



CBI 学会 2021 年大会 企業セッション SS-22
構造生物学研究におけるクラウド活用 の現在と展望

Amazon Web Services で 創（はじ）めるクラウド HPC

Specialist Solutions Architect, HPC
Daisuke Miyamoto
2021/10/27

自己紹介

□ 名前

宮本 大輔 (みやもと だいすけ)

□ 所属

アマゾン ウェブ サービス ジャパン 株式会社

技術統括本部

Specialist Solutions Architect, HPC



本日の概要

- クラウド HPC の特徴・利点
- HPC on AWS 関連サービス

山田先生、守屋先生のご講演に登場する
AWS のサービスを中心にご紹介

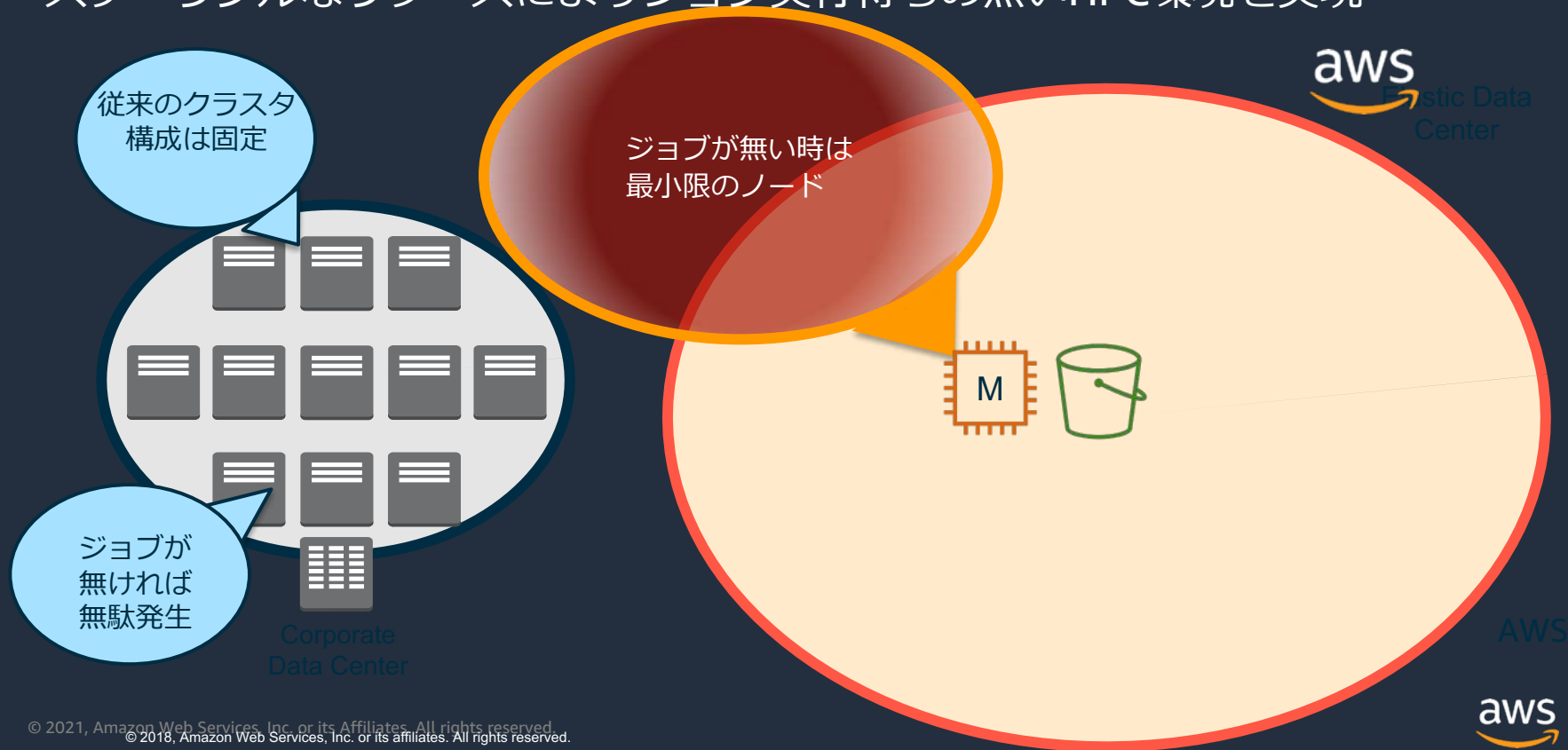
クラウド HPC の特徴・利点

これまでの HPC クラスタの課題

- サーバ台数が限られており、需要が増加する時期には長大なキュー待ち時間が発生する
- 同じ環境を複数メンバーで共有するため、アプリケーションによってはリソースが無駄になることも
- サーバ台数が多く、ハードウェアの保守・管理が煩雑

AWSなら、必要な時に必要なだけ利用可能

スケーラブルなリソースによりジョブ実行待ちの無いHPC環境を実現



AWSなら、必要な時に必要なだけ利用可能

スケーラブルなリソースによりジョブ実行待ちの無いHPC環境を実現

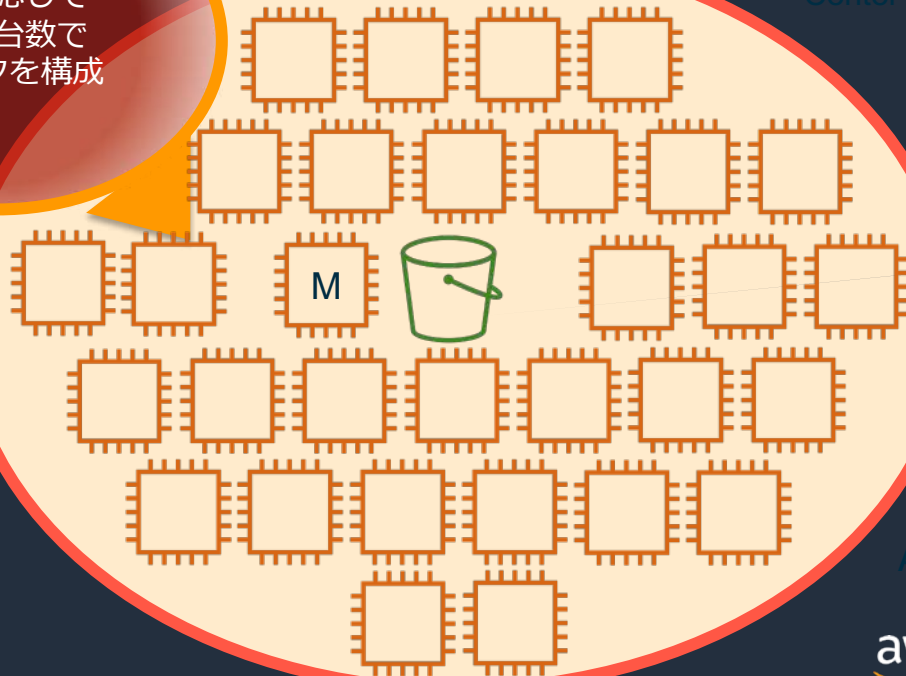
従来のクラスタ
構成は固定



ジョブが
無ければ
無駄発生

Corporate
Data Center

必要に応じて
必要な台数で
クラスタを構成



aws
Elastic Data
Center

AWS

aws

AWSなら、必要な時に必要なだけ利用可能

スケーラブルなリソースによりジョブ実行待ちの無いHPC環境を実現



従来のクラスタ
構成は固定



ジョブが
無ければ
無駄発生

Corporate
Data Center

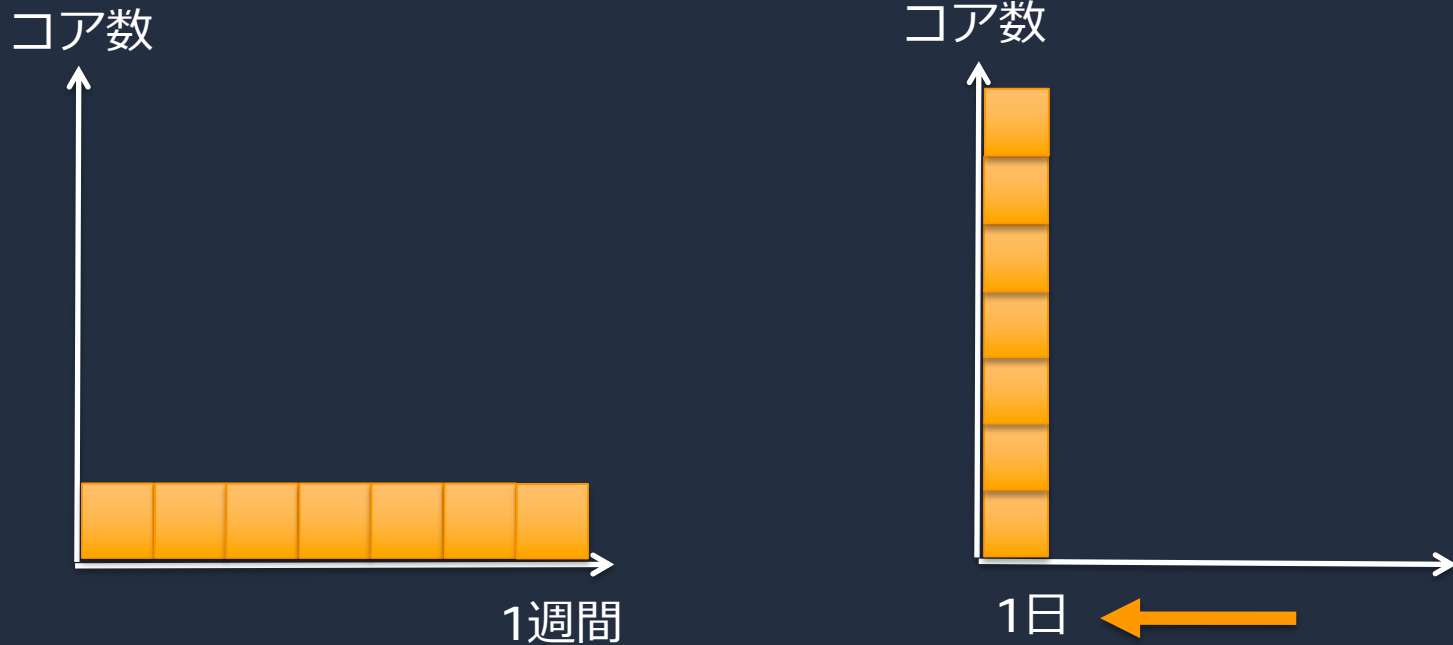
処理が終了すると
インスタンスを終了
課金停止



AWS



スケーラビリティの活用による計算時間短縮

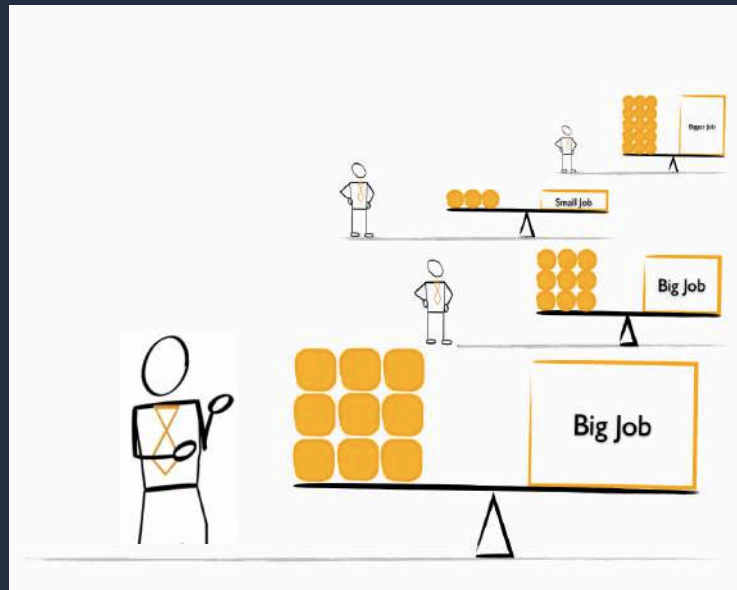


従来は手持ちの限られたリソースで、逐次処理していたジョブも
AWSなら必要な台数、インスタンスを起動して、一斉処理
しかも費用は「時間×台数」なのでどちらも同じ

アプリケーションに合わせた構成のクラスタを構築可能

ユーザやタスク単位で専用のクラスタを構築できるため
要件や規模に合わせて、最適構成のクラスタを作成可能

- CPUコア/メモリ
- ストレージ
- アクセラレータ
- ネットワーク
- インストールするソフトウェア



One size does not fit all!

計算機管理の手間を抑える

- ・ハードウェア保守
- ・ネットワーク管理/保守
- ・電源管理
- ・空調管理
- ・設置場所の費用/運用

計算機の規模が大きくなればなるほど
大変に、、、

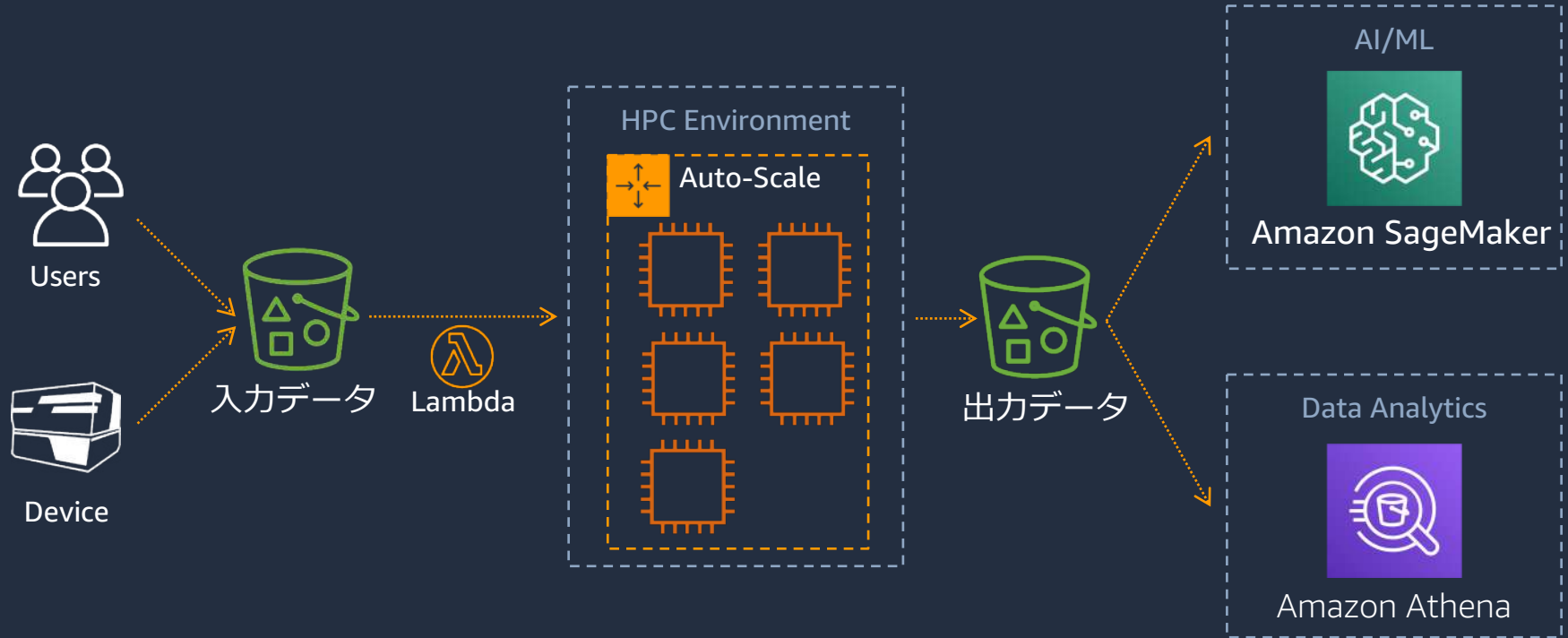


競争優位につながらない物理的管理は全てAWSにお任せ
他社と差別化可能な部分に集中

データドリブンなHPC環境とデータ活用

データのアップロードをトリガーにHPC環境を展開し自動で処理を行う

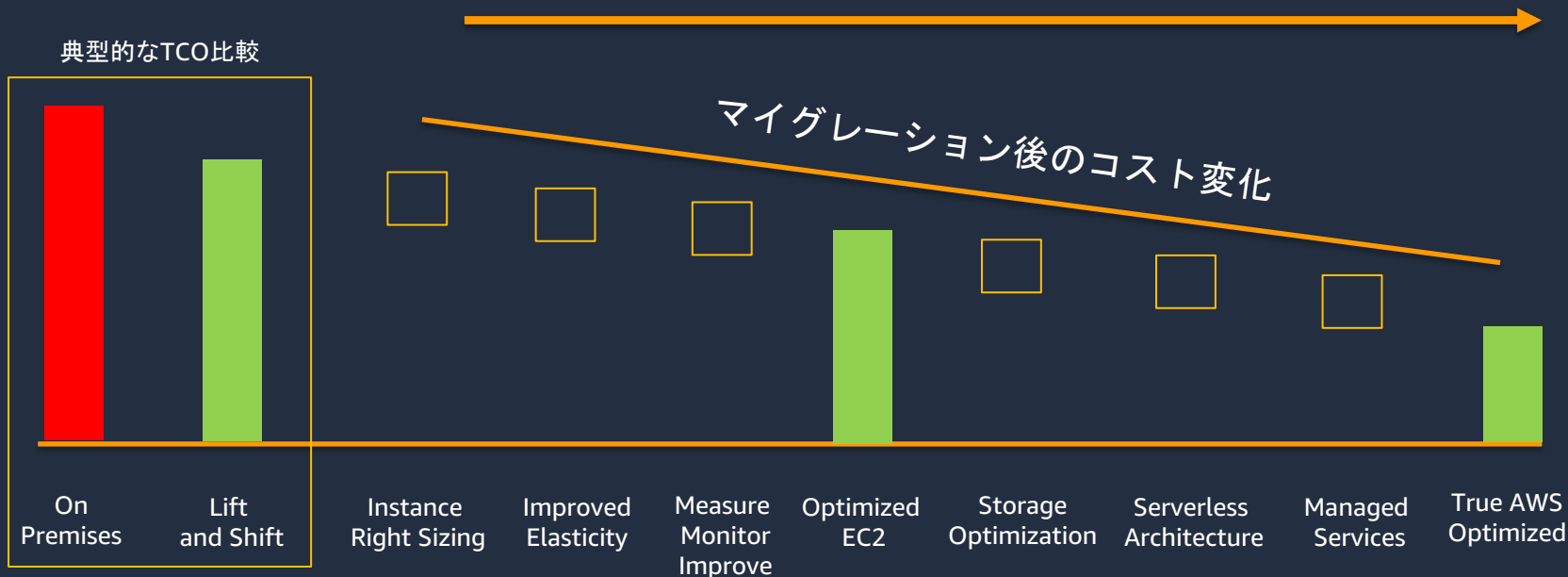
更にS3のデータレイク化により大規模シミュレーション結果を機械学習環境で活用



継続的な改善によるコスト最適化

クラウド活用では、最初から最適を目指すよりも
小さな改善のサイクルを継続的に実施していくことが重要

マイグレーション後の構成の見直しや
新しいサービスの活用



HPC on AWS 関連サービス

AWS における HPC 関連サービス

多様な HPC ワークロードに対応するための数多くのサービス

コンピューート

Amazon EC2



用途に応じて多様なインスタンスを利用可能な仮想サーバサービス



NVIDIA A100 GPU
搭載



Xilinx Virtex
UltraScale+ 搭載



100 Gbps の
ネットワーク帯域

スポットインスタンスの活用で大幅なコスト減も可能

ストレージ

Amazon S3



高い耐久性と低コストでのデータ保管を実現するオブジェクトストレージ

Amazon FSx for Lustre



S3連携可能な高速な分散ファイルシステムをフルマネージドで提供

AWS Snow ファミリー



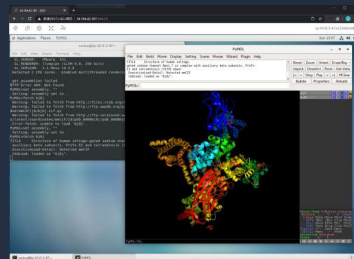
ストレージ内蔵の筐体を郵送することで S3 にデータを Import/Export

可視化

NICE-DCV



GPUアクセラレーションに対応し、インタラクティブなアプリケーションに適したデスクトップ仮想化



管理自動化

AWS ParallelCluster

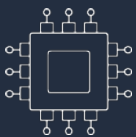


AWS上に HPC クラスタを自動で構築。**Slurm / SGE / Torque**といったジョブスケジューラに対応しており既存HPC環境からの移行が容易

AWS Batch



コンテナベースの大規模バッチジョブコンピューティング環境をフルマネージドで提供



Amazon EC2 (Elastic Compute Cloud)

必要なときに必要な計算リソースを確保可能な仮想サーバサービス

- 数分で起動し、秒単位の従量課金（一部タイプについては1時間単位）
- 独自の仮想化基盤 Nitro System により、仮想化オーバーヘッドを極小化
- ワークロードに応じて様々なインスタンスタイプを選択可能

高性能計算向けインスタンスタイプの例

高性能 CPU の選択肢



Intel Xeon processor
(x86_64 arch)

AMD EPYC processor*
(x86_64 arch)

AWS Graviton Processor
(64-bit Arm arch)

C5 インスタンス

Cascade Lake or Skylake
最大 3.9 GHz 駆動

M5zn インスタンス

Cascade Lake
最大全コア 4.5 GHz 駆動

C5a インスタンス

EPYC Rome
最大 3.3 GHz 駆動

C6g インスタンス

64bit Arm Neoverse N1ベース
Graviton2 CPU 搭載

アクセラレータの選択肢



NVIDIA GPU

Xilinx FPGA

P3 インスタンス

V100 GPU 搭載

P4d インスタンス

A100 GPU 搭載

G4 インスタンス

T4 GPU 搭載

F1 インスタンス

Virtex UltraScale+
VU9P 搭載



Amazon S3

AWS の提供するオブジェクトストレージサービス

- **容量無制限**
(1オブジェクトは最大5TBまで)
- データを3つ以上の AZ (データセンタ群) に保管し
99.999999999% という高い**耐久性**
- **低コスト**
Standard: 0.023 USD/GB※
～ S3 Glacier Deep Archive: 0.00099 USD/GB ※
- **スケーラブルで安定した性能**
データ容量に性能が依存しない
(ユーザが、サーバ台数、媒体本数やRAID、RAIDコントローラを考える必要がない)



※2021年10月現在のus-east-1での価格
価格はストレージクラスによって異なる
<https://aws.amazon.com/jp/s3/pricing/>

Amazon S3 ストレージクラスの選択肢



S3 Standard



S3 Intelligent-Tiering



S3 Standard-IA



S3 One Zone-IA



S3 Glacier



S3 Glacier Deep Archive

高頻度 ← アクセス頻度 → 低頻度

- アクティブ、頻繁にアクセスするデータ
- ミリ秒アクセス
- ≥ 3 AZ
- \$0.0210/GB~

- 変化するアクセスパターンのデータ
- ミリ秒アクセス
- ≥ 3 AZ
- \$0.0210~
- \$0.0125/GB
- オブジェクト毎の管理料金
- 最低保持期限30日

- 低頻度アクセスデータ
- ミリ秒アクセス
- ≥ 3 AZ
- \$0.0125/GB~
- GB毎の取り出し料金
- 最低保持期限30日
- 最小オブジェクトサイズ

- 再作成可能な低頻度アクセスデータ
- ミリ秒アクセス
- 1 AZ
- \$0.0100/GB~
- GB毎の取り出し料金
- 最低保持期限30日
- 最小オブジェクトサイズ

- アーカイブデータ
- 分~時間アクセス
- ≥ 3 AZ
- \$0.0040/GB~
- GB毎の取り出し料金
- 最低保持期限90日
- 最小オブジェクトサイズ

- アーカイブデータ
- 時間アクセス
- ≥ 3 AZ
- \$0.00099/GB~
- GB毎の取り出し料金
- 最低保持期限180日
- 最小オブジェクトサイズ

※ 価格は2021年10月現在の N. Virginia リージョン利用時

© 2021, Amazon Web Services, Inc. or its Affiliates. All rights reserved.



Amazon FSx for Lustre

高速な分散ファイルシステムである Lustre をフルマネージドで提供

LustreはPOSIX準拠のファイルシステムとして利用可能

階層型ストレージの機能もあり、S3と透過的にデータのimport/exportが可能

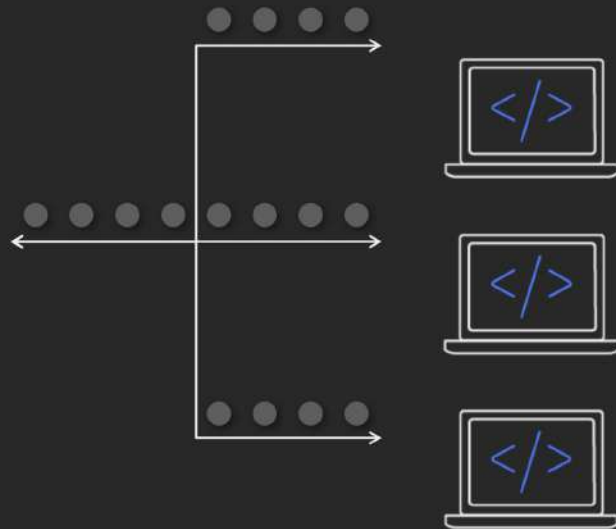
Link your Amazon S3 data set to your Amazon FSx for Lustre file system, then...



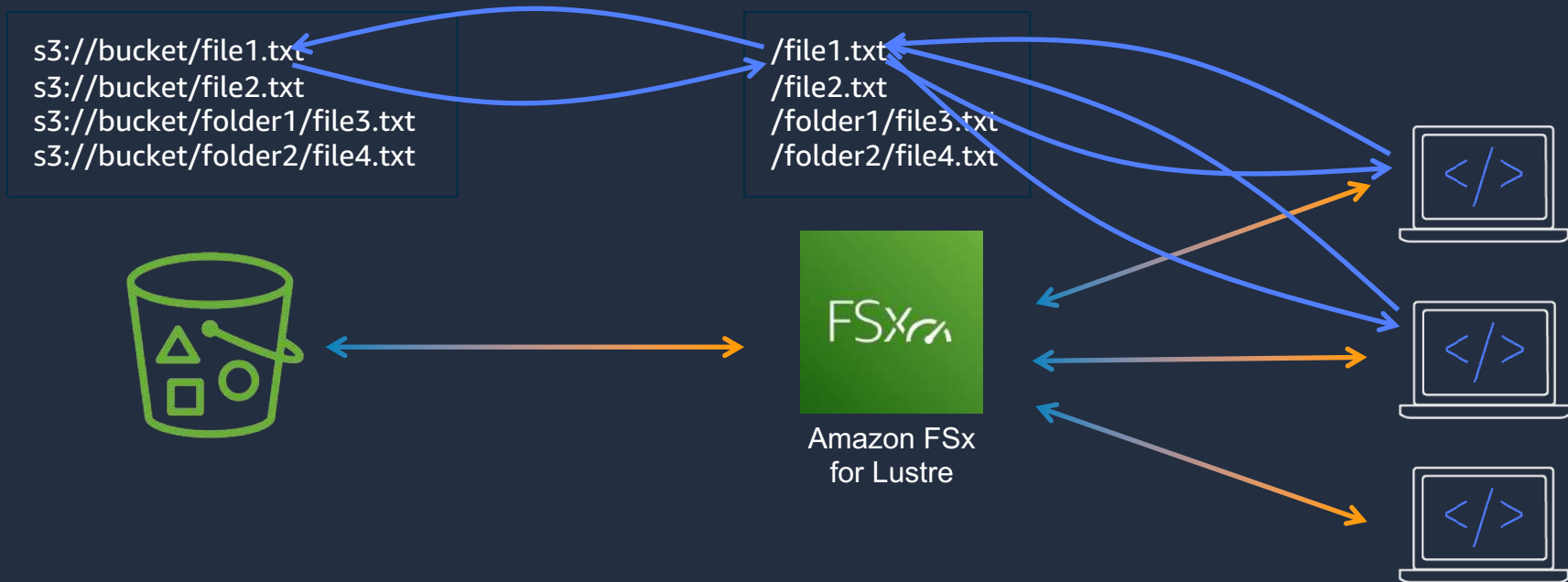
Data stored in Amazon S3 is loaded to Amazon FSx for processing



Output of processing returned to Amazon S3 for retention



Amazon FSx for Lustre の Amazon S3 への透過的アクセス



- ファイルシステム作成時にS3 bucketを関連付け、ファイルのインデックスが作成される
- 各ファイルに初回アクセスがあった時点でデータがS3からファイルシステムに自動でインポートされる（インポートのための若干のレイテンシが発生）
- 二度目のアクセスからは高速なレスポンスが可能

AWS Snow ファミリー

ハードウェアアプライアンスを郵送することでオンプレミス〜クラウド間の大量データ移行を高速化

データの自動暗号化や不正開封防止筐体によるセキュリティ

AWS Snowball Edge

大容量を持つ Storage Optimized (100TB) または Compute Optimized (7.68TB[NVMe SSD] + 42TB) の物理アプライアンスによりデータを郵送

AWS Snowcone

約 2 kg と軽量の筐体に8TB のHDDを搭載

AWS DataSync エージェントに対応し、オンライン転送も可能

※ AWS Snowball Edge/AWS Snowcone 共に東京リージョンで利用可能



AWS Snowball Edge



AWS Snowcone

NICE-DCV

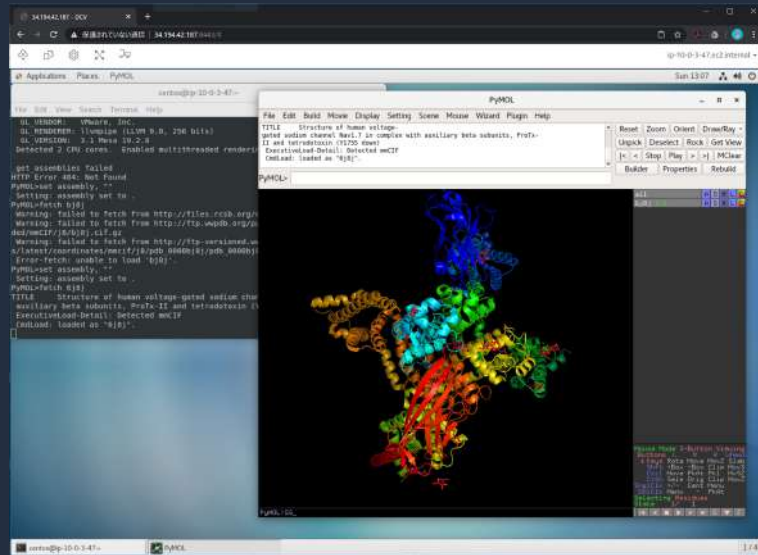
クラウド上のデスクトップ画面をストリーミングするためのソフトウェア

特徴

- 専用のプロトコルによる高速かつスムーズなストリーミング
- サーバはWindows、Linuxの両方に対応
- GPUにも対応し、G4dn/G4adインスタンス等を利用することでより高速な描画が可能
- ネイティブクライアントの他、HTML5対応ブラウザからも利用可能

コスト

- Amazon EC2で利用する場合はライセンスコスト無しで利用可能 (EC2以外での利用は有償)



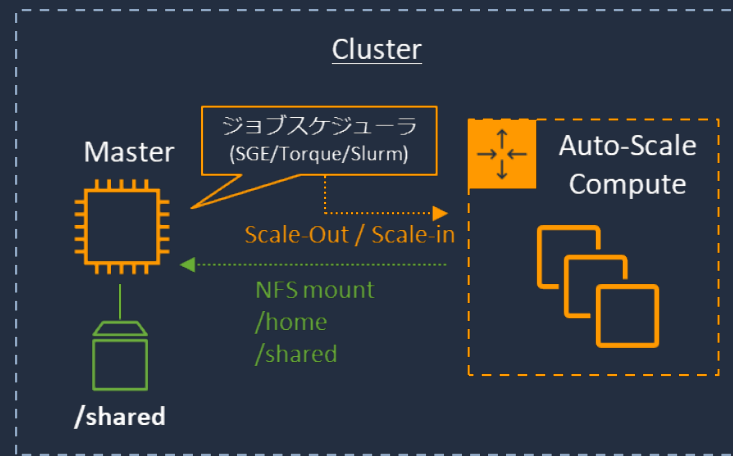
AWS ParallelCluster とは

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向けジョブスケジューラと Auto-Scaling を連携した環境を作成
Slurm / SGE / Torque ※に対応
- 少しのコマンド操作でクラスタ作成可能
- MPI/NCCL 環境がセットアップ済みで、すぐに利用可能
- 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
- 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

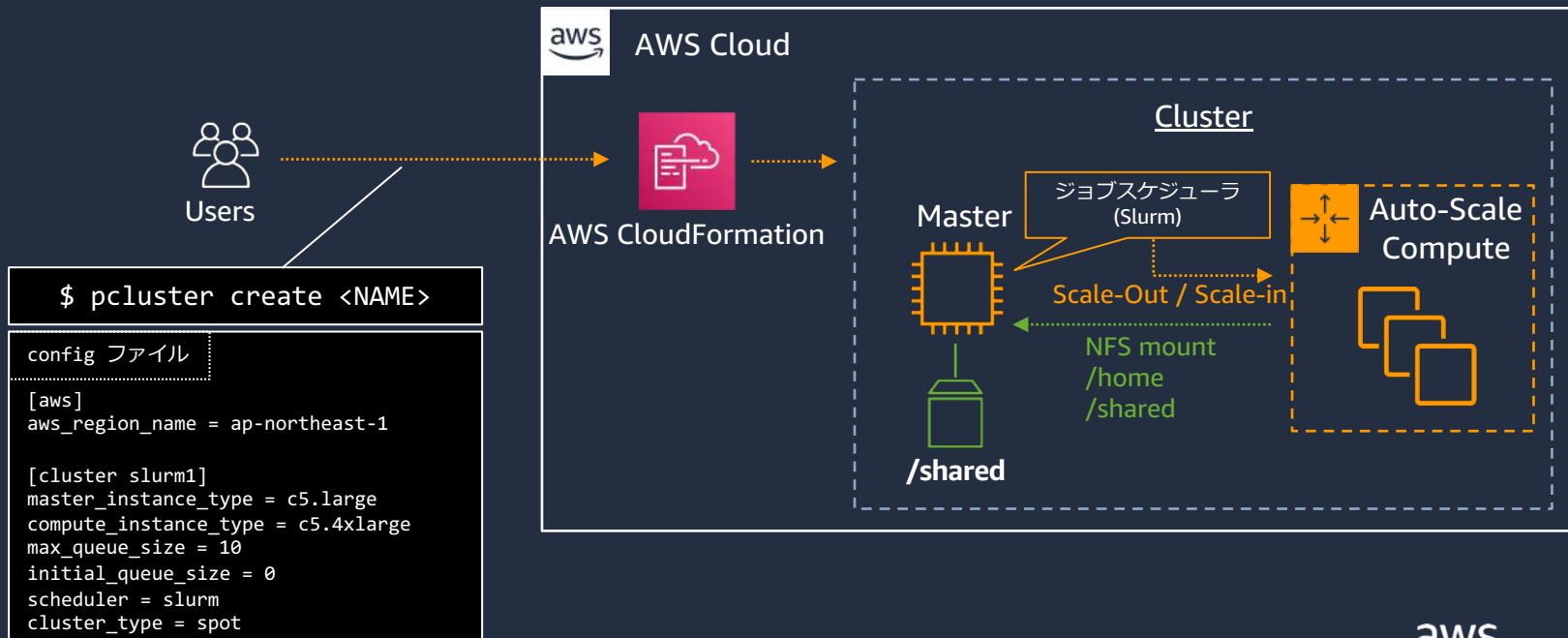
<https://github.com/aws/aws-parallelcluster>



※ 将来的にSGE/Torqueについてはサポートの終了がアナウンスされており、Slurmの利用を推奨

AWS ParallelCluster の利用イメージ

まずは自分のPC等に ParallelCluster ソフトウェアをインストール
config ファイルを記述し、pcluster create コマンドを実行することで、
ジョブ投入に応じて Auto-Scale するクラスタ環境が自動的に作成される



まとめ

- クラウドのスケラビリティを活用し、必要な時に必要な量の必要なタイプの計算リソースを確保
- AWS では仮想サーバ環境だけでなく高速なストレージや、HPC環境構築ツールなど、研究のための大規模計算を行うのに適したサービスを提供している

スケラブルかつ柔軟な計算環境により
研究における、計算機に関連したボトルネックを解消



HPC on AWS