

ORACLE



PCクラスタコンソーシアム HPCクラウド部会 第1回ワークショップ

OCI for HPC 最前線

2023年10月4日

日本オラクル株式会社

クラウド営業統括 松山 慎



Safe harbor statement

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, timing, and pricing of any features or functionality described for Oracle's products may change and remains at the sole discretion of Oracle Corporation.



Oracle Cloud Infrastructure (OCI) 概要

Oracle Cloud Infrastructure(OCI)のリージョン

2023年8月現在：45リージョン提供中、さらに7リージョン計画

デュアル・リージョン：基本的にすべての国/地域で2つ以上のリージョンを提供し、お客様の業務継続要件に対応していく
(日本の場合は東京-大阪)
各リージョンはOracle Backboneで接続

サステナビリティ：2025年までに、全リージョンで100%再生可能エネルギーを使用することを表明(南米/欧州は達成済み)



Oracle Cloud Infrastructure の日本国内展開

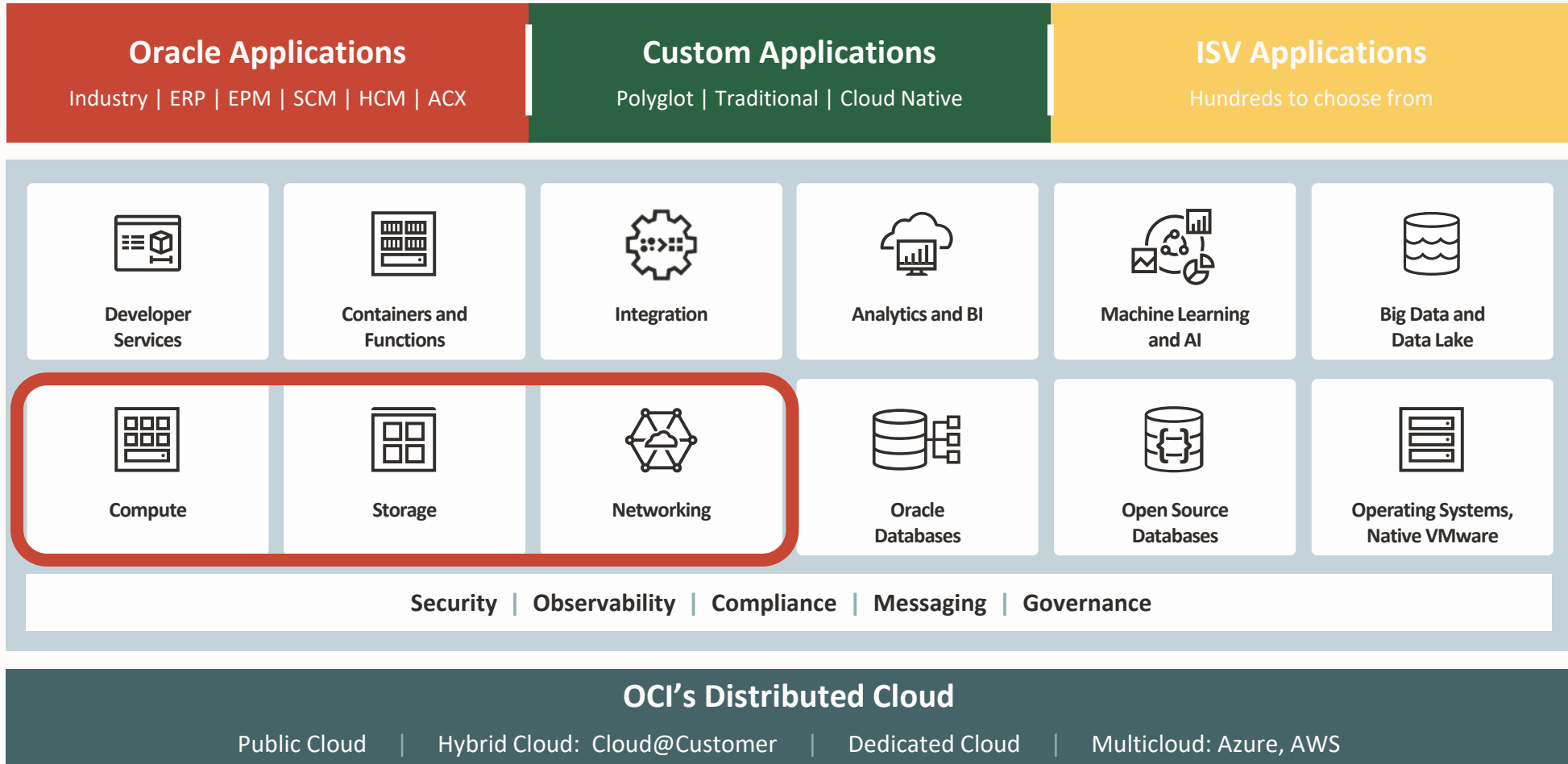
| リージョン | 大阪 (Japan Central) | 東京 (Japan East) |
|--|---|-------------------|
| 開業年 | 2020年2月 | 2019年5月 |
| SINETクラウド接続(*)  | 2020年7月 | 2019年10月 |
| 政府のクラウド施策対応 | ISMAP登録済み、ガバメントクラウド対象クラウドサービス | |
| 準拠法 | 日本国の実体法と手続法を適用 | |
| 裁判管轄 | 東京地方裁判所を第一審の専属的合意管轄裁判所とする | |
| 決済方式 | 現地通貨(日本円) および請求書ベースの支払い | |
| 単価 | ワールドワイドで同一の単価 (月10TB超のインターネットアウトバウンドデータ量を除く) 円建て単価契約 | |

(*) SINET接続/FastConnect仮想回線
経由の通信は上り/下りとも
無制限に転送データ量課金無し



OCI サービス全体像

Infrastructure, platform, and SaaS in **one flat-network**









コンピューート：インスタンスの選択

1コアからスタート、ベアメタル、RDMA、GPUなど、高いコストパフォーマンスで幅広く提供

用途に応じてインスタンスのシェイプを選択可能

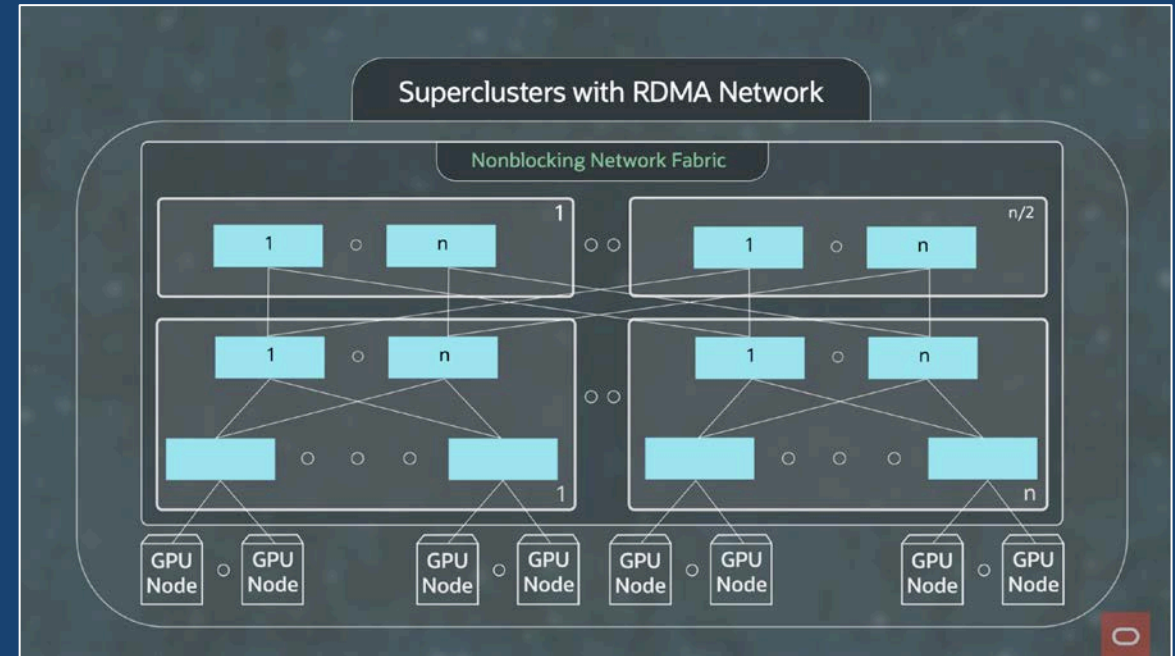
- VMもベアメタルも、同じインフラ（ネットワーク、ハードウェア、ファームウェア、API）上で提供

| | | | | |
|-----------|---|--|--|--|
| CPUの種類 |  |  |  | |
| インスタンスの種類 | VMインスタンス <ul style="list-style-type: none">• 仮想マシン環境<ul style="list-style-type: none">- Standard Flexible VMs: CPU 1コア単位、メモリ1GiB単位で自由に組合せ可能  Virtual Machine | | ベアメタルインスタンス <ul style="list-style-type: none">• 物理サーバー環境<ul style="list-style-type: none">- 性能重視アプリケーション向け- 仮想マシンと同様に容易なデプロイ可能  Bare Metal Compute | |
| シェイプタイプ | Standard <p>汎用シェイプ 一般的なワークロード向け インスタンス停止で課金停止 (他のシェイプタイプは課金停止は終了する必要あり)</p> | Dense I/O <p>NVMe SSD内蔵シェイプ I/O性能重視ワークロード向け</p> | HPC/Optimized <p>HPCクラスタタイプ 高帯域/低遅延RDMAネットワーク HPCクラスタ用インスタンス</p> | GPU <p>GPUクラスタタイプ AI、機械学習、 HPC</p>  |



OCI Supercluster

- 広帯域/低遅延 大規模RDMAネットワーク
- 最大 16,384のコンピュータ・ノード
- 最大 32,768基のGPU

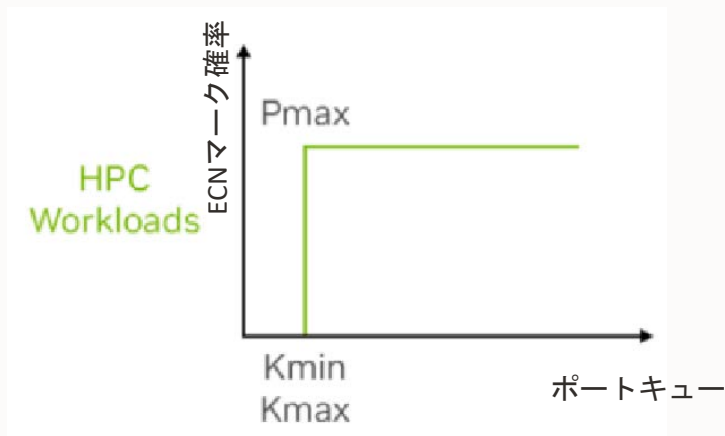


3種のOCI Supercluster/Cluster Network輻輳制御プロファイル

クラスタベアメタルノードにLinux コマンドでユーザが選択投入可能

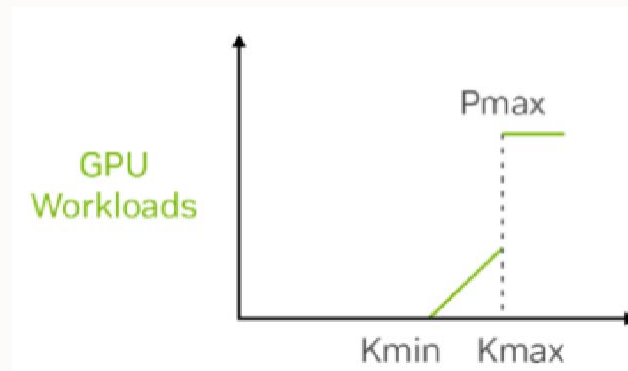
HPC MPI Workloads

- **低遅延要求**
- 科学技術計算
- 高頻度取引



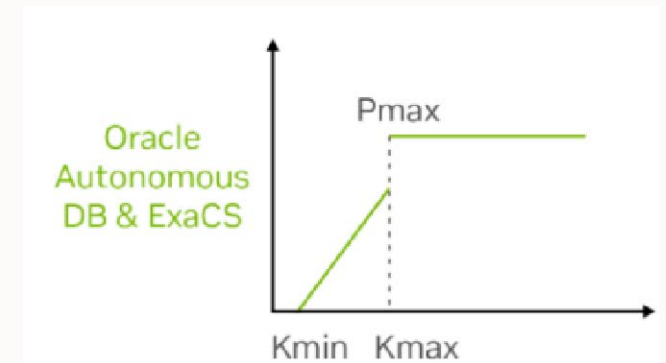
ML Workloads

- **高スループット要求**
- MPIワークロードに対してジッターとRTTに比較的寛容



Oracle Database RAC Workloads

- 帯域-遅延に対する要求はMPI WorkloadsとML Workloadsの間



ECN(Explicit Congestion Notification)

スーパーコンピューティング技術産業応用協議会 (産応協) HPCものづくりワークショップ 大規模FEM構造解析ベンチマーク (FrontISTR)実施結果

産応協: <http://www.icscp.jp/>

FrontISTR: <https://www.frontistr.com/>

OracleのFrontISTRベンチマーク実施の動機

[これまでのパブリッククラウド評]

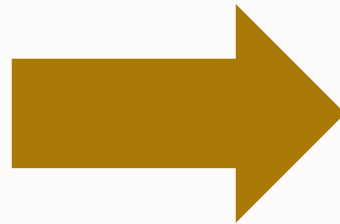
他社パブリッククラウドで
HPCワークロードを実行すると実施する
度に実行時間に大きな差が発生する

↓

実行計画が立たない
本来不要な課金が発生する

↓

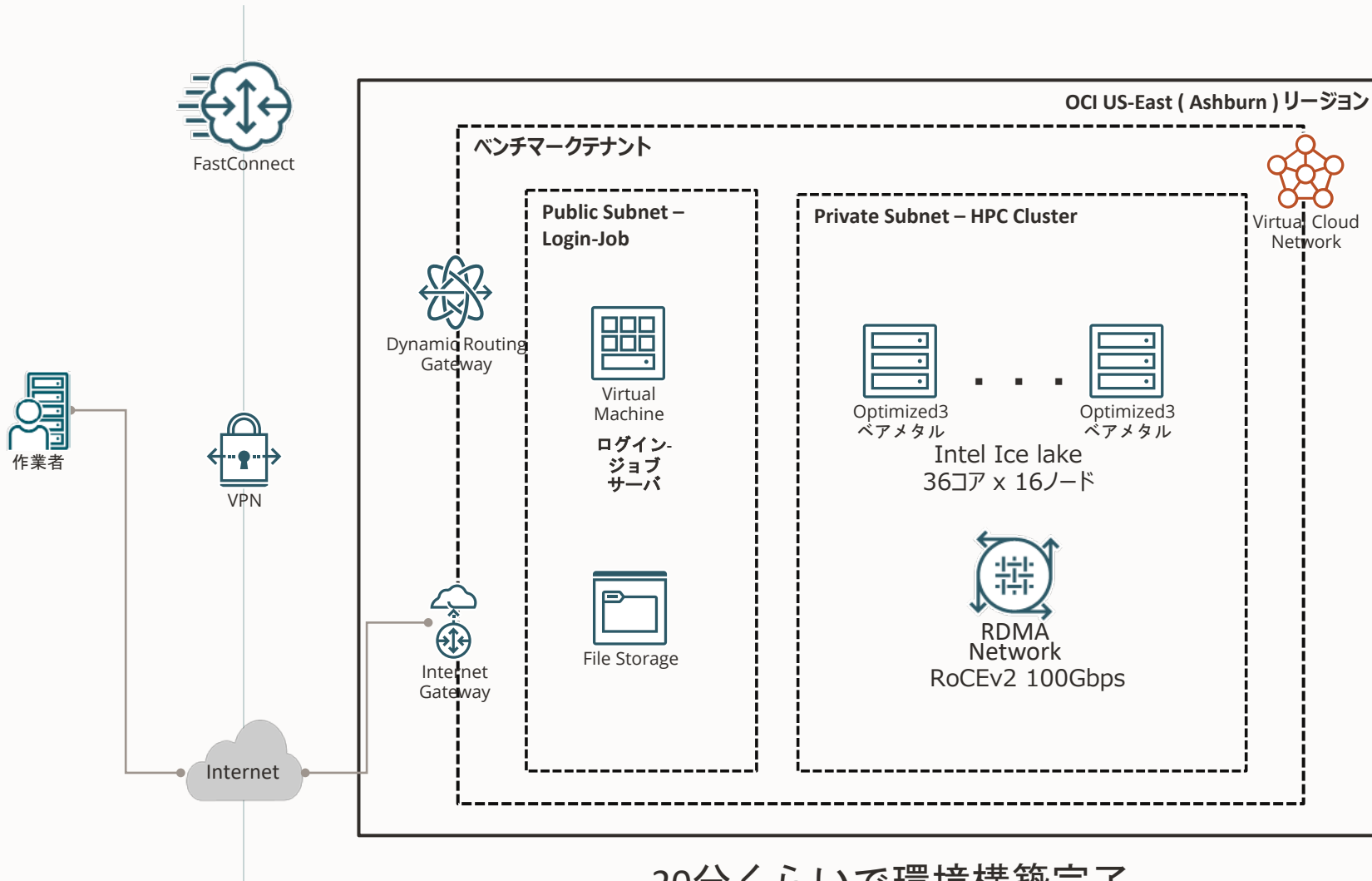
パブリッククラウドは使いにくい



[Oracle Cloud Infrastructure (OCI)]

パブリッククラウドであっても
ベアメタルノードと帯域と遅延を約
束するフラットネットワークを持つ
OCIであれば、オンプレミスのHPCク
ラスタと同様の高い性能再現性が示
せるはず

FrontISTR ベンチマーク実施環境構成図



20分くらいで環境構築完了

OCI US-East (Ashburn) リージョン
バージニア州ラウドン郡 (時差14時間)



ログインジョブサーバ (Slurm)
Standard2 仮想マシン x 1ノード
Intel Xeon Platinum 8167M 4 CPUコア
60GBメモリ
100GB ブロックボリューム

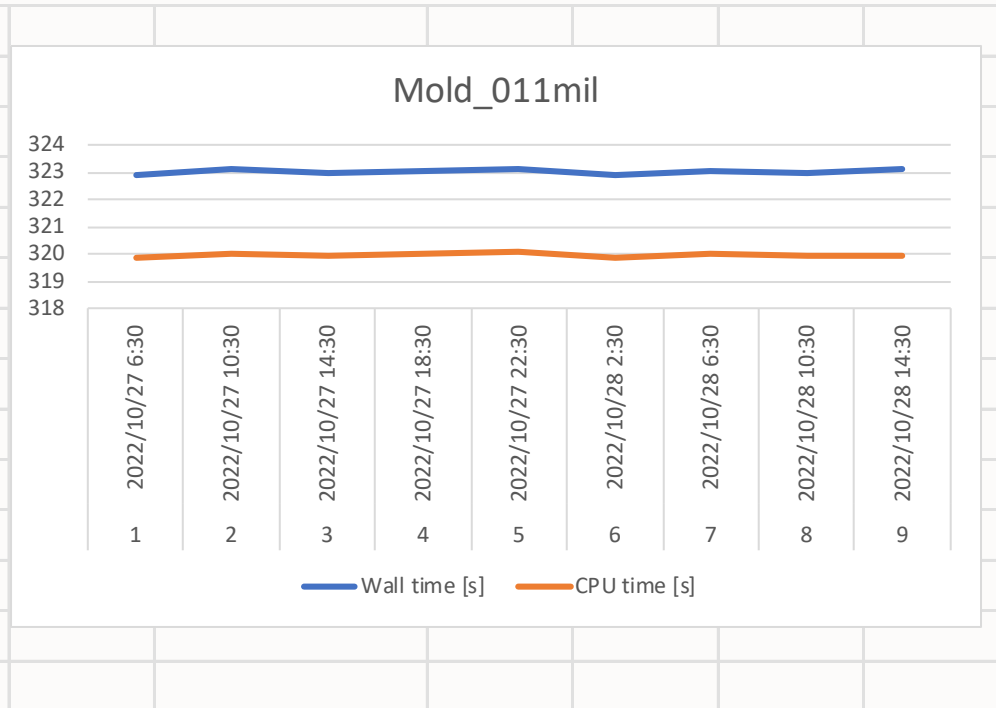
計算ノード
Optimized3 ベアメタル x 16ノード
[1ノードあたり]
Intel Xeon 6354 3.0-3.6GHz 18コアx2
512GBメモリ
100GB ブロックボリューム
3.8TB NVMe SSD
100Gbps RoCEv2 (遅延片方向 1.5μs 程度)
50Gbps TCP/IP

File Storage : NFS v3 マネージドサービス
保存データ量課金
(SIZEは8EBとしてレポートされます)



Mold_011mil (MPIプロセス: 16、SMPスレッド: 4)

| # | 実施時間(現地時間) | Wall time [s] | CPU time [s] |
|---|------------------|---------------|--------------|
| 1 | 2022/10/27 6:30 | 322.88 | 319.85 |
| 2 | 2022/10/27 10:30 | 323.12 | 320.04 |
| 3 | 2022/10/27 14:30 | 322.98 | 319.91 |
| 4 | 2022/10/27 18:30 | 323.05 | 320.01 |
| 5 | 2022/10/27 22:30 | 323.15 | 320.1 |
| 6 | 2022/10/28 2:30 | 322.91 | 319.88 |
| 7 | 2022/10/28 6:30 | 323.07 | 320.02 |
| 8 | 2022/10/28 10:30 | 322.98 | 319.96 |
| 9 | 2022/10/28 14:30 | 323.1 | 319.94 |
| | Max | 323.15 | 320.1 |
| | Min | 322.88 | 319.85 |
| | 誤差 | 0.08% | 0.08% |



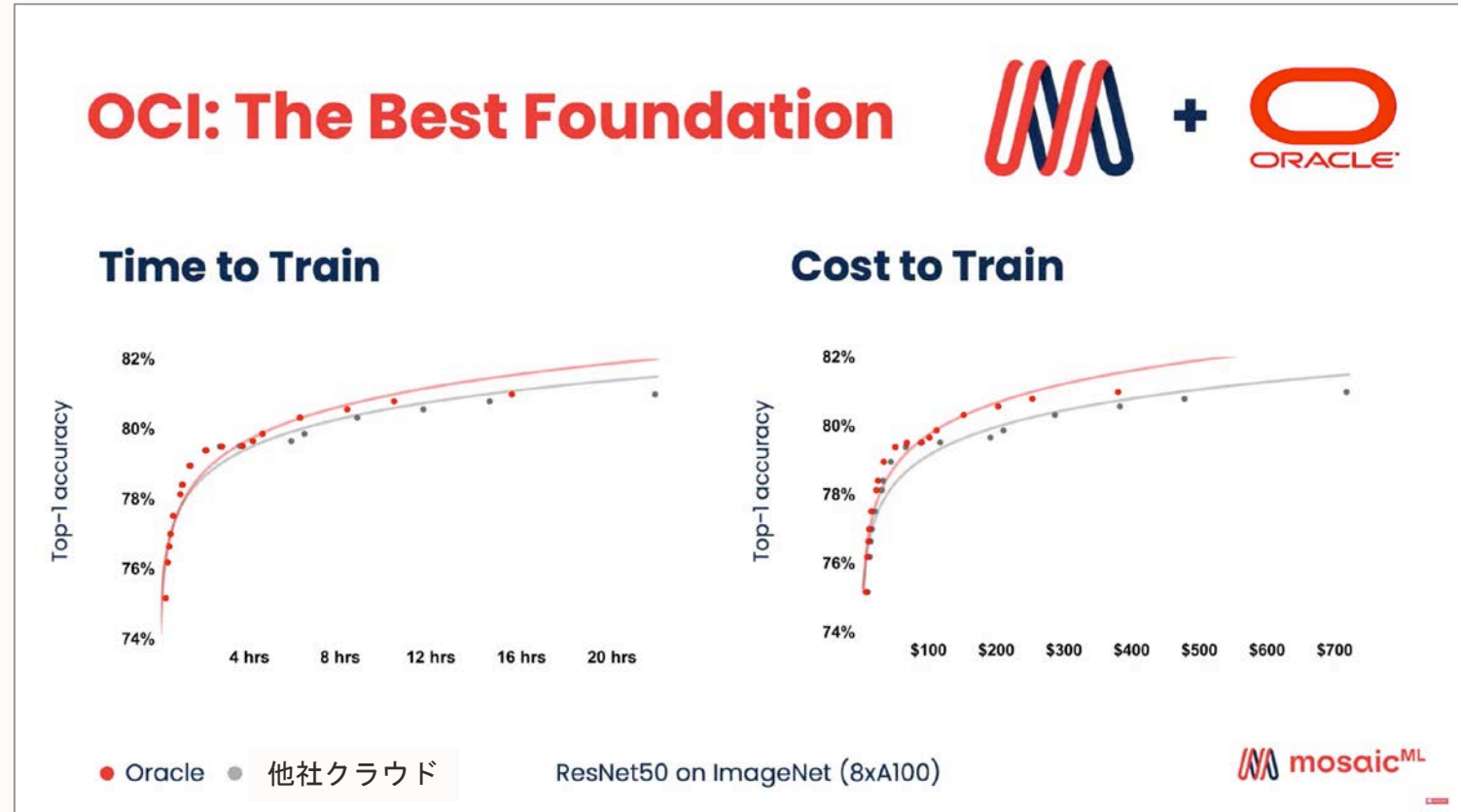
MosaicML社 A100ベンチマーク

MosaicMLによるOCIと他社クラウドとのGPU(A100)パフォーマンス比較

「我々は、OCIがベストな基盤であることを発見した。

精度80%付近でOCIを他社クラウドを比較した場合、**OCIはAWSの半分の時間、半分のコスト。**

これは非常に大きい。
\$200と\$100の場合は気にすることはないかもしれないが、これが\$200,000と\$100,000だと気にする必要がある。」



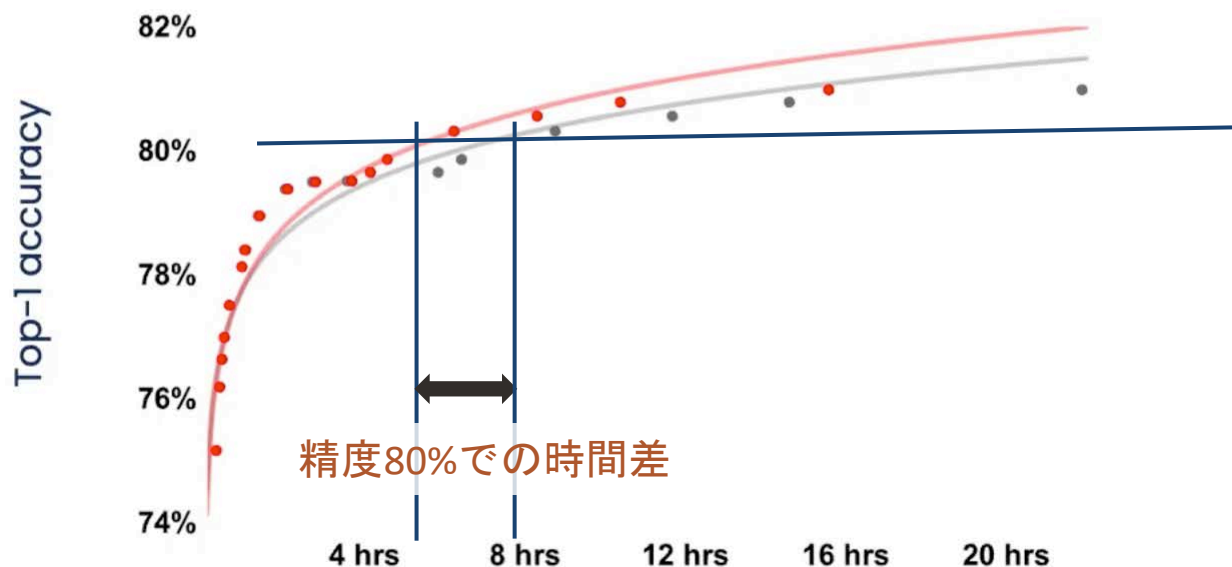
<https://youtu.be/IAZPpUVjNJ0>



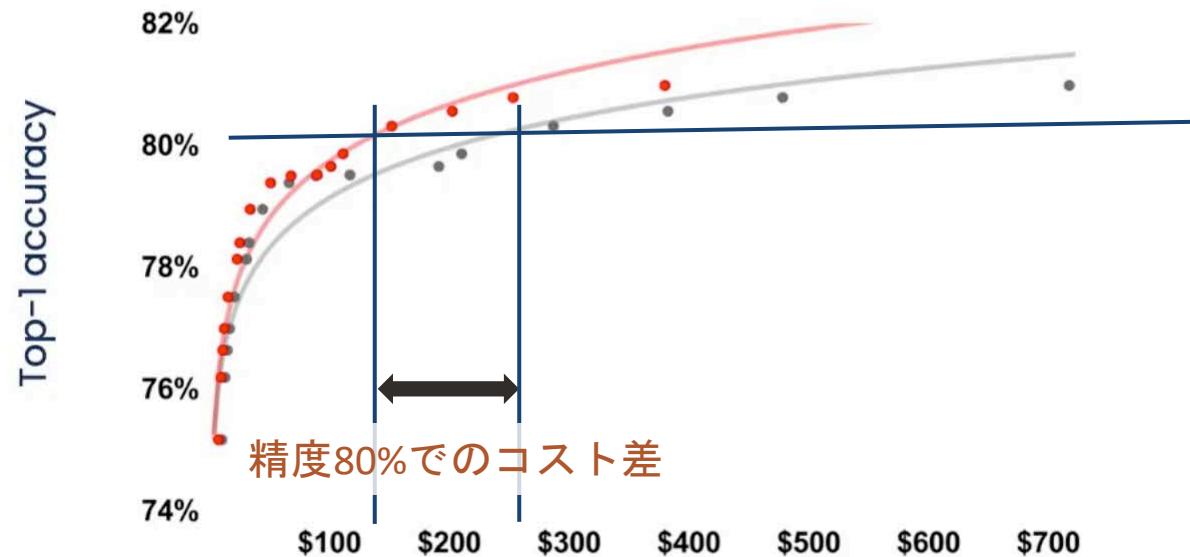
OCI: The Best Foundation



Time to Train



Cost to Train



● Oracle ● 他社クラウド

ResNet50 on ImageNet (8xA100)



Oracle AI

セキュアで高性能な**エンタープライズAI**



学習済み AI Services

API経由で利用、アプリケーションにすぐ組み込み可能





セキュアで高性能な、 企業向け生成AIサービスを提供

- お客様がデータを所有して制御。**お客様のデータが完全に分離された、セキュアな生成AI**を提供します。データの出所やシステムへのアクセスも可能になる予定で、安心してビジネスでお使いいただけます。
- **Oracle独自の業界知識、データを活用**してユースケースにあった最適なAIを使用可能です。また、**お客様のデータをベクターストアとして整備して専用の生成AIをセキュア且つ安価に作成**することも可能です。
- パブリック・クラウドだけではなく、**オンプレミスのデータセンターでも使用できます**。よりセキュアにデータを活用でき、オンプレミスのアプリケーションで高性能な生成AIを活用できます。

「OracleとCohere社は、データ・セキュリティ、モデルのカスタマイズ、企業のビジネス価値創造を実現することに共通の重点を置いています。両社は共に、データの安全性とプライバシーを確保しながら、世界中の企業がAIへの取り組みを加速させ、価値を高め、ビジネスの成功を最大化する新しいレベルの自動化を実現できるよう支援します。」

Martin Kon, President & COO, Cohere

ORACLE AI

Oracle Applications +  cohere

 cohere + ベクターストア

OCI セキュリティ (顧客データの完全な分離)

OCI Supercluster
AI Infrastructure | NVIDIA GPU



Oracle Cloud Lift Services

- OCI ... Oracle Cloud Infrastructure
- ExaDB-D ... Oracle Exadata Database Service
- OCVS ... Oracle Cloud VMware Solution

サービス概要

- お客様のクラウド移行をご支援する**無償**サービス
- 現行システムのクラウドへの移行検討で妨げとなる課題に直面しているお客様をご支援

お客様のベネフィット

1. クラウド移行に精通したオラクルの**専任エンジニア**が、移行のご支援をいたします
2. OCIのノウハウがないお客様でも、環境を**迅速に構築**し、ご利用頂くことができます
3. お客様がクラウド環境の準備に伴う**初期費用や時間を抑える**ことができます
4. 併せてお客様プロジェクトメンバーに対して、クラウド環境の利用・運用するための知識を持てるよう、**スキルトランスファー**のご支援をさせていただきます

主なサービス内容

クラウド移行に向けた“ケーススタディ支援”

クラウドへの移行事例を共有し、新たな観点でのクラウドメリットを可視化

クラウド移行に向けた“フィジビリティスタディ支援”

現行システムの評価を検証し、クラウド移行の課題とロードマップを可視化

クラウド移行に向けた“実機検証 (PoC) の支援”

PoCを通じて、クラウド移行の問題やリスクに対するソリューションを検証

プロジェクト初期フェーズの“早期立ち上げ支援”

クラウドサービスを利用して、検証・開発環境の導入、立ち上げをご支援

主なワークロード

Oracle Database
On Cloud

Oracle DBを
OCIのPaaSに移行する

VMware
on
Cloud

VMware環境を
OCIのOCVSに移行する

High Performance
Computing
applications on
Cloud

HPCアプリケーションを
OCIのHPC環境に移行する

※その他ワークロードに関しましては、担当営業にご相談ください





Fugaku

Oracle Cloudで「富岳」の高度な計算資源の有効活用と研究成果創出を促進

- パブリック・クラウドと連携した「富岳」の柔軟な利活用の支援を目的に、学術情報ネットワークSINETを介した「富岳ダイレクト接続サービス」を整備
- 膨大な研究データの転送コストを気にすることなく、「富岳」のネットワークから、Oracle Cloud Infrastructureの高い性能のコンピュータやストレージ・リソースなどを予測可能なコストで利用可能
- 安全かつ低コストでの接続により、セキュリティ、パフォーマンス、伸縮性に優れたコンピュータやストレージのリソースを「富岳」のネットワークから低コストで利用可能に



大阪大学、日本電気共同、日本オラクル共同研究事例

～ 新型コロナウイルス感染症などに向けクラウドバースティング

を通じたベアメタル計算資源の提供～



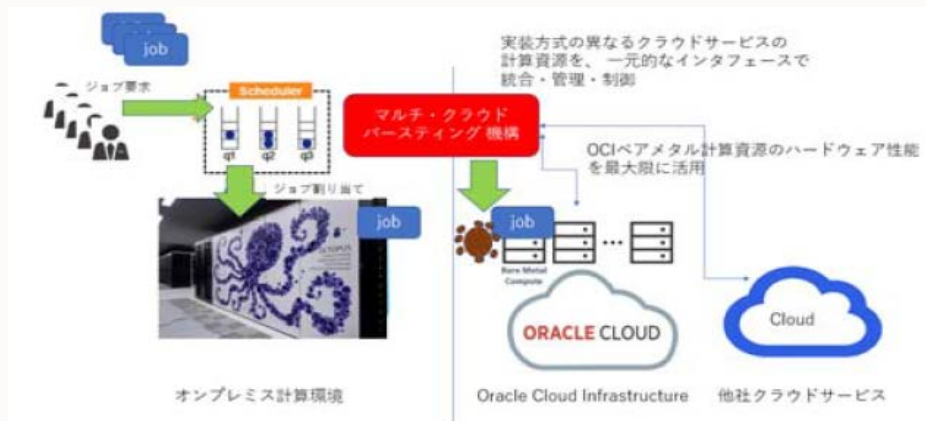
大阪大学
OSAKA UNIVERSITY

研究背景

- CMC のスーパーコンピュータOCTOPUS は、利用者からのスカラ型スーパーコンピュータに対する様々な計算ニーズ・需要を収容可能であり、非常に高い利用率で利用される状況になっています。しかし、その一方で、利用者の**計算要求から計算完了までの待ち時間が定常的に長時間になるという問題が深刻**になりつつあり、利用者からの問い合わせ・相談の声も大きくなっていった。

ソリューション

- Oracle Cloud Infrastructure bare metal



図：マルチクラウドバースティング機構を通じた OCI ベアメタルクラウド計算資源.

Why Oracle Cloud

- ベアメタルマシンも仮想マシンと同様にオンデマンドに必要な時に起動し、**不要な場合は停止するHPC環境を構築**することが可能
- 最新の**CPU、GPU、高性能なノード間通信等**のHPC 関連技術がリリースされた際には**迅速に提供**される。
- 一般的にHPC用途では多くの分野でデータが大きくなる傾向があるが専用ネットワーク接続サービスを利用することで、**転送データ量が無制限に無料**
- OCI のベアメタル計算資源であれば、**オンプレミスのOCTOPUS 計算ノードと同様にユーザの計算要求を実行出来る**ことが確認された
- 多くの計算ノードを利用する**並列計算**においてクラウド計算資源を利用した場合でも**高いスケーラビリティが得られる**

【研究成果のポイント】

大型計算機におけるクラウドバースティングが実現可能であることが実証されると利用者が**計算結果を得られるまでの時間の削減**ができるだけでなく、**COVID-19 対策のような急な計算需要拡大への対応**を含め、**計算機リソースの問題で解決できなかった事象を解析**できるようになる。

これにより、大型計算機を用いた研究分野において、**学術的・教育的に大きな成果**が出ることが期待されます。

また、**低コストでのスケールアウト**を示すことにより、**学術機関と企業との連携による産業利用・産学連携の加速、企業・社会課題の解決が加速**されることが期待されます。



Oracle Cloud Infrastructureを用いた Hybrid HPC システムイメージ

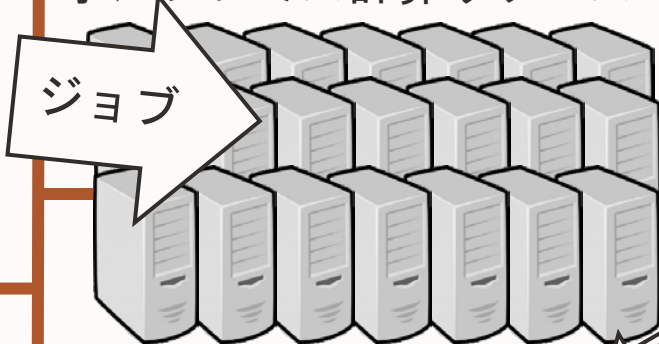


プリポスト、フロントエンド、アプリ開発等の専有利用

小さいジョブやコプロセッサ処理
もしくは一部の一連の処理

オンプレミス計算リソース

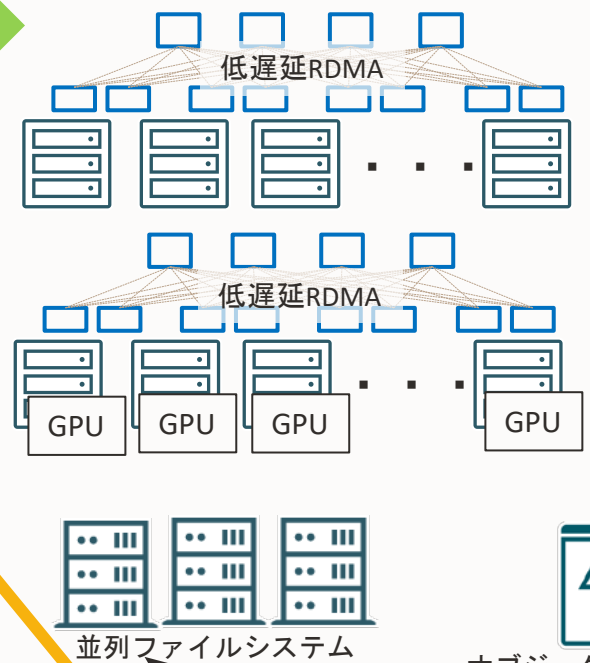
ジョブ



適宜データコピー
or 直接マウント

一定期間アクセスの
無いファイルをオフロード

Cloudbursting



- ジョブスケジューラと連携して必要なクラウドリソースを起動/停止
- 適宜最新HWに更新

IoTデバイスデータ

フィールドAIデバイスデータ



Storage-offload



ユーザはリソースの場所を意識しない

ジョブスケジューラが自動的にジョブ分配/
データステージング



【お問い合わせ先】



日本オラクル株式会社

クラウド事業統括 公共営業本部

松山 慎 (まつやま まこと)

E-Mail: Makoto.Matsuyama@oracle.com

電話: 080-1289-8315

お気軽にご連絡ください!



ORACLE