

フルマネージド型バッチ処理サービスのご紹介

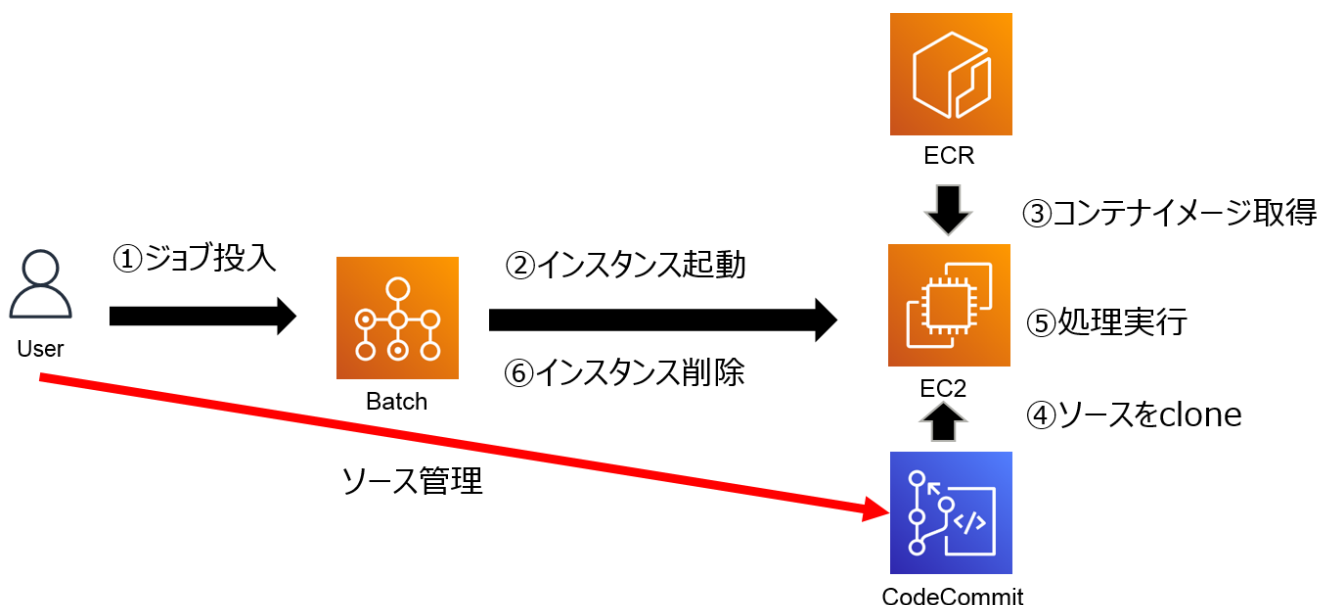
「AWS Batch」ってなに？

- AWS Batchとは、バッチ処理を行うためのマネージドサービスであり、開発者、科学者およびエンジニアが数十万件のバッチコンピューティングをAWSで簡単にかつ効率的に実行ができます。
⇒複数のプログラムによる定時処理をジョブスケジューラによって実行する夜間バッチのようなイメージではありません。
- AWS Batchの特徴は以下になります。
 - ・ 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基に作成し、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。
 - ・ 100万vCPUクラスの大規模な計算にも対応。

構成例

- AWS Batchの基本構成については、[こちら](#)（38ページ目：AWS Batchのアーキテクチャ）をご参照下さい。
- AWS Batchの構成例は以下になります。
⇒ユーザーはCodeCommitでソース管理をしていてAWS Batchにジョブを投入すると自動で実行され、自動でインスタンスが削除されます。

※インスタンス削除は、コンピュートリソースの最小vCPU数の設定が必要になります。またCodeCommitから処理実行するためのスクリプトをGitでCloneしてコンテナで実行する動作についてはAWS Batchの投入するジョブではありますが、EC2のコンピューティング側での処理になります。



■ ユースケース

- 主にシミュレーションやデータ解析の処理で使用されており、変数を変えて大量に流す処理や細切れにして並列に流せる処理に向いています。

- 金融サービス

ハイパフォーマンスコンピューティング	価格付け、市場位置、リスク管理などの分野でハイパフォーマンスコンピューティング集中型のワークロードにおいて、ジョブのリソーシングやスケジューリングを自動化して、コストを削減し、意思決定や市場投入までを高速化できます。
ポストトレード解析	AWS Batchでは、ワークロードのオートメーションが可能なため、翌日の取引サイクルに関連するリスクを把握することや、そのデータに基づいてより理想的な決定を行うことができます
不正調査	Amazon Machine Learning により、よりインテリジェントな方法で、アルゴリズムやモデルを使用してデータを分析し不正の課題に対処することができます。AWS Batch と 併用すると、データ内の、マネーロンダリングや不正な支払いといった不正な活動を示す可能性がある不規則なパターンを検出するために必要な、データ処理や分析を自動化することができます。

- ライフサイエンス

薬剤スクリーニング	AWS Batchは、創薬にかかわる科学的研究員が、より効率的かつ迅速に低分子ライブラリを調査し、一般にタンパク質受容体や酵素といった、目的の薬剤に結合できそうな構造を特定するのに役立ちます。
DNA シーケンス	AWS Batchでは複数の重複した読み取り値と参照配列を比較し、参照とサンプル間の不整合から生じるデータのエラーを潜在的に減らすことで、生のDNA 読み取り値の完全なゲノムシーケンスへの組み立てを、単純化かつ自動化することができます。

- デジタルメディア

レンダリング	AWS Batchは、コンテンツレンダリングワークロードを自動化し、実行時の依存性やリソースのスケジューリングによる人の介入の必要性を減らすことができるツールです。
トランスコーディング	AWS Batchでは、非同期プロセスの実行のスケジューリングおよびモニタリングを行い、条件に応じた応答をトリガーとして必要な場合に指定のワークロードのリソースを スケールすることにより、ワークフローの自動化、リソースのボトルネックの克服、および手動での処理数を削減し、バッチおよびファイルベースのトランスコーディングワークロードを加速させます。
メディアサプライチェーン	AWS Batchにより、プロセスの各段階における異なるジョブや依存関係にあるジョブの実行を調整することにより複雑なメディアサプライチェーンワークフローを簡略化し、メディアサプライチェーンの異なる一員へのコンテンツ準備の管理を行うための共通フレームワークをサポートできます。

- 詳細については、「[AWS Batch のユースケース](#)」をご参照下さい。

■ Lambdaとの相違点

- 比較対象としてLambdaが第一に挙げられますので、今回はLambdaとの相違点を以下にまとめました

	Batch	Lambda
大量のデータ処理	○	×
定期的な実行処理	○	○
複雑なデータ処理	○	×
ある操作をトリガーにした時の実行	○	○
マシンリソースの拡張性	○	○

- 科学技術計算や金融など各種ユースケースでの大規模計算処理のバッチ処理においては、AWS Batchを選択します。
その他のジョブ実行の基盤としてLambdaがありますが、実行時間の制限（15分）がありますので、処理時間などの要件に準じて適切なAWSサービスを使い分けるのが重要です。

■ おわりに

AWS Batchについてご紹介してきましたが、いかがでしたでしょうか。

最後に、今回ご紹介したAWS Batchのまとめです。

- バッチ処理を行うマネージドサービスであり、膨大なバッチコンピューティングをAWSで簡単に効率的に実行可能。
- Batchが自動でインスタンスの起動停止を行うため、スケジューラや計算ノードの管理が不要。
- 大量でかつ複雑なデータ処理に適している。
- 金融サービス、ライフサイエンス、デジタルメディアなど様々な分野で活用可能。

富士通では「AWS Well-Architectedフレームワーク」(※1)を積極的に活用することで、より質の高いシステム基盤導入と運用サービスを提供致します。

(※1) AWSのソリューションアーキテクトが長年の経験やノウハウなどから、5つの柱「運用の優秀性」「セキュリティ」「可用性」「パフォーマンス効率」「コスト最適化」で整理したAWSにおけるベストプラクティス集です。このフレームワークを使うことで、AWSの最適な設計をもとにシステム設計することが可能となります。また、運用についても継続な改善を行うことができ、常に新しく最適なシステム・運用を維持することが出来ます。

最後までお読みいただき、ありがとうございました。

なお、富士通ではAWS BatchやAWS Lambdaなどと連携可能な業務自動運転ソリューションを提供しています。

- [業務自動運転ソリューションについて詳しく見る](#)



記載されている製品名などの固有名詞は、各社の商標または登録商標です。

お問い合わせ先

製品・サービスについてのお問い合わせは[コチラ](#)

富士通株式会社 〒211-8588 神奈川県川崎市中原区上小田中4-1-1

