



HPC ワークロードを加速する HPC インスタンスと構成パターン

夏の Amazon EC2 祭り 2023 最新インスタンス活用編

小林 広志

Solution Architect, HPC

アマゾン ウェブ サービス ジャパン 合同会社

小林 広志

アマゾン ウェブ サービス ジャパン 合同会社
ソリューションアーキテクト, HPC

グローバル製造業にて機械系CAEエンジニアとして
プロダクト開発を経験。その後情シス部門に移動し、
クラウドHPCとデータ分析基盤の導入と運用を担当。

AWS 上で大規模な計算を行われるお客様の技術支援を担当



幅広い業界での HPC の活用



自動運転のシミュレーションと機械学習



ゲノム解析 & 創薬シミュレーション



ビッグデータ解析



機械学習



金融リスク分析



流体解析
半導体設計



再生可能エネルギー
石油 & ガス探索



気象予報

密結合 vs. 疎結合



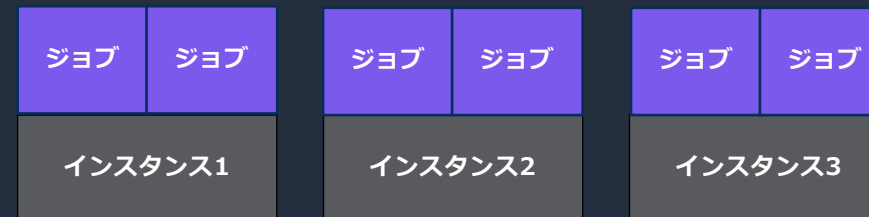
密結合 ワークロード

- 大規模な単一のジョブを高速に実行するために複数のノードをまたいで処理を行う
- 多くのアプリケーションでは、ノード間の通信によるボトルネックのためスケール性能に限界がある
- 例：流体計算(CFD)、分子動力学シミュレーション、気象シミュレーション、ML 分散学習



疎結合 ワークロード

- 小規模な処理を大量に実行するために、複数のノードで独立したジョブを処理する
- 基本的にノード間の通信が発生しないため、スケールさせやすい
- 例：パラメータ最適化、創薬ドッキングシミュレーション、半導体設計シミュレーション (Electronic Design Automation)



疎結合ワークロード 構成パターン



疎結合 ワークロード

- 小規模な処理を大量に実行するために、複数のノードで独立したジョブを処理する
- 基本的にノード間の通信が発生しないため、スケールさせやすい
- 例：パラメータ最適化、創薬ドッキングシミュレーション、半導体設計シミュレーション (Electronic Design Automation)



構成上の注意点

- 広大な計算領域をカバーするための
- 計算資源のキャパシティ確保
- 計算資源のコスト最適化

利用サービス

- Amazon Elastic Compute Cloud (Amazon EC2), Amazon EC2 Spot instances, AWS Batch

大規模データ処理を支える共通基盤に AWS を活用。最大 30,000CPU コアの処理 を実現し、ビジネスと開発の加速に貢献



ビジネス課題

- さまざまなサービスを運用するうえで、広告データ、ログデータ、属性データなどを利用したデータ分析は必要不可欠。社内組織や社内プロジェクトで利用可能なデータ処理基盤として『Crois』を開発。

ソリューション

- コンテナの実行基盤には AWS Batch と Amazon ECS on AWS Fargate を採用。機械学習のような CPU、メモリー、ディスクを大量に消費するタスクや GPU インスタンスを利用するタスクでは AWS Batch。データウェアハウスへのクエリーのような軽量なタスクには短時間でスケールができる Amazon ECS on AWS Fargate を利用

導入効果

- 大規模な処理を複数同時に実行することが可能で、最大 30,000CPU コアの処理に対応
- 『Crois』を導入することで、汎用モジュールの利用やユーザーが自作したモジュールの再利用などにより開発効率が上がり、リードタイムは 2 分の 1 に短縮。管理運用も AWS のマネージドサービスの利用により、年間で最大 10 人月程度の運用工数の削減の見込

“ 社内向け大規模データ処理基盤『Crois』に AWS のマネージドサービスを活用することで、リクルートのビジネスに求められる高いアジリティとスケーラビリティを実現しています ”

大石 壮吾 氏

株式会社リクルート プロダクト統括本部 プロダクト開発統括室 データ推進室長

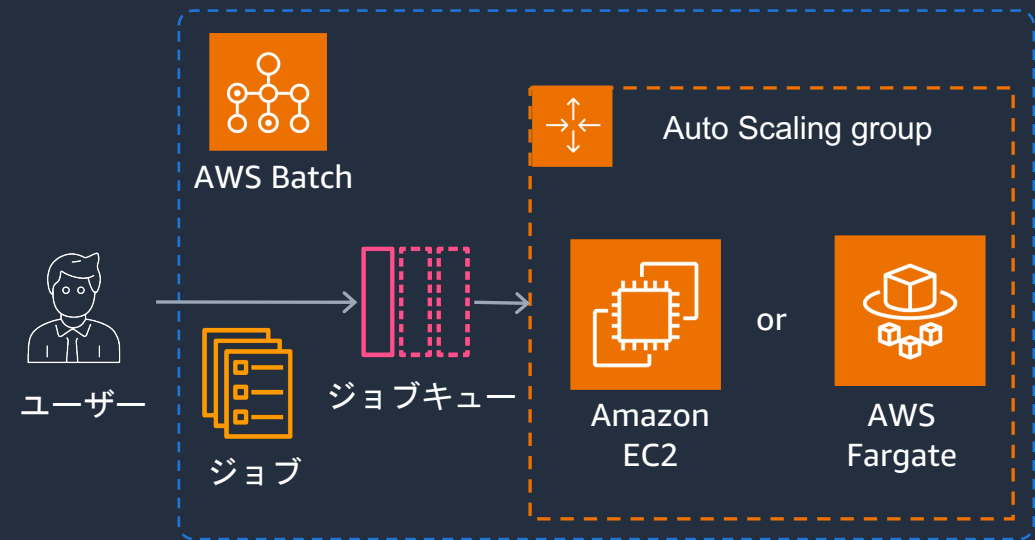
<https://aws.amazon.com/jp/solutions/case-studies/recruit-case-study/>



AWS Batch

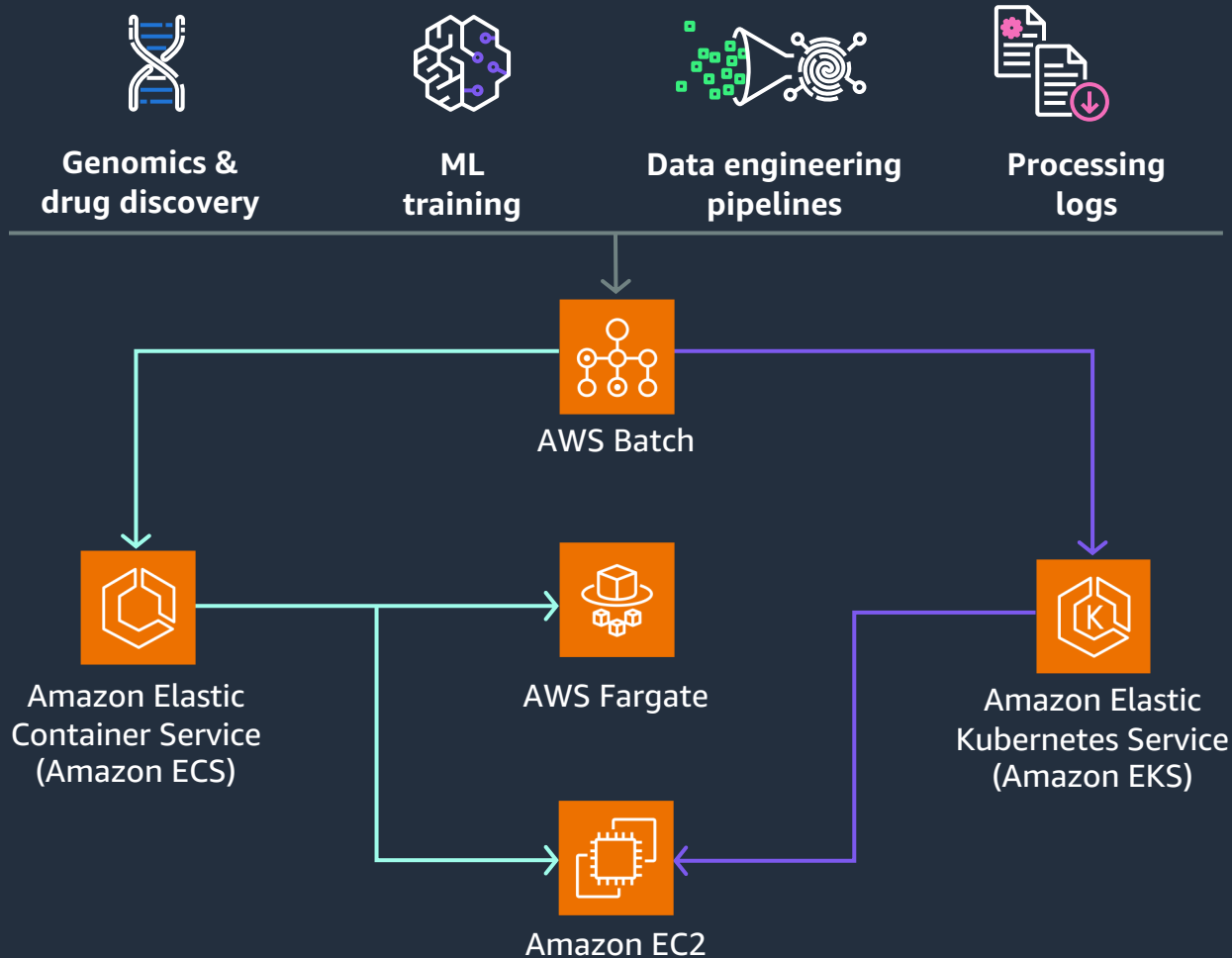
大規模バッチ処理のため環境をフルマネージドで提供

- AWS Batch がインスタンスの起動や停止を行うため、**スケジューラや計算ノードなどの管理が不要**
- ジョブは **Docker コンテナイメージ**を元に作成し、自動でスケールするコンピューティング環境で実行する
- コンピューティング環境ではインスタスタタイプや vCPU 数、スポットインスタンス利用有無などを任意に指定可能



コンテナイメージを用意するだけでスケーラブルな大規模バッチ処理環境が得られる

AWS Batch: Platform



Container Orchestration

- Amazon ECS or Amazon EKS

Fargate

- Small jobs (less than 16vCPU and 128GiB memory)
- require quick response time
- Pure serverless environment

EC2

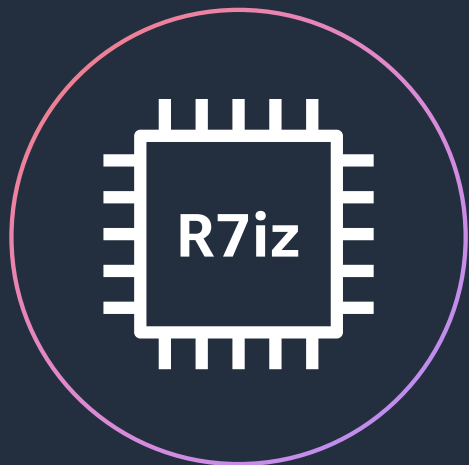
- Large workloads that require high scalability (> 1M +)
- Customizable Compute Environments
- Can launch any size instance

Amazon EC2 R7iz インスタンス

PREVIEW

高クロック周波数のCPUを搭載したメモリ最適化インスタンス

- 第4世代 Intel Xeon Scalable Processor(Sapphire Rapids) を搭載し、最大 3.9GHz で動作
- 最大 128vCPU、1,024GiB の DDR5 メモリ。最大 50Gbps のネットワーク帯域
- 高クロック周波数の CPU により、疎結合タイプのジョブを高速に実行



	z1d.12xlarge	m5zn.12xlarge	x2iezn.12xlarge	r7iz.32xlarge
vCPU数	48	48	48	128
最大クロック周波数	4.0 GHz	4.5 GHz	4.5 GHz	3.9 GHz
RAM	384 GiB (8GB/vCPU)	192 GiB (4GB/vCPU)	1536 GiB (32GB/vCPU)	1024 GiB (8GB/vCPU)
ネットワーク帯域幅	25 Gbps	100 Gbps	100 Gbps	50 Gbps
EFA 対応	なし	あり	なし	なし
価格※	\$ 4.464 USD/h	\$3.9641 USD/h	\$10.008 USD/h	TBD



EC2 Spot Instances are how you can do more for less

On-Demand

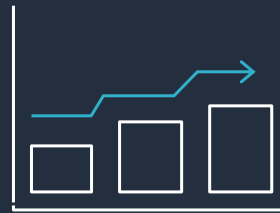
Pay for compute capacity by the second with no long-term commitments



Spiky workloads, to define needs

Savings Plans and Reserved

Make a 1 or 3 year commitment and receive a significant discount off On-Demand prices



Committed and steady-state usage

Spot Instances

Spare Amazon EC2 capacity at savings of up to 90% off On-Demand prices



Fault-tolerant, flexible, stateless workloads

Spot is more than cost savings...

More Compute



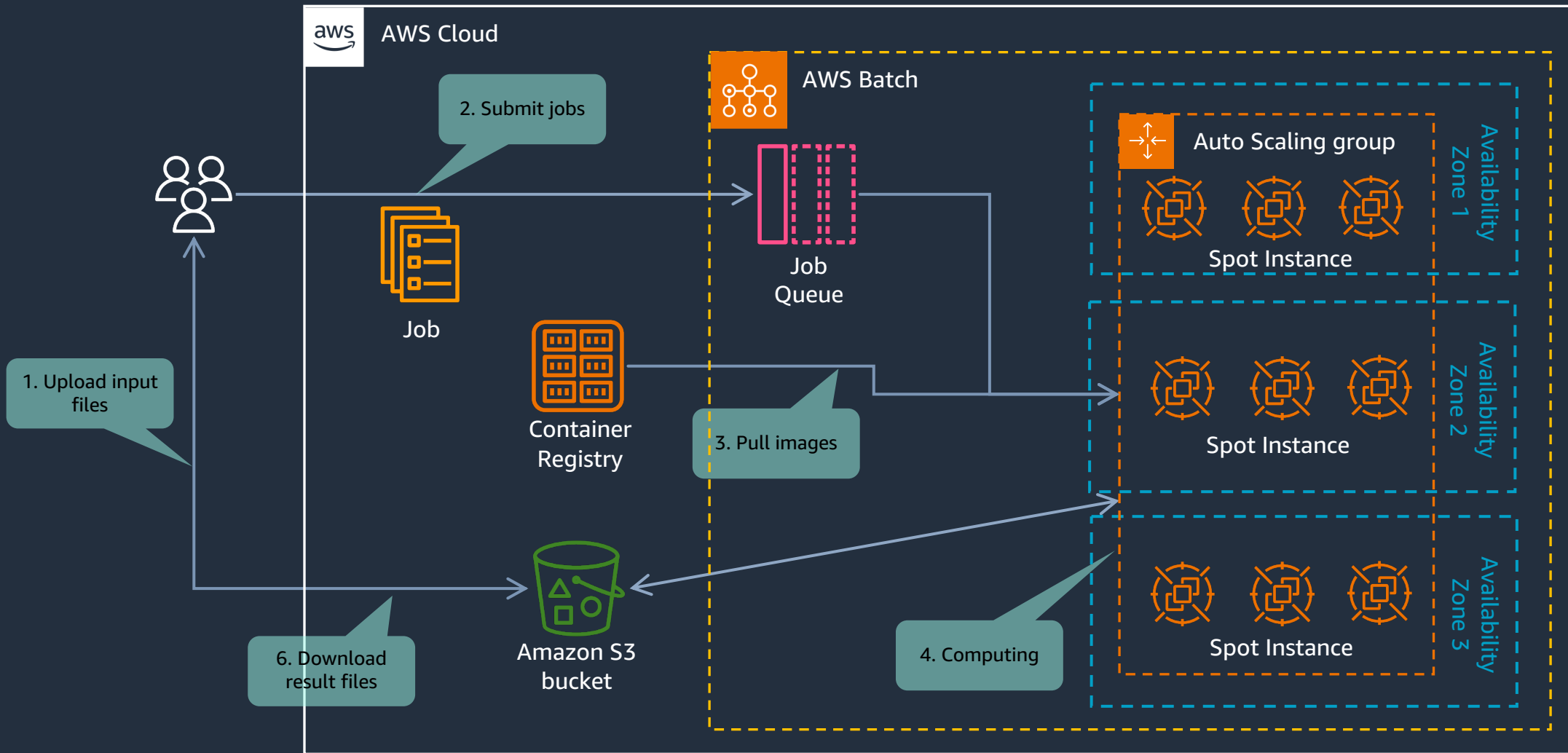
Faster Results



Accelerated Innovation



疎結合ワークロード アーキテクチャ



密結合ワークロード 構成パターン



密結合 ワークロード

- 大規模な単一のジョブを高速に実行するために複数のインスタンスをまたいで処理を行う
- 代表的なワークロード
 - 流体計算(Computational Fluid Dynamics)
 - 気象シミュレーション
 - 大規模 AI モデルの分散学習

構成上の注意点

- 大量のインスタンス間通信を処理するための低遅延・広帯域のネットワーク
- たくさんの計算インスタンスからデータを高速に読み書きできる高スループットなストレージ

利用サービス

- Amazon EC2, Elastic Fabric Adaptor, Amazon FSx for Lustre



Baker Hughes : AWS HPC により、シミュレーション実行時間、CO2排出、HPCコストを削減



Customer: Baker Hughes Inc
Industry: Oil and Gas
Country: United States
Segment: Large Enterprise
Website: Bakerhughes.com

企業プロフィール:

Baker Hughes は120カ国に拠点を持ち、約54,000人の従業員を擁する、エネルギーテクノロジーのリーディング企業です。エネルギーを前進させるための革新的な技術の設計、製造、サービスを提供しています。



課題

- Baker Hughes社のエンジニアは、ガスタービン設計のシミュレーションにオンプレミスのHPCクラスターを使用していましたが、計算容量の拡大が課題となっていました。
- エンジニアは、長いシミュレーションの待ち時間や長い実行時間の問題に直面し、物理的なプロトタイプの必要性が高まっていました。



解決法

- Baker Hughes社は、計算流体力学（CFD）アプリケーションをオンプレミスからAWSへ移行することを選択しました。
- AWSプロフェッショナルサービスは、数週間以内に予算内で概念実証を実施し、最高のランタイムパフォーマンスを証明しました。



結果

- Baker Hughes社は、HPCコストを40%削減し、待機時間を98%、実行時間を26%、HPCソリューションの二酸化炭素排出量を99%削減し、結果までの時間を短縮することに貢献しました。

“AWS上でAnsysシミュレーションを実行することで、タービン開発のスケジュールの短縮と市場投入までの時間の短縮を実現”

- David Meyer, Director of Digital Operations for HPC and Remote Visualization, Baker Hughes



Elastic Fabric Adapter

MPI/NCCL で利用可能な低レイテンシネットワークアダプタ

c6in.32xlarge, c7gn.16xlarge, p4dn.24xlarge 等の EFA 対応インスタンスで利用可能

OpenMPI, Intel MPI, MVAPICH2 が対応

通信プロトコルとして、TCP/IP ではなく、独自の SRD (Scalable Reliable Datagram) を使用

インスタンス間通信速度

C6in	<u>c6in.32xlarge</u> 200 Gbps
C7gn	<u>c7gn.16xlarge</u> 200 Gbps
P4d	<u>p4d.24xlarge</u> 400 Gbps (GPUDirect RDMA対応)

計算ノード間の通信レイテンシが重要となる処理例



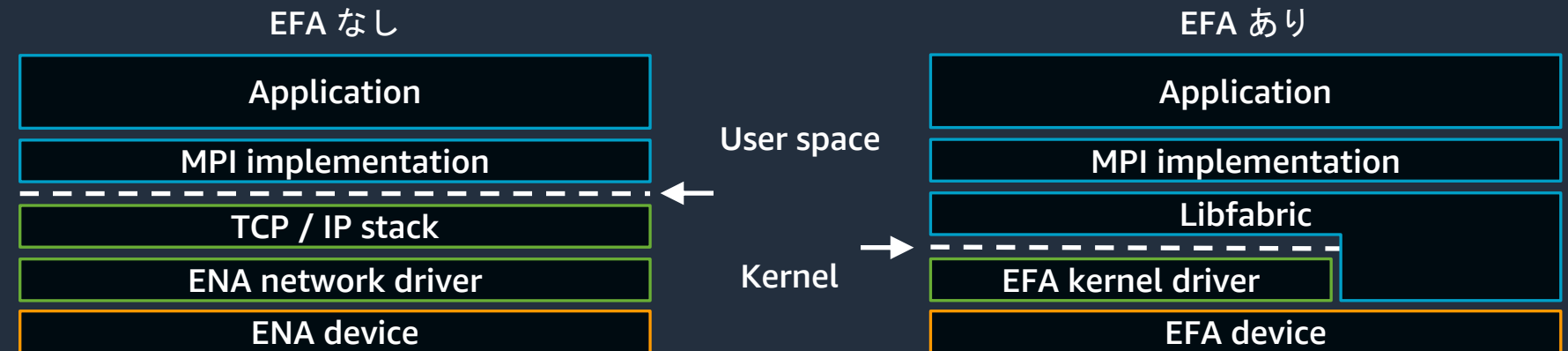
流体解析



地震
シミュレーション



気象予測



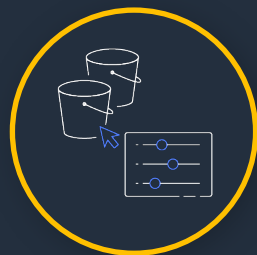
Amazon FSx for Lustre



ハイパフォーマンスなワークロードに対応するフルマネージドLustreファイルシステム



大規模にスケールする
パフォーマンス



データリポジトリへの
シームレスなアクセス



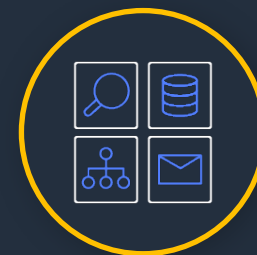
シンプルで
フルマネージド



ネイティブファイル
システムインターフェース



コンピューティンテンシブ
ワークロードに最適化されたコスト



セキュア
且つコンプライアンス
にも容易に準拠



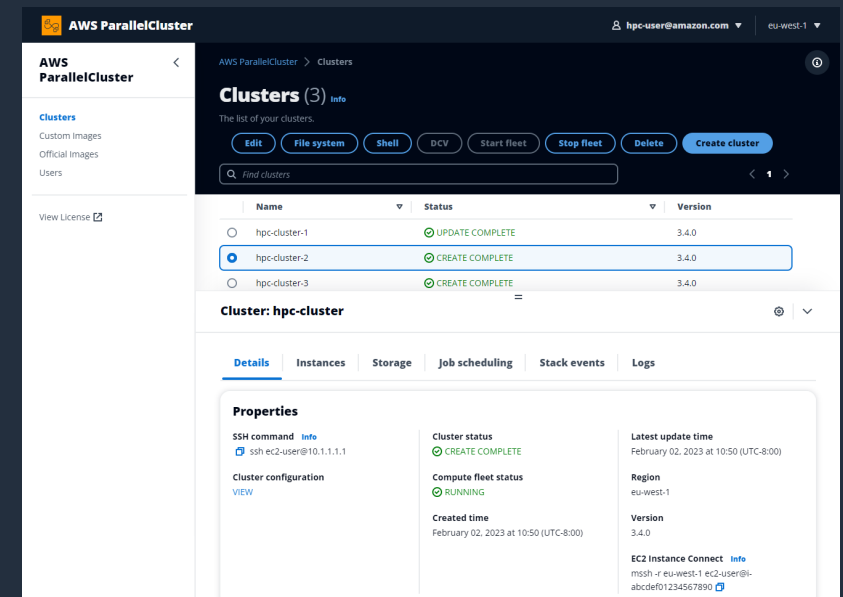
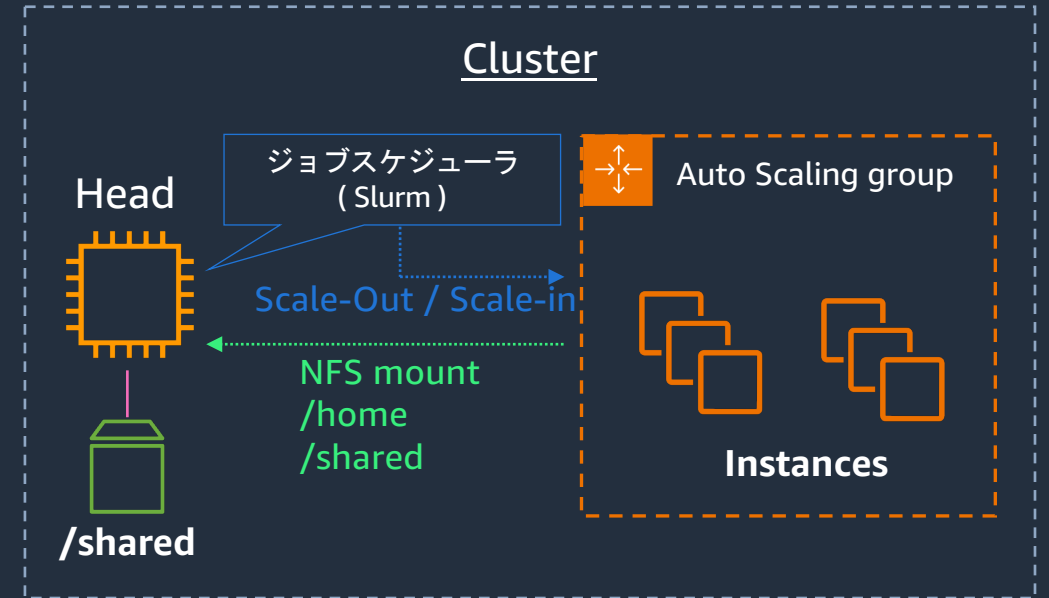
AWS ParallelCluster

ジョブ投入に応じて自動でスケールするクラスタを AWS 上に構築可能な AWS 公式のオープンソースソフトウェア

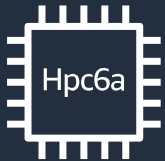
AWS ParallelCluster の特徴

- 既存の HPC 向けスケジューラ（Slurm）と連動しスケラブルな計算環境を作成
- MPI / NCCL 環境がセットアップ済み
- x86 / Arm 両方の環境に対応
- 使用するOSやネットワーク環境、ストレージ構成などを柔軟にカスタマイズ可能
- コマンド / グラフィカル のインターフェース
- オープンソースプロジェクトであり、誰でもソースコードを入手可能

<https://github.com/aws/aws-parallelcluster>



HPC特化型 Amazon EC2 インスタンス ラインナップ



Powered by third-generation AMD EPYC processors

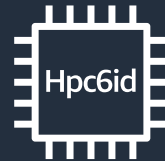
Amazon EC2 Hpc6a instances

Compute-intensive applications like Computational Fluid Dynamics and Numerical Weather Prediction

3.6Ghz 96 cores AMD Milan

384GB RAM

100Gbps EFA



Powered by third-generation Intel Xeon Scalable processors

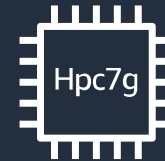
Amazon EC2 Hpc6id instances

Memory-bound and data-intensive workloads like Finite Element Analysis and seismic simulations

3.5Ghz 64 cores Intel IceLake

1024GB RAM | 15.2TB NVMe

200Gbps EFA



Powered by the next generation AWS Nitro System

Amazon EC2 Hpc7g instances

Based on custom AWS Graviton3E processors with low latency, and high network performance for MPI-based applications

2.6Ghz 64 cores Graviton3E

128GB RAM

200Gbps EFA



Amazon EC2 HPC6a インスタンス

Amazon EC2で実行されるHPCワークロードに対して最高の価格パフォーマンスを提供するように設計



AMD

AMD 第3世代 EPYC Milanプロセッサを 96コア、および 384GB のメモリを搭載

同等の AmazonEC2 x86 ベースのコンピューティング最適化インスタンスよりも最大 65% 優れた価格パフォーマンス、**1物理コアあたり \$0.03/時間***の低コストで提供

100 Gbps のネットワーク帯域幅、EFA 対応

気象、金融、流体力学、ヘルスケアおよびライフサイエンスなど大規模な計算ワークロードに最適

オハイオ、ストックホルム、GovCloud (米国西部および東部)ならびに東京リージョン (※個別対応) で提供

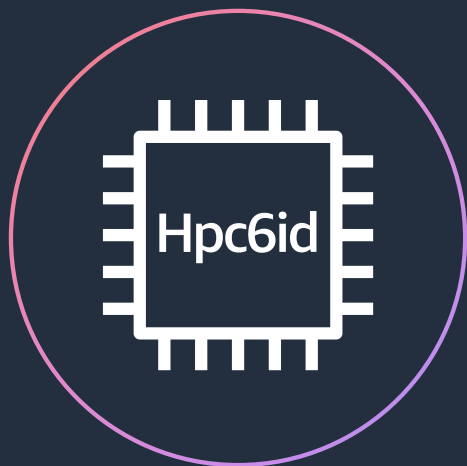
* Ohio リージョンにおける 2023 年 7 月時点のオンデマンド価格

<https://aws.amazon.com/jp/blogs/news/new-amazon-ec2-hpc6a-instance-optimized-for-high-performance-computing/>

Amazon EC2 HPC6id インスタンス

AVAILABLE

メモリ・データ集約的な HPC ワークロードに適したインスタンス

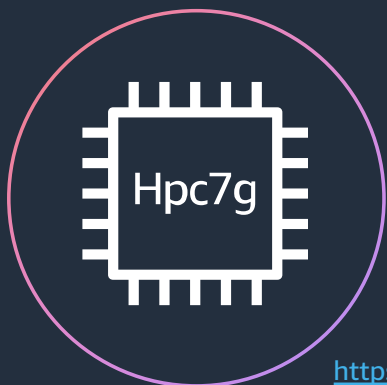
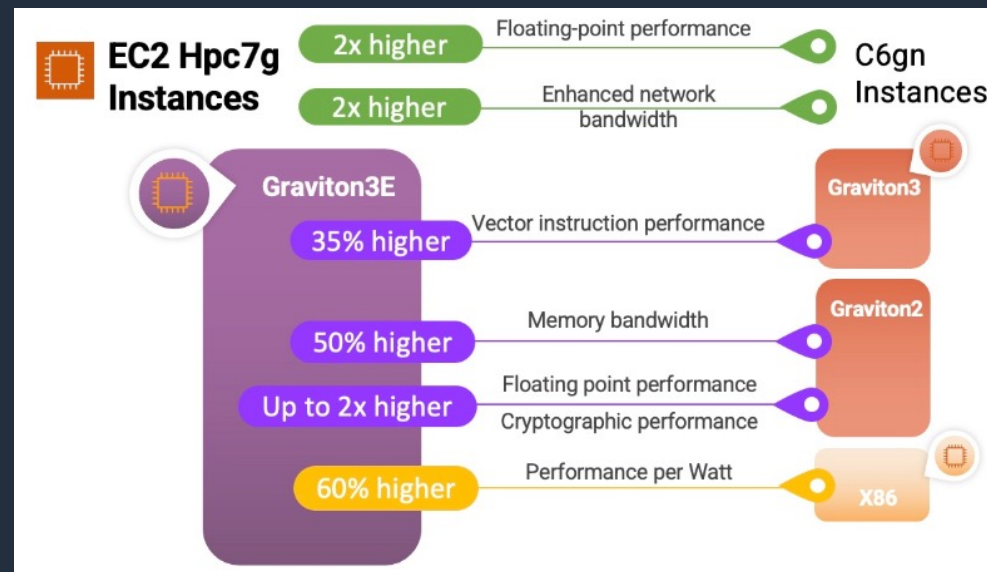


- 第三世代 Intel Xeon Scalable Processor (Ice Lake) を搭載し、最大 3.5GHz で動作。64 物理コア、1024GiB メモリ、15.2TB の SSD を搭載。200Gbps の EFA ネットワークに対応
- 有限要素法解析のような広いメモリバンド幅が求められるワークロードに最適
- 1物理コアあたり \$0.089/ 時間*
- オハイオ、フランクフルト、GovCloud (米国西部) で一般利用開始

Amazon EC2 HPC7g インスタンス

AVAILABLE

- AWS **Graviton3E** 搭載、最新の Nitro v5 プラットフォームによる第二世代 EFA 採用し **200 Gbps** のネットワーク帯域
- 前世代のC6gn インスタンスと比較して**2倍の浮動小数点演算性能**、同等のx86インスタンスと比較して最大**60%高いエネルギーあたりの計算性能**
- 気象予測、流体解析、金融リスク計算等のワークロードに最適
- N.Virginia リージョンにて2023年6月より一般提供開始

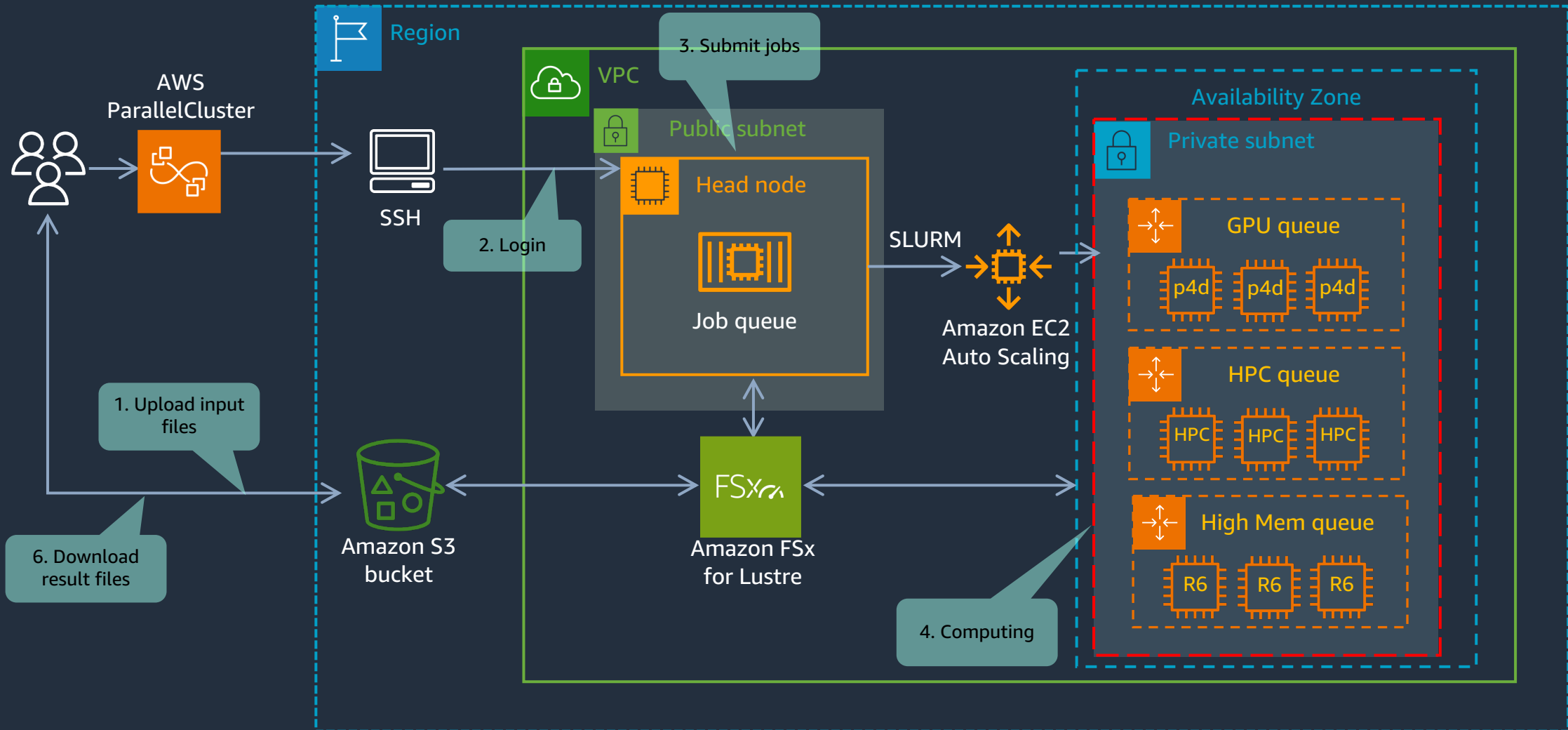


Instance	vCPU	memory	CPU Freq	EFA Network	On-Demand Hourly rate
hpc7g.4xlarge	16	128GB	2.6GHz	200Gbps	\$1.6832
hpc7g.8xlarge	32				
hpc7g.16xlarge	64				

<https://aws.amazon.com/blogs/aws/new-amazon-ec2-hpc7g-instances-powered-by-aws-graviton3e-processors-optimized-for-high-performance-computing-workloads/>



密結合ワークロード アーキテクチャ



まとめ

- HPC ワークロードの特徴とそれに合わせた AWS の構成パターンと EC2 インスタンスを紹介
- お客様のワークロードに合わせた構成をとっていただくことで、効率的に HPC ワークロードを実行しビジネスを加速



Thank you!

Hiroshi Kobayashi
hiroskb@amazon.co.jp