



このコンテンツは公開から3年以上経過しており内容が古い可能性があります  
最新情報については[サービス別資料](#)もしくはサービスのドキュメントをご確認ください

# [AWS Black Belt Online Seminar]

## HPC on AWS

ソリューションカットシリーズ



Specialist Solutions Architect, HPC  
Daisuke Miyamoto  
2020/12/09

AWS 公式 Webinar  
<https://amzn.to/JPWebinar>



過去資料  
<https://amzn.to/JPArchive>



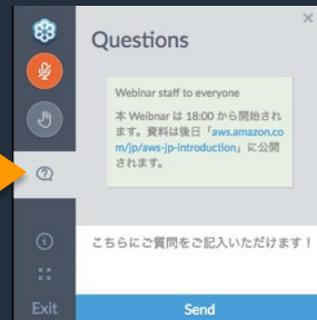
# AWS Black Belt Online Seminar とは

「サービス別」「ソリューション別」「業種別」のそれぞれのテーマに分かれて、アマゾンウェブ サービス ジャパン株式会社が主催するオンラインセミナーシリーズです。

## 質問を投げることができます！

- 書き込んだ質問は、主催者にしか見えません
- 今後のロードマップに関するご質問はお答えできませんのでご了承下さい

- ① 吹き出しをクリック
- ② 質問を入力
- ③ Sendをクリック



Twitter ハッシュタグは以下をご利用ください  
#awsblackbelt

# 内容についての注意点

- 本資料では2020年12月09日時点のサービス内容および価格についてご説明しています。最新の情報はAWS公式ウェブサイト(<http://aws.amazon.com>)にてご確認ください。
- 資料作成には十分注意しておりますが、資料内の価格とAWS公式ウェブサイト記載の価格に相違があった場合、AWS公式ウェブサイトの価格を優先とさせていただきます。
- 価格は税抜表記となっております。日本居住者のお客様には別途消費税をご請求させていただきます。
- AWS does not offer binding price quotes. AWS pricing is publicly available and is subject to change in accordance with the AWS Customer Agreement available at <http://aws.amazon.com/agreement/>. Any pricing information included in this document is provided only as an estimate of usage charges for AWS services based on certain information that you have provided. Monthly charges will be based on your actual use of AWS services, and may vary from the estimates provided.

# 本セミナーの概要

## □ 本セミナーで学習できること

- ❖ HPC環境をクラウド化するメリット
- ❖ HPC on AWSに関連するサービスの概要
- ❖ HPC on AWSの事例

## □ 対象者

- ❖ クラウドHPCについて知りたい方
- ❖ これから HPC 環境を AWS 上で構築したいと考えている方
- ❖ AWS上のHPC環境をより最適化したい方
- ❖ 次の AWS のサービスの概要レベルの知識が前提になります

Amazon VPC / Amazon EC2 / Amazon S3 などのAWS基礎サービス

# 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

## □ 好きな AWS サービス

- ❖ AWS ParallelCluster
- ❖ Amazon FSx for Lustre
- ❖ AWS Snowball シリーズ



# 本セミナーの目的

以下のようなHPC (High Performance Computing) 環境が活用できるワークロードを大規模に実施する際に役立つサービス、事例、注意事項についてご紹介

## 製造

- 流体解析、構造解析
- EDA (Electronic Design Automation)

## 創薬

- 分子動力学、ドッキングシミュレーション、量子化学計算
- ゲノム解析

## 金融

- リスク計算

## その他

- メディア処理
- 気象予測
- 分散機械学習



流体解析



分子動力学計算

# 本日の流れ

## 入門編：これからHPC on AWSを始める、検討中の方向け

- HPC on AWS の特徴・利点
- HPC on AWS の利用イメージ
- HPC on AWS を支えるサービス
- HPC on AWS 事例紹介

## 応用編：既に利用中でより良い構成を考えたい方向け

- HPC on AWS ベンチマーク例
- HPC on AWS の構成ポイント

## アップデート編

- HPC on AWS 関連最新アップデート紹介

# 入門編

# HPC on AWS 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

従来のクラスタ  
構成は固定

ジョブが  
無ければ  
無駄発生

ジョブが無い時は  
最小限のノード



aws  
Elastic Data  
Center

Corporate  
Data Center

AWS

# AWSなら、必要な時に必要なだけ利用可能

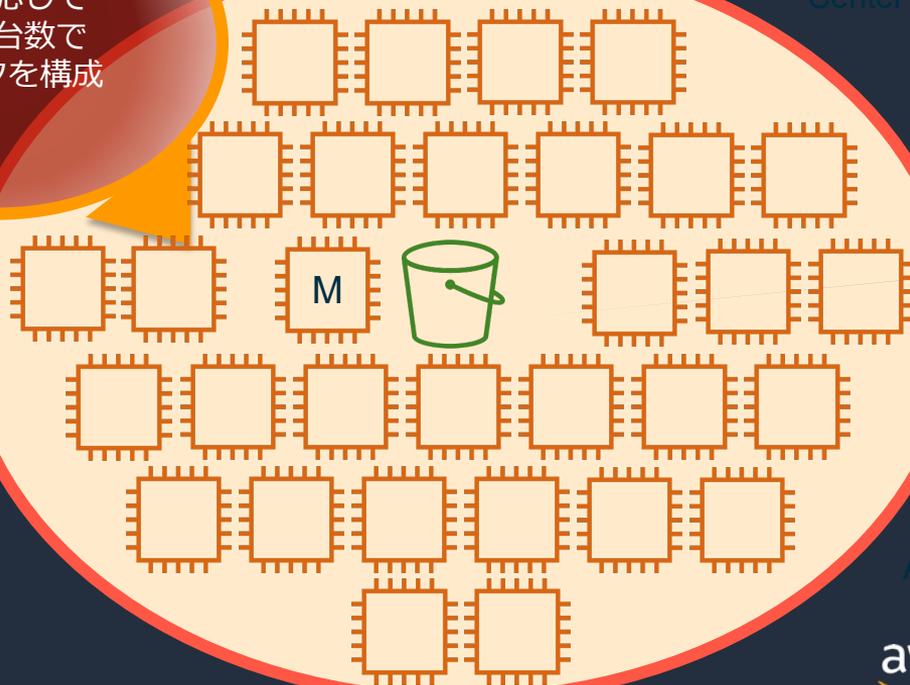
スケーラブルなリソースによりジョブ実行待ちの無いHPC環境を実現

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

必要に応じて  
必要な台数で  
クラスタを構成

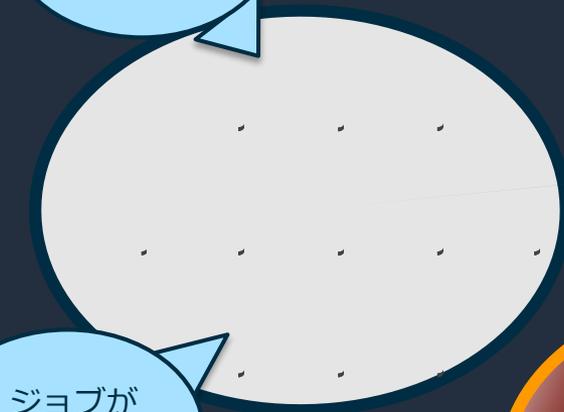


# AWSなら、必要な時に必要なだけ利用可能

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

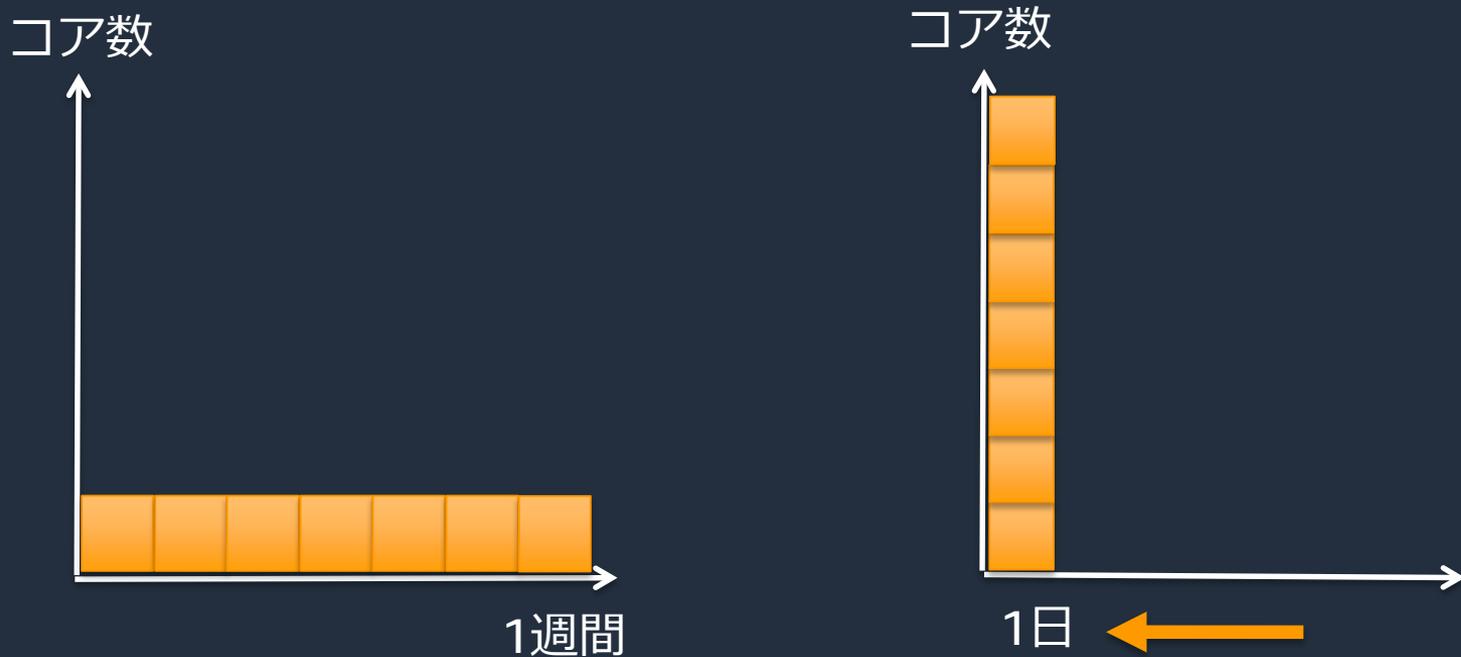
処理が終了すると  
インスタンスを終了  
課金停止



AWS



# スケーラビリティの活用による計算時間短縮

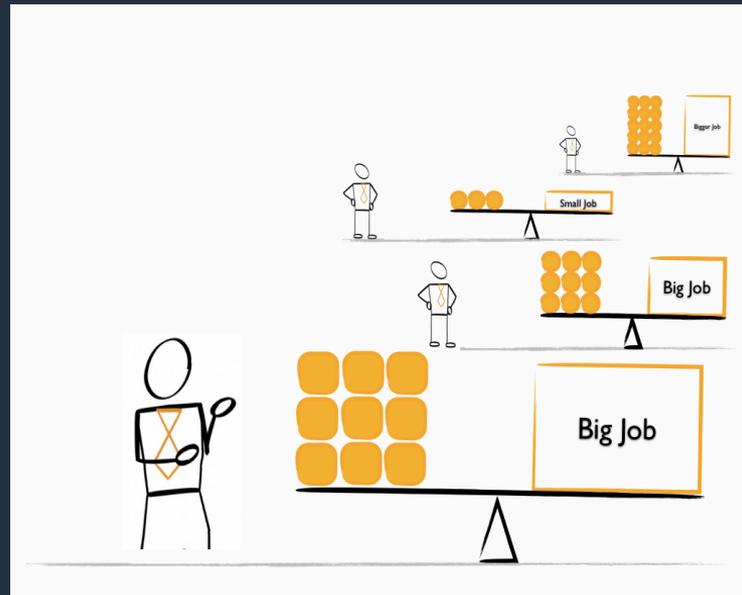


従来は手持ちの限られたリソースで、逐次処理していたジョブも  
AWSなら必要な台数、インスタンスを起動して、一斉処理  
しかも費用は「時間×台数」なのでどちらも同じ

# 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、、、

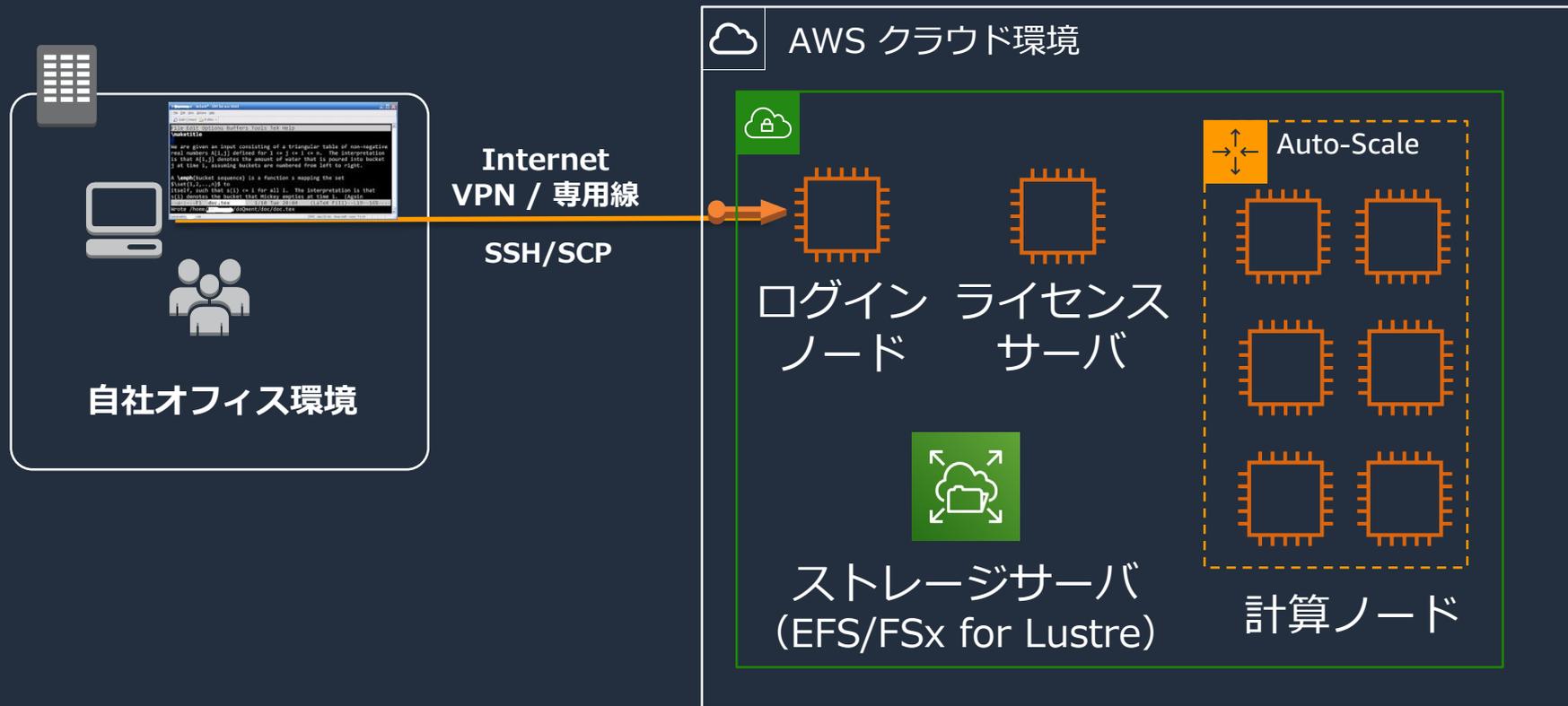


競争優位につながらない物理的管理は全てAWSにお任せ  
他社と差別化可能な部分に集中

# HPC on AWS の利用イメージ

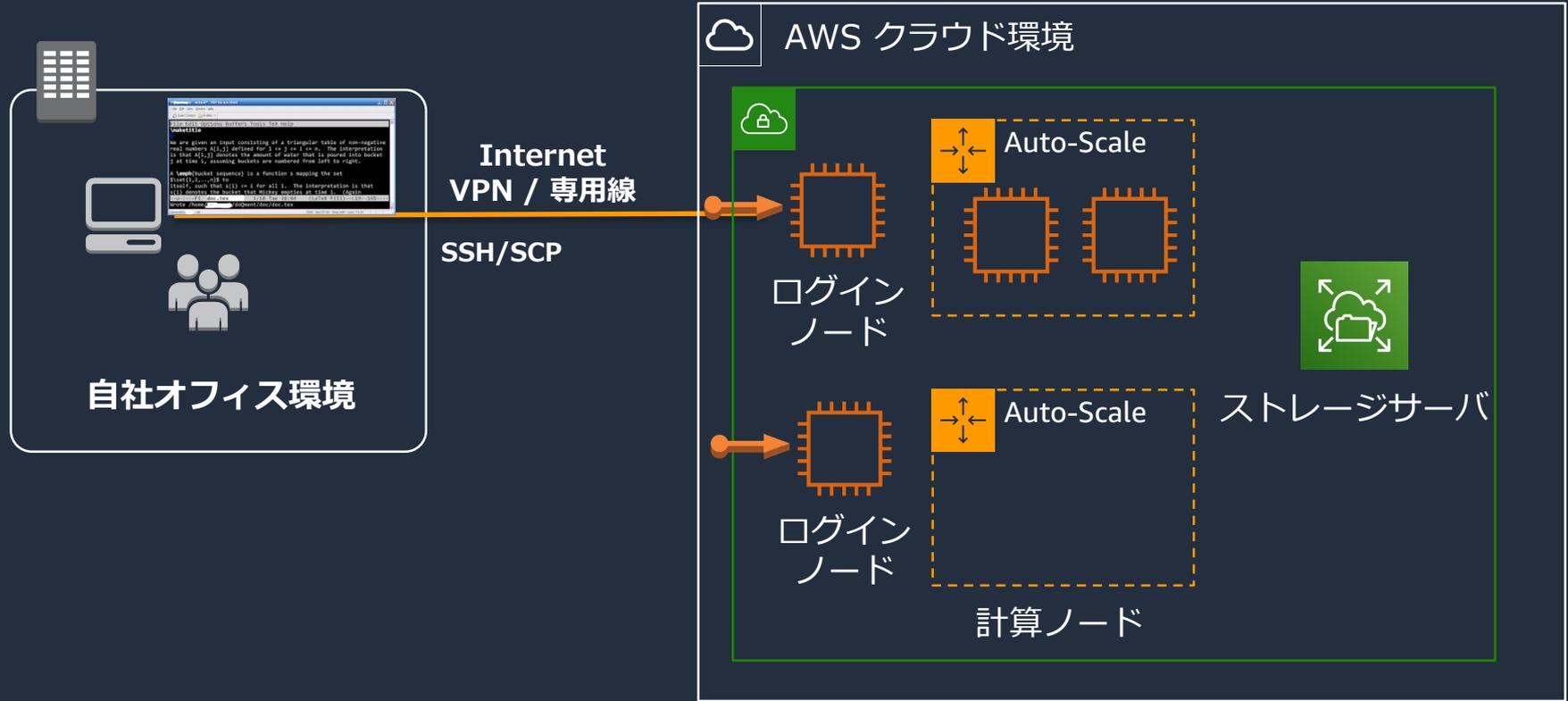


# AWSでも基本的なシステム構成は同じ



需要に応じて伸縮する計算環境・マネージドサービスの活用

# AWSでは、1人1クラスタなどより柔軟性の高い構成も可能

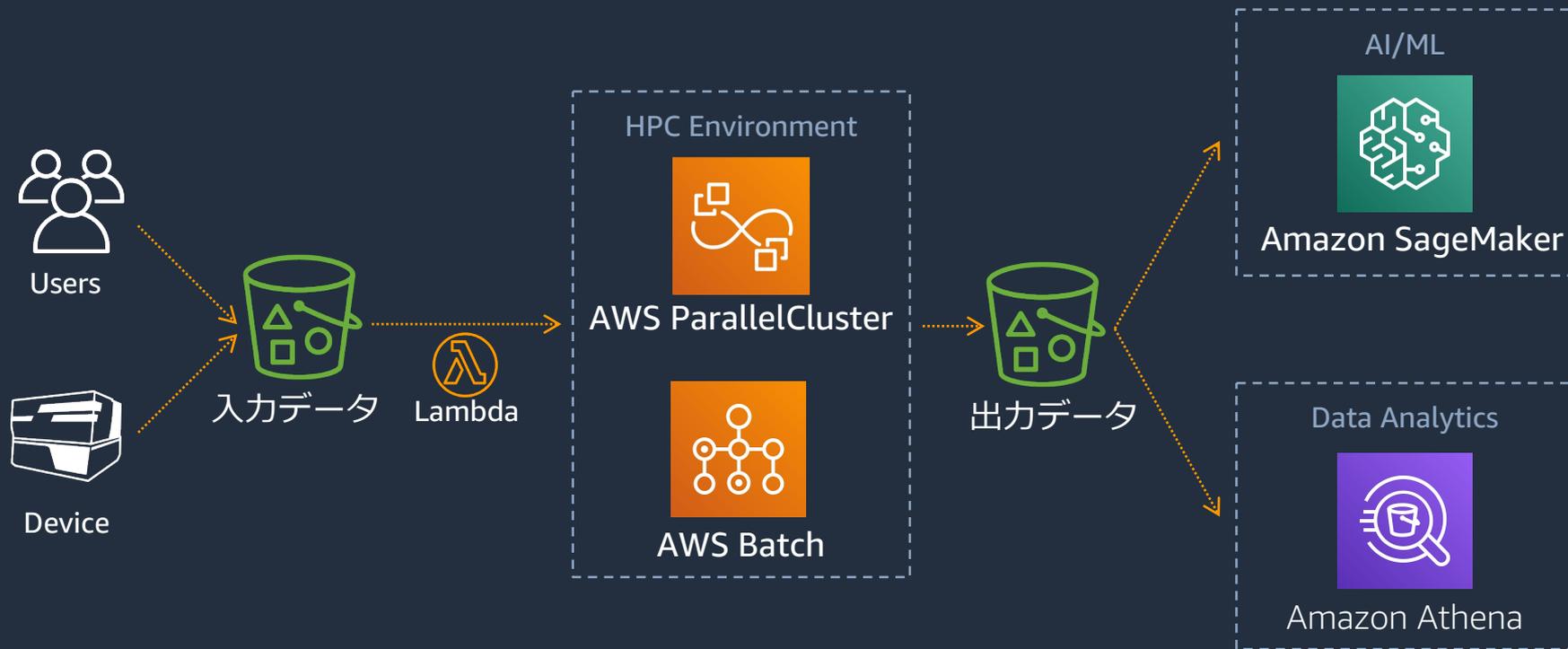


計算ノードは自動でスケールするため、複数クラスタを作成しても維持コストが低い

# 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

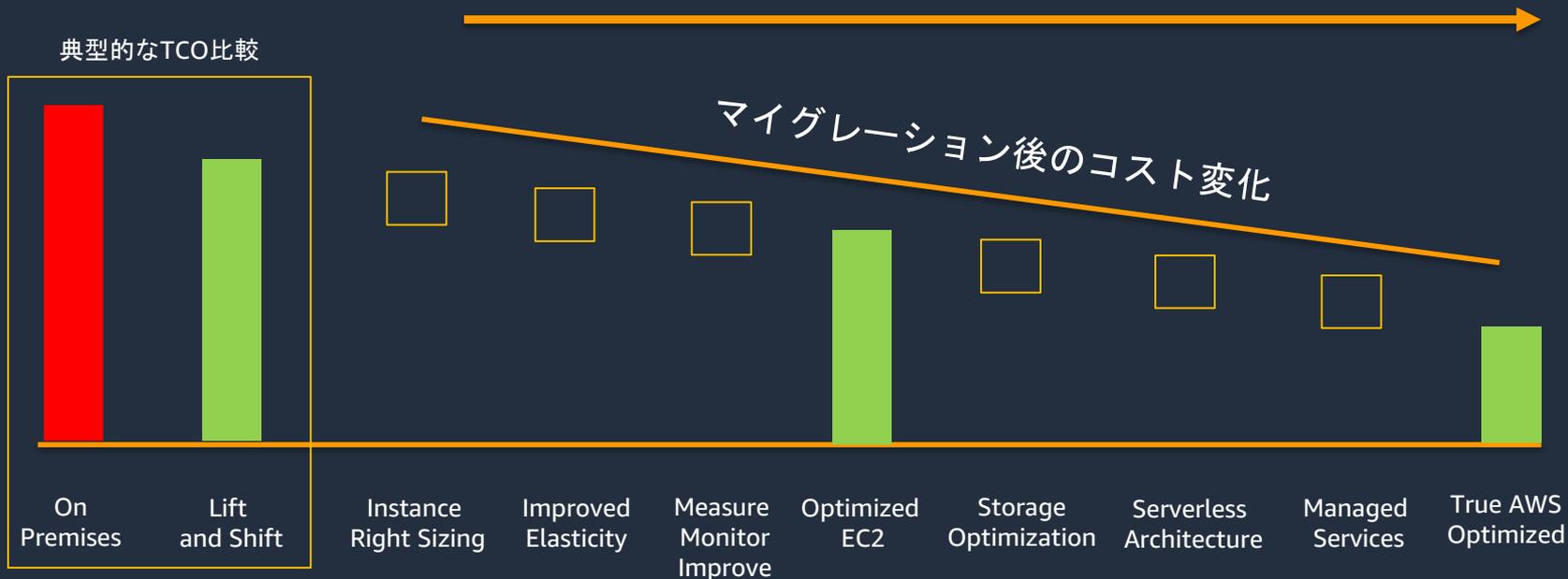
更にS3のデータレイク化により大規模シミュレーション結果を機械学習環境で活用



# 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 を支えるサービス

# 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

## ネットワーク

### Placement Group

EC2インスタンスの基盤上の配置を制御してネットワークを高速化

### Elastic Fabric Adapter

MPI/NCCL 専用の低レイテンシ  
ネットワークアダプタ

## ストレージ

### 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

## 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

## 可視化

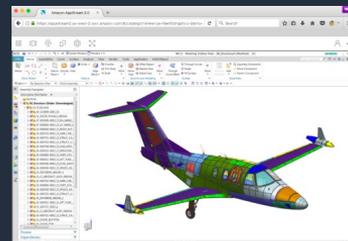
### NICE-DCV

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

### Amazon AppStream 2.0



フルマネージドのアプリ  
ケーションストリーミン  
グサービス



# 仮想サーバサービス 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）
- 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
- インスタンスを停止するだけでマシンスペック変更が可能

## インスタンスタイプ一覧と分類



# Amazon EC2 で選択できる高性能CPUの選択肢



**Intel Xeon processor**  
(x86\_64 arch)

最大3.9GHz駆動  
Cascade Lake  
or Skylake 搭載  
**C5インスタンス**



**AMD EPYC processor**  
(x86\_64 arch)

最大3.3GHz駆動  
Romeコア搭載  
**C5aインスタンス**



**AWS Graviton Processor**  
(64-bit Arm arch)

64bit Arm Neoverse N1ベース  
Graviton2 CPU搭載  
**C6gインスタンス**

アプリケーションとワークロードに応じて  
最適なコンピューティング環境を選択

# EC2で選択できるアクセラレータ搭載インスタンス



Xilinx  
UltraScale+ FPGA  
**F1**インスタンス



NVIDIA  
V100 GPU搭載 **P3**  
A100 GPU搭載 **P4d**  
T4 GPU搭載 **G4**



AWS  
Inferentia  
**Inf1**インスタンス

アプリケーションとワークロードに応じて  
多様なアクセラレータを提供

# EC2 購入オプション

## オンデマンドインスタンス

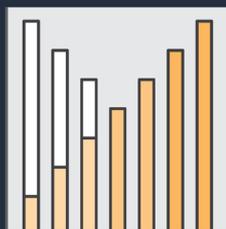
長期コミット無し、使用分への支払い(秒単位/時間単位)。Amazon EC2の定価



スパイクするようなワークロード

## リザーブドインスタンス (Savings Plans)

1年/3年の長期コミットをする代わりに大幅なディスカウント価格



一定の負荷の見通しがあるワークロード

## スポットインスタンス

Amazon EC2の空きキャパシティを活用し、**最大90%値引き**。中断が発生することがある



中断に強く、かつ様々なインスタンスタイプを活用できるワークロード

HPC 等では特にスポットインスタンスを活用することでコストパフォーマンスの良い計算が可能

# EC2の高性能ネットワーク技術

## • 拡張ネットワーキング

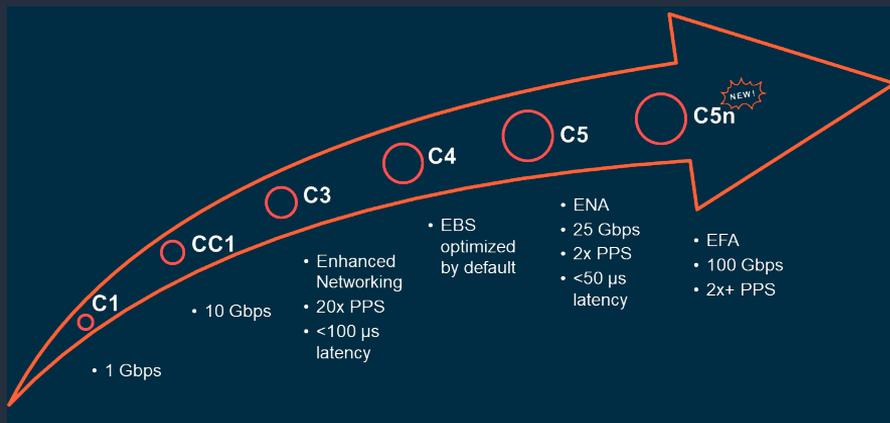
- SR-IOVに対応し、仮想化オーバーヘッドを低減することで低レイテンシでの通信が可能

## • Cluster Placement Group

- **インスタンスの配置を最適化**することで広帯域/低レイテンシ/フルバイセクション通信を実現

## • Elastic Fabric Adapter

- **HPC向けに、MPI (Message Passing Interface) やNCCL (NVIDIA Collective Communications Library) などの libfabric 対応のアプリケーションでの通信をより低レイテンシ化**



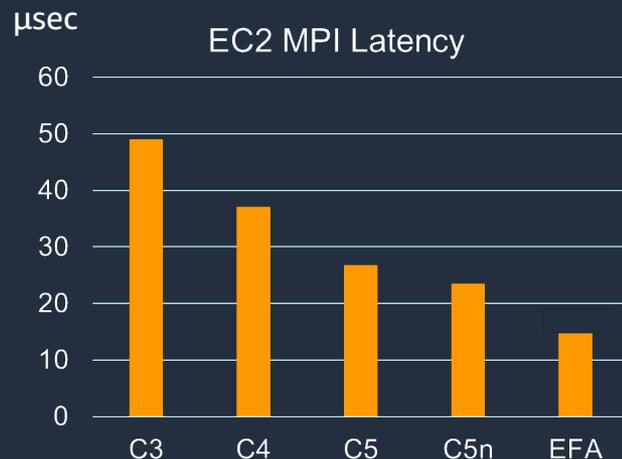
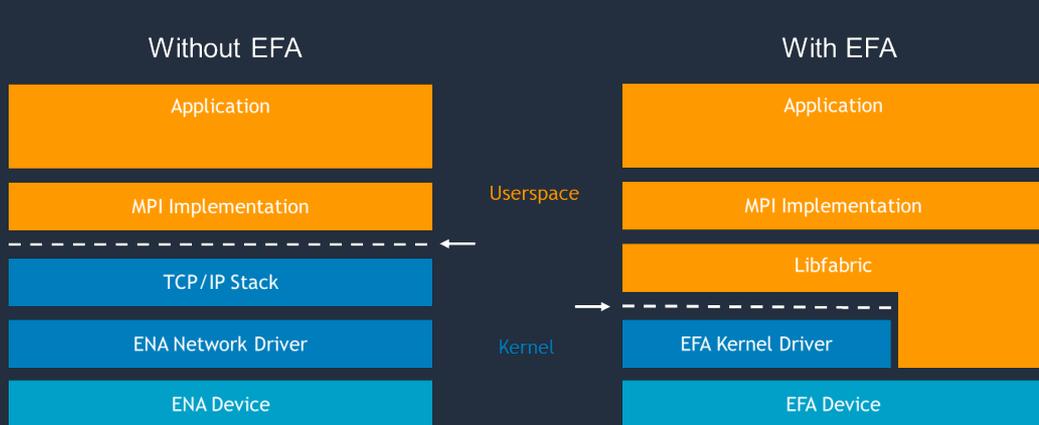
EC2のネットワークも進化を続け  
現在はEC2インスタンスあたり  
最大 400Gbps (P4dインスタンス) まで  
サポート※

※インスタンスタイプとサイズによって通信帯域は異なります

# EFA (Elastic Fabric Adapter)

MPI/NCCL専用のネットワークアダプタ Elastic Fabric Adapter により  
低レイテンシでのノード間通信を実現

- 利用には対応したMPI環境 (Intel MPI/OpenMPI) が必要だが、プログラムの変更は原則不要
- EFA対応インスタンス : c5n.18xlarge, m5n.24xlarge, p4d.24xlarge , p3dn.24xlarge etc.



L. Shalev, H. Ayoub, N. Bshara and E. Sabbag, "Supercomputing on Nitro in AWS Cloud," in IEEE Micro, doi: 10.1109/MM.2020.3016891.

<https://ieeexplore.ieee.org/document/9167399>

# EFAで使用されているSRD (Scalable Reliable Datagram)

**AWS のデータセンターネットワーク向けに  
新たに開発されたトランスポートプロトコル**

**配信保証** : EC2のリソースを使用せずに保証を行う

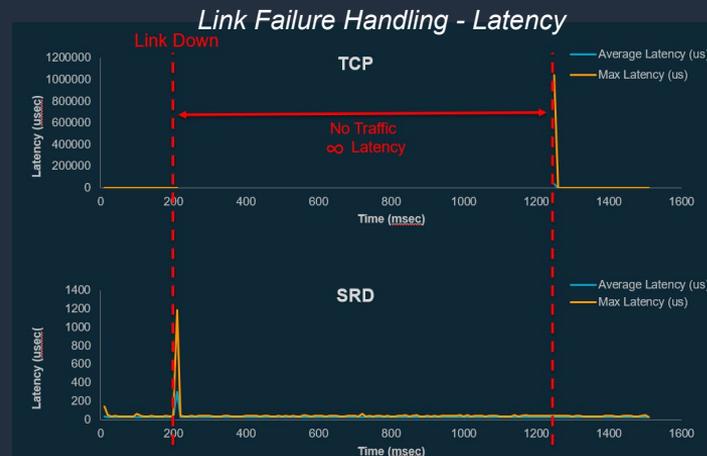
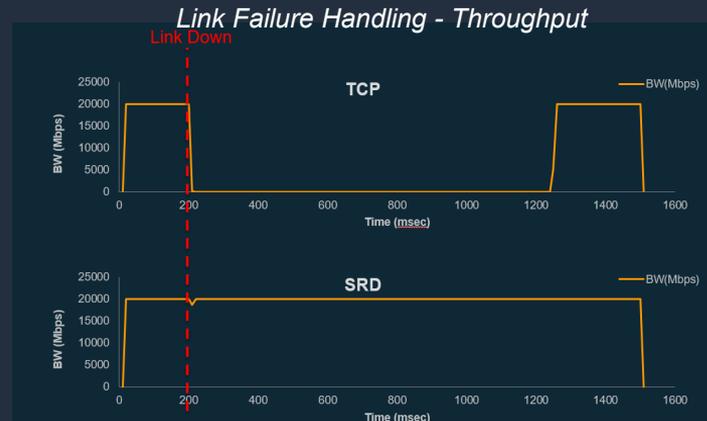
**マルチパスルーティング** : データセンターの複数の  
ネットワーク経路を活用

**レイテンシ・ジッターの低減** : 独自の link/switch ダウ  
ン検出、輻輳制御

**アウトオブオーダーでの転送** : ブロックを抑制

L. Shalev, H. Ayoub, N. Bshara and E. Sabbag, "Supercomputing on Nitro in  
AWS Cloud," in IEEE Micro, doi: 10.1109/MM.2020.3016891.

<https://ieeexplore.ieee.org/document/9167399>



# 参考：InfiniBandとSRDの比較

TCP	InfiniBand	SRD
Stream	Messages	Messages
In-order	In-order	Out-of-order
Single path	Single (ish) path	ECMP spraying with load balancing
High limit on retransmit timeout (>50ms)	Static user-configured timeout (log scale)	Dynamically estimated timeout ( $\mu$ s resolution)
Loss-based congestion control	Semi-static rate limiting (limited set of supported rates)	Dynamic rate limiting
Inefficient software stack	Transport offload with scaling limitations	Scalable transport offload (same number of QPs regardless cluster size)

# Amazon FSx for Lustre とは

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

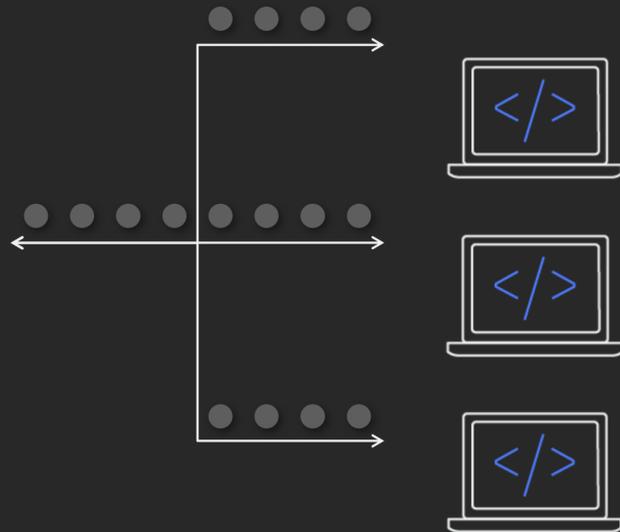
Link your Amazon S3 data set to your Amazon FSx for Lustre file system, then...



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



Output of processing returned to Amazon S3 for retention



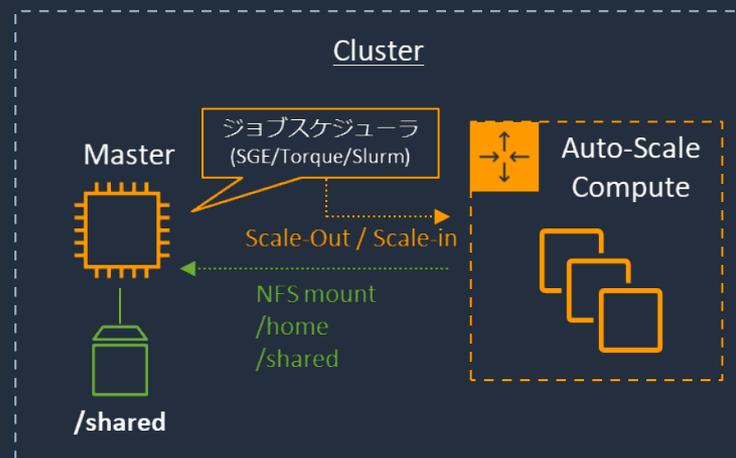
# AWS ParallelCluster とは

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

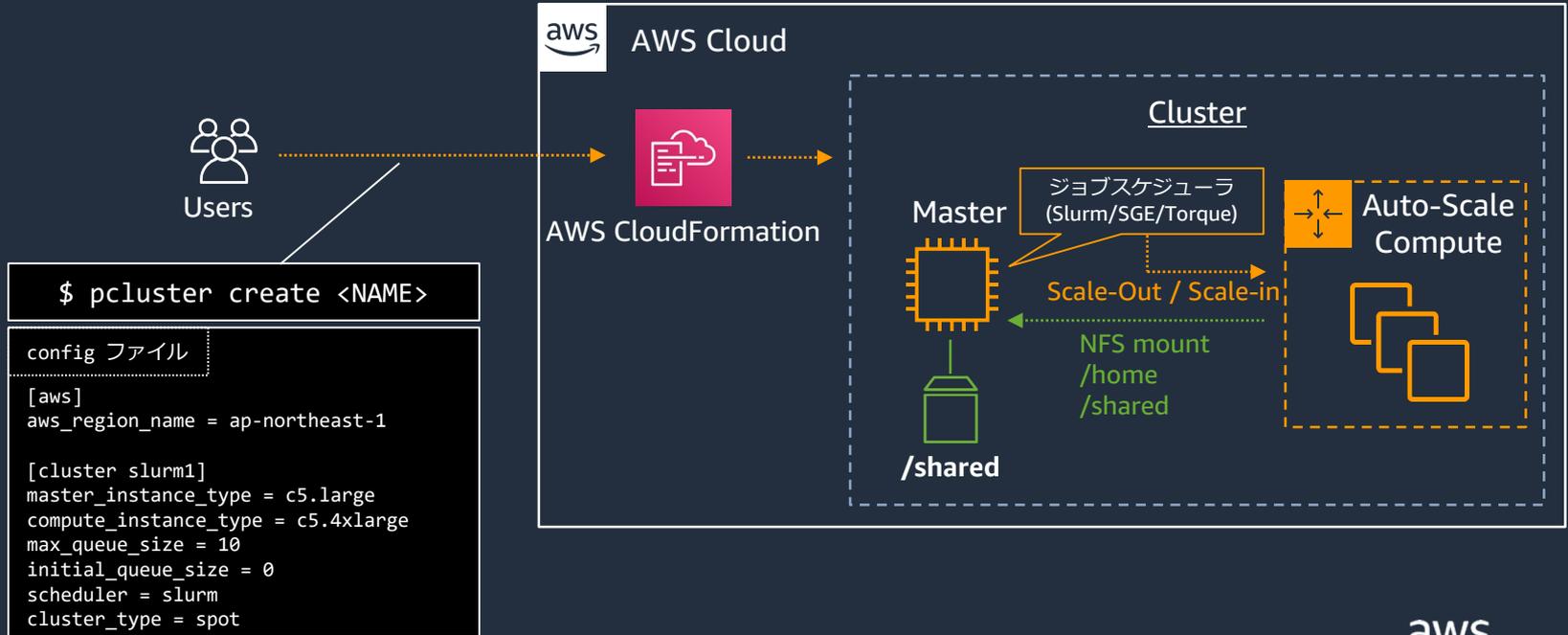
<https://github.com/aws/aws-parallelcluster>



※ 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

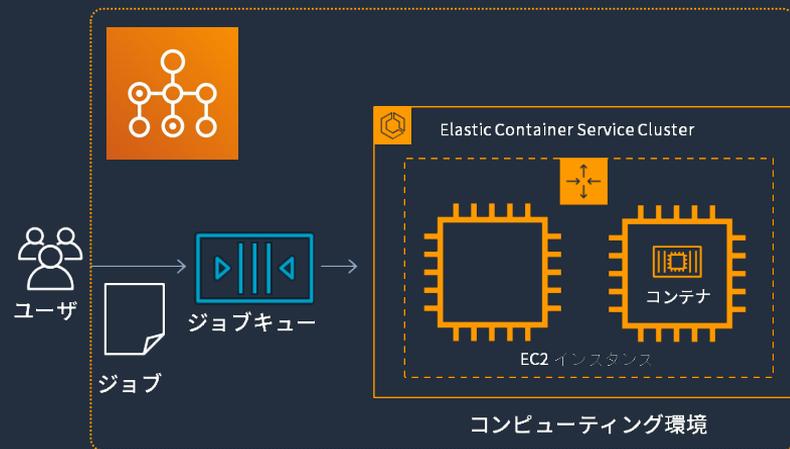
まずは自分のPC等に ParallelCluster ソフトウェアをインストール  
config ファイルを記述し、pcluster create コマンドを実行することで、  
ジョブ投入に応じて Auto-Scale するクラスタ環境が自動的に作成される



# AWS Batch とは

## 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ンスタイプや 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

# NICE-DCV とは

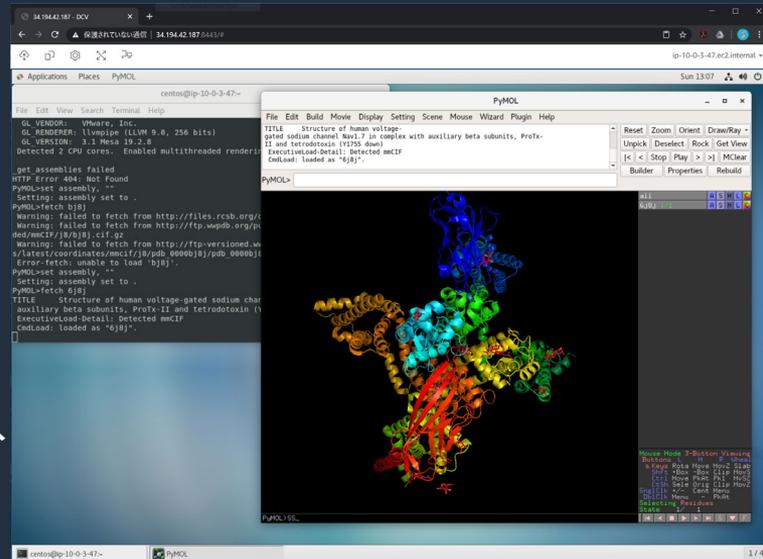
クラウド上のデスクトップ画面をストリーミングするためのソフトウェア

## 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 における HPC 関連サービス (再掲)

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

## ネットワーク

### Placement Group

EC2インスタンスの基盤上の配置を制御してネットワークを高速化

### Elastic Fabric Adapter

MPI/NCCL 専用の低レイテンシ  
ネットワークアダプタ

## ストレージ

### 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

## 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

## 可視化

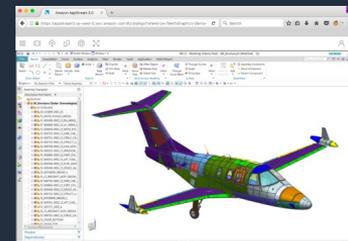
### NICE-DCV

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

### Amazon AppStream 2.0



フルマネージドのアプリ  
ケーションストリーミン  
グサービス



# HPC on AWS 事例紹介

# Western Digital

## 次世代ハードディスクのドライブヘッドシミュレーションをクラウド上で実行



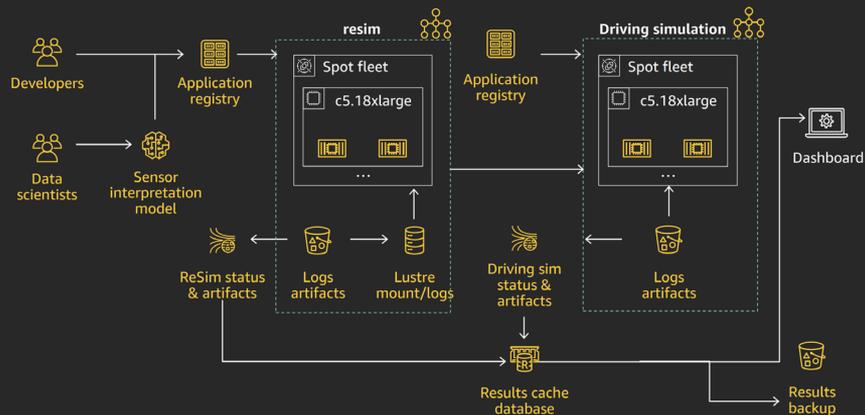
- 250万強におよぶパラメータスタディを最大100万vCPU以上同時利用して処理
- 従来であれば20日掛かる処理を約8時間で全てのタスクを終えることに成功
- スポットインスタンスを活用することでコストを最適化、100万vCPUを1時間あたり\$17,164 (1vCPUあたり\$0.017)で利用

<https://aws.amazon.com/jp/blogs/news/western-digital-hdd-simulation-at-cloud-scale-2-5-million-hpc-tasks-40k-ec2-spot-instances/>

# Mobileye

大規模な自動運転のシミュレーション環境をAWS Batchを用いて構成  
最大同時 500,000 CPUコアを利用して、1時間あたり70年分のコンピューティングを実行、S3上のデータを1ヶ月あたり100PB処理

## Large scale AV simulations with AWS Batch



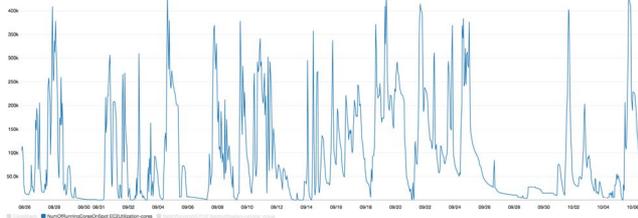
## resim at scale: compute

### Scale

Processing 100s PB in S3/month  
500,000 concurrent cores  
1 hour = 70 years of compute  
99.6% jobs success rate

### Cost

99% Spot Instances  
90% utilization  
90% EBS usage reduction



[https://d1.awsstatic.com/events/reinvent/2019/Navigating\\_the\\_winding\\_road\\_toward\\_driverless\\_mobility\\_AUT307.pdf](https://d1.awsstatic.com/events/reinvent/2019/Navigating_the_winding_road_toward_driverless_mobility_AUT307.pdf)

# Descartes Labs

Amazon EC2 C5インスタンスを使用したクラスタで約2PFLOPSのLINPACK性能を達成

- Top500(2019/06)で136位にランクイン
- 41,472コア(Xeon Skylake 3.0GHz)利用
- 2.6h稼働、コストは約\$5,000で実現

“One of the more interesting aspects of the story is that **we didn't ask Amazon to give our engineers any special dispensation, discount, or custom planning or setup.** We wanted to see if we could do this on our own, which if completed successfully, would also be a testament to the self-service model of AWS.”

<https://medium.com/descarteslabs-team/thunder-from-the-cloud-40-000-cores-running-in-concert-on-aws-bf1610679978>

Amazon EC2 C5 Instance cluster us-east-1a - Amazon EC2 Instance Cluster C5, Xeon Platinum 8124M 18C 3GHz, 25G Ethernet

Site:	Descartes Labs
System URL:	<a href="https://aws.amazon.com/ec2/instance-types/c5/">https://aws.amazon.com/ec2/instance-types/c5/</a>
Manufacturer:	Amazon Web Services
Cores:	41,472
Memory:	157,824 GB
Processor:	Xeon Platinum 8124M 18C 3GHz
Interconnect:	25G Ethernet
<b>Performance</b>	
Linpack Performance (Rmax)	1,926.4 TFlop/s
Theoretical Peak (Rpeak)	3,981.31 TFlop/s
Nmax	2,985,984
<b>Power Consumption</b>	
Power:	
<b>Software</b>	
Operating System:	Amazon Linux 2
Compiler:	gcc 7.3.1
Math Library:	Intel MKL
MPI:	OpenMPI 4.1.0

# 日本国内でも多様な分野で利用が広がる HPC on AWS

AWS Summit Online Japan 等で公開になった最新の HPC on AWS 事例をご紹介します

- CAE・製造
- 半導体デバイス設計
- 創薬
- ゲノミクス
- 気象シミュレーション

# AGC株式会社 様

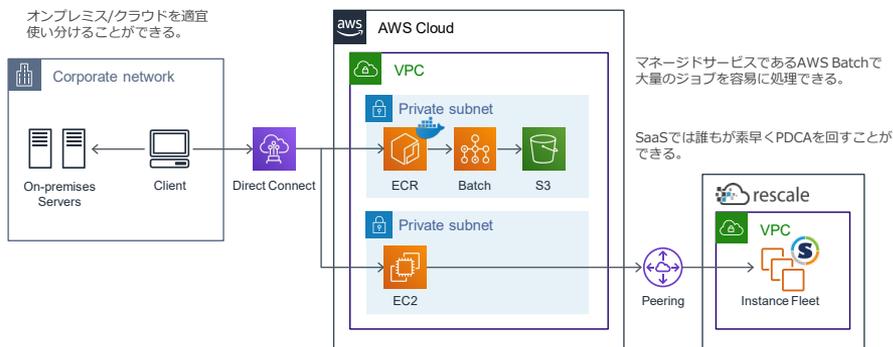
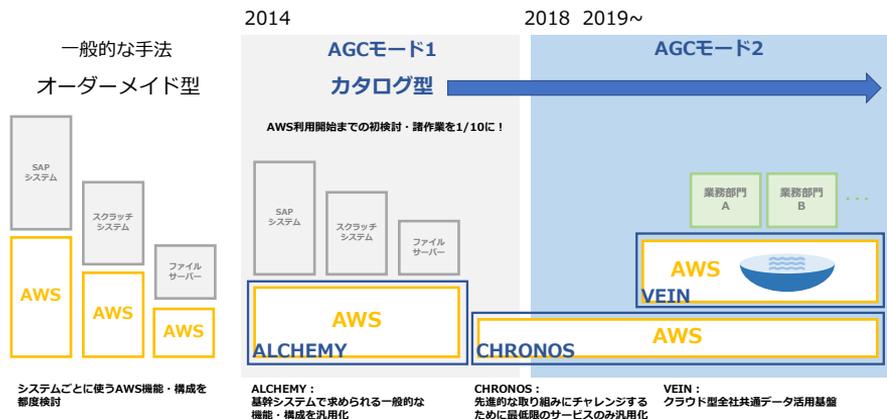
**AGC** AGCはAWSを活用し、製品・プロセス開発に関わるCAE\*を進化させています。  
情報管理部門と開発部門の連携、マネージドサービスとコンテナの活用が成功要因です。

\*CAE: Computer Aided Engineering

情報システム部の管轄の下、最新のAWSサービスを取り込み、業務に活用できる仕組みがあります。

IaaS\*1とSaaS\*2の両方を利用しています。  
IaaS利用にDx接続、SaaS利用にVPC Peeringを活用しています。

\*1IaaS: Infrastructure as a Service, \*2SaaS: Software as a Service



CAE用途では、CHRONOSという社内向けサービスの枠組みの中でエンタープライズレベルのセキュリティ・ガバナンスを担保し、仮想サーバー・コンテナサービス等を利用しています。

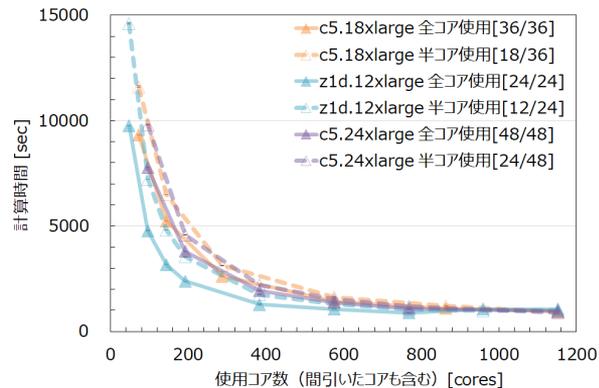
自社開発ソフトウェアを用いた解析は、AWS Batchで計算時間の短縮・高度化が実現しました。仮想コンテナを活用することで、様々なワークロードを迅速にクラウドに展開できています。

# 三菱電機株式会社 先端技術総合研究所 様

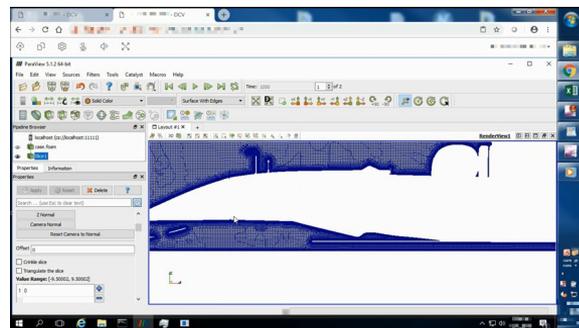
解析システムのクラウド利用推進を目的として、汎用熱流体解析ソフトウェアの検証を実施

- オープンソースベースの汎用熱流体解析ソフトウェア (iconCFD)を用いてAWS上で並列性能ベンチマークテストを実施
- Amazon EC2 の通常のイーサネットでも600コア程度まで性能がスケールすることを確認(今後100GbpsのネットワークやElastic Fabric Adapter(EFA)を使用することで更なるスケーラビリティ向上も期待できる)
- 計算結果の可視化処理もAWS上で実行、GPU搭載インスタンスを利用することで1億メッシュ規模のデータも問題なく扱えた
- クラウドの利用により、初期投資を小さくし、大規模な解析をすぐに実行することができる

• iconCFDは、Icon Technology & Process Consulting Ltd.とIDAJIによって開発されました。  
• iconCFDは、GPLに準じたオープンソースベースの汎用熱流体解析プログラムです。



Amazon EC2上での並列処理性能検証結果



NICE-DCVを利用した1億メッシュ可視化検証

# 株式会社リコー 様

9/8~30のAWS Summit Onlineにて公開された、リコー様によるEDAワークロードのクラウド化事例セッション動画と資料が公開中です。是非ご覧ください。

リコーのデバイス設計システム クラウド化への道  
~評価して分かった効果と注意点~

株式会社リコー  
イノベーション本部 先端デバイス研究センター  
第2研究室 第7研究グループ  
西村行雄

aws SUMMIT ONLINE

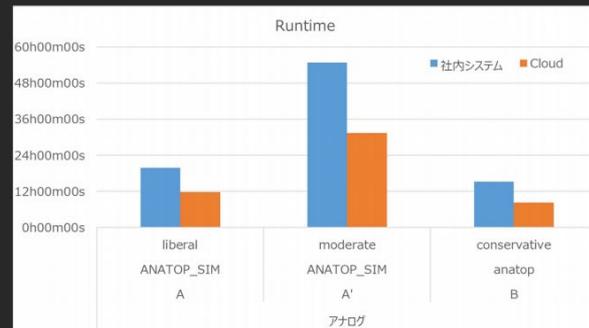
CUS-34：リコーのデバイス設計システムクラウド化への道~評価して分かった効果と注意点~

## 評価結果

シミュレーションの実行時間に差が確認出来た。  
AWSでの実行時間が約30~50%短い



シミュレーション時間  
短縮に効果あり。



同時に使用されるコアの効率が向上した？！

⇒最大同時使用コア数は16コアで同じだがAWSの方が同時に動くコアの割合が多かった？

シングルコアで動作するツールについては短縮効果は出ていない。

AWS Summit Online Japan 2020

リコーのデバイス設計システム クラウド化への道 ~評価して分かった効果と注意点~

<https://resources.amazonaws.com/aws-summit-online-japan-2020-on-demand-industry-1-09272/cus-34-aws-summit-online-2020-ricoh-2>

## 「一日でも早く薬を届ける」ためにクラウドHPCを活用

多様な創薬ワークロードに対して、適したAWSのサービスを選択

### ゲノム分析

- **AWS Batch** を活用し、ゲノム分析に必要な解析パイプラインを構築
- TB級のデータを必要な時間内での処理を実現
- **スポットインスタンス**の利用でEC2コストを約50%～節約



**AWS Batch**

### タンパク質立体構造解析

- **AWS ParallelCluster** を使用し既存のジョブスケジューラから変えずに移行
- P3インスタンスによるGPU活用
- **NICE-DCV**により計算結果をクラウド上で即座に確認



**AWS ParallelCluster**

# 国立大学法人 京都大学大学院医学研究科附属ゲノム医学センター 様

大規模なヒトゲノム解析による難病研究向けプラットフォーム  
クラウドとスパコンの両方の利点を生かすため、ハイブリッドのシステムを構築

## CGMハイブリッドクラウドシステムの実装（1）

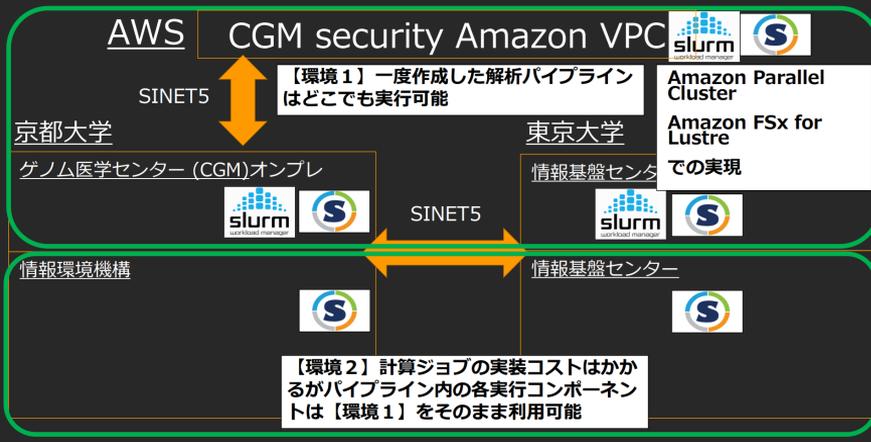
### バッチジョブシステム

 全世界のTOP500 supercomputers  
の半数以上で利用

コンテナの利用による再現性・再利用性・信頼性の担保

 Singularity v3を用いたパイプライン構築

## CGMハイブリッドクラウドシステムの実装（2）



AWS Summit Online Japan 2020

ヒトゲノム情報統合解析に向けた京都大学ゲノム医学センターのハイブリッドクラウドシステム構築について

<https://resources.awscloud.com/aws-summit-online-japan-2020-on-demand-industry-2-55513/cus-12-aws-summit-online-2020-japan-cgm-kyoto-university>

# 株式会社ウェザーニューズ様

台風やゲリラ豪雨など気象リスクに対して  
短い予報間隔・高い更新頻度の気象予報をユーザーに配信するためクラウドを活用

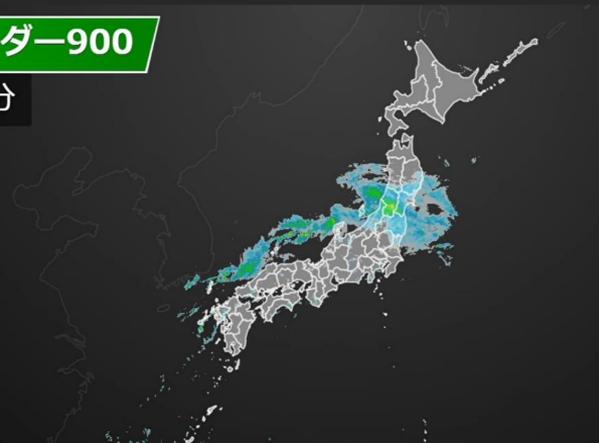
## 気象予測が出来るまで



## 成果物

AIレーダー-900

10時30分



AWS Summit Online Japan 2020

新たな気象リスクへの挑戦を可能にした HPC on AWS

<https://resources.awscloud.com/aws-summit-online-japan-2020-on-demand-industry-3-12034/cus-98-aws-summit-online-2020-weathernews>

© 2020, Amazon Web Services, Inc. or its Affiliates. All rights reserved.



# 株式会社ウェザーニューズ様

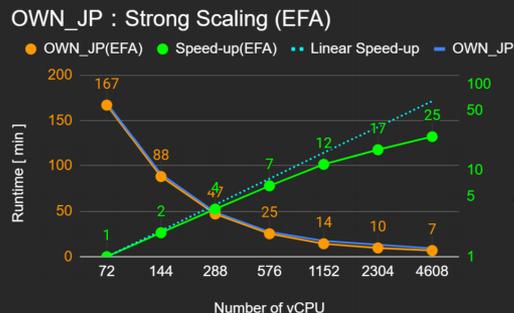
AWS ParallelCluster + Elastic Fabric Adapter によるスケラビリティ  
一定期間内に計算を完了させるためのmulti-regionクラスタ構成  
スポットインスタンス活用

## EFAの利用

Parallel Clusterの設定  
でEFAを用いた場合

特に並列数を増やした  
場合に**高速化**された

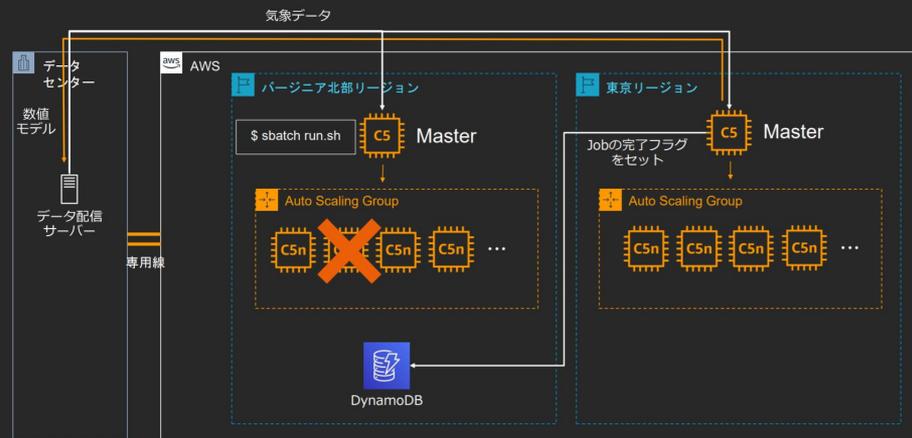
vCPU数が少ない場合は  
1~3%の改善だが  
4068まで並列化させた場合は  
25%の改善が見られた



改善率(%)

1	3	4	7	18	25
---	---	---	---	----	----

## 副系での計算処理の流れ



AWS Summit Online Japan 2020

新たな気象リスクへの挑戦を可能にした HPC on AWS

<https://resources.awscloud.com/aws-summit-online-japan-2020-on-demand-industry-3-12034/cus-98-aws-summit-online-2020-weathernews>

© 2020, Amazon Web Services, Inc. or its Affiliates. All rights reserved.



# 最新事例に含まれる HPC on AWS トレンド

- これまでは、製造や自動車などCAE領域での活用が多かったが、**創薬や気象予測といった幅広い分野での本番環境**としての利用が進んでいる
- AWSの活用方法については、個人向け環境であったり共通基盤であったり自動的なジョブ実行環境であったり様々
  - 使用しているサービスも、Amazon EC2（オーケストレーション無し）、AWS Batch、Amazon ECS、AWS ParallelCluster 等 **用途に応じて使い分け**ていただいている
  - 既存のHPCクラスタにとらわれない多様な計算環境の活用
- Elastic Fabric Adapterを活用することで、流体や気象のシミュレーションでも数百～数千 vCPU規模までスケール
- NICE-DCVのようなリモートデスクトップ環境により、データをクラウドから持ち出さない運用パターンが見られる

# 應用編

# HPC on AWS ベンチマーク例

# 密結合ワークロードでのEFA活用例

## 計算流体力学

- Solvers - OpenFOAM, Fluent, Star-CCM+, LS-DYNA, OVERFLOW, FUN3D

## 数値気象予測

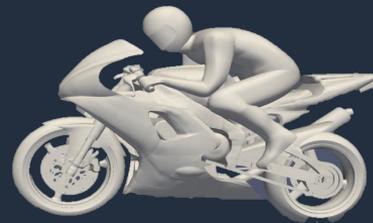
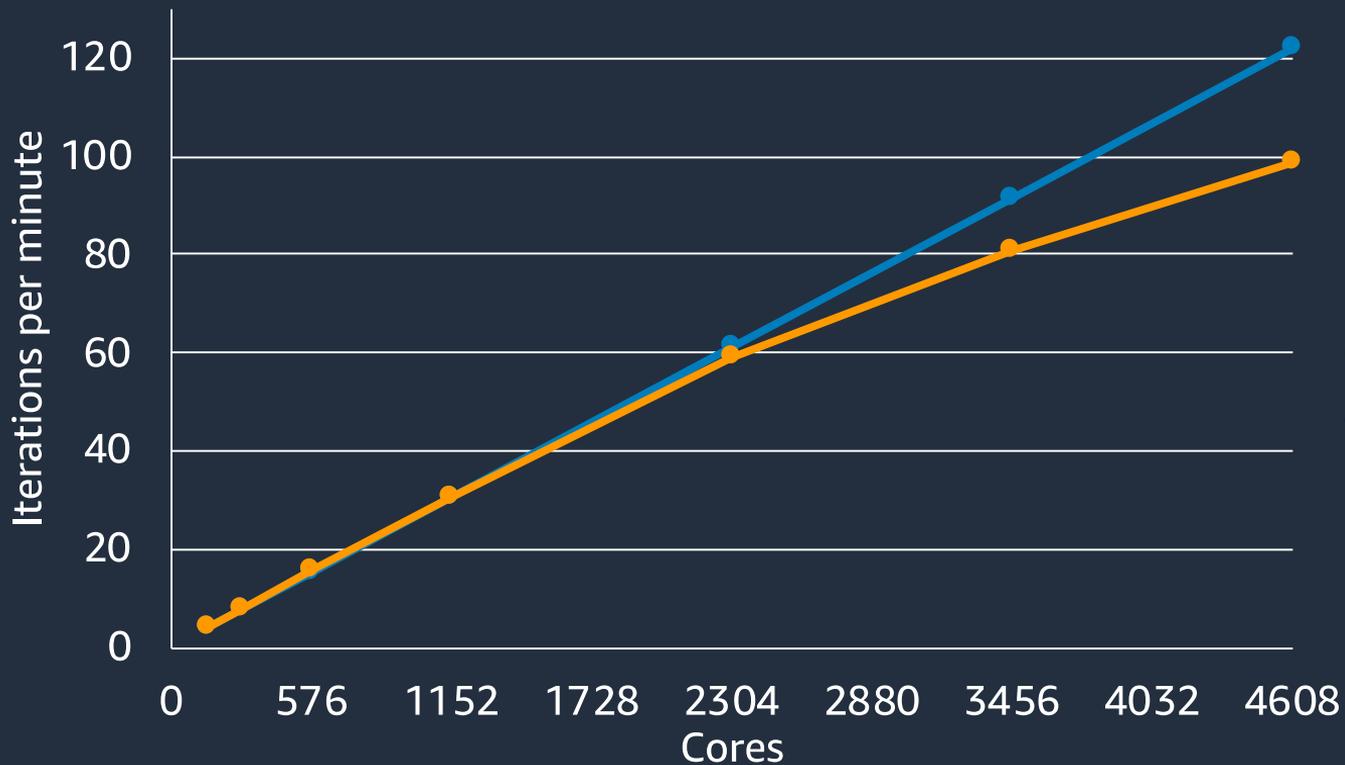
- Weather Models - WRF, FV3, MPAAS

## 分散機械学習

- ML Models - BERT, Seq-2-seq

# Scaling on AWS - OpenFOAM

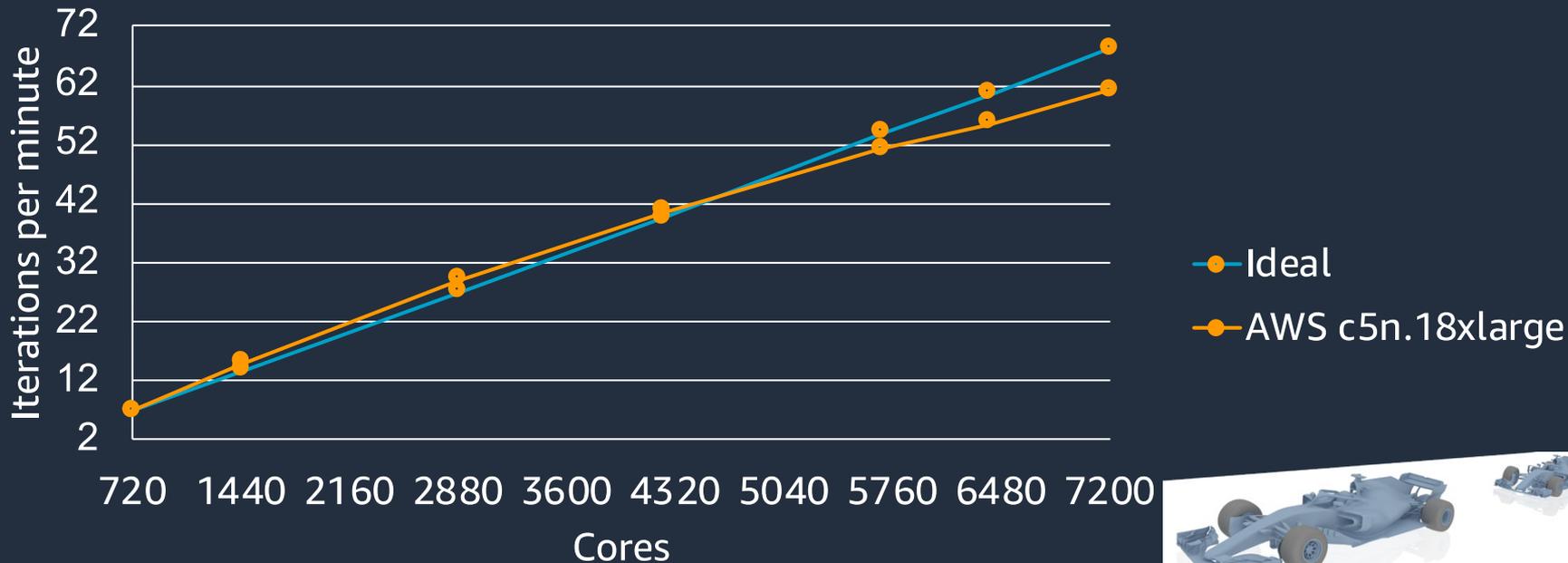
OpenFOAM v1912 - MotorBike (222M cells) - IntelMPI 2019.6 - AL2 - PC2.6.1



—●— Ideal  
—●— AWS c5n.18xlarge

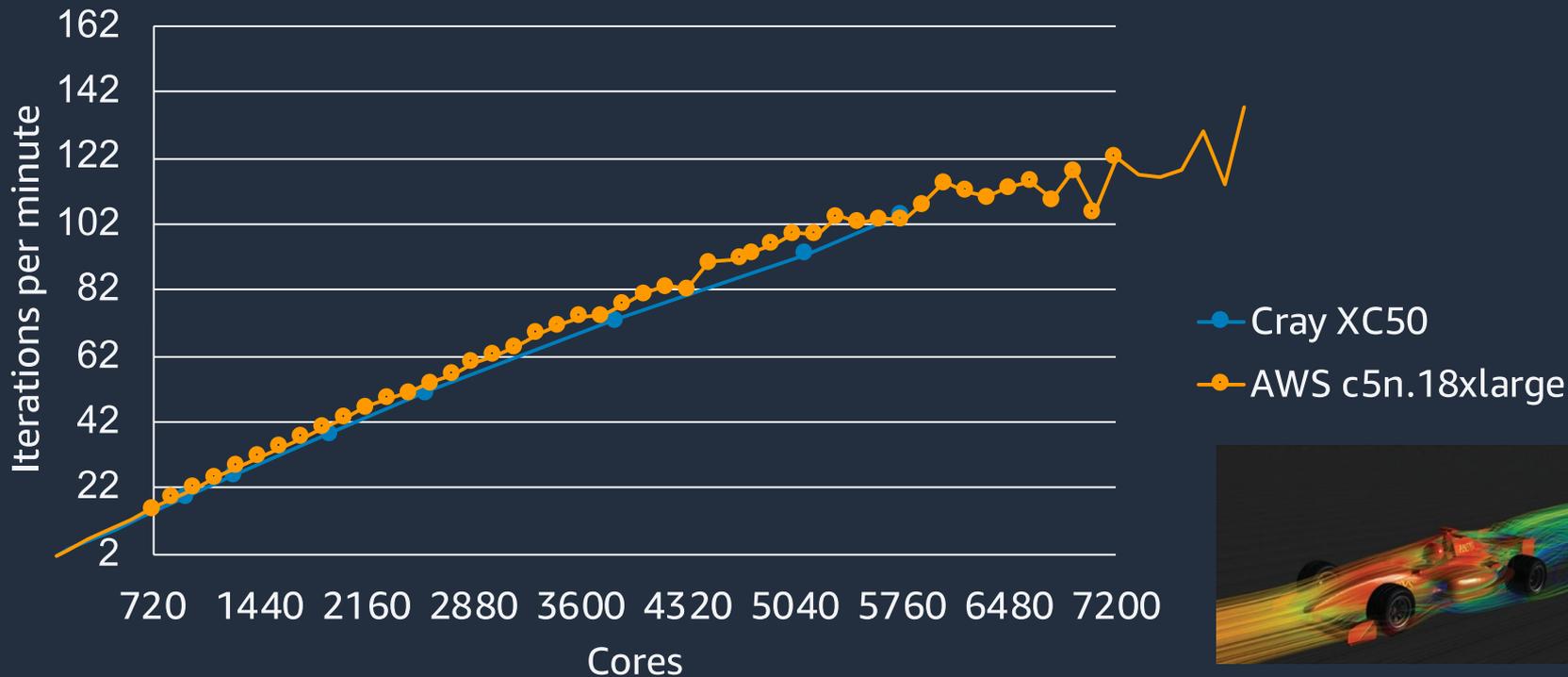
# Scaling on AWS – STAR-CCM+

Simcenter STAR-CCM+ 2020.1 - F1 (403M cells) - IntelMPI 2019.6 - AL2 - PC2.6.1



# Scaling on AWS – Fluent

ANSYS Fluent 19.5 - F1 (140M cells) - IntelMPI 2019.5 - AL2 - PC2.5.1



# Facebook AI Research (FAIR) & EFA

32 * P3dn nodes, NVLINK, all-reduce via NCCL		
256 V100 GPUs	AWS ENA	AWS EFA
<b>BERT (NLP)</b>	1.0x	1.5x
<b>TDS-Seq2Seq (ASR)</b>	1.0x	1.3x

ML Training is sensitive to interconnect

- All-reduce is most common collective
- Large messages O(1-10MB) are common

P3dn with EFA provides up to 100 Gbps Model parallelism

- Different communication patterns
- Needs high bisection bandwidth

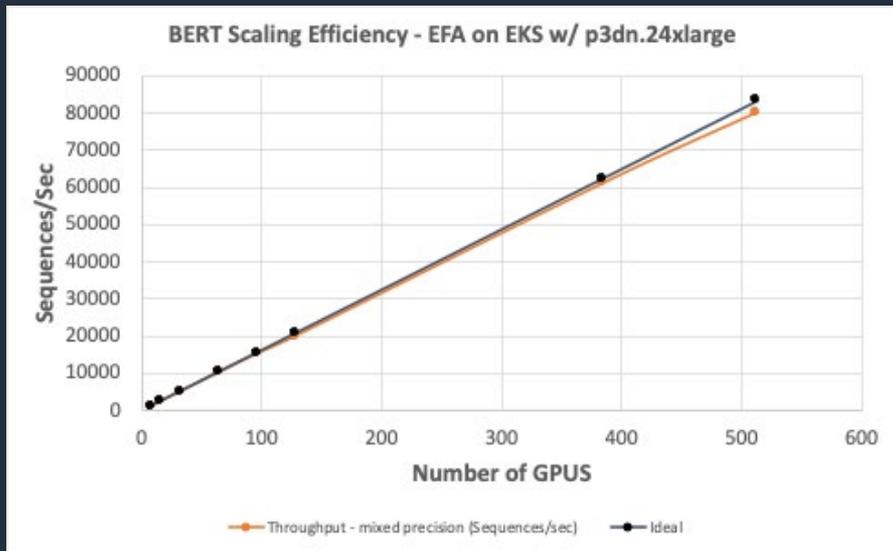
New emerging workloads: graph learning

ENA - Elastic Network Adapter

EFA - Elastic Fabric Adapter



# 参考 : Elastic Kubernetes Service (EKS) & EFA



Number of Nodes	Number of GPUS	Batch Size / GPU	Max Accumulation Steps	Sequence length	Throughput - mixed precision (Sequences/sec) - EFA
1	8	64	128	128	1303
2	16	64	128	128	2561
4	32	64	128	128	5114
8	64	64	128	128	10254
12	96	64	128	128	15315
16	128	64	128	128	20053
48	384	64	256	128	61161
64	512	64	256	128	80190

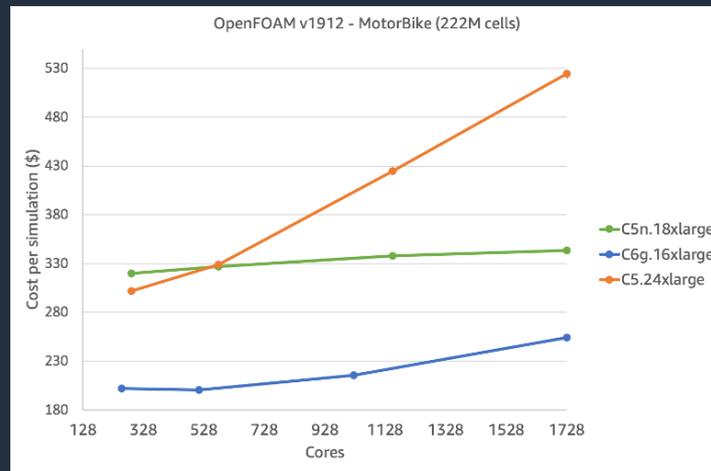
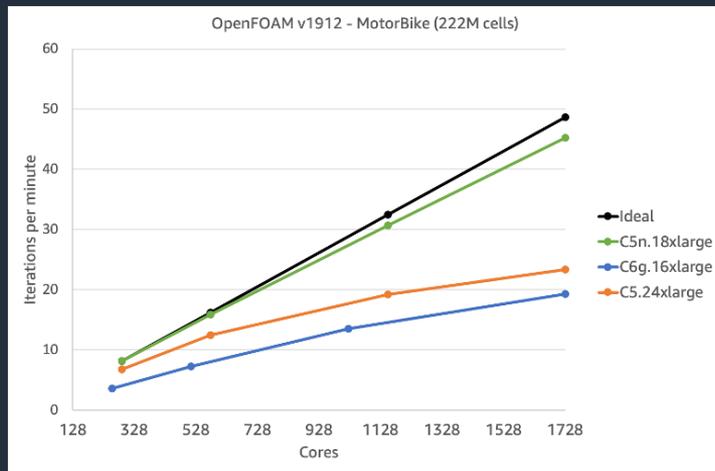
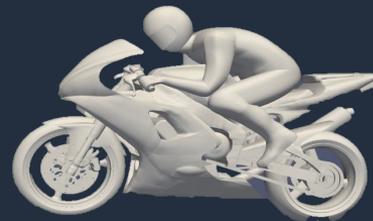
BERT (128 sequence length). Dataset is the Wikipedia/Books Corpus prepared from NVIDIA Deep Learning examples

**At 16 nodes of p3dn.24xl (128 V100 GPUs), we achieve ~96% scaling efficiency**

For more, visit <https://github.com/aws-samples/eks-efa-examples>

# C6g による OpenFOAM Benchmarks

- 222M cell Motorbike モデルを用いて評価
- 計算性能としては、EFAを搭載したC5nが最も高い
- シミュレーションあたりのコストでは、C6g が最も安価



<https://aws.amazon.com/blogs/compute/c6g-openfoam-better-price-performance/>  
<https://gitlab.com/arm-hpc/packages/-/wikis/packages/openfoam>

# Arm 公式 Blog での HPC 利用例

- Evaluation of the NEMO Ocean Model on Arm Neoverse-based AWS Graviton2
  - <https://community.arm.com/developer/tools-software/hpc/b/hpc-blog/posts/evaluation-of-the-nemo-ocean-model-on-aws-graviton2>
- Demonstration of low mach-number CFD modeling with Nalu on AWS Graviton2 M6g instances
  - <https://community.arm.com/developer/tools-software/hpc/b/hpc-blog/posts/low-mach-number-cfd-modeling-with-nalu-on-graviton2-aws-m6g>
- Seismic Modeling with Arm Neoverse N1 and AWS Graviton2
  - <https://community.arm.com/developer/tools-software/hpc/b/hpc-blog/posts/seismic-modeling-with-arm-neoverse-n1-and-aws-graviton2>

# HPC on AWS の構成ポイント

# HPC on AWS におけるパフォーマンスの基本

まずはシンプルな構成で試してみる、動かしてみる事が重要

CPU/メモリ/ネットワーク/ストレージ等、ボトルネックに応じて対処  
(現実的なコスト・計算時間で収まっているのであれば、深追いしないのも一つの戦略)

パフォーマンスチューニングにおける基本的な考え方は AWS 上でも大きく変わるわけではない

# パフォーマンス：コンピュート

## • vCPU 表記

- AWS での vCPU 表記は Intel Hyper-Threading Technology における 1 thread (= 1 論理コア) を示しており、物理コア数は vCPU 数の半分となる点に注意 (例: 72 vCPU = 36 物理コア)

## • CPU設定

- Intel Hyper-Threading Technology の無効化や、NUMA設定、Pinning等は **オンプレミス環境と同様**に行うことができる
- 特に最大のインスタンスサイズでは 複数 Socket 構成となっている事が多いため注意

## • インスタンス選択

- C5/C4 (コンピュート最適化) を中心に、より多くのメモリを必要とする場合に、M5/M4 (汎用) や R5/R4 (メモリ最適化) を選択
- 1 コアあたりのパフォーマンスを求める場合は高いクロック周波数に対応した Z1d/M5zn も検討
- C5a/M5a/R5a といった AMDタイプインスタンスによりコストメリットが生まれる場合もある
- 世代の新しいインスタンスタイプの方が性能は良いが、旧世代は Spot インスタンス価格が下がっていることもあるため、**コストパフォーマンスを考慮して選択**

## • リージョン選択

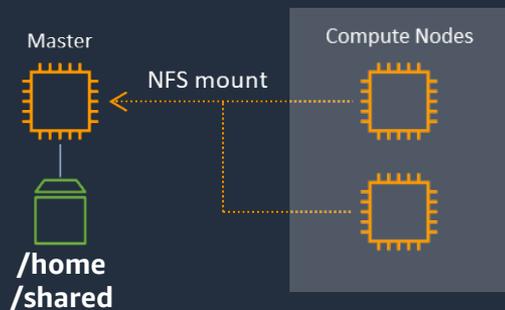
- 10,000 vCPU ~ など大規模利用を行う場合は、インスタンスコストの安価な北米リージョン (us-east-1 等) の利用を検討

# パフォーマンス：ストレージ

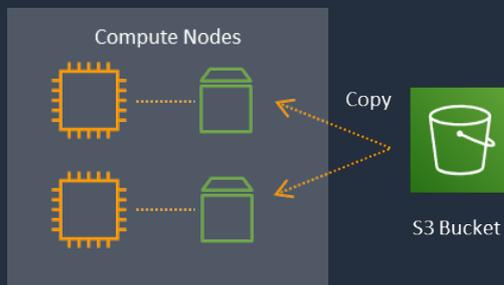
データの移動パターン等によって選択

- **NFS on Master**: ParallelClusterのデフォルト構成、扱いやすいが、Masterのインスタンスサイズや、EBS の性能に注意
- **S3 to local disk**: ジョブ開始時にS3等からローカルディスクにデータをコピー  
データ転送が最初と最後のみの場合に選択
- **FSx for Lustre**: 高速かつスケーラブルな共有ディスクが必要な場合に選択

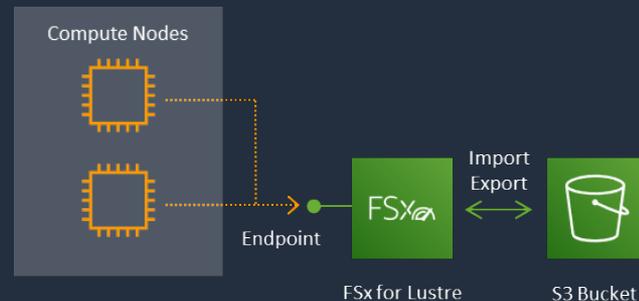
## NFS on Master (default)



## S3 to local disk



## FSx for Lustre



# パフォーマンス : ネットワーク

MPI による通信が頻繁に発生するワークロードでは、Cluster Placement Group や、Elastic Fabric Adapter の利用を検討

- **Cluster Placement Group**: AZ 内の、できるだけ物理的に近いハードウェア上でインスタンスを起動するように指定  
インスタンス間通信が低レイテンシ化
- **Elastic Fabric Adapter**: MPI/NCCL に特化した専用の低レイテンシネットワークアダプタ
  - EFA 対応 MPI: OpenMPI 4.0.2 、 Intel MPI 2019 Update 6 が対応
  - EFA 利用可能インスタンスタイプ: c5n.18xlarge, c5n.metal, i3en.24xlarge, i3en.metal, inf1.24xlarge, m5dn.24xlarge, m5n.24xlarge, r5dn.24xlarge, r5n.24xlarge, p3dn.24xlarge, p4d.24xlarge etc..

※ そもそもノード間通信を発生させないために、c5.24xlarge (96 vCPU) 等大きなインスタンスサイズを使用することも検討

[https://docs.aws.amazon.com/ja\\_jp/AWSEC2/latest/UserGuide/placement-groups.html](https://docs.aws.amazon.com/ja_jp/AWSEC2/latest/UserGuide/placement-groups.html)

[https://docs.aws.amazon.com/ja\\_jp/AWSEC2/latest/UserGuide/efa.html](https://docs.aws.amazon.com/ja_jp/AWSEC2/latest/UserGuide/efa.html)

# オーケストレーション：サービス選択



## AWS ParallelCluster

- **Slurm/SGE/Torque** 等の利用者に馴染みのあるジョブスケジューラを利用可能
- 単一のジョブが大量の CPU core を使用する密結合ワークロードに向いている
- **Pros:** 既存の HPC クラスタからの移行が容易
- **Cons:** 単一の AZ、単一のインスタンスタイプでの利用となるため、可用性確保には工夫が必要



## AWS Batch

- **フルマネージド**の独自ジョブスケジューラ
- 少数の CPU core を使用するジョブを大量に実行する疎結合・High Throughput Computing ワークロードに向いている
- **Pros:** フルマネージドサービスでありテナンサさえ用意すれば計算基盤の管理は不要
- **Pros:** 複数 AZ や複数インスタンスタイプの利用が可能であり、可用性を確保しやすい
- **Cons:** コンテナ化やスケジューラ対応が必要

# オーケストレーション：ユーザ認証

## AWS ParallelCluster

- ParallelCluster はデフォルトでは単一ユーザーのみ利用可能な設定
- 複数ユーザ向けの共通基盤として利用したい場合には、以下のドキュメントが参考になる
  - シンプルなマルチユーザ構成  
<https://github.com/aws/aws-parallelcluster/wiki/MultiUser-Support>
  - Active Directory (AWS Directory Service) との連携  
<https://aws.amazon.com/jp/blogs/opensource/aws-parallelcluster-aws-directory-services-authentication/>
  - OpenLDAPとの連携  
<https://aws.amazon.com/jp/blogs/opensource/managing-aws-parallelcluster-ssh-users-with-openldap/>
- ユーザごとに独立したクラスタを構築するパターンもある

## AWS Batch

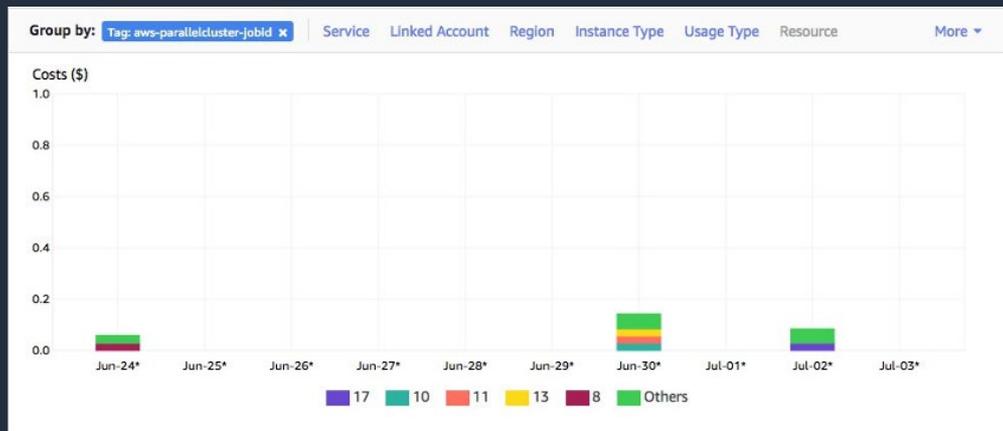
- ジョブ送信など各API呼び出し権限はIAMでの管理となる
- IAM PolicyのConditionでテナ内での利用可能POSIXユーザの制限が可能

# Blog: Using cost allocation tags with AWS ParallelCluster

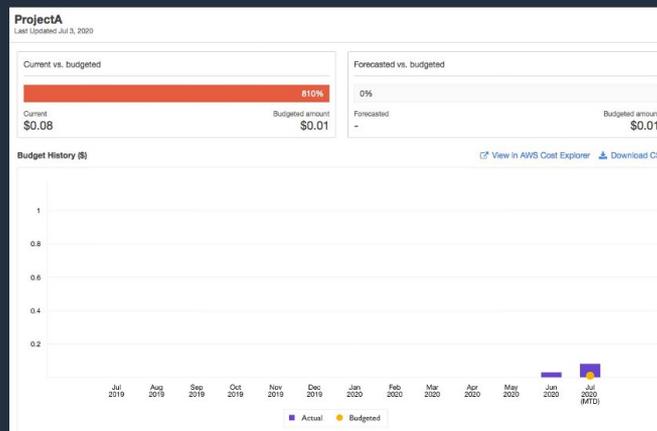
AWS ParallelClusterで、ジョブIDやプロジェクトごとにコストを可視化するための参考実装

- ジョブごとにタグを追加
- Cost Explorer や AWS Budgets と連携して、予算管理が可能（一定金額を超えていたらジョブを受け付けない等）

Cost Explorer



AWS Budgets

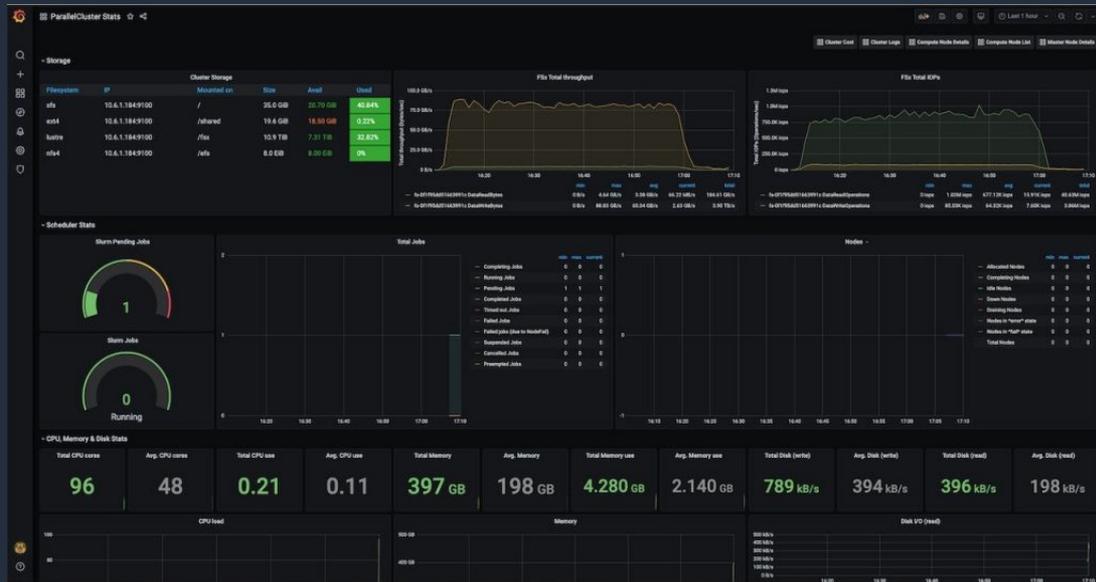


<https://aws.amazon.com/jp/blogs/compute/using-cost-allocation-tags-with-aws-parallelcluster/>

# Blog: Monitoring dashboard for AWS ParallelCluster

## Grafana/Prometheusによる高度なParallelCluster用ダッシュボード作成

- ParallelCluster 2.10 以降で post\_install スクリプトを指定するだけで利用可能
- ブラウザ経由で Master ノードで動作する Grafana のダッシュボードにアクセス
- CPU 使用率、メモリ使用量に加え、Prometheus-Slurm-Exporter を使用することでジョブやノードの情報も収集可能



<https://aws.amazon.com/jp/blogs/compute/monitoring-dashboard-for-aws-parallelcluster/>

# アップデート編

# HPC on AWS 関連サービス主要アップデート

# Amazon EC2 M5zn インスタンス

Cascade Lake世代のIntel Xeon Scalable Processorを搭載し、全コア同時に Turbo boostによる **4.5GHz** で動作する M5zn をローンチ

- 同様に高CPUクロックに対応していた Z1d インスタンスと比較して、より高いCPUクロック、高速なネットワーク帯域に対応
- m5zn.12xlarge/m5zn.metalは Elastic Fabric Adapter にも対応し、低レイテンシでの MPI通信が可能

	z1d.12xlarge	m5zn.12xlarge
vCPU数	48	48
最大クロック周波数	4.0 GHz	<b>4.5 GHz</b>
RAM	384 GiB	192 GiB
ネットワーク帯域幅	25 Gbps	<b>100 Gbps</b>
EFA 対応	なし	<b>あり</b>
価格※	\$ 4.464 USD/h	\$3.9641 USD/h

※ N.Virginia リージョンにおける2020年12月時点のオンデマンド価格

AWS US East (N. Virginia) 、 US East (Ohio) 、 US West (N.California) 、 US West (Oregon) 、 EU (Ireland) 、 EU (Frankfurt) 、 およびAsia Pacific(東京)で利用可能

<https://aws.amazon.com/jp/about-aws/whats-new/2020/12/introducing-amazon-ec2-m5zn-instances/>

# Amazon EC2 AWS Graviton2 プロセッサ搭載インスタンス

Arm ベースの AWS Graviton2 プロセッサを搭載し、  
同等の現行世代x86ベースのインスタンスよりも**最大40%優れた価格性能**を提供

## M6g

汎用

4GB DRAM/vCPU

## C6g

コンピューティング最適化

2GB DRAM/vCPU

## R6g

メモリ最適化

8GB DRAM/vCPU

### 全てGA済み!

ローカルNVMe SSD ストレージを搭載したインスタンス  
汎用 (**M6gd**), コンピューティング最適化 (**C6gd**), メモリ最適化(**R6gd**)  
全てのインスタンスタイプにおいてベアメタルの選択肢を提供

バーストタイプである **T4g** も提供済み (2021年3月まで無料枠も提供中)

# Amazon EC2 C6gn インスタンス (Coming soon)

AWS Graviton2 搭載インスタンスについて、より高速なネットワークを搭載した C6gn インスタンスを発表

- C6g と比較して、4倍高速な 100 Gbpsのネットワーク帯域、4倍のパケット処理性能、2倍高速な 38 GbpsのEBS帯域を提供
- C6gn.16xlarge は Elastic Fabric Adapter にも対応し、低レイテンシでのMPI通信が可能

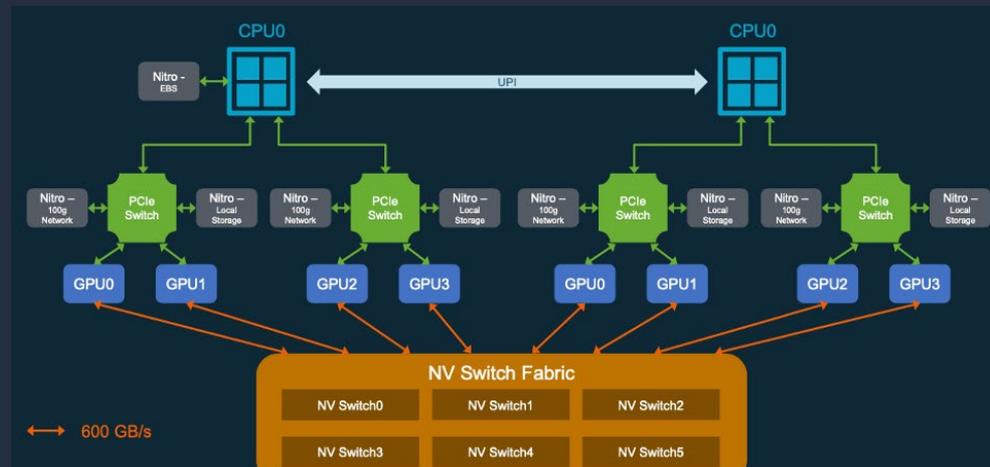
	c6g.16xlarge	c6gn.16xlarge
vCPU数	64	64
RAM	128 GiB	128 GiB
EBS 帯域	19 Gbps	<u>38 Gbps</u>
ネットワーク帯域幅	25 Gbps	<u>100 Gbps</u>
EFA 対応	なし	<u>あり</u>

<https://aws.amazon.com/jp/blogs/aws/coming-soon-ec2-c6gn-instances-100-gbps-networking-with-aws-graviton2-processors/>

# Amazon EC2 P4d インスタンス

**NVIDIA A100 Tensor Core GPU** を搭載した P4d インスタンスがローンチ

- GPU間は 600 Gbps の NVSwitch/NVLink で接続 (下図)
- インスタンスあたり **400 Gbps** の高速なネットワークインターフェースを提供
- EFA 及び GPUDirect RDMA over EFA に対応し、マルチノードでの通信性能が向上
- 現在は p4d.24xlarge (A100 x 8枚搭載) の 1サイズのみ提供



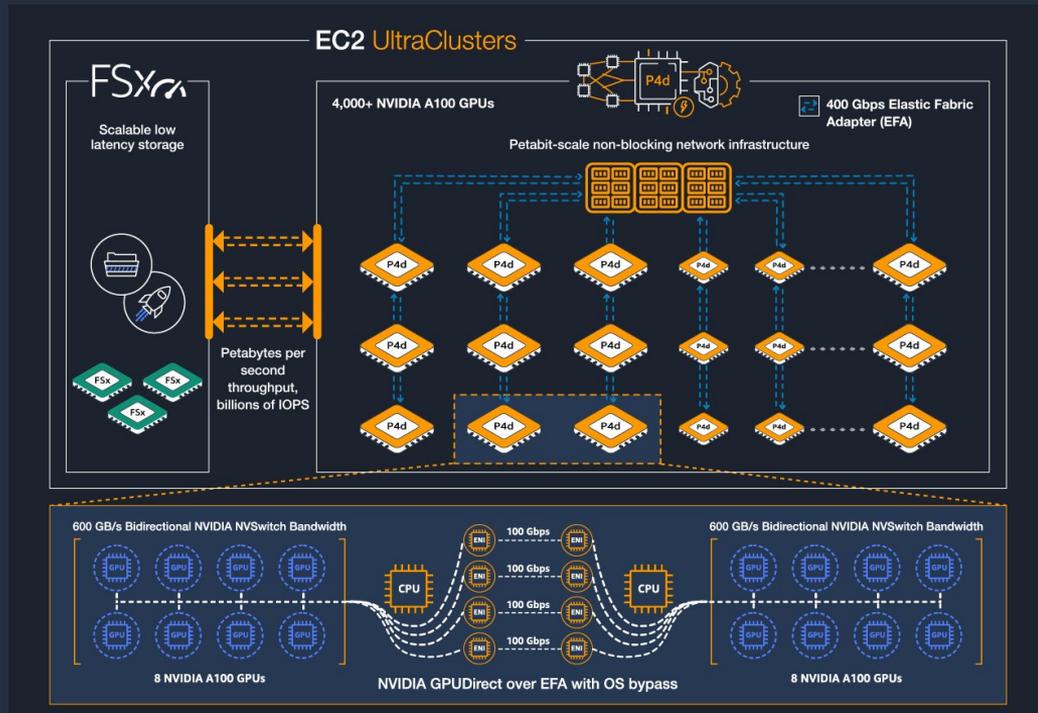
US East (N. Virginia) 、 US West (Oregon) で利用可能

<https://aws.amazon.com/jp/blogs/compute/amazon-ec2-p4d-instances-deep-dive/>

# EC2 Ultra Clusters of P4d Instances

クラウド上でスーパーコンピュータクラスの性能を実現

- インスタンスあたり400 Gbpsの高速なネットワーク
- GPUDirect RDMA over EFAによる低レイテンシ通信
- 4000を超えるA100 Tensor Core GPUを接続
- FSx for Lustre による高速な分散ストレージとS3への永続化
- 必要なときに起動して、使用した分のみの支払い



# Amazon EC2 G4adインスタンス (Coming soon)

AMDの **Radeon Pro V520 GPU** とEPYC Roma プロセッサを搭載したインスタンスを発表

- ゲームストリーミング、リモートグラフィックスワークステーション、レンダリングなど、グラフィックスを多用するワークロードに最適
- グラフィックス用GPU搭載インスタンスとしては G4dn (NVIDIA T4搭載) もローンチ済みであり、多様なGPUの選択肢を提供
- CUDA、CudNN、NVENC等のNVIDIAのライブラリを使用しない場合は G4ad により、より高いコストパフォーマンスが得られる可能性

<https://aws.amazon.com/jp/blogs/aws/new-amazon-ec2-g4ad-instances-featuring-amd-gpus-for-graphics-workloads/>

# その他の Amazon EC2 インスタンス関連アップデート

## D3/D3en インスタンス ローンチ

- クラウドで最高密度のローカルHDDストレージ（最大 336 TB）を提供
- 分散/クラスター化されたファイルシステム、ビッグデータと分析、大容量のデータレイクなどのワークロードに最適

## R5b インスタンス ローンチ

- 最大 60Gbps の EBS 帯域幅と 260K IOPS をサポートした、次世代のAWS Nitro Systemに基づいたインスタンス
- 大規模なリレーショナルデータベースワークロードに最適

# FSx for Lustre: ストレージタイプ追加、スナップショット

これまでの SCRATCH\_1 に加えて、より高いバースト性能を持った **SCRATCH\_2** や 永続領域として利用可能な **PERSISTENT** (SSDタイプ、HDDタイプ、HDD+SSDキャッシュタイプの3種類にそれぞれにパフォーマンスの異なるバリエーションが存在) が登場  
PERSISTENT については、スナップショットの保存・復元にも対応

ストレージオプション	Persistent (GB あたり月額料金 USD)	Scratch (GB あたり月額料金 USD)
<b>HDD ベースのファイルシステム</b>		
12 MB/秒/TiB ベースライン、最大 375 MB/秒/TiB バースト	0.025USD	-
12 MB/秒/TiB ベースライン、最大 375 MB/秒/TiB バースト (SSD キャッシュ使用)	0.041USD	-
40 MB/秒/TiB ベースライン、最大 1.3 GB/秒/TiB バースト	0.083USD	-
40 MB/秒/TiB ベースライン、最大 1.3 GB/秒/TiB バースト (SSD キャッシュ使用)	0.099USD	-
<b>SSD ベースのファイルシステム</b>		
50 MB/秒/TiB ベースライン、最大 1.3 GB/秒/TiB バースト	0.140USD	-
100 MB/秒/TiB ベースライン、最大 1.3 GB/秒/TiB バースト	0.190USD	-
200 MB/秒/TiB ベースライン、最大 1.3 GB/秒/TiB バースト	0.290USD	0.14USD

※ N.Virginia リージョンにおける2020年12月時点の価格

<https://aws.amazon.com/jp/blogs/news/amazon-fsx-for-lustre-persistent-storage/>

<https://aws.amazon.com/about-aws/whats-new/2020/06/amazon-fsx-lustre-provides-highly-durable-file-system-backups/>

# AWS ParallelCluster: Multiple queues 対応

1クラスタに複数のジョブキュー（パーティション）を設定可能に

- ジョブキューごとにインスタンスタイプや購入オプションを指定可能
- On-demand/SpotやCPU/GPUの併用が容易に
- ジョブスケジューラは Slurm のみ対応
- ParallelCluster 2.9以降で利用可能

```
[queue od-queue]
compute_resource_settings = c5n-od
compute_type = ondemand
```

```
[queue spot-queue]
compute_resource_settings = c5n-spot
compute_type = spot
```

```
[compute_resource c5n-od]
instance_type = c5n.18xlarge
min_count = 0
max_count = 10
```

```
[compute_resource c5n-spot]
instance_type = c5n.18xlarge
min_count = 0
max_count = 10
```

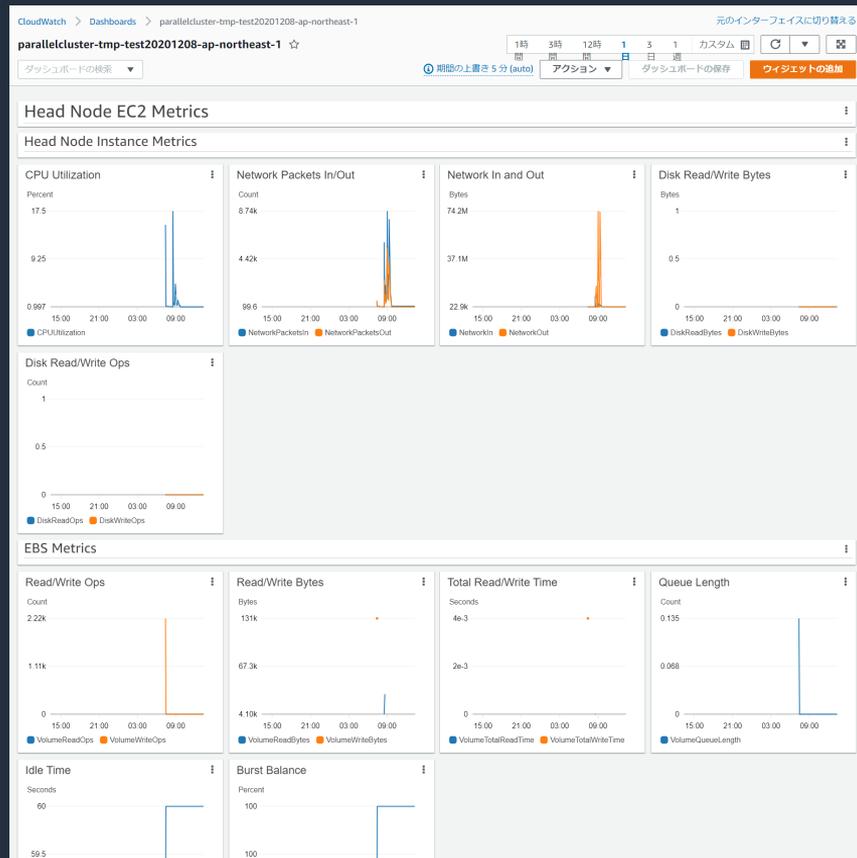
ParallelClusterでの設定例

<https://aws.amazon.com/jp/blogs/opensource/using-multiple-queues-and-instance-types-in-aws-parallelcluster-2-9/>

# AWS ParallelCluster: CloudWatch ダッシュボード自動作成

クラスタの各種メトリクスを集中監視可能なダッシュボードが自動で作成

- CPU使用率の他、EBSの各種メトリクス、Slurmのログなどを監視
- ParallelCluster 2.10以降で対応（デフォルトで有効）



# AWS Batch: Fargate/Fargate Spot 対応

Compute Environment が計算基盤としてFargate/Fargate Spotに対応

- 0.25 vCPUなどの小さい単位でのリソース指定が可能に
- ジョブ定義はFargate環境専用で作成する必要がある

## Provisioning model

- On-demand**  
EC2 instances that are billed per second.

- Spot**  
Save money by using Spot instances but be aware your instances can be interrupted with a two minute notification when EC2 needs the capacity back.

- Fargate**  
Run AWS Batch jobs on serverless compute for containers with Fargate; no need to manage EC2 instances or AMI's

- Fargate Spot**  
Fargate Spot allows you to run interruption tolerant AWS Batch jobs at up to a 70% discount off the Fargate price.

<https://aws.amazon.com/jp/blogs/aws/new-fully-serverless-batch-computing-with-aws-batch-support-for-aws-fargate/>

# まとめ

- クラウドのスケー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することが可能
- AWSでは、仮想サーバだけでなく、ネットワーク、ストレージ、オーケストレーション、可視化など多種多様な HPC 関連サービスを提供中
  - 特にEFAを活用することで、密結合なワークロードでも高いスケーラビリティを得ることが可能に
- 日本国内でも製造、半導体設計、創薬、ゲノミクス、気象予測等様々な分野での本番活用が行われている
  - オンプレミスの環境をそのまま移行するだけではなく、より「クラウドに向けた」使い方に改善していくことで、より良いコストパフォーマンスを得ることが可能

# HPC on AWS 関連資料

- HPC on AWS  
<https://aws.amazon.com/jp/hpc/>
- HPC on AWS ハンズオン  
<http://bit.ly/aws-hpc>
- HPC on AWS @ 2019 30分でわかるクラウドHPCの現在  
<https://www.slideshare.net/DaisukeMiyamoto6/hpc-on-aws-2019>
- クラウド規模での Western Digital HDD シミュレーション – HPC タスク 250 万件、EC2 スポットインスタンス 4 万個  
<https://aws.amazon.com/jp/blogs/news/western-digital-hdd-simulation-at-cloud-scale-2-5-million-hpc-tasks-40k-ec2-spot-instances/>
- Saving Koalas Using Genomics Research and Cloud Computing  
<https://aws.amazon.com/jp/blogs/aws/saving-koalas-using-genomics-research-and-cloud-computing/>

# AWS ParallelCluster 関連 Blog: Architecture

- Deploying an HPC cluster and remote visualization in a single step using AWS ParallelCluster <https://aws.amazon.com/blogs/opensource/deploy-hpc-cluster-remote-visualization-single-step-parallelcluster/>
- AWS API Gateway for HPC job submission <https://aws.amazon.com/blogs/opensource/aws-api-gateway-hpc-job-submission/>
- AWS ParallelCluster with AWS Directory Services Authentication <https://aws.amazon.com/blogs/opensource/aws-parallelcluster-aws-directory-services-authentication/>
- Enabling job accounting for HPC with AWS ParallelCluster and Amazon RDS <https://aws.amazon.com/blogs/compute/enabling-job-accounting-for-hpc-with-aws-parallelcluster-and-amazon-rds/>
- Building an HPC cluster with AWS ParallelCluster and Amazon FSx for Lustre <https://aws.amazon.com/blogs/storage/building-an-hpc-cluster-with-aws-parallelcluster-and-amazon-fsx-for-lustre/>

# AWS ParallelCluster 関連 Blog: Software

- Best Practices for Running Ansys Fluent Using AWS ParallelCluster  
<https://aws.amazon.com/jp/blogs/opensource/best-practices-running-ansys-fluent-aws-parallelcluster/>
- Running Simcenter STAR-CCM+ on AWS with AWS ParallelCluster, Elastic Fabric Adapter and Amazon FSx for Lustre  
<https://aws.amazon.com/blogs/compute/running-simcenter-star-ccm-on-aws/>
- Scale HPC Workloads with Elastic Fabric Adapter and AWS ParallelCluster  
<https://aws.amazon.com/blogs/opensource/scale-hpc-workloads-elastic-fabric-adapter-and-aws-parallelcluster/>
- Building an interactive and scalable ML research environment using AWS ParallelCluster  
<https://aws.amazon.com/blogs/machine-learning/building-an-interactive-and-scalable-ml-research-environment-using-aws-parallelcluster/>

# HPC 関連 Black Belt Online Seminar

- Amazon Elastic Compute Cloud (EC2)  
<https://www.slideshare.net/AmazonWebServicesJapan/20190305-aws-black-belt-online-seminar-amazon-ec2>
- Amazon Elastic Block Store (EBS)  
<https://www.slideshare.net/AmazonWebServicesJapan/20190320-aws-black-belt-online-seminar-amazon-ebs>
- AWS Batch  
<https://www.slideshare.net/AmazonWebServicesJapan/20190911-aws-black-belt-online-seminar-aws-batch>
- Amazon FSx for Lustre  
<https://www.slideshare.net/AmazonWebServicesJapan/20190319-aws-black-belt-online-seminar-amazon-fsx-for-lustre>
- Amazon EC2 スポットインスタンス  
<https://www.slideshare.net/AmazonWebServicesJapan/20190306-aws-black-belt-online-seminar-amazon-ec2>

# Q&A

お答えできなかったご質問については

AWS Japan Blog 「<https://aws.amazon.com/jp/blogs/news/>」にて  
後日掲載します。

# AWS の日本語資料の場所「AWS 資料」で検索



日本担当チームへお問い合わせ サポート 日本語 ▾ アカウント ▾

コンソールにサインイン

製品 ソリューション 料金 ドキュメント 学習 パートナー AWS Marketplace その他 🔍

## AWS クラウドサービス活用資料集トップ

アマゾン ウェブ サービス (AWS) は安全なクラウドサービスプラットフォームで、ビジネスのスケールと成長をサポートする処理能力、データベースストレージ、およびその他多種多様な機能を提供します。お客様は必要なサービスを選択し、必要な分だけご利用いただけます。それらを活用するために役立つ日本語資料、動画コンテンツを多数ご提供しております。(本サイトは主に、AWS Webinar で使用した資料およびオンデマンドセミナー情報を掲載しています。)

[AWS Webinar お申込 »](#)

[AWS 初心者向け »](#)

[業種・ソリューション別資料 »](#)

[サービス別資料 »](#)

<https://amzn.to/JPArchive>



# AWS Well-Architected 個別技術相談会

毎週“W-A個別技術相談会”を実施中

- AWSのソリューションアーキテクト(SA)に  
対策などを相談することも可能

- **申込みはイベント告知サイトから**

(<https://aws.amazon.com/jp/about-aws/events/>)

AWS イベント

で[検索]



AWS Well-Architected



# ご視聴ありがとうございました

AWS 公式 Webinar  
<https://amzn.to/JPWebinar>



過去資料  
<https://amzn.to/JPArchive>

