

クラウドサービス利用料算出機能の設計と試作

正村 駿† 小川 康志† 松浦 陽平†

三菱電機株式会社 情報技術総合研究所†

1. はじめに

近年、システムを短時間で構築可能なパブリッククラウドが注目されている。しかし、クラウド利用時は、利用料に関する考慮事項が多いため、利用料の見積もりは難しい。例えば、オンプレミス利用時に固定費となるサーバー費用がクラウド利用時には変動費となる場合や、クラウドサービスによってはオートスケール機能等で動的にリソースが変更される場合等がある。オンプレミス利用時の利用料と比較してクラウド利用時ではどの程度かを把握するため、設計等の早い段階で利用料を見積もる必要がある。しかし、クラウドベンダが提供する利用料算出ツールでは概算費用しか算出できない問題がある。

そこで、クラウドサービスの課金方式をモデル化し、システムの負荷に応じたクラウドサービス利用料を算出可能とする機能について、設計と一部サービスに対する試作を実施した。

2. 課題

クラウドベンダの AWS や Microsoft は、利用料見積もりツールを提供している。しかし、これらの利用料算出ツールでは、ワークロードピーク時の期間や、ワークロードピーク時に必要なインスタンスの数といった大まかな入力しかできないため、複雑で不規則なパターンのワークロードに応じた利用料を算出できない。また利用サービスによっては、オートスケール機能でワークロードに応じて動的にスループットを調整可能であり、一般に、オートスケール機能利用時にはスループットが利用料に影響を与える。よって、オートスケール機能を利用したクラウドサービスの利用料を見積もるためには、時間的に変化するワークロードの情報をもとに各時刻のスループットを算出する必要があり、クラウドベンダが提供する利用料算出ツールでは不十分である。

AWS が提供する Amazon DynamoDB(以下、DynamoDB)を例に挙げる。DynamoDB では事前に使用キャパシティを指定するプロビジョニングモードがある。プロビジョニングモードはオートスケールに対応しており、リクエスト数に応じて、キャパシティユニットを増減できる。キャパシティユニットは、1秒あたりに消費可能なユニット数を示し、リクエストの種別(読み込み/書き込み)とリクエストで扱うデータ量に応じて消費される。AWS の利用

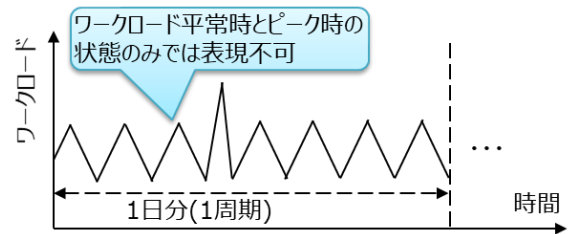


図1 時間的に変化するワークロード

料算出ツールである Pricing Calculator[1]では、DynamoDB のプロビジョニングモードのワークロードの情報として、ワークロード平常時とピーク時の2つの状態を対象とし、それぞれにおける期間とワークロードの値のみを入力し、1か月あたり(使用時間を730時間で計算)のキャパシティユニットの利用料を見積もることができる。しかし、図1のように、ワークロード平常時とピーク時の情報のみで表現できないワークロードを想定する場合は、Pricing Calculator で対応できない。一方、プロビジョニングモードでは、1時間ごとの最大キャパシティユニットをもとに課金される。そのため、オートスケール機能利用時に、図1のようなワークロードを想定して利用料を算出する場合、ワークロードをもとに時刻ごとのキャパシティユニットを予測する必要がある。

また DynamoDB では、プロビジョニング済みのユニットのうち利用されなかった分を過去300秒までリザーブするバーストキャパシティ機能が存在するが、これは Pricing Calculator では未考慮である。

よって、既存ツールの Pricing Calculator では、DynamoDB のオートスケール機能利用時での時間で変化するワークロードやバーストキャパシティを考慮した利用料算出に対応しておらず、これらに対応可能な利用料算出ツールを開発する必要がある。本稿では、DynamoDB を具体的な題材として、利用料算出のためのモデルを検討した。

3. クラウドサービス利用料算出機能の設計

クラウドサービスにおいて、利用料に影響を与えるリソース使用量は「総量」と「単位時間あたりの最大使用量」に分類される。例えば「総量」は処理件数の和で利用料が決定される。本稿では、DynamoDB のキャパシティユニットに関する「単位時間あたりの最大使用量」を対象とする。

外部リクエストに対して、サービス毎の利用料に影響のあるリソース使用量の割合を性能プロファイルとして定義する。事前に利用サービスに負荷を与え、そのサービスの課金体系にもとづいて性能プロファイルを取得しておく。

DynamoDB のオートスケール機能で利用するキャパシティユニットの課金については、プロビジョニング済みのユニットに対して行われる。しかし、各時刻のプロビジョニング済みのユニットは、実際に消費されたユニットとスケール条件をもとに決まるため、性能プロファイルには、間接的に課金に影響を与える、消費されたユニットの情報を含めるようにする。この性能プロファイルと想定ワークロードパターン、スケール条件等を含む DynamoDB のサービスオプションを利用することで、時刻ごとのプロビジョニング済みのユニットを算出する。プロビジョニング済みのユニットの時間分布をもとに、利用料を算出可能である。

以上を踏まえて検討した、クラウドサービス利用料算出機能の構成を図 2 に示す。図 2 ではパブリッククラウドとして AWS を利用することを前提としている。利用料算出機能のポイントを以下に示す。

(1) 性能プロファイルの取得

CloudWatch メトリクスや、AWS が公開する API を利用して性能プロファイルのパラメータに対応する値を取得する。DynamoDB のオートスケール機能利用時のキャパシティユニットを対象とする場合は、負荷掛け時、1 秒ごとに API を利用して消費されたキャパシティユニットを取得し、その中の最大値を、負荷掛けの間に最大とした 1 秒あたりのリクエスト数で除算した値をパラメータとする。

(2) ワークロードパターンの定義

表 1 に示すフォーマットの CSV ファイルを用いて、時間的に変化するワークロードを表現する。表 1 の 2 行目以降に記載のワークロードのデータについて、ある時刻に対応するデータが欠損する場合は、その時刻に最も近い前後の時刻に応じたデータを用いて、ワークロードの値を直線で 2 つの間を補完し、欠損時のワークロードの値を算出可能とする。

(3) リソース使用量の算出

性能プロファイルとワークロードパターン、サービスオプションをもとに、月あたりのリソース使用量を算出する。DynamoDB のオートスケール機能利用時では、毎秒、バーストキャパシティを考慮して消費されたユニットを算出し、それをもとにプロビジョニング済みのユニットをスケールする。

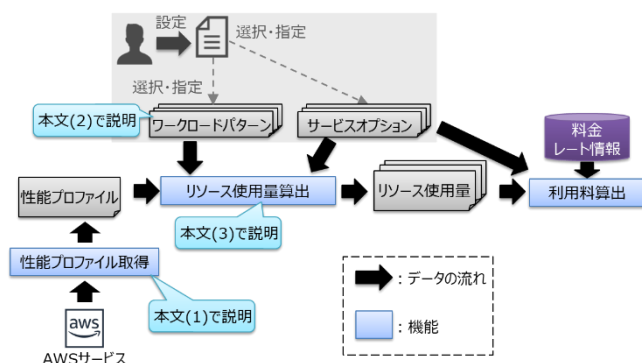


図 2 クラウドサービス利用料算出機能の構成

表 1 ワークロードパターンのフォーマット

1 行目	ワークロードの周期 (0:時間周期, 1:日周期)
2 行目以降	各時刻のワークロードのデータ 1 列目:時刻(DDhhmmss 形式) 2 列目:ワークロードの値

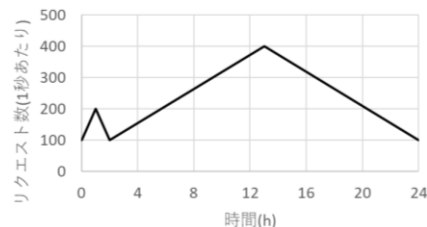


図 3 入力するワークロードパターン

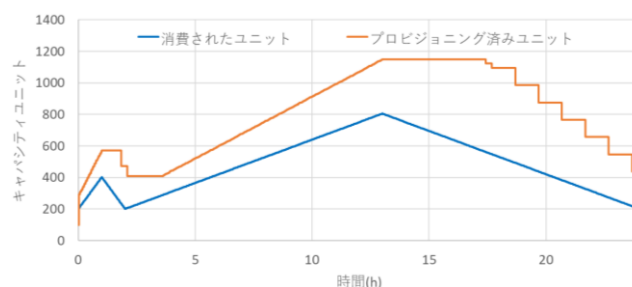


図 4 キャパシティユニットの推移

1 時間ごとのプロビジョニング済みユニットの最大値を月あたりで合計した値がリソース使用量となる。

4. クラウドサービス利用料算出機能の評価

DynamoDB のオートスケール機能を題材に、クラウドサービス利用料算出機能を実装した。性能プロファイルのキャパシティユニットのパラメータの値を 2.01, 最初のプロビジョニング済みユニットを 100, 図 3 に示す日周期のワークロードパターンを入力したときのキャパシティユニットの推移結果を図 4 に示す。図 4 より、消費されたユニットはプロビジョニング済みユニットに追従して推移しており、オートスケールが働いていることがわかる。また開始 0 時付近では、プロビジョニング済みユニットより消費されたユニットのほうが多いが、バーストキャパシティによって賄っている。図 4 において 1 時間ごとのプロビジョニング済みのユニットの最大値の合計を一か月換算で 30 倍した値をリソース使用量とする。これとオプション、料金レート情報をもとに、利用料を算出できることを確認した。

5. おわりに

本稿では、時間的に変化するワークロードに対応可能なクラウドサービス利用料算出機能の設計を行った。また DynamoDB のオートスケール機能を題材としてツール実装を行い、実現性を検証した。今後は、DynamoDB 以外にもより多くのサービスに対応する利用料算出機能を実現する。

参考文献

[1]Amazon Web Services, Inc., “AWS Pricing Calculator”, <<https://calculator.aws/>>, (参照 2020-12-17)