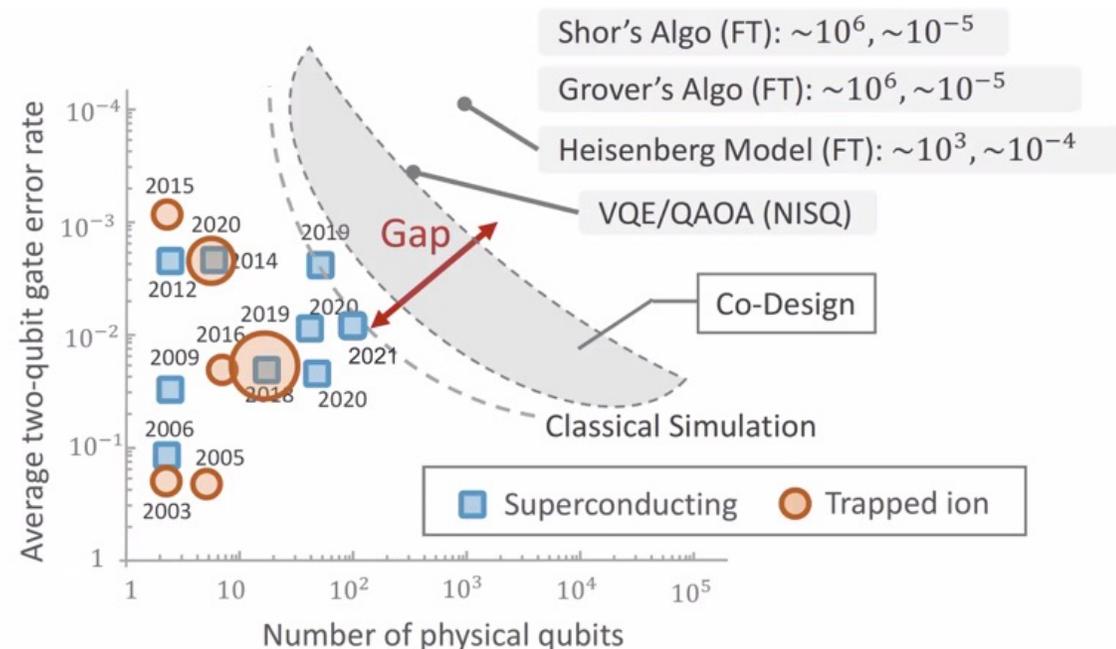
- トランスフォーマー
- 基盤モデル
- リファインメント
- Self-Supervision
- RLHF
- 生成AI

- ここ数年で量子コンピュータを「使う」時代になった。
 - IBM Q Experience (2019)、Amazon Braket (2019)等
- 実用化はまだ先だが、先行者利益を確保しようと研究開発が熱を帯びている。

量子コンピュータの開発ロードマップ



量子アルゴリズムとハードウェアのギャップ



“Closing the Gap between Quantum Algorithms and Machines by Breaking Systems Abstractions,” Fred Chong@Univ. of Chicago

- 量子アニーリング
 - イジングモデルを物理の法則等を利用して解く。
 - 量子インスパイアデバイス（富士通、NEC、東芝、日立等）はビジネス実証レベルに。
 - 例）組合せ最適化、ボルツマンサンプリング
- NISQ (Noisy Intermediate-Scale Quantum)
 - 問題の分割や、小さい量子プログラムと統計処理・最適化を組み合わせて、能力が限定的な現在の量子コンピュータを使う方法が模索される。
 - 量子コンピュータを使うメリットがあるかは懐疑的。
 - 例）VQE（変分量子固有値ソルバ）、QAOA（量子近似最適化）、QNN（量子ニューラルネットワーク）
- FTQC (Fault Tolerant Quantum Computer)
 - エラー耐性がある量子コンピュータ。
 - 例）Shorの素因数分解アルゴリズム、Groverの探索アルゴリズム、QPE（量子位相推定）、HHLアルゴリズム

- 近年、エラー訂正技術の研究が熱い。
- Harvard大、QuEraらによる48論理量子ビットの実現が話題に。
- FTQCアルゴリズムの研究が加速し、スパコンを用いた量子回路シミュレーションの重要性が高まるのではないか。
 - 48量子ビット→4PiBメモリが必要（GPU数万基に相当）

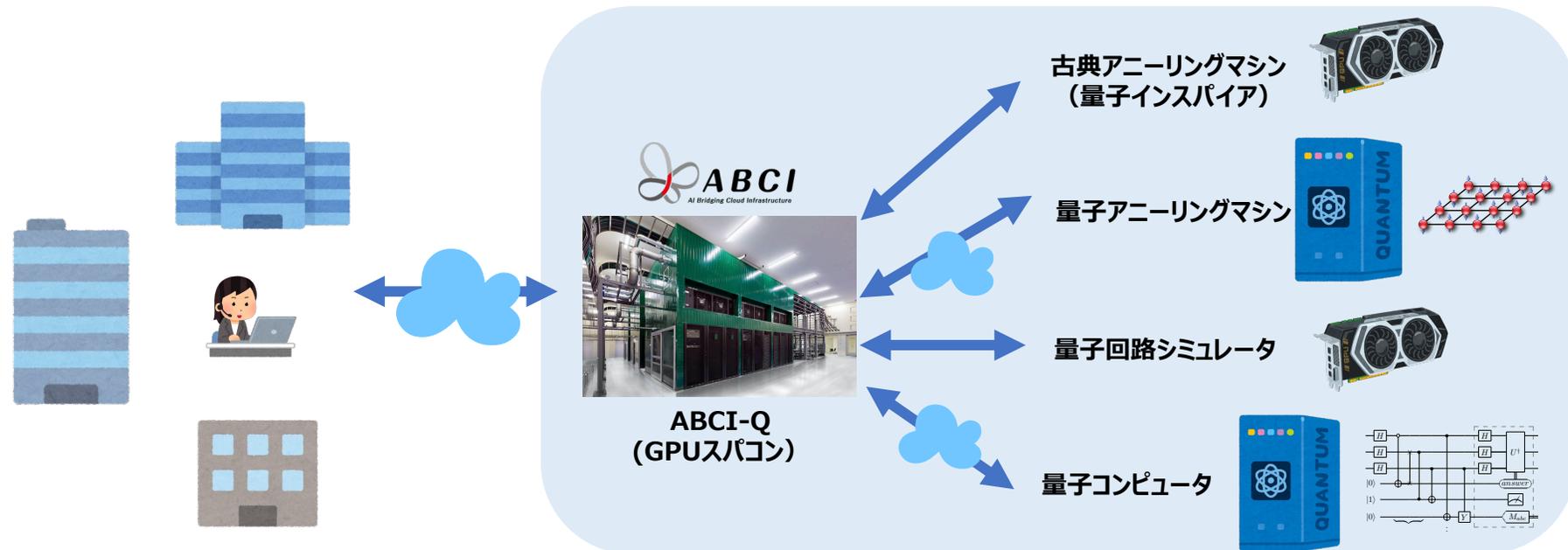


The screenshot shows the top portion of a Nature article page. At the top left is the 'nature' logo. Below it are navigation links: 'Explore content', 'About the journal', and 'Publish with us'. The breadcrumb trail reads 'nature > articles > article'. The article title is 'Logical quantum processor based on reconfigurable atom arrays'. Below the title is a list of authors: Dolev Bluvstein, Simon J. Evered, Alexandra A. Geim, Sophie H. Li, Hengyun Zhou, Tom Manovitz, Sepehr Ebadi, Madelyn Cain, Marcin Kalinowski, Dominik Hangleiter, J. Pablo Bonilla Ataides, Nishad Maskara, Iris Cong, Xun Gao, Pedro Sales Rodriguez, Thomas Karolyshyn, Giulia Semeghini, Michael J. Gullans, Markus Greiner, Vladan Vuletić & Mikhail D. Lukin. At the bottom of the article preview, it says 'Nature 626, 58–65 (2024) | Cite this article' and '69k Accesses | 2 Citations | 816 Altmetric | Metrics'.

<https://www.nature.com/articles/s41586-023-06927-3h>

ABCI-Q : 量子・AI融合処理向け大規模クラウド基盤

- 産総研AI橋渡しクラウドABCIの構築・運用ノウハウを活かし、量子・古典アニーリングマシン・量子コンピュータを融合した量子・AIクラウドを構築・運用
 - ▶ 古典アニーリングマシン商用化に向けたユースケース探索・評価
 - ▶ 大規模量子回路シミュレータ、イジングマシンを活用した量子AIライブラリ開発
 - ▶ 様々な量子コンピュータ技術を試用・評価できる場の創出



量子・AIクラウドを企業等に提供し、AI、エネルギー、材料、金融、バイオなどの様々なユースケースを創出→事業化支援・スタートアップ創出

オフィスエリア

- 2F：連携スペース
 - ・ 連携者の滞在
 - ・ オフィス、分室等の設置
- 3F：産総研研究者
 - ・ 本部機能
 - ・ 連携支援人材
 - ・ 多様な専門人材



インキュベーション機能

- 多様なステイクホルダーの結節点
- ・ ユーザ
 - ・ ベンダー
 - ・ サプライヤー
 - ・ 産学官
 - ・ VCなど

中心部に設置した交流スペース

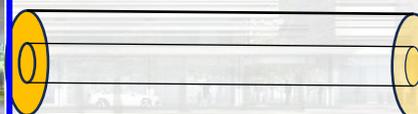
NEC/産総研 量子アニーリングマシン

計算能力 エクサAIFLOPS級

- ・ イジングマシン (アニーリング)
- ・ 量子回路シミュレータ
- ・ AI, 古典ソルバなど



ABCI-Qを設置



超高速インターフェース
古典・量子ハイブリッド計算基盤



量子コンピュータ

- ・ 量子アニーリング
- ・ 量子ゲート
 - ✓ 超伝導
 - ✓ イオントラップ
 - ✓ 中性原子
 - ✓ 光量子など

量子コンピュータの設置

- 「量子未来社会ビジョン」では、量子技術を社会経済システム全体に取り込み、従来型（古典）技術システムとの融合（ハイブリッド）を図ることにより、我が国の産業の成長機会の創出・社会課題の解決につながると提言されている。
- 量子コンピュータを、古典コンピュータのみでは量的・質的に解けない問題を現実的な時間で解く「アクセラレータ」として活用し、従来のAI技術と組み合わせることで、計算能力の飛躍的な向上やデータ利用の高度化を実現することが可能になる。
- 量子・AIハイブリッド技術を活用したアプリケーションの創出を促進するには、さまざまな用途に再利用可能な形でアルゴリズムを実装した共通ライブラリの開発と、継続的な管理体制の構築・整備が必要である。



(出典) 内閣府 量子未来社会ビジョン概要

研究開発項目①：量子・AIアプリケーション開発・実証

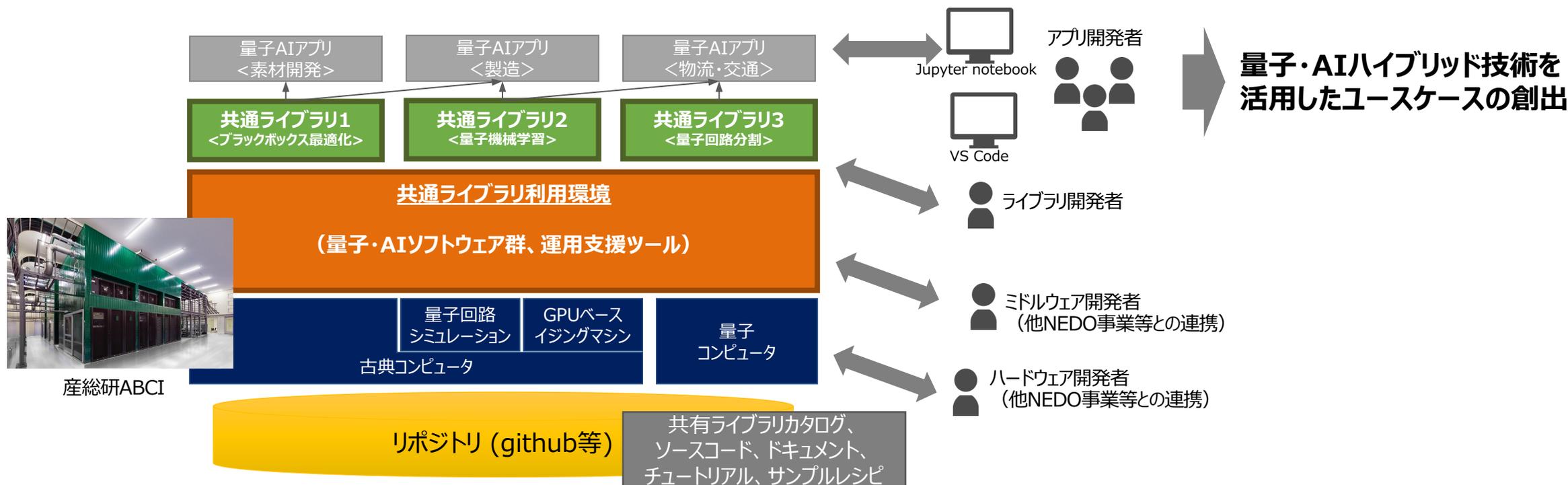
テーマ名	開発分野
量子・AI次世代創薬	素材開発
量子・AIハイブリッドによる創薬向け大規模Virtual Screening法の開発	素材開発
高柔軟性薬剤分子の結晶構造予測システム	素材開発
量子・AI支援による機能タンパク質最適化技術の研究開発	素材開発
量子生成AIによる半導体製造用新材料開発	素材開発
高次リサイクルシステム構築を志向する解体性接着技術開発	素材開発
物流現場における人間機械協調作業のためのデジタルツイン量子最適化	物流・交通
量子+古典AIによる物流業務効率化のアプリケーション開発	物流・交通
製造における異常検知技術の研究開発	製造
量子・AIによるポスト5G・6G用メタサーフェスデバイスの研究開発	製造

研究開発項目②：量子・AI最適化等に向けたライブラリ開発

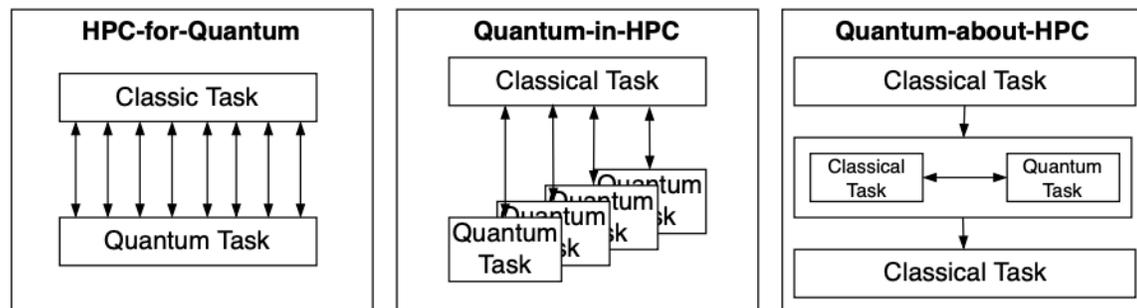
テーマ名	実施機関
量子・AIハイブリッド技術の活用を加速する共通ライブラリ基盤の研究開発 (ブラックボックス最適化共通ライブラリの研究開発)	学校法人慶應義塾
量子・AIハイブリッド技術の活用を加速する共通ライブラリ基盤の研究開発 (量子機械学習共通ライブラリの研究開発)	国立大学法人東北大学
量子回路分割ライブラリ	PwCコンサルティング合同会社
量子・AIハイブリッド技術の活用を加速する共通ライブラリ基盤の研究開発	国立研究開発法人産業技術総合研究所、株式会社長大

量子・AI共通ライブラリの開発と管理体制の整備

- 共通ライブラリを試行してその効果を実感できる「場」をABCI上に構築し、ユースケース創出から共通ライブラリの開発・維持を持続的に実施する体制構築を目指す。
- 一般的なAI技術者が使い慣れた環境（Jupyter Notebook、VS Code等）から、量子・AIハイブリッド技術を利用・評価できる共通ライブラリ利用環境を提供することで、社会実装フェーズへの移行の円滑化を目指す。



- 量子・AI融合アプリケーションをHPCと量子コンピュータから構成されるシステム上で効率的に実行するミドルウェア含むソフトウェアスタックが必要。

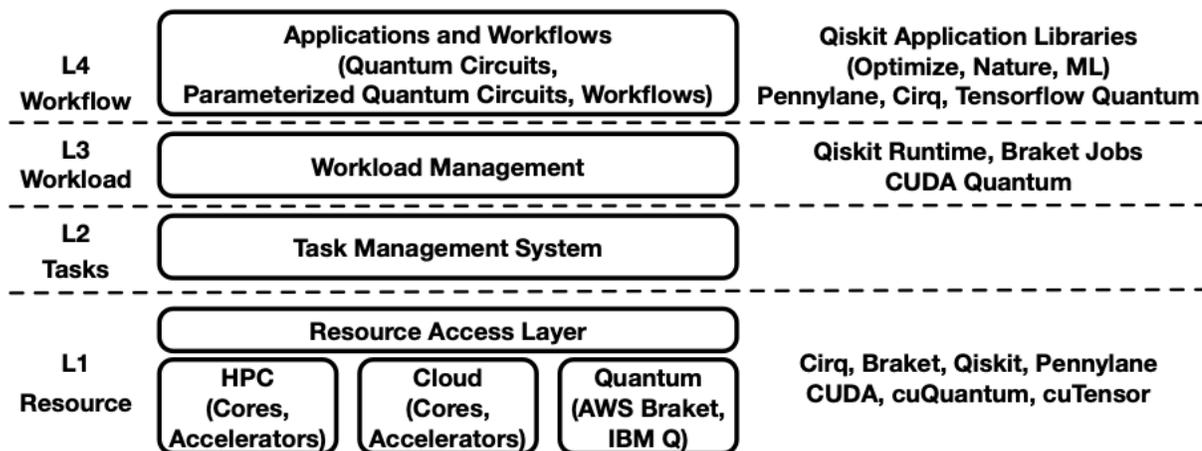


Examples:

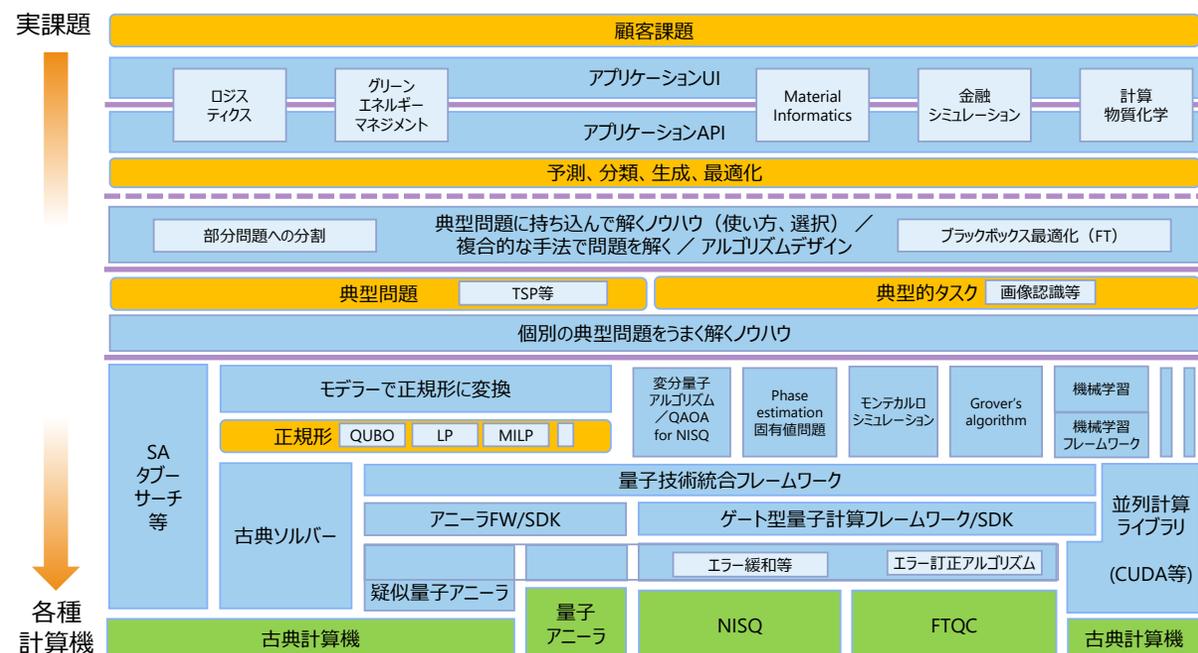
Quantum Error Correction

Variational Algorithms

Workflows



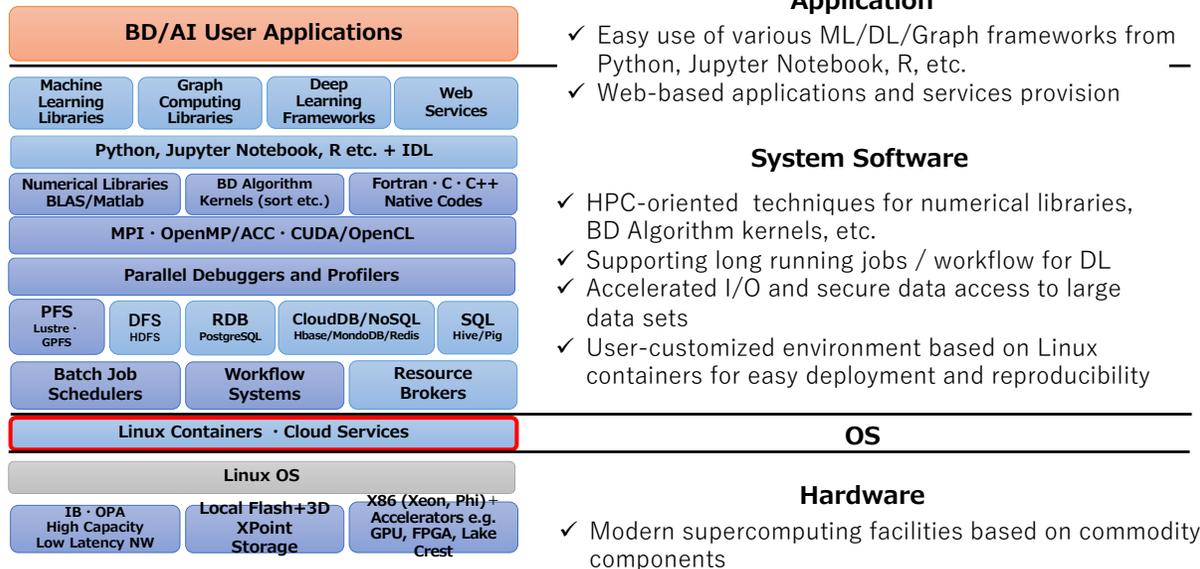
量子コンピューティングエコシステム構築に向けたソフトウェアオープンプラットフォームの検討 (Q-STAR)



5年前はAIとHPCの融合と言っていた、、、

- 2010年代にはいろいろな深層学習フレームワークが存在したが、現在はPyTorchにほぼ集約。
- 量子コンピューティングも同じ流れを辿っていくのでは。
- ユーザ環境をコンテナ、環境モジュール、Spackでカスタマイズできれば、大抵のニーズは満たせる。

Basic Requirements for AI Cloud System





ABCIIソフトウェアスタック

基本ソフトウェア

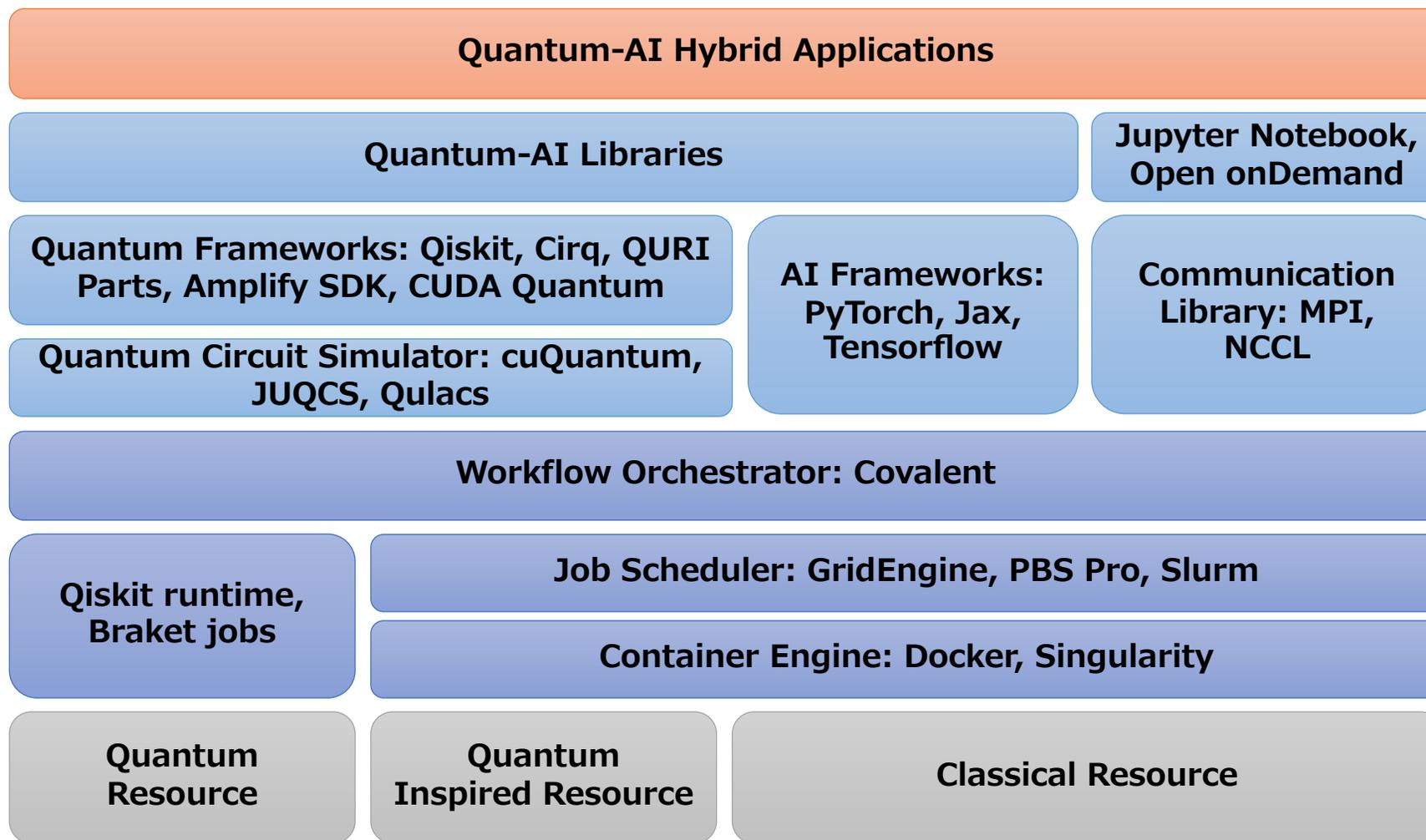
Operating System	CentOS, RHEL
Job Scheduler	Univa Grid Engine
Container Engine	Docker, Singularity
MPI	OpenMPI, MVAPICH2, Intel MPI
Development tools	Intel Parallel Studio XE Cluster Edition, PGI Professional Edition, NVIDIA CUDA SDK, GCC, Python, Ruby, R, Java, Scala, Perl
Deep Learning	Caffe, Caffe2, TensorFlow, Theano, Torch, PyTorch, CNTK, MXnet, Chainer, Keras, etc.
Big Data Processing	Hadoop, Spark

コンテナサポート

- 2つのコンテナエンジンをサポート
 - Docker: 最もユーザベースが大きい
 - Singularity: ローレンスバークレー国立研究所で開発されHPC分野での利用が広まりつつある
- コンテナをベースとしたソフトウェアエコシステムの活用
 - グローバルコミュニティで開発された最新の成果をすぐにABCII上で試せる
 - ABCIIで開発したコンテナをグローバルコミュニティに還元することも容易
- ABCIIに最適化されたコンテナの開発・提供
 - ABCIIに最適化されたChainerMN、NNablaをはじめとする分散ディープラーニングフレームワークをコンテナ化することで、誰でも大規模データを使ったディープラーニング応用に取り組むことが可能に。

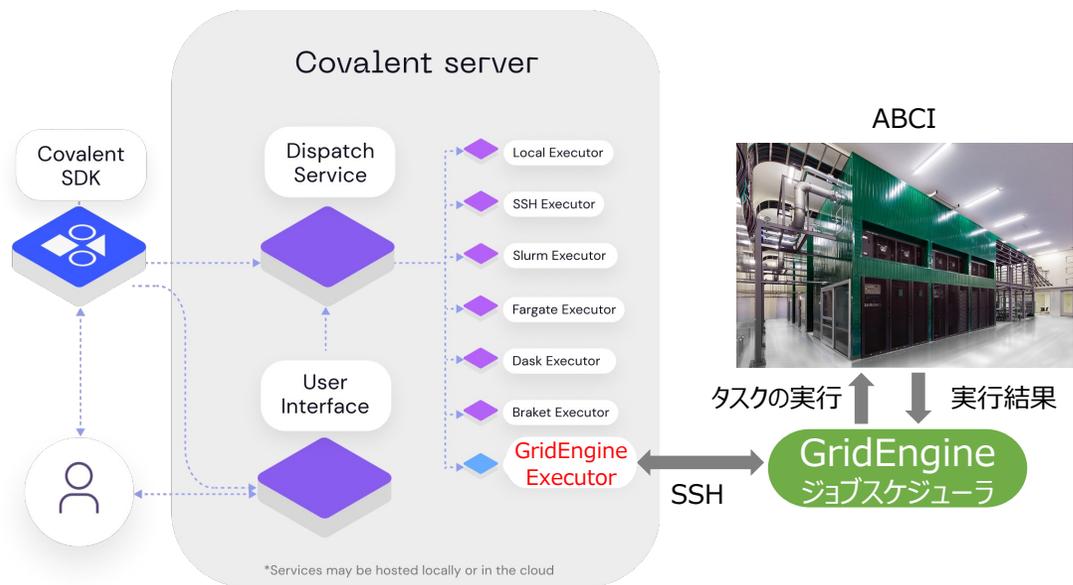


- 量子・AI融合アプリケーションをHPCと量子コンピュータから構成されるシステム上で効率的に実行するミドルウェア含むソフトウェアスタックが必要。

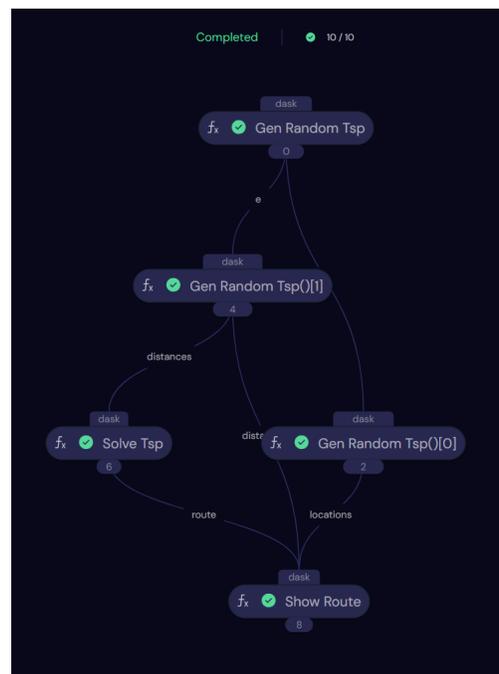


ワークフローオーケストレータ Covalent

- Agnostiq社を中心に開発されているOSSワークフローオーケストレータ
- IBM Quantum、Amazon Braket、Slurmなどに対応
- GridEngine用Executorプラグインを新規開発。OSS化予定
 - 巡回セールスマン問題、ブラックボックス最適化



CovalentのABCI対応



TSPのワークフローグラフ

```
In [11]: import covalent

ABCI上でタスクを実行するために、GridEngineExecutorオブジェクトを作成します。

In [12]: abci_executor = covalent.executor.GridEngineExecutor(
          username="username", # ABCIのユーザー名を入力してください,
          address="localhost",
          port=10022,
          ssh_key_file="~/ssh/id_rsa", # ABCIにSSH接続するために使用するSSH鍵ファイルのパスを入力してください,
          remote_workdir="$HOME/ampLify_env",
          poll_freq=30,
          cleanup=True,
          embedded_qsub_args={
              "rl": ["rt_F=1", "h_rt=1:00:00"],
          }, # ジョブスクリプト内に書かれるqsubのオプション

In [19]: dispatch_id = covalent.dispatch(workflow)(10)

In [20]: print(dispatch_id)

Out [20]: f68c9ec2-0f1f-4077-945f-f6e1ddab1abe

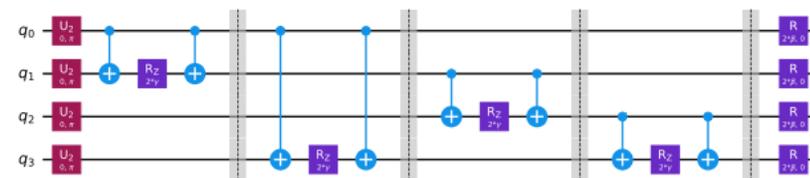
In [21]: result = covalent.get_result(dispatch_id, wait=True)
          print(result)

Out [21]: Lattice Result
          =====
          status: COMPLETED
          result: 3.089814832129642
          input args: ['10']
          input kwargs: {}
          error: None

          start_time: 2023-12-27 06:10:45.715374
          end_time: 2023-12-27 07:00:37.814164
```

Jupyter Notebookでの実行

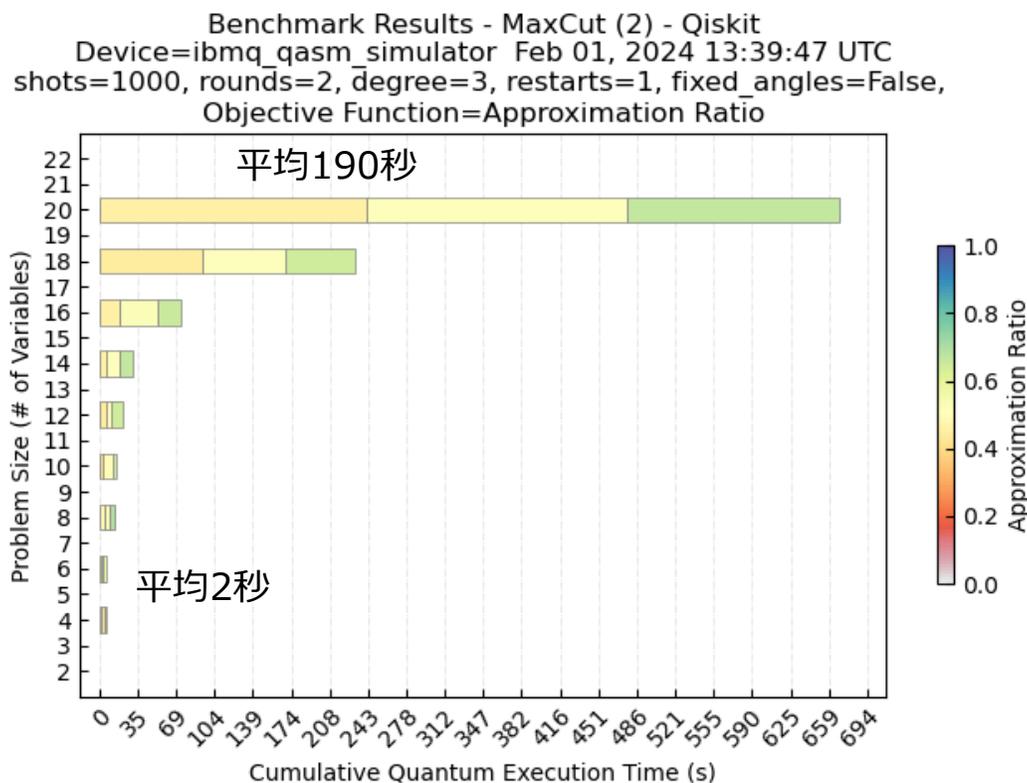
- 量子・古典アルゴリズム (VQE、QAOA) では、量子ジョブの待ち時間により実行時間が律速される。
- Qiskitのハイブリッドジョブ
 - セッションを利用することで、後続するジョブの優先度が上がり、待ち時間を削減可能。
 - ただしIBM Quantumを使う場合のみ有効。



QAOA (MaxCut問題) 4量子ビット、2イテレーションの実行時間 [秒]

QPU	iter.	Classic	Quantum	job wait	quantum seconds
		Subtotal	subtotal		
osaka	1	2.3	2630.0	2531.6	17.0
osaka	2	-	-	64.4	17.0
kyoto	1	1.7	508.3	404.8	17.0
kyoto	2	-	-	69.5	17.0
brisbane	1	6.1	5242.4	5153.8	17.0
brisbane	2	-	-	54.6	17.0

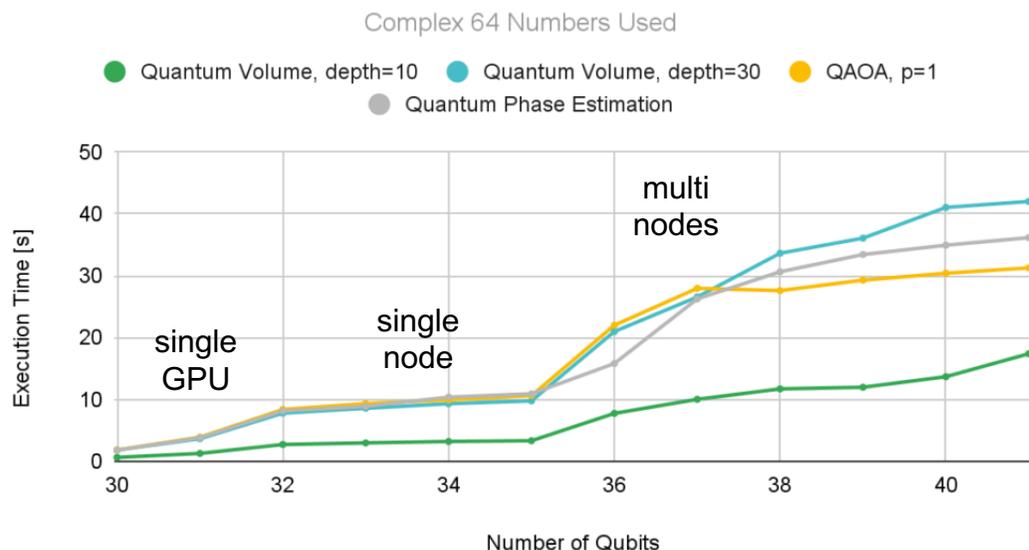
- 量子回路シミュレーションでは、量子ビット数を増やすと指数関数的に実行時間、及びメモリ使用量が増加する。
- NVIDIA cuQuantumによる加速が有効。
 - GPUの活用のみならず、マルチGPU・マルチノード対応も可能。



世界トップレベルの高速量子回路シミュレーション

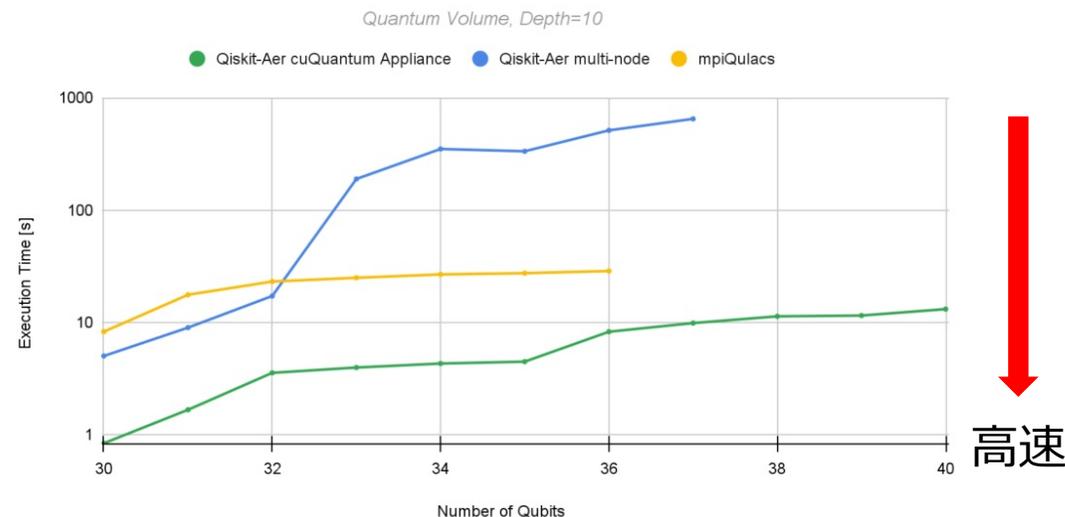
- ABCI上で512GPUを用いた**世界最速の41量子ビット量子回路シミュレーション**を実施
- NVIDIA合同会社と産総研の共同グループによるABCI GRANDチャレンジの成果

Weak Scaling Performance of cuQuantum Appliance on ABCI



cuQuantum Appliance multi-node weak scaling performance from 30 to 40 qubits for complex 64

Weak Scaling Comparison of Multi-node Simulators



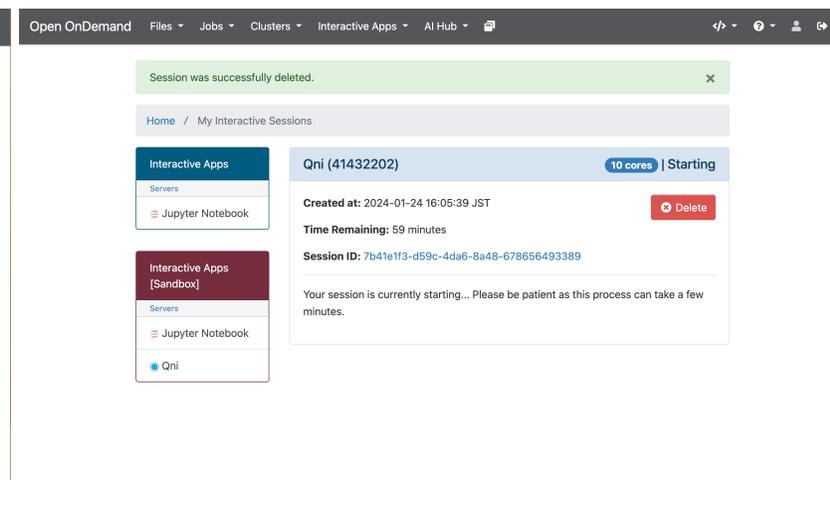
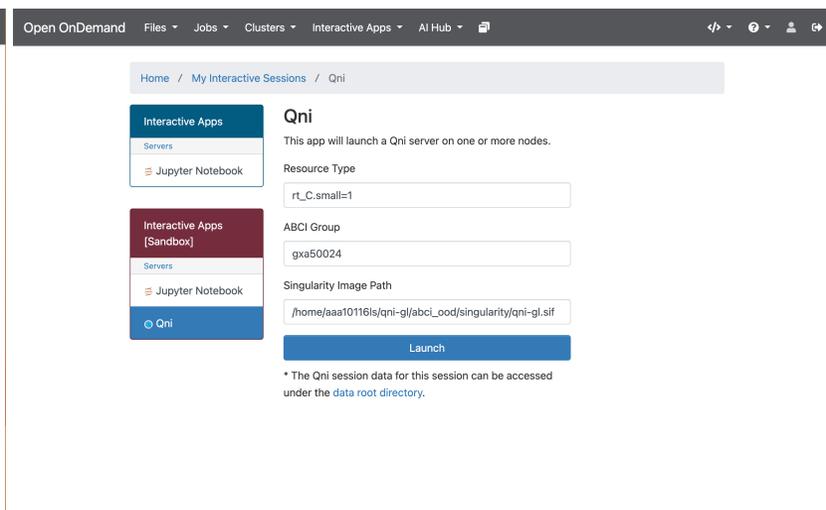
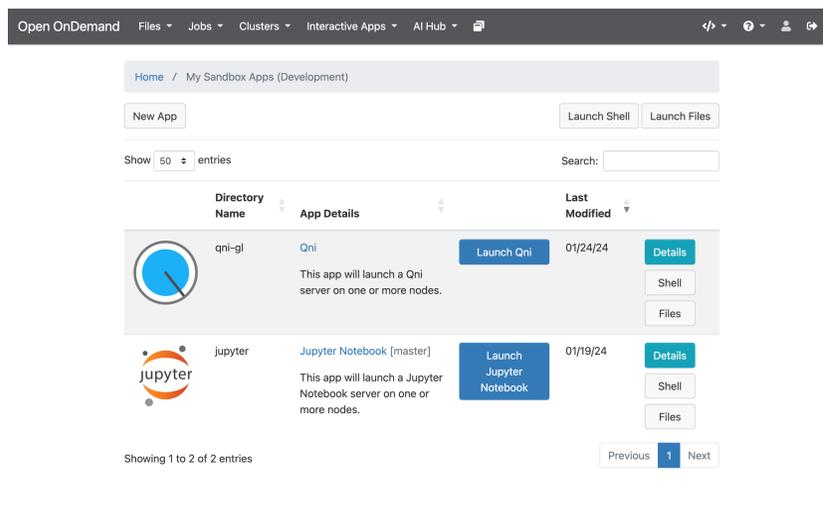
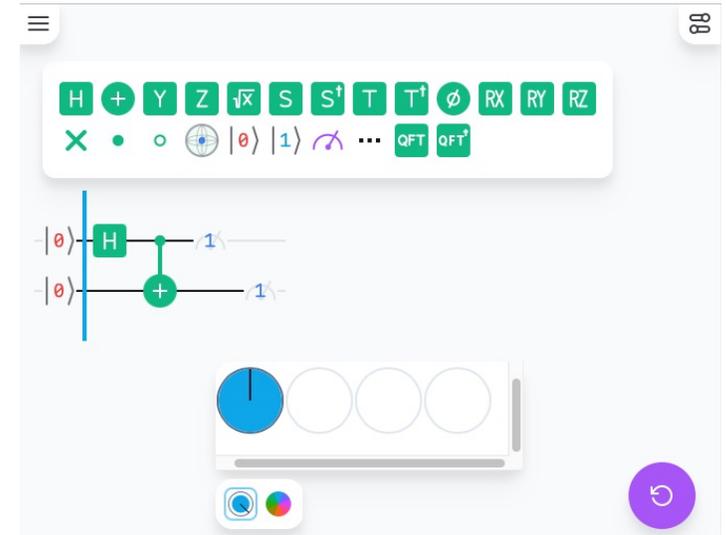
More than 3x faster than SoTA (mpiQulacs on A64FX-based cluster)

開発された量子回路シミュレータはNVIDIAより公開、ABCI上で誰もが利用可能

<https://www.digiarc.aist.go.jp/publication/quantum/20221219.html>

<https://developer.nvidia.com/blog/best-in-class-quantum-circuit-simulation-at-scale-with-nvidia-cuquantum-appliance/>

- Webブラウザで動作する量子シミュレータ（ライブプログラミングアプリ）QniのバックエンドとしてABCIを利用可能にした。
 - <https://github.com/qniapp/>
- Open onDemandアプリ化も実施した。
- 3/18のHPC研究会で発表予定（中田他「AI橋渡しクラウドにおけるWebUIと量子コンピューティングサービスの導入」）



- 産総研G-QuATは、量子プロセッサ製造、システム化、サプライチェーン、計算基盤、ユースケース、標準化に至る一貫した**量子工学(Quantum Engineering)**を推進し、**国内外の産業界とともにグローバル産業エコシステムの構築**を実現する「**世界で唯一無二の研究センター**」を目指す。
- 量子コンピュータ技術の実用化・商用化を加速するための戦略的計算基盤としてABCI-Qを整備する。
- 量子コンピュータ分野においてもオープンソースソフトウェアの活用は重要であり、HPCソフトウェアスタックとの融合は重要な研究課題になる。



謝辞：この成果の一部は、国立研究開発法人新エネルギー・産業技術総合開発機構（NEDO）の「量子・AIハイブリッド技術のサイバー・フィジカル開発事業」（JPNP23003）の委託事業の結果得られたものです。