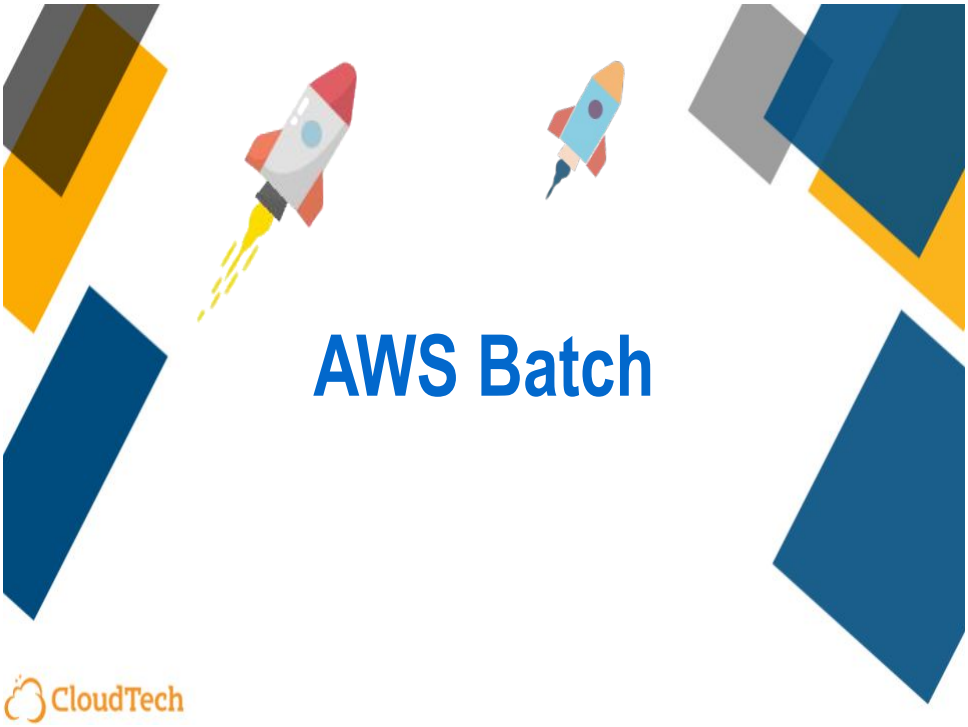
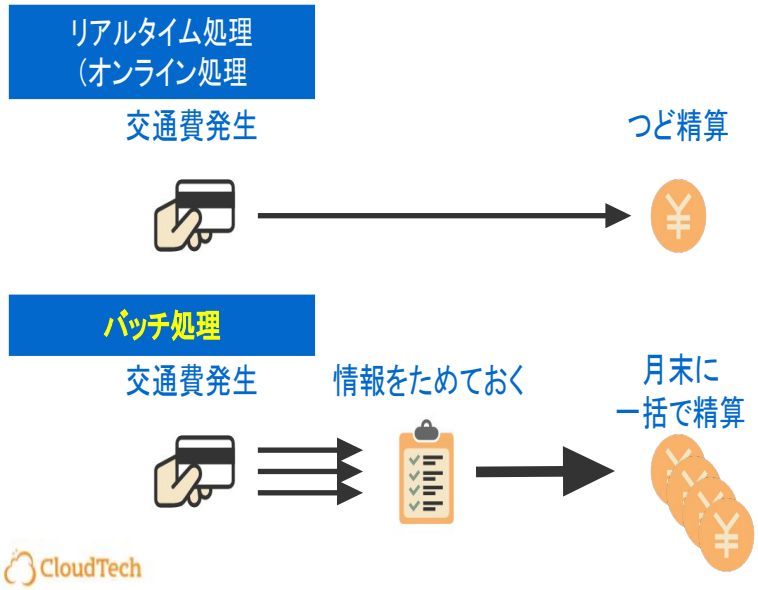


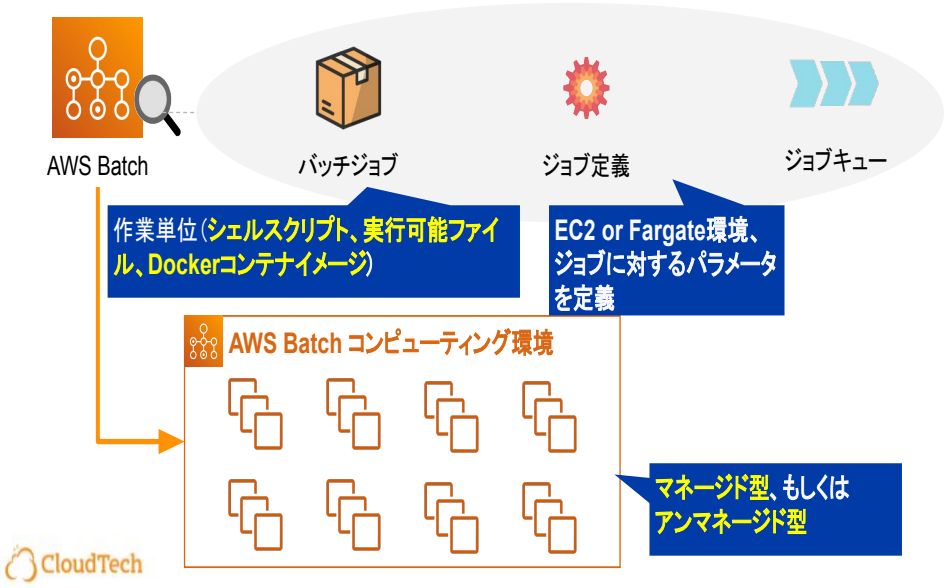
- AWS Batchはバッチ処理のワークロードを AWS上で実行できるサービスです。

バッチ処理とは？



- 前提のおさらいですが、バッチ処理とは何でしょうか？
  - これは、コンピュータでひとまとまりのデータを、手動介入なしに実行することです。
    - 例えば、会社の交通費の精算をすぐさま振り込むリアルタイム処理、に対し
    - 月末に一括で処理されるバッチ処理、このように一連のデータを一括で処理するものだとイメージしてください。

AWS Batch



- さて、AWS Batchがバッチ環境を構築するまでに、いくつかの基本用語がありますので、押さえていきましょう。
  - なお、用語自体が試験で問われることはありませんので、ざっくりとした理解で大丈夫です。
- まずは、バッチジョブです。
  - これは、AWS Batch に送信する作業単位を表します。
  - 例えば、(シェルスクリプト、 実行可能ファイル、Docker コンテナイメージを表します。
- 次に、ジョブ定義
  - ジョブ定義は、バッチジョブの実行方法を定義する、JSON形式のテンプレートのようなものです。
    - バッチジョブはジョブ定義で指定したパラメーターを使用して、後述のコンピューティング環境でひとまとめにされます。
    - バッチジョブの実行方法、例えば、コンピューティング環境を、サーバーレスコンテナ実行基

        - 盤のFargate か、または EC2 で実行するかを指定します。
            - 補足ですが、大規模なワークロードが必要な場合はEC2での実行が推奨されます。
                - そして、特定のプロセッサ、GPU、アーキテクチャの要件などがある場合も EC2 での実行が必要となるでしょう。
            - ジョブ定義は、他にも、バッチジョブに対するパラメータ、環境変数、コンピューティング要件、などの情報を事前に定義します。
- ジョブキューについて、
    - バッチジョブはジョブキューに送信されてジョブキューにとどまります。
    - 複数のジョブキューで優先度を設定でき、この順位に基づいて実行順序を判断します。
        - 加えて、たとえば、優先度の高いジョブには EC2 オンデマンドインスタンスを使用するキューを作成し、
        - 優先度の低いジョブには EC2 スポットインスタンスを使用する、といったEC2の購入方法を分ける方法もあります。
- 最後に、コンピューティング環境
    - AWS Batch でのコンピューティング環境とは、バッチジョ ブが実行されるコンピューティングリソースの集合です。
        - 適切なインスタンスの選定や環境のスケーリングなどをAWSがマネージメントしてくれる マネージド型 と、自分で設定する アンマネージド型 が存在します。

- 簡単でしたが、AWS Batchの解説でした。
    - AWS認定試験では出題頻度はそれほど高くないと思われますが、サービス概要はしっかりと理解していきましょう。