
コスト最適化の柱

AWS Well-Architected Framework

コスト最適化の柱: AWS Well-Architected Framework

Copyright © 2023 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon's trademarks and trade dress may not be used in connection with any product or service that is not Amazon's, in any manner that is likely to cause confusion among customers, or in any manner that disparages or discredits Amazon. All other trademarks not owned by Amazon are the property of their respective owners, who may or may not be affiliated with, connected to, or sponsored by Amazon.

Table of Contents

要約とはじめに	1
要約	1
はじめに	1
コスト最適化	3
設計原則	3
定義	3
クラウド財務管理を実践する	5
COST01-BP01 コスト最適化の所有権を設定する	6
実装のガイダンス	7
リソース	8
COST01-BP02 財務とテクノロジーの連携を確立する	8
実装のガイダンス	7
リソース	8
COST01-BP03 クラウドの予算と予測を確立する	12
実装のガイダンス	7
リソース	8
COST01-BP04 組織のプロセスにコスト意識を採り入れる	14
実装のガイダンス	7
リソース	8
COST01-BP05 コスト最適化に関して報告および通知する	15
実装のガイダンス	7
リソース	8
COST01-BP06 コストをプロアクティブにモニタリングする	17
実装のガイダンス	7
リソース	8
COST01-BP07 新しいサービスリリースに関する最新情報を把握しておく	18
実装のガイダンス	7
リソース	8
COST01-BP08 コスト意識を持つ文化を生み出す	19
実装のガイダンス	7
リソース	8
COST01-BP09 コスト最適化によるビジネス価値を数値化する	20
実装のガイダンス	7
リソース	8
経費支出と使用量の認識	23
ガバナンス	23
COST02-BP01 組織の要件に基づいてポリシーを策定する	23
COST02-BP02 目標およびターゲットを策定する	26
COST02-BP03 アカウント構造を実装する	28
COST02-BP04 グループとロールを実装する	31
COST02-BP05 コストコントロールを実装する	32
COST02-BP06 プロジェクトのライフサイクルを追跡する	34
コストと使用量のモニタリング	35
COST03-BP01 詳細情報ソースを設定する	36
COST03-BP02 コストと使用状況に組織情報を追加する	37
COST03-BP03 コスト属性カテゴリを特定する	38
COST03-BP04 組織のメトリクスを確立する	41
COST03-BP05 請求およびコスト管理ツールを設定する	42
COST03-BP06 ワークロードメトリクスに基づいてコストを配分する	44
リソースを削除する	45
COST04-BP01 ライフタイム全体にわたってリソースを追跡する	45
COST04-BP02 削除プロセスを実装する	46
COST04-BP03 リソースを削除する	48
COST04-BP04 自動的にリソースを削除する	49

COST04-BP05 データ保持ポリシーを適用する	50
コスト効率を考慮しながらリソースを利用する	52
サービスを選択する際にコストを評価する	52
COST05-BP01 組織のコスト要件を特定する	52
COST05-BP02 ワークロードのすべてのコンポーネントを分析する	53
COST05-BP03 各コンポーネントの詳細な分析を実行する	54
COST05-BP04 コスト効率の高いライセンスを提供するソフトウェアを選択する	56
COST05-BP05 組織の優先順位に従ってコストが最適化されるようにこのワークロードのコン ポーネントを選択する	57
COST05-BP06 異なる使用量について経時的なコスト分析を実行する	59
正しいリソースタイプ、リソースサイズ、リソース数を選択する	60
COST06-BP01 コストモデリングを実行する	60
COST06-BP02 データに基づいてリソースタイプ、リソースサイズ、リソース数を選択する	62
COST06-BP03 メトリクスに基づいて自動的にリソースタイプ、リソースサイズ、リソース数を 選択する	63
最適な料金モデルを選択する	65
COST07-BP01 料金モデルの分析を実行する	65
COST07-BP02 コストに基づいてリージョンを選択する	67
COST07-BP03 費用対効果の高い条件を提供するサードパーティーの契約を選択する	69
COST07-BP04 このワークロードのすべてのコンポーネントに対して料金モデルを実装する	70
COST07-BP05 管理アカウントレベルで料金モデル分析を実行する	72
データ転送を計画する	73
COST08-BP01 データ転送モデリングを実行する	74
COST08-BP02 データ転送コストを最適化するコンポーネントを選択する	75
COST08-BP03 データ転送コストを削減するサービスを実装する	77
需要を管理しリソースを供給する	79
COST09-BP01 ワークロードの需要に関する分析を実行する	79
実装のガイダンス	7
リソース	8
COST09-BP01 需要を管理するためのバッファまたはスロットルを実装する	81
実装のガイダンス	7
リソース	8
COST09-BP03 リソースを動的に供給する	83
実装のガイダンス	7
実装手順	7
リソース	8
継続的最適化	89
レビュープロセスを定義し、ワークロードを定期的に分析	89
COST10-BP01 ワークロードレビュープロセスを開発する	89
COST10-BP02 このワークロードを定期的に見直し、分析する	91
オペレーションの自動化	92
COST11-BP01 オペレーションのオートメーションを実行する	92
まとめ	95
寄稿者	96
その他の資料	97
改訂履歴	98

コスト最適化の柱 - AWS Well-Architected フレームワーク

公開日: 2023 年 12 月 6 日 ([改訂履歴 \(p. 98\)](#))

要約

このホワイトペーパーでは、アマゾン ウェブ サービス (AWS) Well-Architected フレームワークのコスト最適化の柱に焦点を当てます。本書は、お客様が AWS 環境の設計、提供、メンテナンスを行う際にベストプラクティスを適用できるようにするためのガイダンスを提供します。

コストが最適化されたワークロードでは、すべてのリソースがフル活用され、コストの最小化を達成して、機能要件が満たされます。このホワイトペーパーは、組織内での機能の構築、ワークロードの設計、サービスの選択、サービスの構成と運用、およびコスト最適化手法の適用に関する詳細なガイダンスを提供します。

はじめに

[AWS Well-Architected フレームワーク](#)は、AWS でワークロードを構築する際の決定事項を理解するのに役立ちます。このフレームワークは、信頼性が高く、安全かつ効率的で、コスト効率に優れ、持続可能なワークロードをクラウド内で設計および運用するためのアーキテクチャ設計のベストプラクティスを提供します。アーキテクチャをベストプラクティスに照らして評価し、改善すべき領域を特定する一貫した方法を紹介します。Well-Architected ワークロードを備えることによって、ビジネス成功の可能性が大幅に高まると私たちは確信しています。

このフレームワークは次の 6 つの柱に基づいています。

- オペレーショナルエクセレンス
- セキュリティ
- 信頼性
- パフォーマンス効率
- コスト最適化
- サステナビリティ

本ホワイトペーパーでは、コスト最適化の柱と、サービスやリソースを最も効果的に活用したワークロードの設計方法、最小限のコストでビジネス成果を達成する方法に焦点を当てています。

コスト最適化の柱のベストプラクティスを組織に適用する方法について学習します。従来のオンプレミスソリューションでは、コストを最適化するためのハードルにぶつかる可能性があります。これは将来のキャパシティやビジネスニーズを予測すると同時に、複雑な調達プロセスを進める必要があるためです。このホワイトペーパーのプラクティスを適用すると、皆さんの組織が以下の目標達成を支援します。

- クラウド財務管理を実践する
- 経費支出と使用量の認識
- コスト効率を考慮しながらリソースを利用する
- 需要を管理しリソースを供給する

- 継続的最適化

本ホワイトペーパーは、最高技術責任者 (CTO)、最高財務責任者 (CFO)、設計者、デベロッパー、財務管理者、ファイナンシャルプランナー、ビジネスアナリスト、オペレーションチームメンバーなどの技術担当者および財務担当者を対象としています。本ホワイトペーパーでは、運用の詳細やアーキテクチャのパターンについては説明しておりませんが、該当するリソースへの参照先が記載されています。

コスト最適化

コスト最適化とは、システムのワークロードのライフサイクル全体にわたって改良、改善する継続的プロセスです。本ホワイトペーパーのプラクティスは、組織がコストを最小限に抑えて投資利益率を最大化すると同時に、ビジネス成果の達成につながるようなコストを意識したワークロードの構築および運用を支援します。

トピック

- [設計原則 \(p. 3\)](#)
- [定義 \(p. 3\)](#)

設計原則

コスト最適化のため、以下の設計原則を検討する：

クラウド財務管理を実装する：クラウドで財務面での成功を実現し、ビジネス価値の実現を加速するには、クラウドの財務管理に投資する必要があります。組織は、テクノロジーと使用量管理の新たなドメインで機能を構築するために必要な時間とリソースを投入する必要があります。コスト効率の高い組織になるためには、セキュリティまたはオペレーションの能力と同様、知識の積み上げ、プログラム、リソース、プロセスを通じて能力を構築する必要があります。

消費モデルを導入する：コンピューティングリソースの使用分のみを支払い、ビジネス要件に応じて使用量を増減することができます。たとえば、通常、1 週間の稼働日に開発環境とテスト環境を使用するのは、1 日あたり 8 時間程度です。未使用時にこのようなリソースを停止することで、コストを 75% 削減できる可能性があります (168 時間から 40 時間に減少)。

全体的な効率を測定する：ワークロードのビジネス成果と、提供にかかったコストを測定します。このデータを利用すると、生産性および機能性の向上とコスト削減から得られるメリットを理解することができます。

差別化につながらない高負荷の作業に費用をかけるのをやめる：サーバーラックの準備や機器の設置、電源確保などの面倒なデータセンター運用作業は AWS が行います。また、マネージドサービスを使用することで、オペレーティングシステムやアプリケーションの管理に伴う運用上の負担も解消されます。この結果、IT インフラストラクチャよりも顧客と組織のプロジェクトに集中できるようになります。

費用を分析し帰属関係を明らかにする：クラウドを使用すると、ワークロードの使用量とコストを正確に特定しやすくなり、IT コストと収益ストリームや、個別のワークロードの所有者との帰属関係が明瞭になります。これによって投資収益率 (ROI) を把握できるため、ワークロードの所有者はリソースを最適化してコストを削減する機会が得られます。

定義

クラウドでのコスト最適化には、5 つの重点分野があります。

- クラウド財務管理を実践する
- 経費支出と使用量の認識
- 費用対効果の高いリソース
- 需要の管理とリソースの提供
- 継続的最適化

この柱にも、Well-Architected フレームワーク内の他の柱と同様に、コスト最適化のために検討すべきトレードオフがあります。たとえば、市場投入に要する期間と、コストのどちらを優先すべきでしょうか。市場投入に要する期間の短縮、新機能導入、納期順守といったケースでは、前払いコストの投資を最適化するよりも、スピードを重視して最適化することが最善ことがあります。

設計上の決定は、データではなく時間的制約によって下されることがあります。また、#期的にコストを最適化できるワークロードのベンチマークの選定に時間を掛けるよりも、「万#の場合」の備えを過度に重視してしまう傾向が常にあります。過大な見積もりの結果、オーバプロビジョニングになり、最適化が不十分なデプロイを行ってしまいます。ただし、オンプレミス環境からクラウド環境に「リフトアンドシフト」式にリソースを移行してから最適化する必要がある場合は、選択肢としては妥当なこともあります。

適切な労力を当初からコスト最適化戦略に投入すると、ベストプラクティスが一貫して適用され、不要なオーバプロビジョニングも回避できるため、クラウドの経済的メリットをより早く実感できます。以下のセクションでは、クラウドの財務管理の初期および継続的な実装と、ワークロードのコスト最適化のための手法とベストプラクティスを提供します。

クラウド財務管理を実践する

クラウド財務を管理するには、コストの透明性、制御、計画、AWS 環境の最適化を確立して運用するために、既存の財務プロセスを進化させる必要があります。

従来の静的なウォーターフォール計画と IT 予算およびコスト評価モデルを動的なクラウドの使用に適用すると、リスクが生じ、計画が不正確になり、可視性が低くなります。最終的には、コストを効果的に最適化、管理し、長期的なビジネス価値を実現する機会を失うことにつながります。これらの落とし穴を回避するために、クラウドでネイティブにアプリケーションを構築する場合でも、ワークロードをクラウドに移行する場合でも、クラウドサービスの採用を拡大する場合でも、クラウドジャーニー全体でコストを積極的に管理する必要があります。

クラウド財務管理 (CFM) は、財務、製品、技術、ビジネス組織が、AWS での利用拡大やスケールの拡張に伴い、コストの管理、最適化、計画を実施できるようにします。CFM の主な目的は、お客様が最もコスト効率の高い方法でビジネス成果を達成し、俊敏性とコントロールの適切なバランスを見つけながら、経済的価値とビジネス価値の創造を加速できるようにすることです。

CFM ソリューションは、コスト透明性、コントロール、予測、最適化により、ビジネスを変革するのに役立ちます。これらのソリューションは、すべてのチームと機能にわたる説明責任を担うためのコスト意識の高い文化を実現することもできます。財務チームは、コストの発生源を確認し、予期しない費用を最小限に抑えながら運用を実施し、動的なクラウドの使用を計画して、チームがクラウドの導入を拡大しながらクラウド費用を節約することができます。これをエンジニアリングチームと共有することで、リソースの選択、使用、最適化に必要な財務状況を提供できます。

AWS CFM は、ビジネスの俊敏性を維持しながら、クラウドコストを管理、最適化、および計画するための一連の機能を提供します。CFM は、コストを効果的に管理するだけでなく、投資を期待できるビジネス成果を促進しているかを検証するうえで最も重要です。これらは、AWS クラウドでのクラウド財務管理フレームワークの 4 つの柱である観察、節約、計画、実行です。これらの柱にはそれぞれ、アクティビティと機能が設定されています。

クラウド財務管理の 4 つの柱。

- **観察:** 現在、クラウドの支出をどのように測定、監視し、説明責任を果たしていますか? AWS を初めて利用する場合や AWS を利用しようと計画している場合、コストと使用状況の可視性を確立する計画がありますか?

AWS のコストを理解し、支出を最適化するには、それらのコストがどこから来ているのかを知る必要があります。そのためには、アカウントやリソースを計画的に管理し、財務部門が支出の流れを把握し、各チームが利益に対する責任を果たせるようにすることが必要です。

AWS のサービス: AWS Control Tower、AWS Organizations、コスト配分タグ、タグポリシー、AWS Resource Groups、AWS Cost Categories、AWS Cost Explorer、AWS Cost and Usage Report、RI および SP

リソース: AWS タグ付けに関するベストプラクティス、AWS Cost Categories

- **節約:** 現在、どのようなコスト最適化の手段を用いていますか? AWS を使用していない場合、一般的な使用ベースおよび価格モデルベースの最適化についてご存知でしょうか?

節約の基本方針では、価格やリソースの推奨によりコストを最適化します。コストの最適化は、新しいクラウド運用モデルに対して明確に定義された戦略を持つことから始まります。クラウドジャーニーのできるだけ早い段階から始め、適切なプロセスと動作によって強化されたコスト意識の高い文化の準備を整えるのが理想的です。

クラウドのコストを最適化するには、さまざまな方法があります。その 1 つは、適切な購入モデル (RI および SP) の選択、または Amazon EC2 スポットインスタンスを採用できるようにワークロードが不

変かつコンテナ化されているかどうかです。さらに、Amazon EC2 Auto Scaling グループを使用してワークロードをスケールします。

AWS のサービス: RI および SP、Amazon EC2 Auto Scaling グループ、スポットインスタンス

リソース: リザーブドインスタンス、Savings Plans、Amazon EC2 を処理するためのベストプラクティス

- 計画する: 現在、将来のクラウドの利用や支出についてどのように計画していますか? 新たな移行に伴う価値創出を定量化する方法論をお持ちですか? クラウドの可変的な利用を採用するために、現在の予算編成と予測のプロセスを発展させましたか?

計画の基本方針を定めることは、柔軟な予算編成と予測によって計画を改善することを意味します。可視化とコスト管理を確立した後は、クラウドプロジェクトの支出を計画し、期待値を設定したいと思うことでしょう。AWS は、動的な予測と予算編成のプロセスを柔軟に構築できるため、コストが予算の上限を守っているか、あるいは超えているかを常に把握することができます。

AWS のサービス: AWS Cost Explorer、AWS Cost and Usage Report、AWS Budgets

リソース: 使用量ベースの予測、AWS 予算レポート、アラート

- 実行: クラウドの支出を管理するために現在使用している運用プロセスやツールにはどのようなものがあり、誰がその取り組みを主導していますか? AWS を実際に使い始めてから、日々の運用面で工夫されたことはありますか?

実行の基本方針とは、請求とコスト管理を実際に行うことです。ガードレールを設定し、ガバナンスを確立することで、経費が予算内に収まるようにすることができます。AWS には、開始するにあたり役に立ついくつかのツールが用意されています。

AWS のサービス: AWS Billing and Cost Management コンソール、AWS Identity and Access Management、サービスコントロールポリシー (SCP)、AWS Service Catalog、AWS Cost Anomaly Detection、AWS Budgets

リソース: AWS Billing コンソールの使用を開始する

以下は、クラウドの財務管理におけるベストプラクティスです。

ベストプラクティス

- [COST01-BP01 コスト最適化の所有権を設定する \(p. 6\)](#)
- [COST01-BP02 財務とテクノロジーの連携を確立する \(p. 8\)](#)
- [COST01-BP03 クラウドの予算と予測を確立する \(p. 12\)](#)
- [COST01-BP04 組織のプロセスにコスト意識を採り入れる \(p. 14\)](#)
- [COST01-BP05 コスト最適化に関して報告および通知する \(p. 15\)](#)
- [COST01-BP06 コストをプロアクティブにモニタリングする \(p. 17\)](#)
- [COST01-BP07 新しいサービスリリースに関する最新情報を把握しておく \(p. 18\)](#)
- [COST01-BP08 コスト意識を持つ文化を生み出す \(p. 19\)](#)
- [COST01-BP09 コスト最適化によるビジネス価値を数値化する \(p. 20\)](#)

COST01-BP01 コスト最適化の所有権を設定する

組織全体のコスト認識を確立し、維持する責任を持つチーム (クラウドビジネスオフィス、Cloud Center of Excellence (CCoE) または FinOps チーム) を作成します。コスト最適化の所有者には、組織全体およびクラウド財務を理解している個人またはチーム (財務、テクノロジー、およびビジネスチームの人材が必要) を指定することができます。

このベストプラクティスが確立されていない場合のリスクレベル: 高

実装のガイダンス

これは、クラウドコンピューティングにおけるコスト意識の文化を確立して維持する責任を負う、クラウドビジネスオフィス (CBO) または Cloud Center of Excellence (CCoE) チーム確立についての概要です。これは社内の個人でも、チームでもかまいません。組織全体から財務、テクノロジーなどの主要な関係者を集めてチームを新規編成することもできます。

担当者 (個人またはチーム) は、コスト管理とコスト最適化活動に必要な時間を優先順序を付けて配分します。小規模な組織の場合、大企業のフルタイムの担当と比較すると、費やす時間の割合は少ない場合があります。

担当者 (個人またはチーム) は、コスト管理とコスト最適化活動に必要な時間を優先順序を付けて配分します。小規模な組織の場合、大企業のフルタイムの担当と比較すると、費やす時間の割合は少ない場合があります。

プロジェクトマネジメント、データサイエンス、財務分析、ソフトウェアやインフラストラクチャ開発など、複合的な能力が求められます。担当者は次の 3 つの異なる所有権内でコスト最適化を実行することにより、ワークロードの効率を高めることができます。

- 一元化: FinOps チーム、クラウド財務管理 (CFM) チーム、クラウドビジネスオフィス (CBO)、Cloud Center of Excellence (CCoE) などの指定チームを通じて、お客様はガバナンスの仕組みを設計、導入し、ベストプラクティスを全体的に推進することができます。
- 分散型: テクノロジーチームに影響を与え、コスト最適化を実行します。
- ハイブリッド: 一元化されたチームと分散されたチームの両方が協力して、コストの最適化を実行することができます。

この担当者は、コスト最適化目標 (ワークロード効率メトリクスなど) に対する実行および提供能力を評価されることになります。

この担当者のためにエグゼクティブスポンサーシップを確保しなければならず、これが重要な成功要因です。エグゼクティブスポンサーは、クラウド利用のコスト効率を判断する最高責任者として、チームの考え方を上長にエスカレーションして、組織が定める優先事項としてコスト最適化活動が扱われるようサポートします。そうでなければ、ガイダンスは無視され、コスト削減の機会が優先されなくなります。エグゼクティブスポンサーとチームは協力して、組織のクラウド利用を効率化し、ビジネスバリューを実現できるようにします。

ビジネス、Enterprise-On-Ramp、またはエンタープライズ [サポートプラン](#) を利用していて、このチームまたは担当者の構築に支援が必要な場合は、アカウントチームを通じてクラウド財務管理 (CFM) のエキスパートにご連絡ください。

実装手順

- 主要なメンバーを定義する: コスト管理には、組織内のすべての関係部署が貢献し、関心を持つ必要があります。一般的な組織内チームには、通常、財務、アプリケーションまたはプロダクトの所有者、管理、技術チーム (DevOps) が含まれています。一部は専属 (財務、技術) で、その他は必要に応じて定期的に関与します。CFM を実行する個人またはチームには、以下のスキルセットが必要です。
 - ソフトウェア開発: スクリプトとオートメーションが構築される場合。
 - インフラストラクチャエンジニアリング: スクリプトをデプロイし、プロセスを自動化して、サービスまたはリソースのプロビジョニング方法を理解するためです。
 - 運用の洞察: CFM とは、クラウドの効率的な利用を測定、モニタリング、修正、計画、スケールすることで、クラウドで効率的に運用することです。
- 目標とメトリクスを定義する: この担当者は、さまざまな方法で組織に価値をもたらす必要があります。これらの目標は定義され、組織が進化するにつれて継続的に進化します。一般的な活動には、組織全体

のコスト最適化に関する教育プログラムの作成と実行、コスト最適化のためのモニタリングやレポート作成などの組織全体の標準策定、最適化に関するワークロード目標の設定などがあります。この担当者は、組織のコスト最適化機能について定期的に組織に報告する必要もあります。

価値ベースまたはコストベースの重要業績指標 (KPI) を定義できます。KPI を定義すると、効率性と予想されるビジネス成果の観点から、予想されるコストを計算できます。価値ベースの KPI は、コストおよび使用量のメトリクスをビジネスバリュー要因に結び付け、AWS の費用の変化を合理化するうえで役立ちます。価値ベースの KPI を導き出す最初のステップは、組織横断的に協力し、KPI の標準セットを選択し、合意することです。

- 定期的なミーティングを設定する: グループ (財務、技術、およびビジネスチーム) は定期的にミーティングを開いて、目標とメトリクスをレビューする必要があります。一般的なミーティングでは、組織の状態の確認、現在実行中のプログラムの確認、全体的な財務および最適化メトリクスの確認を行います。その後、主要なワークロードの詳細が報告されます。

このような定期的なミーティングで、ワークロードの効率性 (コスト) とビジネス成果をレビューできます。例えば、ワークロードのコストが 20% 増加したが、顧客使用量も増加したかもしれません。この場合、この 20% のコスト増加を投資と解釈できます。このような定期的なミーティングにより、チームは組織全体にとって有意義な価値ベースの KPI を特定できます。

リソース

関連するドキュメント:

- [AWS CCOE ブログ](#)
- [クラウドビジネスオフィスの作成](#)
- [CCOE - Cloud Center of Excellence](#)

関連動画:

- [Vanguard CCOE 成功事例](#)

関連する例:

- [Cloud Center of Excellence \(CCoE\) を活用した企業全体の変革](#)
- [企業全体を変革する CCOE の構築](#)
- [CCOE を構築するときに回避すべき 7 つの落とし穴](#)

COST01-BP02 財務とテクノロジーの連携を確立する

クラウドジャーニーのすべての段階で、コストと使用状況に関するディスカッションに財務チームとテクノロジーチームを参加させます。チームは、定期的に集まり、組織の最終目的や目標、コストと使用状況の現状、財務や会計のプラクティスなどのトピックについて話し合います。

このベストプラクティスを活用しない場合のリスクレベル: 高

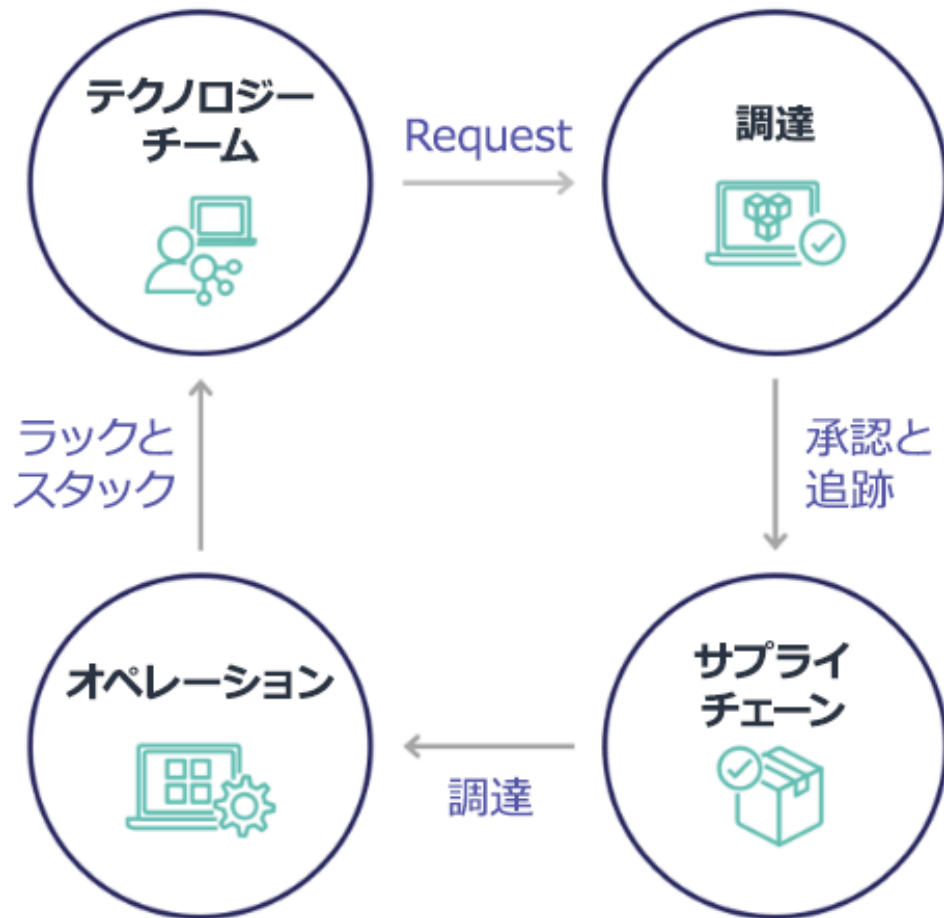
実装のガイダンス

承認、調達、インフラストラクチャのデプロイサイクルが短縮されるため、テクノロジーチームは、クラウドでのイノベーションを迅速に行うことができます。ファイナンス組織はこれまでプロジェクト承認時

のデータセンターやオンプレミス環境の調達に大量に使用されていた時間とリソースを調整することができます。

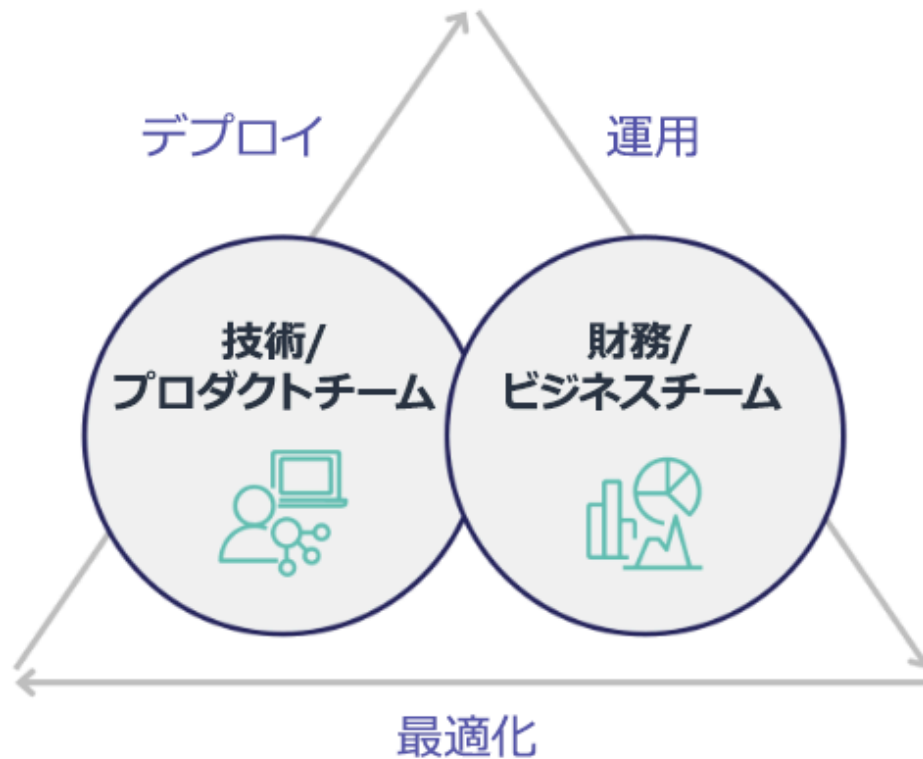
財務および調達組織の観点から見ると、資本予算、資本要求、承認、調達、物理的インフラストラクチャの設置のプロセスは、何十年にもわたって学習され、標準化されてきたプロセスの一つです。

- エンジニアリングチームや IT チームは、通常、依頼主である
- さまざまな財務チームが承認者、調達者として機能する
- 運用チームは、すぐに使えるインフラストラクチャのラック、スタック、および引き渡しを行う



クラウドの採用により、インフラストラクチャの調達と消費は依存関係の連鎖と切り離されます。クラウドモデルでは、テクノロジーチームと製品チームは単なる構築者ではなく、製品の運用者兼所有者となり、調達とデプロイなど、従来は財務および運用チームに割り当てられていたアクティビティのほとんどを担当します。

クラウドリソースのプロビジョニングに必要なものは、ユーザーアカウントと適切なアクセス許可のセットだけです。これは IT および財務リスクを軽減することにもなります。つまり、チームは常に数回のクリックまたは API コールで、アイドル状態または不要なクラウドリソースを停止することができるのです。また、テクノロジーチームがより速くイノベーションを起こすことができるのも、実験を立ち上げては破棄する俊敏性と能力があってこそです。クラウド消費には変動的な性質があるため、資本支出予算と予測の観点から予測可能性に影響することがある一方で、クラウドによって組織は、オーバープロビジョニングのコストを削減し、控えめなアンダープロビジョニングに伴う機会コストも削減できます。



主要な財務部門とテクノロジー部門のステークホルダー同士のパートナーシップを確立し、組織の目標の共通理解を深め、クラウドコンピューティングのさまざまな消費モデルにおいて財務上の成功を実現するためのメカニズムを開発します。クラウドジャーニーのすべてのステージにおいて、コストと使用量に関するディスカッションに参加する必要がある組織内の関連するチームは、以下の。

- ファイナンシャルリード: CFO、財務管理者、フィナンシャルプランナー、ビジネスアナリスト、調達、ソーシング、支払担当は、クラウドの消費モデル、購入オプション、月次請求プロセスを理解する必要があります。財務部門は、テクノロジーチームと連携して、IT バリューの事例を作成およびソーシカル化し、テクノロジー支出がビジネスの成果にどのように関連しているかをビジネスチームが理解できるようにする必要があります。このように、技術支出はコストではなく投資と見なされます。クラウド運用にはオンプレミスのオペレーションと比べて根本的な違い (使用量の変動率、従量課金制やティア別課金制、料金モデル、請求明細と使用量情報など) があるため、クラウド利用が調達プロセス、インセンティブ追跡、コスト配分、財務諸表などのビジネス局面に与えるインパクトをファイナンス部門で理解することが不可欠です。
- テクノロジーリード: テクノロジーリード (製品およびアプリケーションの所有者を含む) は、財務要件 (予算の制約など) やビジネス要件 (サービスレベルアグリーメントなど) を認識する必要があります。これにより、組織が目指すビジネス目標を達成するワークロードの導入が可能になります。

財務とテクノロジーのパートナーシップには、以下のような利点があります。

- 財務チームとテクノロジーチームは、コストと使用量をほぼリアルタイムで把握できます。
- 財務チームとテクノロジーチームは、クラウドへの支出の変動に対応するための標準となる運用手順を確立します。
- 財務部門のステークホルダーは、コミットメント割引 (リザーブドインスタンスや AWS Savings Plans など) の購入に資金がどう使用されるか、また組織を拡大するためにクラウドがどのように利用されるかに関して、戦略アドバイザーとして行動します。

- 既存の支払いアカウントと調達プロセスは、クラウドとともに使用されます。
- 財務チームとテクノロジーチームは、協力して将来的な AWS のコストと使用量を予測し、組織の予算を調整および構築します。
- 両者の共通言語により組織間のコミュニケーションが向上し、財務の概念への共通理解が得られます。

コストと使用量のディスカッションについて、組織内で関わるべきその他のステークホルダーは以下のとおりです。

- **事業部門オーナー:** 事業部門オーナーは、事業部門と会社全体の両方に方向性を提供できるように、クラウドのビジネスモデルを理解する必要があります。こうしたクラウド知識は成長とワークロード使用量を予測する際に、またリザーブドインスタンスや Savings Plans などの長期購入オプションを検討する際に重要な役割を果たします。
- **エンジニアリングチーム:** エンジニアがクラウド財務管理 (CFM) に取り組むよう促す、コストを意識した企業文化の構築には、財務チームとテクノロジーチームのパートナーシップの確立が欠かせません。CFM や財務業務の実務担当者や財務チームが共通して抱える問題の 1 つは、エンジニアにクラウド上のビジネス全体を理解させ、ベストプラクティスに従わせ、推奨されるアクションを取らせることです。
- **サードパーティー:** サードパーティー (コンサルタントやツールなど) を利用する場合、こうしたサードパーティーが財務目標に適合し、エンゲージメントモデルと投資収益率 (ROI) を通じて、どちらの整合性も実証できるようにします。通常、サードパーティーは自社管理のワークロードのレポートと分析を担当したり、自社設計のワークロードのコストを分析したりします。

CFM を導入し、成功させるには、財務、テクノロジー、ビジネスの各チームが協力し、組織全体におけるクラウド費用の伝達と評価の方法を変える必要があります。エンジニアリングチームを巻き込み、あらゆる段階でコストと使用に関する議論に参加させ、ベストプラクティスに従って合意されたアクションを取るよう奨励します。

実装手順

- **主要なメンバーを定義する:** 財務チームとテクノロジーチームのすべての関連メンバーがこの連携に関与していることを確認します。関連する財務メンバーは、クラウドの請求書に関する業務に従事するメンバーです。通常は、CFO、財務コントローラー、財務プランナー、ビジネスアナリスト、購買管理です。テクノロジーチームのメンバーは通常、プロダクトおよびアプリケーションの所有者、テクニカルマネージャー、およびクラウド上に構築するすべてのチームの代表者です。他のメンバーとしては、製品の使用に影響するマーケティングなどのビジネスユニットの所有者、および貴社の目標やメカニズムとの整合性を確保し、報告をサポートするコンサルタントなどのサードパーティーが含まれることがあります。
- **ディスカッションのためのトピックを定義する:** チーム間で共通する、または理解を共有する必要があるトピックを定義します。作成から支払いまでにかかるコストを追います。関連するメンバー、および適用する必要がある組織のプロセスを書き留めます。各ステップまたはプロセスのほか、利用可能な料金モデル、階層化された料金、割引モデル、予算、財務要件などの関連情報を理解します。
- **定期的なミーティングを設定する:** 財務とテクノロジーのパートナーシップを構築するために、定期的なコミュニケーションの機会を設け、連携を維持します。グループは目標とメトリクスに照らして定期的に集まる必要があります。一般的なミーティングでは、組織の状態の確認、現在実行中のプログラムの確認、全体的な財務および最適化メトリクスの確認を行います。その後、主要なワークロードが詳細に報告されます。

リソース

関連するドキュメント:

- [AWS ニュースブログ](#)

COST01-BP03 クラウドの予算と予測を確立する

既存の組織の予算作成および予測プロセスを調整し、非常に変動しやすいクラウドのコストと使用状況の性質に対応できるようにします。プロセスは、トレンドベースまたはビジネスドライバーベースのアルゴリズム、またはそれらの組み合わせを使用して、動的なものとする必要があります。

このベストプラクティスが確立されていない場合のリスクレベル: 高

実装のガイダンス

お客様は効率性、スピード、俊敏性を求めてクラウドを利用しますが、コストと使用量は大きく変動するものです。コストは、ワークロードの効率性の向上や、新規ワークロードや新機能のデプロイにより削減可能です (場合によっては増加することもあります)。ワークロードをスケーリングすると、サービスを提供する顧客が増えますが、その分クラウドの使用量とコストが増加します。リソースは以前より容易にアクセスできるようになります。クラウドの伸縮性は、コストと予測の伸縮性にもつながります。こうした変動を折り込めるように、組織の既存の予算編成プロセスを変える必要があります。

通常、予算は 1 年単位で設定され、固定されるため、関係者全員が厳守する必要があります。一方、予測はより柔軟であり、年間を通じて再調整が可能で、1 年、2 年、または 3 年の期間にわたる動的な予測を提供できます。予算編成と予測はどちらも、さまざまなテクノロジーやビジネスの関係者の間で財務上の期待事項を確立する上で重要な役割を果たします。正確な予測と実装は、第一にプロビジョニングコストに直接責任を負う関係者に説明責任をもたらし、全体的なコスト意識を高めることにもつながります。

トレンドベースのアルゴリズム (コスト履歴を入力値として使用)、動的および変動支出の環境に最適なドライバーベースのアルゴリズム (新製品の発売や営業地域の拡大、ワークロードの新しい環境など)、またはこの 2 つのアルゴリズムを組み合わせ、既存の予算編成と予測プロセスをより動的なものに調整します。

過去の支出に基づいて、定義された将来の期間のトレンドベースの予測を行うには、[AWS Cost Explorer](#) を使用できます。AWS Cost Explorer の予測エンジンは、料金タイプ (リザーブドインスタンスなど) に基づいて履歴データをセグメント化し、機械学習とルールベースのモデルの組み合わせを使用して、すべての料金タイプでの費用を個別に予測します。

使用コストに影響を与える可能性のあるビジネスドライバーを特定し、それぞれについて個別に予測して、予想される使用量が事前に計算されるようにします。組織内の IT チームや製品チームに関連するドライバーもあります。マーケティングイベント、プロモーション、合併、買収など、その他のビジネスドライバーについては、営業、マーケティング、ビジネス部門のリーダーが把握しているため、協力して、これらすべての需要ドライバーについても考慮することが重要です。新しい内部ドライバーへの影響を理解するには、こうした関係者と緊密に連携する必要があります。

Cost Explorer、またはその他のツールを使用してトレンドベースの予測を決定したら、[AWS Pricing Calculator](#) を使用し、予想される使用量 (トラフィック、1 秒あたりのリクエスト数、必要な Amazon EC2 インスタンス) に基づいて、AWS のユースケースと今後のコストを見積もります。これは、AWS を利用する際に、支出方法のプランニング、コスト削減機会の発見、情報に基づいた意思決定にも役立てることができます。予算はこうした予測計算と見積もりに基づいて設定する必要があるため、その予測の正確性を追跡することが重要です。

[AWS Budgets](#) を使用して、期間、繰り返し、または金額 (固定費または可変費) を指定し、サービス、AWS リージョン、タグなどのフィルターを追加することで、カスタム予算を詳細レベルで設定します。既存予算のパフォーマンスについて常に情報を入手するには、[AWS Budgets レポート](#) を作成して、自分と関係者に定期的に E メールで送信されるようにスケジュールします。また、[AWS Budgets アラート](#) は、実際のコストに基づいて作成することも (これは反応型です)、予測コストに基づいて作成することも可能で、潜在的なコスト超過に対する緩和策を実施する時間を確保することができます。コストまたは使用量が予算額を超えた場合や、超えると予測された場合、アラートを受け取ることができます。

[AWS Cost Anomaly Detection](#) を使用して、予想外のコストを防止または削減し、イノベーションを遅らせることなく制御を強化します。AWS Cost Anomaly Detection は、機械学習を利用して、異常な支出と根本

原因を特定するため、迅速な対応が可能になります。[3つのシンプルなステップ](#)で、状況に応じて独自のモニターを作成し、異常な支出が検出されたときにアラートを受け取ることができます。

Well-Architected コスト最適化の柱の[財務とテクノロジーの連携](#)セクションで説明されているように、IT、財務、その他の関係者がパートナーシップと連携を維持し、一貫性を保つために同じツールやプロセスを使用していることを確認することが重要です。予算を変更する必要がある場合、ミーティングの回数を増やすと、それらの変更により迅速に対応できます。

実装手順

- **トレンドベースの予測を分析する:** AWS Cost Explorer や Amazon Forecast など推奨されるトレンドベースの予測ツールを使用してください。サービス、アカウント、タグ、コストカテゴリなどのさまざまなディメンションで使用コストを分析します。高度な予測が必要な場合は、AWS Cost and Usage Report データを Amazon Forecast にインポートします (これにより、機械学習の一形式として線形回帰が適用され、予測が行われます)。
- **ドライバーベースの予測を分析する:** ビジネスドライバーがクラウドの使用に与える影響を特定し、それぞれについて個別に予測して、予想される使用コストを事前に計算します。ビジネスユニットのオーナーや関係者と緊密に連携して新しいドライバーへの影響を把握し、予想されるコストの変化を計算して正確な予算を決定します。
- **既存の予測および予算編成プロセスを更新する:** トレンドベース、ビジネスドライバーベースなど採用されている予測方法、または両方の予測方法の組み合わせに基づいて、予測および予算編成プロセスを定義します。予算は、これらの予測プロセスに基づいて計算され、現実的なものである必要があります。
- **アラートと通知を設定する:** アラートと通知を受け取るには、AWS Budgets アラートと AWS Cost Anomaly Detection を使用します。
- **主要関係者と定期的なレビューを行う:** 例えば、IT、財務、プラットフォームチームや、その他のビジネス分野の関係者が、ビジネスの方向性や使用状況の変化を認識し、連携を図ります。

リソース

関連するドキュメント:

- [AWS Cost Explorer](#)
- [AWS Cost and Usage Report](#)
- [Amazon QuickSight を使用した予測](#)
- [Amazon Forecast](#)
- [AWS Budgets](#)
- [AWS ニュースブログ](#)

関連動画:

- [How can I use AWS Budgets to track my spending and usage](#) (AWS Budgets を使用して支出と使用量を追跡するにはどうすればよいですか?)
- [AWS コスト最適化シリーズ: AWS Budgets](#)

関連する例:

- [Understand and build driver-based forecasting](#)
- [How to establish and drive a forecasting culture](#)
- [How to improve your cloud cost forecasting](#)
- [Using the right tools for your cloud cost forecasting](#)

COST01-BP04 組織のプロセスにコスト意識を採り入れる

コスト意識を高め、透明性を強化し、使用量に影響する新規または既存のプロセスにコストの説明責任を採り入れ、コスト意識に関する既存のプロセスを活用します。従業員のトレーニングにコスト意識の要素を採り入れます。

このベストプラクティスを活用しない場合のリスクレベル: 高

実装のガイダンス

コスト意識は、組織の新規および既存のプロセスに採り入れる必要があります。他のベストプラクティスの基盤となる前提条件の機能の 1 つです。可能な限り既存のプロセスを再利用し、修正することが推奨されます。これにより、俊敏性と速度への影響を最小化することができます。クラウドのコストをテクノロジーチームと、ビジネスチームおよび財務チームの意思決定者に報告して、コスト意識を高め、効率性についての重要業績評価指標 (KPI) を財務およびビジネス部門関係者向けに確立します。次の推奨事項は、ワークロードにコスト意識を実装するのに役立ちます。

- 変更管理に、変更による財務への影響を数値化するコスト測定が含まれていることを確認します。これは、コスト関連の懸念に積極的に対処し、コスト削減を強調するのに役立ちます。
- コストの最適化が、業務能力の中核をなす要素であることを確認します。例えば、既存のインシデントマネジメントプロセスを活用して、コストと使用量に関する異常値 (コスト超過) の根本原因を調査、特定することができます。
- オートメーションやツールにより、コスト削減とビジネス価値の実現を加速します。導入コストを考える場合、時間や費用の投資を正当化するために、投資収益率 (ROI) の要素を含むように話を組み立てます。
- コミットメントベースの購入オプション、共有サービス、マーケットプレイスでの購入を含むクラウド使用に対してショーバックまたはチャージバックを実施し、最もコストを意識したクラウド消費を促進することで、クラウドのコストを配分します。
- 既存のトレーニングおよび開発プログラムを拡張し、コスト意識向上のためのトレーニングを組織全体で実施します。これには継続的なトレーニングと認定が含まれることをお勧めします。これにより、コストと使用量を自己管理できる組織が育成されます。
- 次のような無料の AWS ネイティブツールを利用します。 [AWS Cost Anomaly Detection](#)、[AWS Budgets](#)、および [AWS Budgets Reports](#)。

組織が一貫して [クラウド財務管理](#) (CFM) 実践を採用すると、そのような行動は、仕事の進め方や意思決定方法に定着していきます。その結果、新しいクラウドで生まれたアプリケーションを設計するデベロッパーから、これらの新しいクラウド投資の ROI を分析する財務マネージャーまで、よりコストを意識した文化が生まれます。

実装手順

- 関連する組織のプロセスを把握する: 各組織単位は、そのプロセスをレビューし、コストと使用状況に影響を与えるプロセスを把握します。リソースの作成または終了につながるすべてのプロセスは、レビューの対象とする必要があります。インシデント管理やトレーニングなど、ビジネスにおけるコスト認識の維持につながるプロセスを探します。
- コストを意識した自立的な企業文化を確立する: すべての関係者がクラウドのコストを理解できるように、変更原因と影響をコストとして認識するようにします。そうすることで、組織がコストを意識したイノベーションの文化を自立的に確立することができます。
- コスト意識の要素を採り入れてプロセスを更新する: 各プロセスは、コスト意識が浸透するように変更されます。このプロセスでは、コストの影響の評価などの追加の事前チェック、またはコストと使用状況の予想された変化が発生したかどうかを検証する事後チェックが必要になる場合があります。トレーニング

ングやインシデント管理などのサポートプロセスは、コストと使用状況の項目を含むように拡張できます。

ご不明な点がございましたら、アカウントチームを通じて CFM のエキスパートにお問い合わせいただくか、以下のリソースや関連ドキュメントをご覧ください。

リソース

関連するドキュメント:

- [AWS によるクラウド財務管理](#)

関連する例:

- [効率的なクラウドコスト管理の戦略](#)
- [Cost Control Blog Series #3: How to Handle Cost Shock \(コスト管理ブログシリーズ #3: コストショックの扱い方\)](#)
- [AWS Cost Management 初心者ガイド](#)

COST01-BP05 コスト最適化に関して報告および通知する

クラウド予算を設定して、使用量の異常を検出するメカニズムを設定します。関連ツールで、事前定義済み目標に対するコストと使用量に関連するアラートを設定し、使用量がそれらの目標を超えた場合に通知を受け取るようにします。定期的にミーティングを開催して、ワークロードのコスト効率を分析し、社内コスト意識を浸透させます。

このベストプラクティスが確立されていない場合のリスクレベル: 低

実装のガイダンス

組織内のコストと使用量の最適化について、定期的に報告する必要があります。コストパフォーマンスについて話し合う専用セッションの運用や、ワークロードの定期的な運用レポートサイクルにコスト最適化を盛り込むことも意味があるでしょう。サービスとツールを使用して、コストパフォーマンスを定期的にモニタリングし、コスト節減機会を実現します。

[AWS Cost Explorer](#) を使用すると、複数のフィルターと粒度でコストと使用量を確認できます。これにより、サービス別またはアカウント別のコスト、日次コスト、マーケットプレイスコストなどのダッシュボードとレポートが表示されます。設定された予算に対するコストと使用量の進行状況を追跡するには、[AWS Budgets レポート](#) を使用します。

[AWS Budgets](#) を使用してカスタム予算を設定し、コストと使用量を追跡し、しきい値を超えた場合は E メールまたは Amazon Simple Notification Service (Amazon SNS) 通知でアラートを受信して、迅速に対応します。[優先予算](#) 期間を日次、月次、四半期ごと、または年次に設定し、特定の予算限度を設けて、予算しきい値に対する実際または予測されたコストと使用量の進捗状況に関して、常に情報を入手します。また、[アラートとアクション](#) を設定し、アラートに対してアクションを自動的に実行するか、予算目標を超えたときに承認プロセスを通じて実行するように設定することもできます。

コストと使用量に関する通知を実装して、コストと使用量の変化が予想外の場合はすばやく対応できるようにします。[AWS Cost Anomaly Detection](#) では、イノベーションを遅らせることなく、予想外のコストを削減し、管理を強化できます。AWS Cost Anomaly Detection は異常な支出と根本原因を特定するため、予

想外の請求のリスクを軽減できます。3 つのシンプルなステップで、状況に応じて独自のモニターを作成し、異常な支出が検出されたときにアラートを受け取ることができます。

また、[Amazon QuickSight](#) と AWS Cost and Usage Report (CUR) のデータを使用して、高度にカスタマイズされ、きめ細かなデータを含んだレポートを提供することもできます。Amazon QuickSight では、過去のコストと使用量またはコスト削減機会に関するレポートをスケジュールし、定期的なコストレポートを E メールで受信できます。高度な可視性を実現する、Amazon QuickSight の [Cost Intelligence Dashboard](#) (CID) ソリューションビルドをぜひチェックしてみてください。

[AWS Trusted Advisor](#) を使用すると、プロビジョニングされたリソースがコスト最適化のための AWS のベストプラクティスに準拠しているかどうかを確認するためのガイダンスが提供されます。

視覚的なグラフで、詳細なコストと使用量に関する Savings Plans のレコメンデーションを確認できます。時間単位のグラフには、オンデマンドの支出と推奨される Savings Plans のコミットメントが表示され、推定削減額、Savings Plans の適用範囲、Savings Plans 使用率に関するインサイトが得られます。これにより、組織は、支出分析モデルの構築に時間とリソースを費やすことなく、毎時の支出にどのように Savings Plans が適用されているかを理解できます。

Savings Plans、リザーブドインスタンス、および Amazon EC2 の適切なサイズ設計に関する AWS Cost Explorer からのレコメンデーションを含んだレポートを定期的に作成して、定常状態のワークロード、アイドルおよび使用量の少ないリソースに関するコストの削減を開始します。デプロイされているリソースのうち、クラウドの無駄に関する費用を特定し、回収します。クラウドの無駄は、サイズ設定が正しくないリソースが作成されたときや、使用量のパターンが予想とは異なるときに発生します。AWS のベストプラクティスに従って無駄を減らすか、クラウドコストを[最適化して節約](#)するため、アカウントチームやパートナーの協力を仰ぎます。

定期的にレポートを作成し、リソースの購入オプションを改善することで、ワークロードの単価を下げることができます。Savings Plans、リザーブドインスタンス、Amazon EC2 スポットインスタンスなどの購入オプションは、耐障害性の高いワークロードのコストを最も低く抑え、関係者（ビジネス所有者、財務チーム、テクノロジーチーム）がこれらのコミットメントの議論に参加できるようにするものです。

クラウドの総保有コスト (TCO) の削減に役立つ可能性のある機会または新しいリリースの発表を含んだレポートを共有します。新しいサービス、リージョン、機能、ソリューション、またはさらにコスト削減を達成する新しい方法を採用します。

実装手順

- AWS Budgets を設定する: ワークロードのすべてのアカウントで AWS Budgets を設定します。タグを使用して、アカウント全体の支出の予算とワークロードの予算を設定します。
 - [Well-Architected ラボ: コストと使用量のガバナンス](#)
- コスト最適化に関して報告する: ワークロードの効率について話し合い、分析する定期的なミーティングを設定します。確立されたメトリクスを使用して、達成されたメトリクスとそれを達成するためにかったコストを報告します。好ましくない傾向を特定して修正するとともに、組織全体で推進できるような改善傾向を特定します。報告には、アプリケーションチームと所有者、財務、およびクラウド支出に関する主要な意思決定者の代表者が参加する必要があります。

リソース

関連するドキュメント:

- [AWS Cost Explorer](#)
- [AWS Trusted Advisor](#)
- [AWS Budgets](#)
- [AWS Cost and Usage Report](#)
- [AWS Budgets のベストプラクティス](#)

- [Amazon S3 分析](#)

関連する例:

- [Well-Architected ラボ: コストと使用量のガバナンス](#)
- [Key ways to start optimizing your AWS cloud costs](#)

COST01-BP06 コストをプロアクティブにモニタリングする

ツールとダッシュボードを実装して、ワークロードのコストをプロアクティブにモニタリングします。通知を受けたときだけコストやカテゴリを見るのではなく、設定されたツールや既存のツールで定期的にコストを見直しましょう。コストをプロアクティブにモニタリングし、分析することで、ポジティブな傾向を把握し、組織全体で推進することが可能になります。

このベストプラクティスを活用しない場合のリスクレベル: ミディアム

実装のガイダンス

例外や異常がある場合に限らず、組織内のコストと使用量を事前にモニタリングすることを推奨します。オフィスや職場環境全体を高度に可視化したダッシュボードにより、主な担当者が必要とする情報にアクセスできるようになります。また組織がコスト最適化を重視していることを示すことができます。可視化されたダッシュボードにより、成功した事例を積極的に推進し、組織全体で実践することができます。

日次または頻繁なルーチンを作成するには、[AWS Cost Explorer](#) または [Amazon QuickSight](#) など、その他のダッシュボードを使用して、コストを確認し、プロアクティブに分析します。AWS サービスの使用量とコストを AWS アカウントレベル、ワークロードレベル、または特定の AWS サービスレベルでグループ化とフィルタリングを使用して分析し、予想通りかどうかを検証します。時間単位、リソース単位の詳細度やタグを使用して、上位リソースの発生コストをフィルタリングし、特定します。また、[Cost Intelligence Dashboard](#) という [Amazon QuickSight](#) ソリューション (AWS ソリューションアーキテクトが構築) を使用して独自のレポートを作成して、予算と実際のコストおよび使用量を比較することもできます。

実装手順

- コスト最適化に関して報告する: ワークロードの効率について話し合い、分析する定期的なミーティングを設定します。確立されたメトリクスを使用して、達成されたメトリクスとそれを達成するためにかったコストを報告します。ネガティブな傾向を特定して修正し、ポジティブな傾向を特定して組織全体に普及させます。報告には、アプリケーションチームおよび所有者、財務、および管理の担当者が必要があります。
- コストと使用量について日次の詳細度の [AWS Budgets](#) を作成し、有効にすることで、潜在的なコスト超過を防ぐためのタイムリーなアクションを取ることができます。AWS Budgets ではアラート通知を設定できるため、いずれかの予算タイプが事前設定のしきい値を超えた場合に通知を受けることができます。AWS Budgets を活用する最善の方法は、予想されるコストと使用量を限度として設定することです。そうすれば、予算を超えたものは使い過ぎとみなすことができます。
- コストモニターとして AWS Cost Anomaly Detection を作成します。[AWS Cost Anomaly Detection](#) は、高度な機械学習テクノロジーを使用して、異常な支出と根本原因を特定するため、迅速に対策を講じることができます。評価したい支出セグメント (例えば、個々の AWS サービス、メンバーアカウント、コスト配分タグ、コストカテゴリ) を定義するコストモニターを設定することができ、アラート通知をいつ、どこで、どのように受け取るかを設定することが可能です。各モニターには、ビジネスオーナーやテクノロジーチーム向けの複数のアラートサブスクリプションをアタッチし、各サブスクリプションの名前、コスト影響しきい値、アラート頻度 (個別アラート、日次サマリー、週次サマリー) などを設定します。

- AWS Cost Explorer を使用するか、AWS Cost and Usage Report (CUR) データを Amazon QuickSight ダッシュボードに統合して、組織のコストを視覚化します。AWS Cost Explorer には、使いやすいインターフェイスがあり、AWS のコストと使用量を時系列で可視化し、理解し、管理することができます。最も [Cost Intelligence Dashboard](#) カスタマイズ可能でアクセスしやすいダッシュボードで、独自のコスト管理、最適化ツールの基礎作りを支援します。

リソース

関連するドキュメント:

- [AWS Budgets](#)
- [AWS Cost Explorer](#)
- [日次コストおよび使用量の予算](#)
- [AWS Cost Anomaly Detection](#)

関連する例:

- [Well-Architected ラボ: 可視化](#)
- [Well-Architected ラボ: 高度な可視化](#)
- [Well-Architected ラボ: Cloud Intelligence Dashboards](#)
- [Well-Architected ラボ: コストの可視化](#)
- [Slack による AWS Cost Anomaly Detection アラート](#)

COST01-BP07 新しいサービスリリースに関する最新情報を把握しておく

エキスパートや AWS パートナーに定期的に相談して、コストの低いサービスと機能を検討します。AWS のブログやその他の情報ソースを確認します。

このベストプラクティスを活用しない場合のリスクレベル: ミディアム

実装のガイダンス

AWS は常に新しい機能を追加しているため、最新のテクノロジーを利用して、実験とイノベーションをより迅速に行うことができます。新しい AWS のサービスや機能を実装することでワークロードのコスト効率を改善できる場合があります。新しいサービスと機能のリリースに関する情報を入手するには、[AWS コスト管理](#)、[AWS ニュースブログ](#)、[AWS コスト管理ブログ](#)、および [AWS の最新情報](#) を定期的に確認してください。「最新情報」では、AWS サービス、機能、リージョン拡大の発表があった際に、その概要をお知らせしています。

実装手順

- ブログをサブスクライブする: AWS ブログのページにアクセスし、最新情報ブログやその他の関連ブログをサブスクライブします。E メールアドレスを記載して、[通信方法](#) ページでサインアップできます。
- AWS ニュースを購読する: 新しいサービスと機能のリリースに関する情報を入手するには、[AWS ニュースブログ](#) と [AWS の最新情報](#) を定期的に確認してください。RSS フィード、または E メールでお知らせやリリースを購読することができます。
- AWS の値下げをフォローする: すべてのサービスにおいて定期的な値下げを行うことは、AWS がその規模から得られる経済的な効率性をお客様に還元するための標準的な方法となっています。2022 年 4 月時点で、AWS は 2006 年のサービス提供開始以来、115 回の料金引き下げを行ってきました。料金面

の懸念から経営判断を保留しているものがあれば、値下げや新サービス統合後に再度見直すことも可能です。Amazon Elastic Compute Cloud (Amazon EC2) インスタンスも含め、以前の値下げについては、[AWS ニュースブログの料金引き下げカテゴリを参照してください](#)。

- AWS のイベントおよび交流: ローカルの AWS サミットや、地域内の他の組織との交流に参加しましょう。直接参加できない場合は、バーチャルイベントに参加して、AWS のエキスパートや他のお客様のビジネスケースから情報を得るようにしてください。
- アカウントチームとのミーティングを設ける: アカウントチームとの定期的なミーティングを設定し、チームと会い、業界の動向と AWS のサービスについて話し合います。アカウントマネージャー、ソリューションアーキテクト、サポートチームに相談する。

リソース

関連するドキュメント:

- [AWS コスト管理](#)
- [AWS の最新情報](#)
- [AWS ニュースブログ](#)

関連する例:

- [Amazon EC2 - IT コストの最適化と削減に取り組んだ 15 年間](#)
- [AWS ニュースブログ - 料金引き下げ](#)

COST01-BP08 コスト意識を持つ文化を生み出す

コストを意識した企業文化を醸成するために、組織全体で改革やプログラムを実施しましょう。まず小さく始めて、続いて機能や組織でのクラウド利用の増加に合わせて、規模を拡大していき、さまざまなプログラムを運用していくことを推奨します。

このベストプラクティスが確立されていない場合のリスクレベル: 低

実装のガイダンス

コスト意識を持つ文化があると、組織全体で有機的かつ分散的に実行されるベストプラクティスを通じて、コストの最適化とクラウド財務管理 (財務運用、Cloud Center of Excellence、クラウド運用チームなど) の規模を拡大することができます。コスト意識を持つことで、トップダウンで集中的に行う厳格なアプローチと比較して、最小限の労力で組織全体に高いレベルの能力を生み出すことができます。

クラウドコンピューティング、特にクラウドコンピューティングの主なコスト要因についてのコスト意識を持つことで、チームは予想される変更の成果をコスト面で理解できます。クラウド環境にアクセスするチームは、料金モデルと、従来のオンプレミスデータセンターとクラウドコンピューティングの違いを認識する必要があります。

コストを意識する文化の主な利点は、技術チームが必要に応じて消極的なコスト最適化を行うのではなく、積極的かつ継続的にコストを最適化する (たとえば、新しいワークロードを設計する際や、既存のワークロードを変更する際に機能要件でないと見なされる) ことにあります。

この文化のわずかな変化で、現在や将来のワークロードの効率に大きな影響を与える可能性があります。これには、次のような例があります。

- エンジニアリングチームに可視性を与え、意識を高めることで、自分たちが何をしているのか、コスト面でどのような影響があるのかを理解させることができます。

- 組織全体のコストと使用量にゲーム的要素を取り入れる。これは、公開ダッシュボードや、チーム間の標準コストと標準使用量 (例えば、ワークロードあたりのコストとトランザクションあたりのコストなど) を比較するレポートによって実行できます。
- コスト効率を認識する。自発的または独断で行なったコスト最適化の成果を公開または非公開で評価して、間違いから学び、今後繰り返さないようにします。
- あらかじめ設定された予算でワークロードを実行するために、トップダウンの組織的要件を作成します。
- 変更のビジネス要件と、アーキテクチャインフラストラクチャまたはワークロード設定について要求された変更がコストにどのように影響するかを探索して、必要な分だけ支払うようにします。
- 変更計画者は、予想される変更がコストに影響することを意識し、コスト効率的にビジネス成果をもたらすように関係者に確認してもらう必要があります。

実装手順

- クラウドコストをテクノロジーチームに報告する: これによりコスト意識を高め、財務およびビジネス関係者にとって効率的な KPI を確立します。
- 予定されている変更に関係者やチームメンバーに通知する: 予定されている変更とワークロードに対する費用対効果を週次の変更ミーティングで協議するための議題を作成します。
- アカウントチームとのミーティングを設ける: アカウントチームとの定期的なミーティングを設定し、業界の動向と AWS のサービスについて話し合います。アカウントマネージャー、アーキテクト、サポートチームと話します。
- 成功事例を共有する: ワークロード、AWS アカウント、または組織のコスト削減に関する成功事例を共有して、コスト最適化に関する前向きな姿勢と励みにします。
- トレーニング: 技術チームやチームメンバーが AWS クラウド に関するリソースコストについて認識するためのトレーニングを受けるようにします。
- AWS のイベントおよび交流: ローカルの AWS サミットや、地域内の他の組織との交流に参加します。
- ブログを登録する: AWS ブログのページにアクセスし、[最新情報ブログ](#) やその他の関連ブログを購読して、AWS が共有する新しいリリース、実装、例、および変更に従います。

リソース

関連ドキュメント:

- [AWS ブログ](#)
- [AWS コスト管理](#)
- [AWS ニュースブログ](#)

関連する例:

- [AWS によるクラウド財務管理](#)
- [AWS Well-Architected ラボ: クラウド財務管理](#)

COST01-BP09 コスト最適化によるビジネス価値を数値化する

コスト最適化でビジネス価値を数値化することで、組織に対するメリットの全体像を把握できます。コスト最適化は必要な投資であるため、ビジネス価値を数値化することで、各ステークホルダーに投資利益率を説明できます。ビジネス価値の数値化は、将来のコスト最適化投資に関してステークホルダーからより

多くの賛同を得ることができます。また、組織のコスト最適化活動の結果を測定するためのフレームワークを提供します。

このベストプラクティスが確立されていない場合のリスクレベル: 中

実装のガイダンス

ビジネス価値の数値化とはつまり、企業の行動や決断がもたらす利益を測定することです。ビジネス価値は、有形（経費の削減や利益の向上など）の場合もあれば、無形（ブランドの評判や顧客満足度の向上など）の場合もあります。

コスト最適化によるビジネス価値の数値化とはすなわち、支出を効率化する取り組みがもたらす価値や利益を割り出すことです。例えば、ある企業が AWS へのワークロードのデプロイに 100,000 USD を費やし、後日最適化した結果、品質や成果を損なうことなく、わずか 80,000 USD にコストダウンしたとします。この場合、コスト最適化がもたらすビジネス価値を数値化すると、20,000 USD の節約になります。しかし、単純な節約額だけでなく、納期の短縮や顧客満足度の向上、コスト最適化の取り組みの成果が現れたその他の指標を踏まえて、価値を数値化することもできます。ステークホルダーは、コスト最適化の潜在的価値、ワークロードの最適化にかかるコスト、投資収益率について決定する必要があります。

コスト最適化により削減された金額を報告することに加えて、実現された付加価値を数値化することをお勧めします。コスト最適化のメリットは通常、ビジネス成果に対して削減されたコストという観点で数値化されます。例えば、Savings Plans を購入した場合は Amazon Elastic Compute Cloud (Amazon EC2) のコスト削減を数値化できます。Savings Plans により、コストを削減し、ワークロードの出力レベルを維持できます。アイドル状態の Amazon EC2 インスタンスが削除された場合や、アタッチされていない Amazon Elastic Block Store (Amazon EBS) ボリュームが削除された場合は、AWS の利用料削減を数値化できます。

コスト最適化のメリットは、コスト削減やコスト回避にとどまりません。効率性向上とビジネス価値を測定するために、その他のデータを追加で取得することを検討してください。

実装手順

- ビジネス上の利点を評価する: これは、AWS クラウド のコストを分析し、支出した分、最大限の利益を生み出せるように調整していくプロセスです。ビジネス価値を顧みずコスト削減にばかり着目するのではなく、コスト最適化がもたらすビジネス上の利点と投資収益率を検討してください。支出した金額からもっと価値を引き出せるようになります。賢く支出し、最大の収益率を見込める分野に投資および出費することが大切です。
- AWS の予測コストを分析する: 予測により、財務関係者は、組織内外の他の利害関係者と見通しを立て、組織の財務予測の可能性を向上させることができます。[AWS Cost Explorer](#) を使用して、コストと使用量を予測できます。

リソース

関連するドキュメント:

- [AWS クラウド エコノミクスセンター](#)
- [AWS ブログ](#)
- [AWS コスト管理](#)
- [AWS ニュースブログ](#)
- [Well-Architected フレームワークの信頼性の柱に関するホワイトペーパー](#)
- [AWS Cost Explorer](#)

関連動画:

- [Unlock Business Value with Windows on AWS](#)

関連する例:

- [カスタマー 360 のビジネス価値の測定と最大化](#)
- [The Business Value of Adopting Amazon Web Services Managed Databases](#)
- [The Business Value of Amazon Web Services for Independent Software Vendors](#)
- [Business Value of Cloud Modernization](#)
- [The Business Value of Migration to Amazon Web Services](#)

経費支出と使用量の認識

組織のコストおよびコスト要因を把握することは、コストと使用量を効果的に管理し、コスト削減の機会を特定するうえで極めて重要です。組織では一般に、複数のワークロードが複数のチームによってオペレーションされています。各チームはさまざまな組織単位に属する可能性があり、そのそれぞれに独自の収益の流れがあります。リソースコストの帰属先をワークロード、各組織、製品オーナーのいずれかに割り当てると、リソースを効率的に使用し、無駄を削減できます。コストと使用量を正確にモニタリングすることで、各組織単位や製品の収益性が把握できるようになり、より確かな情報に基づいて組織内のリソース配分を決定できます。使用量が増加するとコストも変動するため、組織内のあらゆるレベルの使用量を認識することは、変化を促進する鍵となります。

使用方法と支出を認識するために多面的なアプローチを取ることを検討してください。チームは、データを収集、分析し、それに続いて報告する必要があります。考慮すべき主な要因は以下のとおりです。

トピック

- [ガバナンス \(p. 23\)](#)
- [コストと使用量のモニタリング \(p. 35\)](#)
- [リソースを削除する \(p. 45\)](#)

ガバナンス

クラウドのコストを管理するには、以下のガバナンス領域を通じて使用量を管理する必要があります。

ベストプラクティス

- [COST02-BP01 組織の要件に基づいてポリシーを策定する \(p. 23\)](#)
- [COST02-BP02 目標およびターゲットを策定する \(p. 26\)](#)
- [COST02-BP03 アカウント構造を実装する \(p. 28\)](#)
- [COST02-BP04 グループとロールを実装する \(p. 31\)](#)
- [COST02-BP05 コストコントロールを実装する \(p. 32\)](#)
- [COST02-BP06 プロジェクトのライフサイクルを追跡する \(p. 34\)](#)

COST02-BP01 組織の要件に基づいてポリシーを策定する

組織によるリソースの管理方法を定義するポリシーを策定し、定期的に検査します。ポリシーでは、リソースのライフタイム全体にわたる作成、変更、削除を含む、リソースとワークロードのコスト面をカバーする必要があります。

このベストプラクティスを活用しない場合のリスクレベル: 高

実装のガイダンス

組織のコストおよびコスト要因を把握することは、コストと使用量を効果的に管理し、コスト削減の機会を特定するうえで極めて重要です。組織では一般に、複数のワークロードが複数のチームによってオペ

レーションされています。各チームはさまざまな組織単位に属する可能性があり、そのそれぞれに独自の収益の流れがあります。リソースのコストをワークロード、それぞれの組織、製品オーナーに帰属させることができると、リソースを効率的に使用し、無駄を削減できます。コストと使用量を正確にモニタリングすることで、ワークロードがどの程度最適化されているか、また組織単位や製品の収益性がどの程度あるかを理解するのに役立ちます。この知識により、組織内のどこにリソースを割り当てるかについて、より多くの情報に基づいた意思決定が可能になります。使用量が増加するとコストも変動するため、組織内のあらゆるレベルの使用量を認識することは、変化を促進する鍵となります。使用方法と支出を認識するために多面的なアプローチを取ることを検討してください。

ガバナンスを実行するための最初のステップは、組織の要件を使用して、クラウド使用に関するポリシーを作成することです。ポリシーでは、組織がクラウドをどのように使用するかや、リソースをどのように管理するかを定義します。ポリシーではコストや使用量に関係するリソースとワークロードのあらゆる局面、つまりリソースのライフタイム全体にわたる作成、変更、削除をカバーする必要があります。クラウド環境でのあらゆる変更について、ポリシーと手順の遵守と実装を徹底してください。IT の変更管理会議では、質問を提起して、計画された変更によるコストへの影響 (増加または減少)、ビジネスの正当性、期待される結果を確認します。

ポリシーを簡単に理解し、組織全体で効果的に実装するには、シンプルなものにする必要があります。また、ポリシーは、(使用されるように) 遵守と解釈が容易で、(チーム間で誤解が発生しないように) 具体的である必要があります。さらに、顧客のビジネス状況や優先順位が変わったときにポリシーが古くなるため、(当社のメカニズムと同様に) 定期的に検査し、更新する必要があります。

使用する地理的リージョンや、リソースを稼働する時間帯など、大局的な幅広いポリシーから始めます。続いてポリシーを徐々に絞り込み、さまざまな組織単位やワークロードに対応させます。一般的なポリシーの例としては、どのサービスと機能を利用できるか (例えば、テスト環境や開発環境では低パフォーマンスのストレージ)、どのタイプのリソースを各グループで利用できるか (例えば、開発用アカウントのリソースの最大サイズを medium にする)、これらのリソースをどの程度のスパンで使用するか (一時的、短期、特定の期間)、などがあります。

ポリシーの例

次に、コスト最適化に焦点を当てた独自のクラウドガバナンスポリシーを作成するために確認できるポリシーの例を示します。組織の要件と関係者の要求に基づいてポリシーを調整するようにしてください。

- **ポリシー名:** リソースの最適化やコスト削減ポリシーなど、明確なポリシー名を定義します。
- **目的:** このポリシーを使用すべき理由と、期待される結果を説明してください。このポリシーの目的は、ビジネス要件を満たすために必要なワークロードのデプロイと実行に必要な最小限のコストがあると確認することです。
- **スコープ:** このポリシーを誰が使用すべきか、いつ使用すべきかを明確に定義してください。例えば、DevOps X Team が X 環境 (本番環境または非本番環境) で米国東部のお客様にこのポリシーを使用する、などです。

ポリシーステートメント

1. ワークロードの環境とビジネス要件 (開発、ユーザー受け入れテスト、本番稼働前、または本番稼働) に基づいて us-east-1 または複数の us-east リージョンを選択します。
2. Amazon EC2 と Amazon RDS インスタンスが、午前 6 時から夜 8 時 (東部標準時 (EST)) に実行されるようにスケジュールを立てます。
3. 8 時間後に未使用の Amazon EC2 インスタンスをすべて停止し、24 時間非アクティブ状態の未使用の Amazon RDS インスタンスをすべて停止します。
4. 非本番環境で非アクティブ状態が 24 時間続いたら、未使用の Amazon EC2 インスタンスをすべて終了します。Amazon EC2 インスタンス所有者に (タグに基づいて) 停止した Amazon EC2 インスタンスを本番環境で確認するように促し、使用していない場合は Amazon EC2 インスタンスが 72 時間以内に終了することを通知します。
5. m5.large などの汎用インスタンスファミリーとサイズを使用し、AWS Compute Optimizer を使用して CPU とメモリの使用率に基づいてインスタンスのサイズを変更します。

6. 自動スケーリングの使用を優先して、トラフィックに基づいて実行中のインスタンスの数を動的に調整します。
7. 重要ではないワークロードにはスポットインスタンスを使用します。
8. キャパシティ要件を確認して、予測可能なワークロードに備えて削減プランやリザーブドインスタンスをコミットし、クラウド財務管理チームに通知してください。
9. Amazon S3 ライフサイクルポリシーを使用して、アクセス頻度の低いデータを安価なストレージ階層に移動できます。保存ポリシーが定義されていない場合は、Amazon S3 Intelligent Tiering を使用してオブジェクトをアーカイブ階層に自動的に移動します。
10. Amazon CloudWatch を使用して、リソースの使用状況をモニタリングし、スケーリングイベントをトリガーするアラームを設定します。
11. それぞれの AWS アカウント について、AWS Budgets を使用して、コストセンターとビジネスユニットに基づいてアカウントのコストと使用量の予算を設定します。
12. AWS Budgets を使用してアカウントのコストと使用量の予算を設定すると、支出を常に把握し、予想しない請求を回避できるため、コストをより適切に制御できます。

プロシージャ: このポリシーを実装するための詳細な手順を提供するか、各ポリシーステートメントの実装方法を説明する他のドキュメントを参照してください。このセクションでは、ポリシー要件を実行するためのステップバイステップの手順を説明する必要があります。

このポリシーを実装するには、さまざまなサードパーティ製ツールや AWS Config ルールを使用してポリシーステートメントの遵守を確認したり、AWS Lambda 関数を使用して自動修復アクションをトリガーしたりできます。AWS Organizations を使用してポリシーを適用することもできます。さらに、リソースの使用状況を定期的に見直し、必要に応じてポリシーを調整して、ビジネスニーズが引き続き満たされていることを確認する必要があります。

実装手順

- 関係者とのミーティングを設ける: ポリシーを策定するには、組織内の関係者 (クラウドビジネスオフィス、エンジニア、またはポリシー実施部門の意思決定者) に、要件を明記して文書化するよう依頼します。幅広く開始し、各ステップで最小単位まで継続的に絞り込んでいくという反復型アプローチを採用します。チームメンバーには、組織単位やアプリケーションの所有者など、ワークロードの直接の関係者に加え、セキュリティチームや財務チームなどのサポートグループを含みます。
- 確認する: 誰が AWS クラウド にアクセスしてデプロイできるかを指定したポリシーに、チームが同意していることを確認します。チームが組織のポリシーに従っているかどうか、同意したポリシーおよび手順に沿ってチームがリソースを作成しているかどうかを確認してください。
- オンボーディングトレーニングセッションを作成する: 新しい組織メンバーに対し、コスト意識を定着させ、組織の要件を特定するために、オンボーディングトレーニングコースを完了するよう求めます。新しいメンバーは、以前の経験から異なるポリシーを想定している場合や、ポリシーについてまったく考えていない場合があります。
- ワークロードの場所を定義する: ワークロードの運用場所 (国や国内のエリアなど) を定義します。この情報は、AWS リージョンとアベイラビリティゾーンへのマッピングに使用されます。
- サービスとリソースを定義およびグループ化する: ワークロードに必要なサービスを定義します。サービスごとに、タイプ、サイズ、必要なリソースの数を指定します。アプリケーションサーバーやデータベースストレージなどの機能別にリソースのグループを定義します。リソースは複数のグループに属することができます。
- 機能別にユーザーを定義およびグループ化する: ワークロードに関係するユーザーについて、当該ユーザーが誰かまたは組織内での地位に焦点を当てるのではなく、何を行うか、またはどのようにワークロードを使用するかに焦点を当てて定義します。類似したユーザーまたは機能をグループ化します。AWS 管理ポリシーをガイドとして使用できます。
- アクションを定義する: 以前に特定した場所、リソース、およびユーザーを使用して、そのライフタイム (開発、運用、削除) にわたってワークロードの成果を得るために、それぞれが必要とするアクションを定義します。各場所で、グループ内の個々の要素ではなく、グループに基づいてアクションを特定し

ます。開始時には読み取りまたは書き込みを幅広く設定し、それぞれのサービスについて、特定のアクションへと絞り込んでいきます。

- レビュー期間を定義する: ワークロードと組織の要件は、時間の経過とともに変化する可能性があります。ワークロードのレビュースケジュールを定義して、組織の優先順位に合わせた状態を維持します。
- ポリシーを文書化する: 組織で、定義されたポリシーが、必要に応じてアクセス可能であることを確認します。これらのポリシーは、環境へのアクセスを実装、保守、監査するために使用されます。

リソース

関連するドキュメント:

- [Change Management in the Cloud \(クラウドにおける変更管理\)](#)
- [AWS Managed Policies for Job Functions](#)
- [AWS 複数アカウントの請求戦略](#)
- [Actions, Resources, and Condition Keys for AWS Services](#)
- [AWS での管理とガバナンス](#)
- [Control access to AWS リージョン using IAM policies](#)
- [グローバルインフラストラクチャリージョンと AZ](#)

関連動画:

- [AWS での大規模な管理とガバナンス](#)

関連サンプル:

- [VMware - What Are Cloud Policies? \(VMware - クラウドポリシーとは\)](#)

COST02-BP02 目標およびターゲットを策定する

ワークロードのコストおよび使用量の両方について、目標およびターゲットを策定します。目標は、期待される結果に基づく方向性を組織に示し、ターゲットは、ワークロードについて具体的に達成すべき測定可能な成果を表します。

このベストプラクティスを活用しない場合のリスクレベル: 高

実装のガイダンス

組織のコスト、目標使用量、ターゲットを設定します。AWS で成長を続ける組織にとって、コスト最適化の目標を設定して追跡することは重要です。これらの目標または [主要業績評価指標 \(KPI\)](#) には、オンデマンドでの支出の割合や、AWS Graviton インスタンスまたは gp3 EBS ボリュームタイプなどの特定の最適化されたサービスの導入などが含まれます。測定可能で達成可能な目標を設定することで、継続的なビジネスオペレーションにとって重要な効率の向上を継続的に測定できるようになります。目標は、期待される成果に関するガイダンスと指示を組織にもたらしめます。ターゲットは、具体的かつ測定可能な達成すべき結果をもたらします。つまり、目標は進みたい方向を示し、ターゲットはその方向へどこまで進み、その目標をいつ達成する必要があるかを表します。このとき、「SMART」つまり具体的 (specific)、測定可能 (measurable)、割り当て可能 (assignable)、現実的 (realistic)、タイムリー (timely) であることをガイダンスとします。「プラットフォームの使用量を大幅に増加させ、コストは微増 (非線形) にとどまるようにする」などは目標の例です。「プラットフォームの使用量を 20% 増加させ、コスト増は 5% 未満に抑える」などはターゲットの例です。ワークロードを 6 か月ごとに効率化する必要があるというケースも、目標としてはよくあります。付随する目標は、ビジネスあたりのコストメトリクスを 6 か月ごとに 5% 削減する必要があるというものです。

コスト最適化の目標は、ワークロードの効率を高めることです。つまり、ワークロードのビジネス成果あたりのコストを経時的に削減することです。この目標と合わせて、6 か月から 1 年ごとに効率を 5% 向上させるなどのターゲットをすべてのワークロードに設定することを推奨します。これは、クラウド内でコスト最適化の機能を構築し、新しいサービスや機能をリリースすることで達成できます。

KPI と関連する節約の機会をほぼリアルタイムで把握し、経時的に進捗状況を追跡することが重要です。KPI 目標の定義と追跡を始めるには、[Cloud Intelligence Dashboards \(CID\) フレームワークの KPI ダッシュボードをお勧めします](#)。KPI ダッシュボードには、AWS Cost and Usage Report から取得したデータに基づいて一連の推奨コスト最適化 KPI が表示され、カスタム目標の設定や、経時的な進捗状況の追跡ができます。

KPI 目標を設定して追跡できる別のソリューションがある場合は、そのソリューションが組織内のすべてのクラウド財務管理関係者に採用されていることを確認してください。

実装手順

- 予想される使用レベルを定義する: まず、使用レベルに焦点を当てます。アプリケーションの所有者、マーケティング、およびより大きなビジネスチームと協力して、ワークロードに対して予想される使用レベルを把握します。顧客の需要が経時的にどのように変化するか、季節要因による増加やマーケティングキャンペーンによる変化が生じるかどうか、などを考慮します。
- ワークロードのリソースとコストを定義する: 使用レベルを定義したうえで、これらの使用レベルを満たすために必要なワークロードリソースの変化を数値化します。ワークロードコンポーネントのサイズまたはリソースの数を増やしたり、データ転送を増やしたり、特定のレベルでワークロードコンポーネントを別のサービスに変更したりすることが必要な場合があります。これらの主な各ポイントにおけるコストと、使用状況が変更された場合のコストの変化を指定します。
- ビジネス目標を定義する: 予想される使用量とコストの変化から結果を取得し、これを、予想されるテクノロジーや実行中のプログラムの変化と組み合わせて、ワークロードの目標を策定します。目標は、使用量、コスト、および両者の関わり方を対象に含めたものである必要があります。目標はシンプルかつ大局的なものにして、そのビジネスで求められる成果を従業員が理解できるような内容にする必要があります (未使用のリソースを特定のコストレベル以下に抑えるなど)。未使用リソースのタイプごとに目標を定義したり、目標とターゲットに対しての損失とするコストを具体的に設定する必要はありません。使用量に変化がない状態でコストの変化が予想される場合は、組織的なプログラム (トレーニングや教育による能力向上など) を用意しておきます。
- ターゲットを定義する: 定義された目標ごとに、測定可能なターゲットを指定します。ワークロードの効率を高めることが目標である場合、ターゲットでは、改善の量 (通常は 1 ドルあたりのビジネス成果で示す) と、その改善をいつ実現できるかを数値化します。例えば、オーバープロビジョニングによる無駄を最小限に抑えるという目標を設定した場合、ターゲットとしては、「実稼働ワークロードの 1 つ目の階層でコンピューティングのオーバープロビジョニングによる無駄を階層コンピューティングコストの 10% 以下に抑え、実稼働ワークロードの 2 つ目の階層でコンピューティングのオーバープロビジョニングによる無駄を階層コンピューティングコストの 5% 以下に抑える」のような内容が考えられます。

リソース

関連するドキュメント:

- [AWS managed policies for job functions](#)
- [AWS Control Tower のランディングゾーンに対する AWS マルチアカウント戦略](#)
- [Control access to AWS リージョン using IAM policies](#)
- [SMART 目標](#)

関連動画:

- [Well-Architected Labs: Goals and Targets \(Level 100\) \(Well-Architected ラボ: 目標とターゲット \(レベル 100\)\)](#)

関連サンプル:

- [Well-Architected Labs: Decommission resources \(Goals and Targets\) \(Well-Architected ラボ: リソースを削除する \(目標とターゲット\)\)](#)
- [Well-Architected Labs: Resource Type, Size, and Number \(Goals and Targets\) \(Well-Architected ラボ: リソースのタイプ、サイズ、数 \(目標とターゲット\)\)](#)

COST02-BP03 アカウント構造を実装する

組織にマッピングされるアカウントの構造を実装します。これは組織全体でのコストの割り当てと管理に役立ちます。

このベストプラクティスが確立されていない場合のリスクレベル: 高

実装のガイダンス

AWS Organizations では、ワークロードを AWS で拡張する際に環境を一元管理するのに役立つ複数の AWS アカウント を作成できます。組織単位 (OU) 構造で AWS アカウント をグループ化し、各 OU の下に複数の AWS アカウント を作成することで、組織階層をモデル化できます。アカウント構造を作成するには、まず、どの AWS アカウント を管理アカウントにするかを決定する必要があります。その後、[管理アカウントのベストプラクティス](#)と[メンバーアカウントのベストプラクティス](#)に従って、設計されたアカウント構造に基づいて、新しい AWS アカウント を作成したり、既存のアカウントをメンバーアカウントとして選択したりできます。

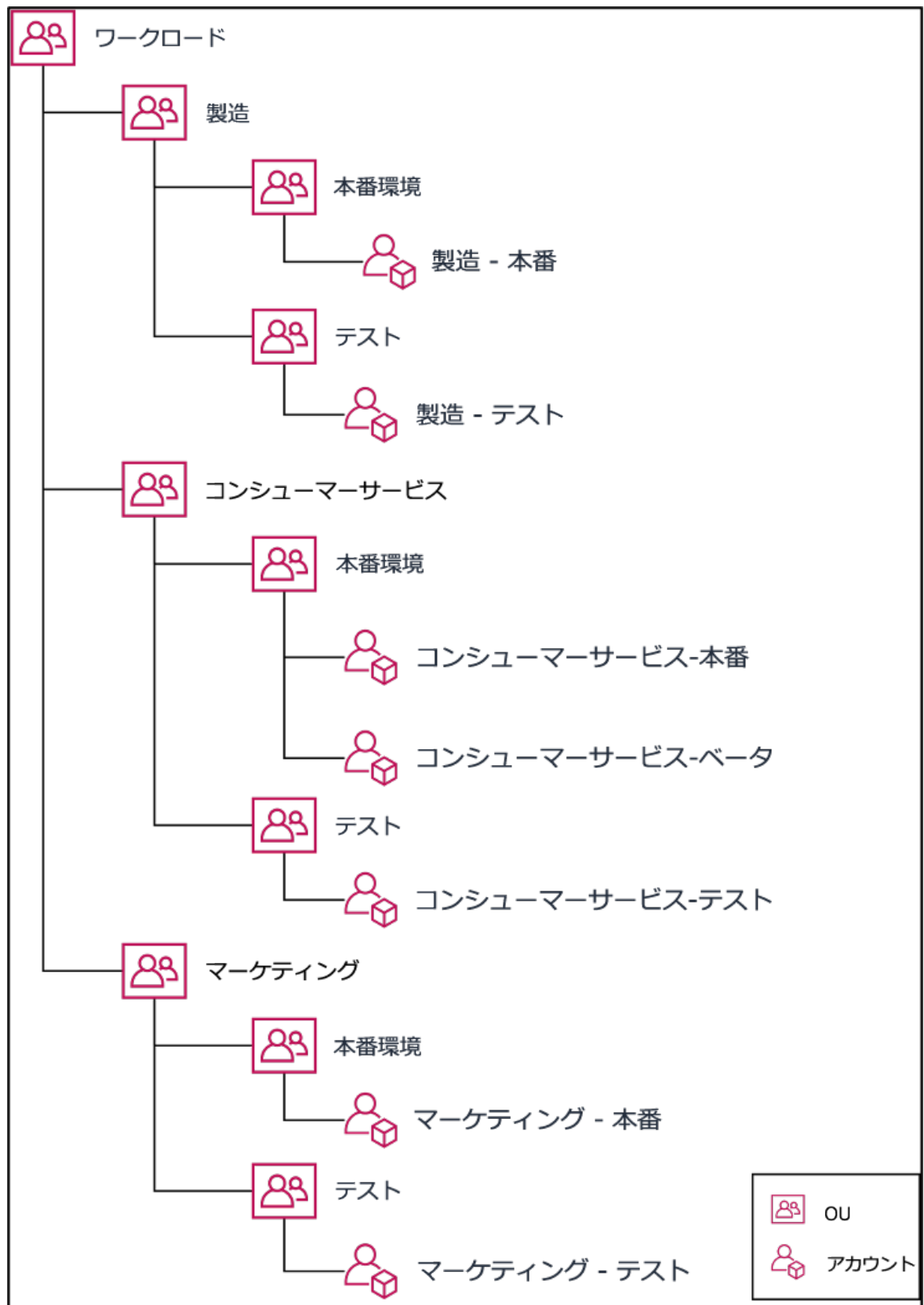
組織の規模や使用状況にかかわらず、少なくとも 1 つの管理アカウントとそれに紐づく 1 つのメンバーアカウントを常に持つことをお勧めします。すべてのワークロードリソースはメンバーアカウント内のみ存在し、管理アカウントにはリソースを作成しないでください。AWS アカウント をいくつ持つべきかということについては、一律に答えることはできません。現在と将来の運用モデルとコストモデルを見積もり、AWS アカウント の構造が組織の目標を反映するようにします。ビジネス上の理由から複数の AWS アカウント を作成する企業もあります。次に例を示します。

- 組織単位、コストセンター、特定のワークロード間で、管理、会計、請求の職務機能を切り離す必要がある場合。
- AWS のサービスの制限が特定のワークロードのみに設定される場合。
- ワークロードとリソース間の隔離と分離には要件があります。

[AWS Organizations](#) 内部では、[一括請求](#)により、1 つ以上のメンバーアカウントと管理アカウントとの間に構造が作成されます。メンバーアカウントを使用すると、コストと使用量をグループ別に区別できます。一般的には、各組織単位 (財務、マーケティング、営業など)、各環境ライフサイクル (開発、テスト、本番など)、各ワークロード (ワークロード a、b、c) にメンバーアカウントをいったん分離したうえで、一括請求 (コンソリデेटッドビルディング) を使用してこれらのアカウントを集約します。

一括請求機能により、複数のメンバー AWS アカウント の支払いを単一の管理アカウントにまとめつつ、リンクされた各アカウントのアクティビティを可視化することができます。コストと使用量が管理アカウントに集計されると、サービスの従量制割引とコミットメント割引 (Savings Plans とリザーブドインスタンス) を最大限に活用し、割引額を最大化できます。

次の図は、AWS Organizations を組織単位 (OU) とともに使用して複数のアカウントをグループ化し、各 OU の下に複数の AWS アカウント を配置する方法を示しています。アカウントを整理するためのパターンを提供するために、さまざまなユースケースやワークロードに OU を使用することをお勧めします。



複数の AWS アカウント を組織単位でグループ化する例。

[AWS Control Tower](#) は、複数の AWS アカウントを迅速に設定、構成し、組織の要件に沿ったガバナンスを確保することができます。

実装手順

- 分離要件を定義する: 分離の要件は、セキュリティ、信頼性、財務構造など、複数の要因の組み合わせです。各要因を順番に確認し、ワークロードまたはワークロード環境を他のワークロードから分離するかどうかを指定します。セキュリティは、アクセス要件とデータ要件への準拠を促進します。信頼性は、環境やワークロードが他の環境に影響を与えないように制限を管理します。Well-Architected フレームワークのセキュリティと信頼性の柱を定期的に見直し、提供されるベストプラクティスに従います。財務構造により、厳格な財務分離 (異なるコストセンター、ワークロードのオーナーシップ、説明責任) が実現します。分離の一般的な例としては、本番稼働用ワークロードとテストワークロードが別々のアカウントで実行されることや、請求書と請求データを、組織内の個々の事業部門や部門、またはアカウントを所有する利害関係者に提供できるように別のアカウントを使用することが挙げられます。
- グループ化要件を定義する: グループ化要件は分離要件を上書きしませんが、管理を支援するために使用されます。分離を必要としない同様の環境またはワークロードをグループ化します。この例としては、1 つ以上のワークロードから複数のテスト環境または開発環境をグループ化することが挙げられます。
- アカウント構造を定義する: これらの分離およびグループ化を使用して、各グループのアカウントを指定し、分離要件を維持します。これらのアカウントは、メンバーアカウントまたは連結アカウントです。これらのメンバーアカウントを単一の管理アカウントまたは支払者アカウントでグループ化することで、使用量が合算されるので、すべてのアカウントでの従量制割引がより大きくなり、すべてのアカウントに対して単一の請求書が提供されます。請求データを分離し、各メンバーアカウントに個別に表示させることも可能です。メンバーアカウントが使用量や請求データを他のアカウントに表示してはならない場合、または AWS から別々の請求書を必要とする場合は、複数の管理アカウントまたは支払者アカウントを定義します。この場合、各メンバーアカウントは独自の管理アカウントまたは支払者アカウントを持つことになります。リソースは常にメンバーアカウントまたは連結アカウントに配置する必要があります。管理アカウントまたは支払者アカウントは、管理のためにのみ使用してください。

リソース

関連するドキュメント:

- [コスト配分タグの使用](#)
- [AWS managed policies for job functions](#) (職務に応じた AWS マネージドポリシー)
- [AWS 複数アカウントの請求戦略](#)
- [Control access to AWS リージョン using IAM policies](#) (IAM ポリシーの使用による AWS Regions へのアクセス制御)
- [AWS Control Tower](#)
- [AWS Organizations](#)
- [管理アカウントとメンバーアカウントに関するベストプラクティス](#)
- [Organizing Your AWS Environment Using Multiple Accounts](#) (複数のアカウントを使用した AWS 環境の組織化)
- [共有リザーブドインスタンスと Savings Plans の割引の有効化](#)
- [一括請求](#)
- [一括請求](#)

関連する例:

- [CUR の分割とアクセスの共有](#)

関連動画:

- [Introducing AWS Organizations](#)(AWS Organizations の紹介)
- [Set Up a Multi-Account AWS Environment that Uses Best Practices for AWS Organizations](#)(AWS Organizations のベストプラクティスを使用するマルチアカウント AWS 環境の設定)

関連する例:

- [Well-Architected ラボ: AWS 組織の作成 \(レベル 100\)](#)
- [Splitting the AWS Cost and Usage Report and Sharing Access](#) (CUR の分割とアクセスの共有)
- [Defining an AWS Multi-Account Strategy for telecommunications companies](#) (通信会社のための AWS マルチアカウント戦略の定義)
- [AWS アカウント の最適化のベストプラクティス](#)
- [AWS Organizations における組織単位のベストプラクティス](#)

COST02-BP04 グループとロールを実装する

ポリシーに沿ったグループおよびロールを実装し、各グループのインスタンスおよびリソースを作成、変更、削除できるユーザーを管理します。たとえば、開発、テスト、本番グループを実装します。これは、AWS のサービスやサードパーティーのソリューションに適用されます。

このベストプラクティスが確立されていない場合のリスクレベル: 低

実装のガイダンス

ユーザーのロールとグループは、安全で効率的なシステムを設計および実装するうえで基本となる構成要素です。ロールとグループは、組織が統制の必要性和柔軟性や生産性の要件とのバランスを図るうえで役立ち、最終的には組織の目標とユーザーのニーズの実現を助けます。「AWS Well-Architected フレームワーク」の「セキュリティの柱」の「[ID とアクセス管理](#)」セクションで推奨されているように、適切なユーザーが適切な条件下で適切なリソースにアクセスできるようにするためには、強固な ID 管理とアクセス許可が必要です。ユーザーには、それぞれの業務を完遂するために必要なアクセス権のみを与えます。そうすることで、不正アクセスや誤用に伴うリスクが最小限に抑えられます。

ポリシーを作成した後で、組織内の論理グループとユーザーロールを作成できます。アクセス許可を割り当て、使用量を制御できるようになり、強固なアクセス制御メカニズムの実装を助け、機密情報への不正アクセスも防止できます。人材のおおまかなグループ化から始めます。通常これは、組織単位と役職 (IT 部門のシステム管理者、会計監査担当者、ビジネスアナリストなど) と合致します。グループによって、類似したタスクに従事し、類似したアクセス権を必要とするユーザーを分類します。ロールとは、グループとして義務付けられた仕事の定義を指します。個々のユーザー単位ではなく、グループやロール単位でアクセス許可を管理する方が簡単です。ロールとグループを通じてユーザー全員に一貫して体系的にアクセス許可を割り当てることで、ミスや不整合を防ぐことができます。

ユーザーのロールが変更された場合、管理者は個々のユーザーアカウントを設定し直さなくても、ロールまたはグループのレベルでアクセス権を調整できます。たとえば、IT のシステム管理者はすべてのリソースを作成するためのアクセスが必要ですが、分析チームのメンバーは分析リソースを作成するアクセスのみで十分です。

実装手順

- グループを実装する: 必要に応じて、組織のポリシーで定義されているユーザーのグループを使用して、対応するグループを実装します。ユーザー、グループ、認証のベストプラクティスについては、「AWS Well-Architected フレームワーク」の「[セキュリティの柱](#)」を参照してください。
- ロールとポリシーを実装する: 組織のポリシーで定義されているアクションを使用して、必要なロールとアクセスポリシーを作成します。ロールとポリシーのベストプラクティスについては、「AWS Well-Architected フレームワーク」の「[セキュリティの柱](#)」を参照してください。

リソース

関連するドキュメント:

- [AWS ジョブ機能の管理ポリシー](#)
- [AWS 複数アカウントの請求戦略](#)
- [AWS Well-Architected フレームワーク - セキュリティの柱](#)
- [AWS Identity and Access Management \(IAM\)](#)
- [AWS Identity and Access Management ポリシー](#)

関連動画:

- [AWS Identity and Access Management \(IAM\)](#)

関連する例:

- [Well-Architected ラボ: 基本的なアイデンティティとアクセス](#)
- [Control access to AWS リージョン using IAM policies](#)
- [Starting your Cloud Financial Management journey: Cloud cost operations](#)

COST02-BP05 コストコントロールを実装する

組織のポリシーと定義済みのグループおよびロールに基づいてコントロールを実装します。これらは、リージョンやリソースタイプへのアクセスコントロールなど、組織の要件によって定義されたコストのみが発生することを保証するものです。

このベストプラクティスが確立されていない場合のリスクレベル: ミディアム

実装のガイドンス

コスト管理を導入する際の一般的な最初のステップは、ポリシー外のコストまたは使用状況イベントが発生した場合に通知するように設定することです。ワークロードや新しいアクティビティを制限したり悪影響を与えたりすることなく、迅速に行動し、是正措置の必要性の有無を確認できます。ワークロードと環境の上限を把握したら、ガバナンスを適用できます。[AWS Budgets](#) では、通知を設定し、AWS コスト、使用状況、コミットメント割引 (Savings Plans とリザーブドインスタンス) の月間予算を定義できます。予算は、集計コストのレベル (たとえば、全コスト)、またはリンクアカウント、サービス、タグ、アベイラビリティゾーンなどの特定のディメンションのみを含む詳細レベルで作成できます。

AWS Budgets で予算限度を設定したら、[AWS Cost Anomaly Detection](#) を使用して予期せぬコストを削減します。AWS Cost Anomaly Detection は、機械学習を使用してコストと使用状況を継続的に監視し、異常な支出を検出するコスト管理サービスです。これにより、異常な支出と根本原因を特定するため、迅速に対策を講じることができます。まず、AWS Cost Anomaly Detection でコストモニターを作成し、ドルのしきい値 (影響が 1,000 USD を超える異常に対してアラートを出すなど) を設定し、アラートの設定を選択します。アラートを受信したら、異常の背後にある根本原因とコストへの影響を分析できます。また、AWS Cost Explorer では独自の異常解析を監視および実行することもできます。

AWS で [AWS Identity and Access Management](#) と [AWS Organizations サービスコントロールポリシー \(SCP\)](#) をととしてガバナンスポリシーを適用します。IAM により、AWS サービスやリソースへのアクセスを安全に管理できます。IAM を使用すると、AWS のリソースを作成または管理できるユーザー、作成できるリソースのタイプ、リソースを作成できる場所を制御できます。これにより、定義されたポリシーの範囲を超えてリソースが作成される可能性が最小限に抑えられます。以前に作成したロールとグループを使用して [IAM ポリシー](#) を割り当て、正しい使用を強制します。SCP は、組織内のすべてのアカウントで利用可能な最大権限を一元管理し、アカウントをアクセス制御ガイドラインの範囲内に維持することができます。SCP はすべての機能が有効になっている組織でのみ使用可能で、デフォルトで SCP によるメンバー

アカウントのアクションの可否を設定できます。アクセス管理の導入の詳細については、[Well-Architected セキュリティの柱に関するホワイトペーパー](#)を参照してください。

[AWSService Quotas](#) を管理することで、ガバナンスを導入することもできます。Service Quotas を最小限のオーバーヘッドで設定し、正確に維持することで、組織の要件以外のリソースの作成を最小限に抑えることができます。これを実現するには、要件がどれだけ速く変化するかを理解し、進行中のプロジェクト (リソースの作成と削除の両方) を理解し、クォータ変更をどれだけすばやく実装できるかを考慮する必要があります。[Service Quotas](#) を使用すると、必要に応じてクォータを引き上げることができます。

実装手順

- 支出に関する通知を実装する: 定義した組織のポリシーを使用して、[AWS Budgets](#) を作成し、ポリシーから外れた支出があった場合に通知します。アカウントごとに複数のコスト予算を設定し、アカウント全体の支出を通知します。アカウント内のより小さな単位について、各アカウント内にコスト予算を追加で設定します。これらの単位は、アカウント構造によって異なります。一般的な例としては、AWS リージョン、ワークロード (タグを使用)、または AWS のサービスがあります。個人の E メールアカウントではなく、E メール配信リストを通知の受信者として設定します。金額を超えたときの実際の予算を設定するか、予測された使用量が通知されたときの予測された予算を使用します。特定の IAM または SCP のポリシーを適用したり、ターゲットの Amazon EC2 または Amazon RDS インスタンスを停止したりできる AWS Budget アクションを事前設定することもできます。予算に関するアクションは、自動的に実行するか、ワークフローの承認を得るようにすることができます。
- 異常な支出に関する通知を実装する: [AWS Cost Anomaly Detection](#) を使用して、組織内の思いがけないコストを削減し、異常な支出の可能性の根本原因を分析します。指定した粒度で異常な支出を特定するコストモニターを作成し、AWS Cost Anomaly Detection で通知を設定すると、異常な支出が検出された際にアラートが送信されます。これにより、異常の背後にある根本的な原因を分析し、コストへの影響を理解できます。AWS Cost Anomaly Detection を設定する際に AWS Cost Categories を使用して、予想外のコストの根本原因を分析し、必要なアクションをタイムリーに実行できるプロジェクトチームまたはビジネスユニットチームを特定します。
- 使用量のコントロールを実装する: 定義した組織のポリシーを使用して、IAM ポリシーとロールを実装し、ユーザーが実行できるアクションと実行できないアクションを指定します。AWS ポリシーには、複数の組織ポリシーを含めることができます。ポリシーを定義するのと同じ方法で、幅広く開始し、各ステップでより詳細なコントロールを適用します。サービスの制限も、使用量に対する効果的なコントロールです。すべてのアカウントに正しいサービス制限を実装します。

リソース

関連するドキュメント:

- [AWS managed policies for job functions](#) (職務に応じた AWS マネージドポリシー)
- [AWS 複数アカウントの請求戦略](#)
- [Control access to AWS リージョン using IAM policies](#) (IAM ポリシーの使用による AWS Regions へのアクセス制御)
- [AWS Budgets](#)
- [AWS Cost Anomaly Detection](#)
- [AWS コストを管理する](#)

関連動画:

- [How can I use AWS Budgets to track my spending and usage](#) (AWS Budgets を使用して支出と使用量を追跡するにはどうすればよいですか?)

関連する例:

- [Example IAM access management policies](#) (IAM アクセス管理ポリシーの例)

- [サービスコントロールポリシーの例](#)
- [AWS Budgets アクション](#)
- [タグを使って Amazon EC2 リソースへのアクセスを管理する IAM ポリシーを作成する](#)
- [IAM アイデンティティのアクセスを特定の Amazon EC2 リソースに制限する](#)
- [Create an IAM Policy to restrict Amazon EC2 usage by family](#) (家族ごとに EC2 の利用を制限する IAM ポリシーを作成する)
- [Well-Architected ラボ: コストと使用量のガバナンス \(レベル 100\)](#)
- [Well-Architected ラボ: コストと使用量のガバナンス \(レベル 200\)](#)
- [Slack integrations for Cost Anomaly Detection using AWS Chatbot](#) (AWS Chatbot を使用した AWS Cost Anomaly Detection のための Slack の統合)

COST02-BP06 プロジェクトのライフサイクルを追跡する

プロジェクト、チーム、環境のライフサイクルを追跡、計測、監査して、不要なリソースの使用やそれに伴う支払いを回避できます。

このベストプラクティスが確立されていない場合のリスクレベル: 低

実装のガイダンス

プロジェクトのライフサイクルを効果的に追跡することで、組織はリソース、時間、品質の計画、管理、最適化を改善し、コスト管理を強化できます。追跡で得られたインサイトは、意思決定に役立つ貴重な情報となり、コスト効率性やプロジェクト全体の成功に寄与します。

ワークロードのライフサイクル全体を追跡すれば、ワークロードやワークロードコンポーネントが不要になった時点でわかります。既存のワークロードとコンポーネントが使用中のように見える場合がありますが、AWS が新しいサービスや機能をリリースした時点で、廃止または刷新される可能性があります。ワークロードの以前のステージに注目してください。ワークロードが本番稼働状態になったら、以前の環境は廃止するか、再び必要になるまでキャパシティを大幅に削減することができます。

AWS には、エンティティのライフサイクル追跡に使用できる管理およびガバナンスサービスが多数用意されています。[AWS Config](#) や [AWS Systems Manager](#) を使用して、AWS のリソースや設定の詳細なインベントリを入手できます。プロジェクトやアセットを管理する既存のシステムを統合して、組織内のアクティブなプロジェクトや製品を追跡することを推奨します。現在のシステムを AWS が提供する豊富なイベントやメトリクスと組み合わせることにより、重要なライフサイクルイベントのビューを作成し、前もってリソースを管理し、不要なコストを削減できます。

[アプリケーションライフサイクル管理 \(ALM\)](#) と同様、プロジェクトライフサイクルの追跡でも、設計と開発、テスト、生産、サポート、ワークロードの冗長性など、複数のプロセス、ツール、チームが連携する必要があります。

プロジェクトのライフサイクルの各段階を注意深く監視することで、組織は重要なインサイトを得て管理を強化し、プロジェクトを計画から実施、完遂に至るまで円滑に進めることができます。入念な監視下で、プロジェクトは品質基準を満たすばかりか、納期通りに予算内で完了し、全体的なコスト効率向上します。

エンティティのライフサイクル追跡実装の詳細については、[AWS Well-Architected フレームワークの運用上の優秀性の柱のホワイトペーパー](#)を参照してください。

実装手順

- プロジェクトのライフサイクル監視プロセスを確立する: [Cloud Center of Excellence チーム](#)は、プロジェクトのライフサイクル監視プロセスを確立する必要があります。ワークロードを監視するための

構造的かつ体系的なアプローチを確立し、プロジェクトの管理、可視性、パフォーマンスを向上させます。モニタリングプロセスの効果と価値を最大限に引き出すために、プロセスの透明性と協調性を高め、継続的に改善していきます。

- ワークロードのレビューを実行する: 組織のポリシーで定められている通りに、既存のプロジェクトを定期的に監査し、ワークロードのレビューを実施します。監査に費やされる労力の量は、組織のおおよそのリスク、価値、またはコストに比例する必要があります。監査に含めるべき主な領域は、インシデントまたは機能停止の組織に対するリスク、価値、組織への寄与 (収益またはブランドに対する評価で測定)、ワークロードのコスト (リソースおよび運用の合計コストとして測定)、およびワークロードの使用量 (時間単位ごとの組織の成果の数で測定) です。これらの領域がライフサイクルを通じて変化する場合、完全または部分的な削除など、ワークロードの調整が必要です。

リソース

関連するドキュメント:

- [Guidance for Tagging on AWS](#)
- [ALM \(アプリケーションライフサイクル管理\) とは何ですか?](#)
- [AWS ジョブ機能の 管理ポリシー](#)

関連する例:

- [Control access to AWS リージョン using IAM policies](#)

関連ツール:

- [AWS Config](#)
- [AWS Systems Manager](#)
- [AWS Budgets](#)
- [AWS Organizations](#)
- [AWS CloudFormation](#)

コストと使用量のモニタリング

ワークロードに詳細な可視化を導入し、チームがコストと使用量に対しアクションを実行できるようにします。コストの最適化は、コストと使用状況の内訳、将来の支出、使用状況、機能をモデル化して予測する機能、コストと使用量を組織の目標に合わせて調整するためのメカニズムの実装をきめ細かく理解することから始まります。コストと使用量をモニタリングするために必要な領域は次のとおりです。

ベストプラクティス

- [COST03-BP01 詳細情報ソースを設定する \(p. 36\)](#)
- [COST03-BP02 コストと使用状況に組織情報を追加する \(p. 37\)](#)
- [COST03-BP03 コスト属性カテゴリを特定する \(p. 38\)](#)
- [COST03-BP04 組織のメトリクスを確立する \(p. 41\)](#)
- [COST03-BP05 請求およびコスト管理ツールを設定する \(p. 42\)](#)
- [COST03-BP06 ワークロードメトリクスに基づいてコストを配分する \(p. 44\)](#)

COST03-BP01 詳細情報ソースを設定する

コスト管理とレポートツールで時間単位の粒度を設定して、コストと使用状況に関する詳細を提供することで、より深い分析と透明性が可能になります。ワークロードが、もたらされるすべてのビジネス成果のログエントリを生成するか持つように設定します。

このベストプラクティスが確立されていない場合のリスクレベル: 高

実装のガイダンス

時間単位の粒度など、コスト管理ツールの詳細な請求情報により、組織は消費量をさらに詳細に追跡でき、コスト増加の原因を特定する手助けとなります。これらのデータソースは、組織全体のコストと使用量の最も正確なビューを提供します。

AWS Cost and Usage Report では、課金されるすべての AWS のサービスについて、日単位または時間単位の使用量の粒度、料金、コスト、使用属性が提供されます。CUR には、タグ付け、場所、リソース属性、アカウント ID など想定可能なすべてのディメンションがあります。

以下のカスタマイズ項目で CUR を設定します。

- リソース ID を含める
- CUR を自動更新する
- 時間単位の詳細
- バージョニング: 既存のレポートを上書きする
- データ統合: Athena (Parquet 形式、圧縮)

[AWS Glue](#) を使用して分析用データを準備し、[Amazon Athena](#) を使用してデータ分析を実行し、SQL を使用してデータをクエリします。[Amazon QuickSight](#) を使用して、カスタムの可視化や複雑な可視化を行い、組織全体に配布することもできます。

実装手順

- コストと使用状況レポートを設定する: 請求コンソールを使用して、少なくとも 1 つのコストと使用状況レポートを設定します。すべての識別子とリソース ID を含む時間単位の粒度でレポートを設定します。粒度が異なる他のレポートを作成して、よりハイレベルな概要情報を提供することもできます。
- Cost Explorer で時間単位の粒度を設定する: 請求コンソールを使用して、[時間単位とリソースレベルのデータ] を有効にすると、時間単位の粒度およびリソースレベルの粒度で過去 14 日間のコストと使用状況のデータにアクセスできます。
- アプリケーションログ記録を設定する: アプリケーションがもたらすビジネスの各成果がログに記録され、追跡および測定が可能であることを確認します。このデータの粒度が少なくとも 1 時間単位であることを確認し、コストと使用状況のデータと一致するようにします。ログ記録とモニタリングの詳細については、[Well-Architected 運用上の優秀性の柱](#)を参照してください。

リソース

関連するドキュメント:

- [AWS Cost and Usage Report](#)
- [AWS Glue](#)
- [Amazon QuickSight](#)
- [AWS コスト管理の料金](#)
- [AWS リソースのタグ付け](#)

- [AWS Budgets を用いてコストを分析する](#)
- [Cost Explorer を用いてコストを分析する](#)
- [AWS Cost and Usage Report の管理](#)
- [Well-Architected 運用上の優秀性の柱](#)

関連する例:

- [AWS Account Setup](#)
- [AWS Cost Explorer's New Look and Common Use Cases](#)

COST03-BP02 コストと使用状況に組織情報を追加する

組織、ワークロード属性、およびコスト配分カテゴリに基づいてタグ付けスキーマを定義します。これによりコスト管理ツールで、フィルター処理によるリソースの検索や、コストおよび使用状況のモニタリングを行うことができます。目的、チーム、環境、またはビジネスに関連するその他の基準によって、可能な限りすべてのリソースに一貫したタグ付けを実装します。

このベストプラクティスが確立されていない場合のリスクレベル: ミディアム

実装のガイダンス

[AWS](#) でタグ付けを実装することにより、リソースに組織の情報を追加します。追加した情報は、コストと使用状況の情報に追加されます。タグはキーと値のペアです。キーは組織全体で一意になるように定義されている必要があります。値はリソースのグループに対して一意になります。キーと値のペアとして、例えばキーを Environment、値を Production にすることができます。本稼働環境のすべてのリソースには、キーと値のペアがあります。タグ付けにより、関連性の高い組織情報を使用して、コストを分類、追跡できます。組織のカテゴリ (コストセンター、アプリケーション名、プロジェクト、オーナーなど) を表すタグを適用し、ワークロードやワークロードの特性 (テストや本番など) を識別して、組織全体のコストと使用状況の帰属先を付与できます。

AWS リソース (Amazon Elastic Compute Cloud インスタンスや Amazon Simple Storage Service バケットなど) にタグを付け、そのタグをアクティブ化すると、AWS はこの情報をコストと使用状況レポートに追加します。タグ付けされたリソースとタグ付けされていないリソースに関するレポート作成および分析を実行することで、社内のコスト管理ポリシーへの準拠を強化し、正確に帰属を特定できます。

組織のアカウント全体に AWS タグ付け基準を作成、導入することで、AWS 環境を一貫性のある統一された方法で管理することができます。AWS Organizations で、アカウントの AWS リソースに対してタグを使用する方法のルールを定義するには、AWS Organizations の [タグポリシー](#) を使用します。タグポリシーを使用すると、AWS リソースにタグを付ける標準アプローチを簡単に導入できます。

[AWS タグエディタ](#) では、複数のリソースのタグを追加、削除、管理できます。タグエディタを使用して、タグ付けするリソースを検索し、検索結果でリソースのタグを管理できます。

[AWS Cost Categories](#) を使用すると、リソースにタグを付けることなく組織としての意味をコストに割り当てることができます。コストと使用量に関する情報を、一意の内部組織構造にマッピングできます。アカウントやタグなどの請求ディメンションを使用して、コストをマッピングおよび分類するカテゴリルールの定義します。これにより、タグ付けに加えて、別のレベルの管理機能が提供されます。また、特定のアカウントとタグを複数のプロジェクトにマッピングすることもできます。

実装手順

- タグ付けスキーマを定義する: ビジネス全体からすべての関係者を集めて、スキーマを定義します。これには通常、技術、財務、および管理ロールの担当者が含まれます。すべてのリソースに必要なタグのリ

ストと、リソースに必要なタグのリストを定義します。タグの名前と値が組織全体で一貫していることを確認します。

- リソースにタグを付ける: 定義したコスト帰属カテゴリを使用し、カテゴリに従ってワークロードのすべてのリソースに[タグを付けます](#)。効率を向上させるには、CLI、タグエディタ、AWS Systems Managerなどのツールを使用します。
- AWS Cost Categories を実装する: タグ付けを実装せずに[コストカテゴリ](#)を作成できます。Cost Categories では、既存のコストと使用量ディメンションを使用します。スキーマからカテゴリルールを作成し、それをコストカテゴリに実装します。
- タグ付けを自動化する: すべてのリソースにわたって大局的なタグ付けを継続するには、タグ付けを自動化して、リソースの作成時にタグ付けが自動的に行われるようにします。リソースの作成時にタグ付けを行うには、[AWS CloudFormation](#)などのサービスを使用します。Lambda 関数を使用して[タグ付けを自動的に行う](#)カスタムソリューションを作成することも、ワークロードを定期的にスキャンして、タグ付けされていないリソースを削除するマイクロサービスを使用することもできます (これは、テスト環境および開発環境に最適です)。
- タグ付けのモニタリングとレポート作成を行う: 組織全体で大局的なタグ付けを継続するには、ワークロード全体のタグについて、レポート作成およびモニタリングを行います。[AWS Cost Explorer](#)を使用して、タグ付けされたリソースとタグ付けされていないリソースのコストを確認することも、[タグエディタ](#)などのサービスを利用することもできます。タグ付けされていないリソースの数を定期的に確認し、必要なレベルのタグ付けになるまでタグを追加するアクションを実行します。

リソース

関連するドキュメント:

- [タグ付けに関するベストプラクティス](#)
- [AWS CloudFormation のリソースタグ](#)
- [AWS Cost Categories](#)
- [AWS リソースのタグ付け](#)
- [AWS Budgets を用いてコストを分析する](#)
- [Cost Explorer を用いてコストを分析する](#)
- [AWS コストと使用状況レポートの管理](#)

関連動画:

- [How can I tag my AWS resources to divide up my bill by cost center or project](#) (AWS リソースにタグを付けて、コストセンターまたはプロジェクト別に請求書を分割する方法)
- [AWS リソースのタグ付け](#)

関連する例:

- [Automatically tag new AWS resources based on identity or role](#) (アイデンティティまたはロールに基づいて新しい AWS リソースへのタグ付けを自動的に行う)

COST03-BP03 コスト属性カテゴリを特定する

組織内のコストを内部消費エンティティに配分するために使用できるビジネスユニット、部門、プロジェクトなどの組織カテゴリを特定します。こうしたカテゴリを活用して、支出の説明責任の徹底、コスト意識の向上、効果的な消費行動の促進を図ります。

このベストプラクティスが確立されていない場合のリスクレベル: 高

実装のガイダンス

コストを分類するプロセスは、予算編成、会計、財務報告、意思決定、ベンチマーキング、およびプロジェクト管理においてきわめて重要です。費用を分類してカテゴリ化することで、チームはクラウドジャーニーで発生するコストの種類をよりよく理解でき、情報に基づいた意思決定を行い、予算を効果的に管理できるようになります。

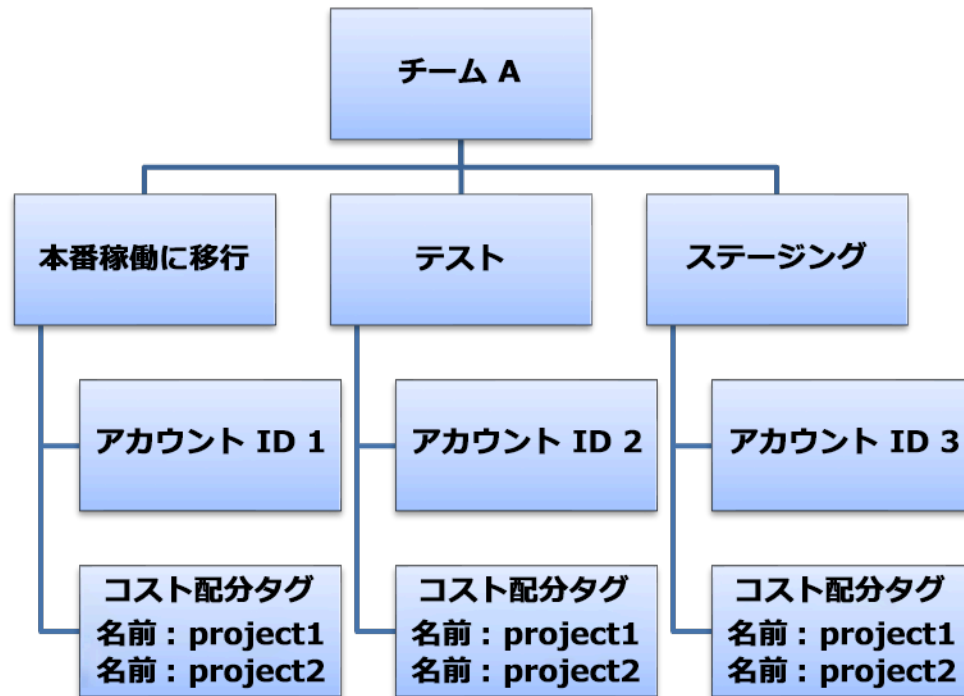
クラウド支出の説明責任は、統制の取れた需要とコスト管理に対する強力なインセンティブを確立します。その結果、クラウド支出の大部分を消費するビジネスユニットやチームに割り当てている組織では、クラウドコストを大幅に節約できます。また、クラウド支出を配分することで、組織は一元化されたクラウドガバナンスのベストプラクティスをさらに採用できるようになります。

定期的なミーティングで、財務チームやその他の関係者と協力し、組織内でコストを配分する方法の要件を理解します。ワークロードのコストは、開発、テスト、本稼働、廃止などライフサイクル全体にわたって配分する必要があります。学習、スタッフ育成、アイデア創出に要したコストが、どのように組織に帰属するかを理解します。これは、この目的で使用される金額を、一般的な IT コスト予算ではなく、トレーニング予算や開発の予算に正しく割り当てらるうで役立ちます。

組織内の関係者と共にコスト属性カテゴリを定義した後、[AWS Cost Categories](#) を使用して、コストと使用状況の情報を AWS クラウド での意味を持つカテゴリ (特定のプロジェクトのコストや部門またはビジネスユニットの AWS アカウント など) にグループ化します。カスタムカテゴリを作成して、アカウント、タグ、サービス、料金タイプなどのさまざまなディメンションを使用して定義したルールに基づき、コストと使用状況の情報をカスタムカテゴリにマッピングすることもできます。コストカテゴリを設定すると、カテゴリごとにコストと使用状況の情報を確認できるようになり、組織の戦略や購入に関する決定をより適切に行うことができます。これらのカテゴリは、AWS Cost Explorer、AWS Budgets、および AWS Cost and Usage Report にも表示されます。

例えば、ビジネスユニット (DevOps チーム) のコストカテゴリを作成し、各カテゴリの下に、複数のルール (各サブカテゴリのルール) を作成します。各ルールでは、定義したグループに基づいて、複数のディメンション (AWS アカウント、コスト配分タグ、サービス、料金タイプ) を使用します。コストカテゴリを使用することにより、ルールベースのエンジンを使用してコストを整理できます。ルールを設定することによって、コストがカテゴリに分類されます。ルール内では、特定の AWS アカウント、AWS サービス、料金タイプなどの各カテゴリについて、複数のディメンションを使用してフィルター処理を行うことができます。これらのカテゴリは、[AWS Billing and Cost Management and Cost Management コンソール](#)内で複数の製品にわたって使用できます。これには、AWS Cost Explorer、AWS Budgets、AWS Cost and Usage Report、AWS Cost Anomaly Detection が含まれます。

例として、次の図は、組織でコストと使用状況の情報をグループ化する方法を示しています。例えば、複数のチーム (コストカテゴリ)、複数の環境 (ルール)、そして複数のリソースまたはアセットを持つ各環境 (ディメンション) にグループができます。



コストと使用状況の組織図

コストカテゴリを使用して、コストのグループを作成することもできます。作成したコストカテゴリは、[AWS Cost Explorer](#)、[AWS Budgets](#)、[AWS Cost and Usage Report](#)、および [AWS Cost Anomaly Detection](#) に表示されます (コストカテゴリの作成後、使用状況レコードの値が更新されるまでに最長で 24 時間がかかります)。AWS Cost Explorer および AWS Budgets では、コストカテゴリが追加の請求ディメンションとして表示されます。これを使用すると、特定のコストカテゴリ値をフィルター処理することも、コストカテゴリ別にグループ化することもできます。

実装手順

- 組織カテゴリを定義する: 社内の関係者およびビジネスユニットとミーティングを行い、組織の構造と要件を反映したカテゴリを定義します。これらのカテゴリは、ビジネスユニット、予算、コストセンター、部門など、既存の財務カテゴリの構造に直接マッピングされます。トレーニングや教育など、クラウドがもたらすビジネスの成果を確認します。これらは組織のカテゴリでもあります。
- 機能カテゴリを定義する: 社内の関係者およびビジネスユニットとミーティングを行い、企業内の機能を反映したカテゴリを定義します。これは、ワークロードまたはアプリケーション名、および実稼働、テスト、開発などの環境のタイプである場合があります。
- AWS Cost Categories を定義する: コストカテゴリを作成し、[AWS Cost Categories](#) を使用してコストと使用状況の情報を整理し、AWS のコストと使用状況を [意味のあるカテゴリ](#) にマッピングします。同じリソースに複数のカテゴリを割り当てることも、同じリソースを複数の異なるカテゴリに含めることもできるため、必要な数のカテゴリを定義します。これにより、AWS Cost Categories を使用して、カテゴリが適用された構造内で [コストを管理](#) できます。

リソース

関連するドキュメント:

- [AWS リソースのタグ付け](#)
- [コスト配分タグの使用](#)

- [AWS Budgets を用いてコストを分析する](#)
- [Cost Explorer を用いてコストを分析する](#)
- [AWS Cost and Usage Report の管理](#)
- [AWS Cost Categories](#)
- [Managing your costs with AWS Cost Categories](#)
- [コストカテゴリを作成する](#)
- [コストカテゴリのタグ付け](#)
- [コストカテゴリ内で料金を分割する](#)
- [AWS Cost Categories Features](#)

関連する例:

- [Organize your cost and usage data with AWS Cost Categories](#)
- [Managing your costs with AWS Cost Categories](#)
- [Well-Architected ラボ: コストと使用状況の可視化](#)
- [Well-Architected Labs: Cost Categories](#)

COST03-BP04 組織のメトリクスを確立する

このワークロード用のメトリクスを組織内で定めます。ワークロードのメトリクスの例として、作成された顧客レポートや顧客に提供されるウェブページが挙げられます。

このベストプラクティスが確立されていない場合のリスクレベル: 高

実装のガイダンス

ワークロードのアウトプットがビジネスの成功に対してどのように測定されるかを理解します。通常、各ワークロードには、パフォーマンスを示す主な成果の小さな組み合わせがあります。多数のコンポーネントを含む高度なワークロードがある場合は、リストに優先順位を付けるか、各コンポーネントのメトリクスを定義して追跡できます。チームと協力して、どのメトリクスを使用するか理解します。この単位は、ワークロードの効率または各ビジネス成果のコストを把握するために使用されます。

実装手順

- ワークロードの成果を定義する: 社内の関係者との会議により、ワークロードの成果を定義します。これらは顧客の使用状況の主要な測定指標であり、技術的メトリクスではなく、ビジネスメトリクスである必要があります。ワークロードごとに少数の概要的なメトリクス (5 つ未満) が存在する必要があります。ワークロードが異なるユースケースで複数の成果を生成する場合は、それらを単一のメトリクスにグループ化してください。
- ワークロードコンポーネントの成果を定義する: 大規模で複雑なワークロードがある場合、または明確に定義された入出力を使用してワークロードをコンポーネント (マイクロサービスなど) に簡単に分割できる場合は、必要に応じて、各コンポーネントのメトリクスを定義します。この作業では、コンポーネントの価値とコストを反映する必要があります。最大のコンポーネントから開始し、大きさの順番で、最小のコンポーネントまで作業します。

リソース

関連するドキュメント:

- [AWS リソースのタグ付け](#)
- [AWS Budgets を用いてコストを分析する](#)

- [Cost Explorer を用いてコストを分析する](#)
- [AWS コストと使用状況レポートの管理](#)

COST03-BP05 請求およびコスト管理ツールを設定する

クラウド支出を管理および最適化するには、組織のポリシーに沿ってコスト管理ツールを設定します。これに含まれるサービス、ツール、リソースを使用すると、コストと使用状況のデータを整理して追跡し、統合された請求とアクセス許可で制御を強化できます。また、予算編成と予測を通じた計画の改善、通知またはアラートの受信、リソースと価格の最適化によるさらなるコスト削減を行うこともできます。

このベストプラクティスが確立されていない場合のリスクレベル: 高

実装のガイダンス

確固たる説明責任を確立するには、まずアカウント戦略をコスト配分戦略の一部として検討する必要があります。これを正しく行えば、それ以上先に進む必要はないかもしれませんが、正しく行えない場合、認識の欠如が発生し、さらに問題点が増える可能性があります。

クラウド支出の説明責任を高めるには、ユーザーがコストと使用状況を可視化するツールにアクセスできるようにする必要があります。すべてのワークロードとチームに、以下の詳細と目的に合わせてツールを設定することをお勧めします。

- 整理: 独自のタグ付け戦略と分類を使用して、コスト配分とガバナンスのベースラインを確立します。
- 整理: 独自のタグ付け戦略と分類を使用して、コスト配分とガバナンスのベースラインを確立します。サポートされている AWS リソースにタグを付け、組織構造 (ビジネスユニット、部門、プロジェクト) に基づいてわかりやすく分類します。特定のコストセンターのアカウント名にタグ付けし、それを AWS Cost Categories とマッピングして、特定のビジネスユニットのアカウントをコストセンターのグループにまとめることで、ビジネスユニットの所有者が複数のアカウントの消費を 1 か所で確認できるようにします。
- アクセス: 組織全体の請求情報を [一括請求](#) で追跡し、適切な関係者とビジネスオーナーがアクセスできることを確認します。
- 制御: 適切なガードレールを使用して、効果的なガバナンスメカニズムを構築し、サービスコントロールポリシー (SCP)、タグポリシー、予算アラートを使用する際の想定外のシナリオを回避します。例えば、効果的な制御メカニズムがあれば、チームが推奨リージョンでのみリソースを作成できるようになります。
- 現在の状態: 現在のコストと使用量のレベルを示すダッシュボードを設定します。ダッシュボードはオペレーションダッシュボードと同様に、作業環境内の目に付きやすい場所で使用できるようにする必要があります。 [Cloud Intelligence Dashboard \(CID\)](#) またはその他のサポート対象製品を使用して、このような可視性を実現します。
- 通知: コストまたは使用量が定義された制限を超えた場合、および AWS Budgets または AWS Cost Anomaly Detection で異常が発生した場合に通知します。
- レポート: すべてのコストと使用状況の情報を要約し、詳細で割り当て可能なコストデータを使用して、クラウド支出の認識と説明責任の意識を高めます。レポートは、それを利用するチームと関連性があり、推奨事項を含めるのが理想的です。
- 追跡: 設定された目標またはターゲットに対する現在のコストと使用量を示します。
- 分析: チームメンバーが、すべての可能なディメンションを使用して、時間単位の詳細度までカスタムおよび詳細な分析を実行できるようにします。
- 調査: リソースのデプロイとコストの最適化を行う機会について最新情報を入手します。組織レベルでのリソースデプロイに関する通知 (Amazon CloudWatch、Amazon SNS、または Amazon SES を使用) を受け取り、コスト最適化の推奨事項 (AWS Compute Optimizer、AWS Trusted Advisor など) を確認します。
- 傾向: 指定の期間におけるコストと使用量の変動を、指定の詳細度で示します。

- 予測: 作成した予測ダッシュボードで、将来の推定コストを示し、リソースの使用量と支出を見積もります。

[AWS Cost Explorer](#)、[AWS Billing and Cost Management](#)、[AWS Budgets](#) などの AWS ツールを必須ツールとして使用するか、CUR データを [Amazon Athena](#) および [Amazon QuickSight](#) と統合して、より詳細なビューにこの機能を提供することもできます。組織に必須のスキルや処理能力がない場合、[AWS ProServ](#)、[AWS Managed Services \(AMS\)](#)、[AWS Partner](#) と連携して、こうした組織のツールを使用することができます。サードパーティ製ツールを使用することもできますが、組織にとって、そのツールの価値がコストに見合っているかどうかを確認する必要があります。

実装手順

- ツールに対するチームベースのアクセスを許可する: アカウントを設定してグループを作成し、必要なコストと使用状況レポート (グループの使用状況に関するもの) へのアクセスを許可します。また、[AWS Identity and Access Management](#) を使用して AWS Cost Explorer などのツールへの [アクセスを制御](#) します。これらのグループには、アプリケーションを所有または管理するすべてのチームの代表者を含める必要があります。これにより、すべてのチームがコストと使用状況の情報にアクセスして、使用を追跡できます。
- AWS Budgets を設定する: ワークロードのすべてのアカウントで [AWS Budgets](#) を設定します。タグを使用して、アカウント全体の支出に対する予算とワークロードに対する予算を設定します。AWS Budgets で通知を設定すると、予算額を超えたときや、推定コストが予算を超えているときにアラートを受信できます。
- AWS Cost Explorer を設定する: ワークロードとアカウントについて [AWS Cost Explorer](#) を設定し、さらに分析を行うためにコストデータを視覚化します。ワークロードのダッシュボードを作成することにより、全体的な支出、ワークロードの主要な使用状況メトリクス、過去のコストデータに基づく将来のコストの予測を追跡できます。
- AWS Cost Anomaly Detection を設定する: アカウント、コアサービス、または作成したコストカテゴリについて [AWS Cost Anomaly Detection](#) を使用することにより、コストと使用状況を監視し、通常と異なる支出を検出できます。集計レポートでアラートを個別に受信したり、E メールまたは Amazon SNS トピックでアラートを受信したりすることで、異常の根本原因を分析および特定し、コストの増加を引き起こしている要因を特定できます。
- 高度なツールを設定する: 必要に応じて、さらなる詳細と粒度を提供する組織用のカスタムツールを作成できます。[Amazon Athena](#) を使用して高度な分析機能を実装し、[Amazon QuickSight](#) を使用してダッシュボードを実装できます。事前設定済みの高度なダッシュボードを備えた [CID ソリューション](#) の使用を検討してください。クラウド支出の監視と最適化を 1 か所で簡単に実現するには、[AWS Partner](#) との連携により、クラウド管理ソリューションを導入することもできます。

リソース

関連するドキュメント:

- [AWS コスト管理](#)
- [AWS リソースのタグ付け](#)
- [AWS Budgets を用いてコストを分析する](#)
- [Cost Explorer を用いてコストを分析する](#)
- [AWS Cost and Usage Report の管理](#)
- [AWS Cost Categories](#)
- [AWS によるクラウド財務管理](#)
- [サービスコントロールポリシーの例](#)
- [AWS APN パートナー - コスト管理](#)

関連動画:

- [Deploying Cloud Intelligence Dashboards \(Cloud Intelligence Dashboards のデプロイ\)](#)
- [Get Alerts on any FinOps or Cost Optimization Metric or KPI \(FinOps またはコスト最適化のメトリクスまたは KPI に関するアラートを受け取る\)](#)

関連する例:

- [Well-Architected ラボ - AWS アカウント設定](#)
- [Well-Architected ラボ: 請求の可視化](#)
- [Well-Architected ラボ: コストと使用量のガバナンス](#)
- [Well-Architected ラボ: コストと使用状況の分析](#)
- [Well-Architected ラボ: コストと使用状況の可視化](#)
- [Well-Architected ラボ: Cloud Intelligence Dashboards](#)
- [サービスコントロールポリシーを使用して、複数アカウントのアクセス許可のガードレールを設定する方法](#)

COST03-BP06 ワークロードメトリクスに基づいてコストを配分する

使用量メトリクスや業績に基づいてワークロードのコストを配分し、ワークロードのコスト効率を測定します。インサイトとチャージバック機能が利用できる分析サービスにより、コストと使用状況データを分析するプロセスを実装します。

このベストプラクティスを活用しない場合のリスクレベル: 低

実装のガイダンス

コスト最適化によって、最低価格でビジネス成果が提供されます。これは、ワークロードメトリクス (ワークロードの効率で測定) ごとにワークロードのコストを配分することによってのみ達成できます。定義されたワークロードメトリクスを、ログファイルまたは他のアプリケーションのモニタリングを使用してモニタリングします。このデータをワークロードのコストと組み合わせます。ワークロードのコストは、特定のタグ値またはアカウント ID でコストを確認することで取得できます。この分析は時間単位で実行することをお勧めします。リクエストレートが変化する静的なコストコンポーネント (恒久的に実行されるバックエンドデータベースなど) がある場合、通常、効率性は変化します (例えば、使用量のピークは午前 9 時から午後 5 時で、夜間のリクエストはほとんどありません)。静的コストと変動コストの関係を理解しておく、最適化のアクティビティに集中する一助となります。

共有リソースのワークロードメトリクスの作成は、Amazon Elastic Container Service (Amazon ECS) および Amazon API Gateway のコンテナ化されたアプリケーションなどのリソースに比べて難しい場合があります。ただし、使用量を分類してコストを追跡する方法はあります。Amazon ECS および AWS Batch の共有リソースを追跡する必要がある場合は、AWS Cost Explorer で分割コスト配分データを有効にできます。分割コスト配分データを使用すると、コンテナ化されたアプリケーションのコストと使用状況を把握して最適化し、共有コンピューティングリソースとメモリリソースの消費状況に基づいてアプリケーションコストを個々のエンティティに配分できます。共有 API Gateway および AWS Lambda 機能がある場合は、[AWS Application Cost Profiler](#) を使用し、以下に基づいて消費量を分類できます。#### ID や ### ID。

実装手順

- ワークロードメトリクスにコストを割り当てる: 定義されたメトリクスと設定されたタグを使用して、ワークロードの出力とワークロードのコストを組み合わせたメトリクスを作成します。Amazon Athena や Amazon QuickSight などの分析サービスを使用して、ワークロード全体や任意のコンポーネントに対する効率性ダッシュボードを作成します。

リソース

関連するドキュメント:

- [AWS リソースのタグ付け](#)
- [AWS Budgets を用いてコストを分析する](#)
- [Cost Explorer を用いてコストを分析する](#)
- [Managing AWS Cost and Usage Reports \(AWS コストと使用状況レポートの管理\)](#)

関連サンプル:

- [AWS 分割コスト配分データにより Amazon ECS および AWS Batch のコストの可視性を向上する](#)

リソースを削除する

プロジェクト、従業員、テクノロジーリソースのリストを管理するようになると、使用されなくなったリソースや所有者のなくなったプロジェクトを特定できるようになります。

ベストプラクティス

- [COST04-BP01 ライフタイム全体にわたってリソースを追跡する \(p. 45\)](#)
- [COST04-BP02 削除プロセスを実装する \(p. 46\)](#)
- [COST04-BP03 リソースを削除する \(p. 48\)](#)
- [COST04-BP04 自動的にリソースを削除する \(p. 49\)](#)
- [COST04-BP05 データ保持ポリシーを適用する \(p. 50\)](#)

COST04-BP01 ライフタイム全体にわたってリソースを追跡する

ライフタイム全体にわたって、リソースや、リソースとシステムとの関係を追跡するメソッドを定義し、実装します。タグ付けにより、リソースのワークロードまたは機能を特定できます。

このベストプラクティスが確立されていない場合のリスクレベル: 高

実装のガイダンス

不要になったワークロードリソースを廃止します。一般的な例としては、テスト用途のリソースがあります。テストが完了したら、リソースは削除できます。タグを使用してリソースを追跡する (およびそれらのタグに関するレポートを実行する) ことで、使用されなくなったり、ライセンスの有効期限が切れたりした場合に、廃止する資産を特定するのに役立ちます。リソース追跡には、タグの使用が効果的な方法です。リソースにその機能が、または廃止可能になる既知の日付をラベリングできます。そうすると、これらのタグでレポートを作成できます。機能タグを付ける場合の一例として、「feature-X testing」という値であれば、ワークロードのライフサイクルの観点からリソースの目的を識別できます。別の例としては、削除するタグのキー名や値などのリソースに LifeSpan または TTL を使用して、廃止する期間や特定の時間を定義する方法があります。

実装手順

- タグ付けスキームを実装する: リソースが属するワークロードを識別するタグ付けスキームを実装し、ワークロード内のすべてのリソースが適切にタグ付けされることを確認します。タグ付けにより、目的、チーム、環境など、ビジネスに関連した基準でリソースを分類することができます。タグ付けのユースケース、戦略、テクニックの詳細については、[AWS タグ付けに関するベストプラクティス](#)を参照してください。

- ワークロードのスループットまたは出力モニタリングを実装する: 入力リクエストまたは出力完了のいずれかをトリガーとして、ワークロードのスループットのモニタリングまたはアラームを実装します。ワークロードのリクエストまたは出力がゼロになったときに、ワークロードのリソースが使用されなくなったことを示す通知を提供するように設定します。ワークロードが通常の条件下で定期的にゼロに低下する場合は、時間要因を組み込みます。未使用または使用率の低いリソースの詳細については、[AWS Trusted Advisor コストの最適化チェック](#)を参照してください。
- AWS リソースをグループ化する: AWS リソースのグループを作成します。[AWS Resource Groups](#) を使用して、同じ AWS リージョンにある AWS リソースを整理および管理できます。ほとんどのリソースにタグを追加して、組織内でのリソースの識別と分類に役立てることができます。[Tag Editor](#)を使用して、サポートされているリソースに一括でタグを追加できます。[AWS Service Catalog](#) を使用して、承認済みの製品のポートフォリオを作成、管理、およびエンドユーザーに配布し、製品ライフサイクルを管理することを検討してください。

リソース

関連するドキュメント:

- [AWS Auto Scaling](#)
- [AWS Trusted Advisor](#)
- [AWS Trusted Advisor コストの最適化チェック](#)
- [AWS リソースのタグ付け](#)
- [カスタムメトリクスの発行](#)

関連動画:

- [AWS Trusted Advisor を使用してコストを最適化する方法](#)

関連する例:

- [AWS リソースを整理する](#)
- [AWS Trusted Advisor を使用してコストを最適化する](#)

COST04-BP02 削除プロセスを実装する

未使用のリソースを特定して削除するためのプロセスを実装します。

このベストプラクティスが確立されていない場合のリスクレベル: 高

実装のガイダンス

組織全体で標準化されたプロセスを導入し、未使用のリソースを特定し、排除します。このプロセスでは、組織のすべての要件が満たされていることを検証するために、検索を実行する頻度と、リソースを削除するプロセスを定義する必要があります。

実装手順

- 削除プロセスを作成して実装する: ワークロードの開発者や所有者と協力して、ワークロードとそのリソースの削除プロセスを構築します。このプロセスでは、ワークロードが使用中であるかどうか、およびワークロードの各リソースが使用中であるかどうかを検証する方法を網羅する必要があります。リソースを削除するために必要なステップを詳述し、サービスから削除すると同時に、規制要件の遵守を確保します。ライセンスやアタッチされたストレージなど、関連するリソースも含める必要があります。削除プロセスが実行されたことをワークロードの所有者に通知します。

プロセスの一部として何を確認する必要があるかについては、以下の削除手順を使用してください。

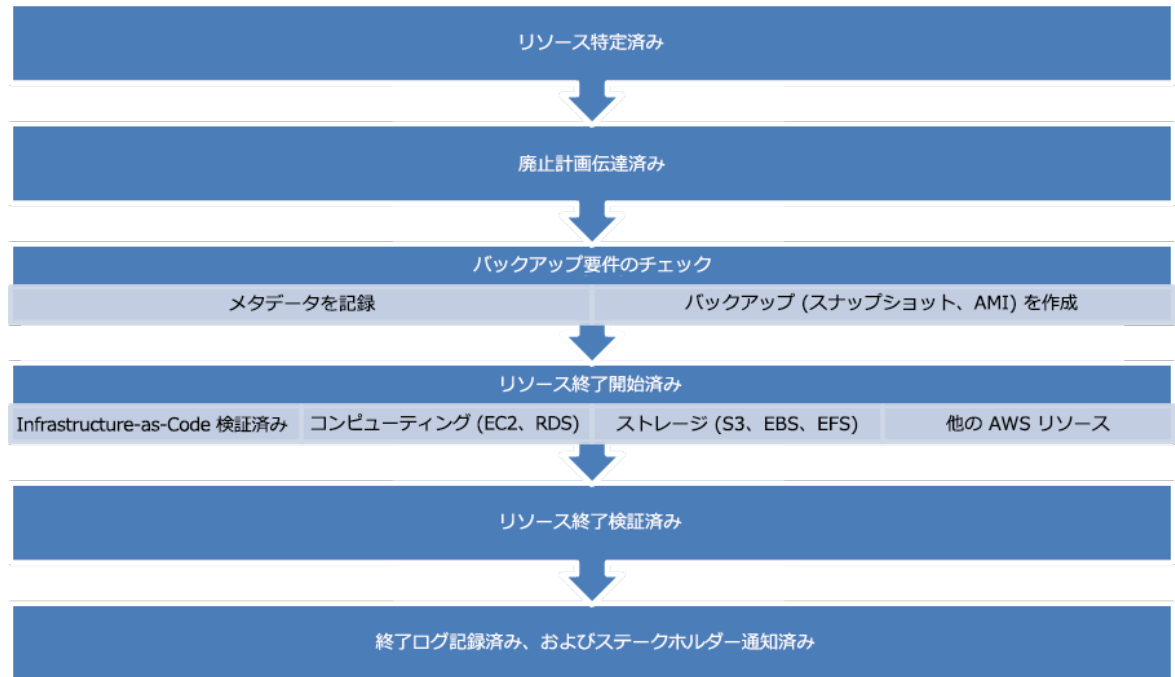
- 削除されるリソースを特定する: AWS クラウド で削除対象となるリソースを特定します。必要な情報をすべて記録し、削除スケジュールを設定します。タイムラインでは、プロセス中に予期せぬ問題が発生した場合 (およびそのタイミング) を考慮してください。
- 調整し、コミュニケーションを取る: ワークロード所有者と協力し、削除されるリソースを確認します。
- メタデータを記録し、バックアップを作成する: 本番環境のリソースに必要な場合、または重要なリソースである場合、メタデータ (パブリック IP、リージョン、AZ、VPC、サブネット、セキュリティグループなど) の記録とバックアップ (Amazon Elastic Block Store スナップショットや AMI の取得、キーのエクスポート、証明書のエクスポートなど) を作成します。
- Infrastructure-as-Code を検証する: 必要に応じて再デプロイできるように、リソースが、AWS CloudFormation、Terraform、AWS Cloud Development Kit (AWS CDK)、またはその他の Infrastructure-as-Code デプロイツールでデプロイされたかどうかを判断します。
- アクセスを禁止する: リソースが必要かどうかを判断する間、一定期間、制限的な制御を適用し、リソースの使用を禁止します。必要に応じて、リソース環境を元の状態に戻せることを確認します。
- 組織内の削除プロセスに従う: 組織のドメインからのリソースの削除、DNS レコードの削除、構成管理ツール、監視ツール、自動化ツール、セキュリティツールからのリソースの削除など、組織の管理タスクと削除プロセスに従います。

リソースが Amazon EC2 インスタンスの場合、次のリストを参照してください。[詳細については、「Amazon EC2 リソースを削除または終了するには」をご参照ください。](#)

- Amazon EC2 インスタンスとロードバランサーをすべて停止または終了します。終了後、Amazon EC2 インスタンスは、終了後、少しの間だけコンソールに表示されます。実行状態にないインスタンスには課金されません。
- Auto Scaling インフラストラクチャを削除します。
- すべての専用ホストを解放します。
- すべての Amazon EBS ボリュームおよび Amazon EBS スナップショットを削除します。
- すべての Elastic IP アドレスを解放します。
- すべての Amazon Machine Images (AMI) の登録を解除します。
- すべての AWS Elastic Beanstalk 環境を終了します。

リソースが Amazon S3 Glacier ストレージ内のオブジェクトであり、最小保存期間を満たす前にアーカイブを削除した場合、日割り計算による早期削除料が課金されます。Amazon S3 Glacier の最小保存期間は、使用するストレージクラスによって異なります。各ストレージクラスの最小保存期間の概要については、「[Amazon S3 ストレージクラス全体のパフォーマンス](#)」をご参照ください。早期削除料金の計算方法の詳細については、[Amazon S3 の料金](#)を参照してください。

次の簡単な削除プロセスのフローチャートは、削除手順を概説しています。リソースを削除する前に、削除対象として特定したリソースが組織で使用されていないことを確認します。



リソースの削除フロー。

リソース

関連するドキュメント:

- [AWS Auto Scaling](#)
- [AWS Trusted Advisor](#)
- [AWS CloudTrail](#)

関連動画:

- [CloudFormation スタックは削除するが、一部のリソースを保持する](#)
- [どのユーザーが Amazon EC2 インスタンスを起動したかを確認する](#)

関連する例:

- [Amazon EC2 リソースの削除および終了](#)
- [どのユーザーが Amazon EC2 インスタンスを起動したかを確認する](#)

COST04-BP03 リソースを削除する

定期監査や使用状況の変化などのイベントを契機として、リソースを削除します。通常、削除は定期的に行われ、手動または自動で実行できます。

このベストプラクティスを確立しない場合のリスクレベル: 中

実装のガイダンス

使用していないリソースを検索する場合は節減額の程度によって検索頻度と投入する労力を決定する必要があります。コスト発生額の小さいアカウントの分析は、コスト発生額が高額のアカウントよりも頻度を下げるべきです。イベントの検索および廃止は、製品が寿命を迎えた場合や交換する場合など、ワークロードの状態の変化によってトリガーされます。イベントの検索および廃止は、市況の変化や製品終了などの外部イベントによってトリガーされる場合もあります。

実装手順

- リソースの削除: これは、不要になった AWS リソースやライセンス契約の終了の減価償却段階です。スナップショットやバックアップの取得などの不要な中断を防ぐために、削除段階に移行してリソースを削除する前に完了したすべての最終チェックを完了します。削除プロセスを使用して、未使用と識別された各リソースを削除します。

リソース

関連するドキュメント:

- [AWS Auto Scaling](#)
- [AWS Trusted Advisor](#)

関連する例:

- [Well-Architected ラボ: リソースを削除する \(レベル 100\)](#)

COST04-BP04 自動的にリソースを削除する

重要度が低いリソース、不要なリソース、使用率が低いリソースを特定して削除する作業を適切に行えるようにワークロードを設計します。

このベストプラクティスが確立されていない場合のリスクレベル: 低

実装のガイダンス

オートメーションを使用して、廃棄プロセスの関連コストを削減または削除します。自動削除するようにワークロードを設計すると、そのライフタイム全体にわたるワークロードコストを削減できます。[AWS Auto Scaling](#) を使用して、削除プロセスを実行できます。また、[API または SDK](#) を使用してカスタムコードを運用して、ワークロードリソースの自動削除もできます。

[モダンアプリケーション](#)は、サーバーレスサービスの採用を優先する戦略であるサーバーレスファーストで構築されています。AWS は、スタックの 3 つのレイヤー (コンピューティング、インテグレーション、データストア) すべてに対応する[サーバーレスサービス](#)を開発しました。サーバーレスアーキテクチャを使用すると、トラフィックの少ない間、自動的にスケールアップおよびスケールダウンしてコストを節約できます。

実装手順

- AWS Auto Scaling を実装する: サポートされているリソースについては、[AWS Auto Scaling](#) を使用して設定します。AWS サービスを利用する際に、AWS Auto Scaling を使用することで利用率とコスト効率を最適化することができます。AWS Auto Scaling は需要が低下すると、余分なリソースを自動的に削除するため、過剰な支出を避けることができます。
- CloudWatch を設定して、インスタンスを終了させる: インスタンスは、[CloudWatch アラーム](#)を使用して終了するように設定できます。削除プロセスのメトリクスを使用して、Amazon Elastic Compute

Cloud アクションでアラームを実装します。ロールアウトする前に、非本番環境でオペレーションを検証します。

- ワークロード内にコードを実装する: AWS SDK または AWS CLI を使用して、ワークロードリソースを削除できます。AWS と統合されるアプリケーション内にコードを実装し、使用されなくなったリソースを終了または削除します。
- サーバーレスサービスを使用する: AWS で、[サーバーレスアーキテクチャ](#)や[イベント駆動型アーキテクチャ](#)を優先的に構築し、アプリケーションの構築と実行を行います。AWS では、本来、自動的に最適化されたリソース使用率と自動削除 (スケールインとスケールアウト) を提供する複数のサーバーレステクノロジーサービスが用意されています。サーバーレスアプリケーションでは、リソース使用率が自動的に最適化され、過剰プロビジョニングの費用が発生しません。

リソース

関連するドキュメント:

- [AWS Auto Scaling](#)
- [AWS Trusted Advisor](#)
- [AWS でのサーバーレス](#)
- [インスタンスを停止、終了、再起動、または復旧するアラームを作成する](#)
- [Amazon EC2 Auto Scaling の開始方法](#)
- [Amazon CloudWatch アラームへの終了アクションの追加](#)

関連する例:

- [Scheduling automatic deletion of AWS CloudFormation stacks](#) (AWS CloudFormation スタックの自動削除のスケジューリング)
- [Well-Architected ラボ – 自動的にリソースを削除する \(レベル 100\)](#)
- [Servian AWS Auto Cleanup](#)

COST04-BP05 データ保持ポリシーを適用する

サポートされるリソースでデータ保持ポリシーを定義し、組織の要件に従ってオブジェクト削除を処理します。不要または孤立したリソースや不要になったオブジェクトを、特定して削除します。

このベストプラクティスを確立しない場合のリスクレベル: 中

データ保持ポリシーとライフサイクルポリシーを使用して、特定されたリソースの廃止プロセスに関連するコストとストレージコストを削減します。データ保持ポリシーとライフサイクルポリシーを定義してストレージクラスの移行や削除を自動で実行すると、ライフタイム期間の全体的なストレージコストが削減されます。Amazon Data Lifecycle Manager を使用して、Amazon Elastic Block Store スナップショットおよび Amazon EBS 搭載 Amazon マシンイメージ (AMI) の作成と削除を自動化できます。また、Amazon S3 Intelligent-Tiering または Amazon S3 ライフサイクル設定を使用して、Amazon S3 オブジェクトのライフサイクルを管理できます。ほかにも、[API または SDK](#) を使用してカスタムコードを実装し、ライフサイクルポリシーとポリシールールを作成して、オブジェクトを自動で削除することもできます。

実装手順

- Amazon Data Lifecycle Manager の使用: Amazon Data Lifecycle Manager でライフサイクルポリシーを使用して、Amazon EBS スナップショットと Amazon EBS 搭載 AMI の削除を自動化します。
- バケットにライフサイクル設定をセットアップする: バケットで Amazon S3 ライフサイクル設定を使用して、ビジネス要件に基づき、オブジェクトのライフサイクル中に Amazon S3 が実行するアクション、およびオブジェクトのライフサイクル終了時の削除を定義します。

リソース

関連するドキュメント:

- [AWS Trusted Advisor](#)
- [Amazon Data Lifecycle Manager](#)
- [Amazon S3 バケットにライフサイクル設定を設定する方法](#)

関連動画:

- [Automate Amazon EBS Snapshots with Amazon Data Lifecycle Manager](#) (DLM を使用して EBS スナップショットを自動化する)
- [Empty an Amazon S3 bucket using a lifecycle configuration rule](#) (ライフサイクル設定ルールを使用して Amazon S3 バケットを空にする)

関連する例:

- [Empty an Amazon S3 bucket using a lifecycle configuration rule](#) (ライフサイクル設定ルールを使用して Amazon S3 バケットを空にする)
- [Well-Architected Lab: Decommission resources automatically \(Level 100\)](#) (Well-Architected ラボ: 自動的にリソースを削除する (レベル 100))

コスト効率を考慮しながらリソースを利用する

ワークロードに適したサービス、リソース、設定の使用は、コスト節減の鍵を握ります。費用対効果の高いリソースを作成する場合、以下の点を考慮してください。

AWS ソリューションアーキテクト、AWS ソリューション、AWS リファレンスアーキテクチャ、APN パートナーを利用すると、学習内容に基づいてアーキテクチャを選択できます。

トピック

- [サービスを選択する際にコストを評価する \(p. 52\)](#)
- [正しいリソースタイプ、リソースサイズ、リソース数を選択する \(p. 60\)](#)
- [最適な料金モデルを選択する \(p. 65\)](#)
- [データ転送を計画する \(p. 73\)](#)

サービスを選択する際にコストを評価する

ベストプラクティス

- [COST05-BP01 組織のコスト要件を特定する \(p. 52\)](#)
- [COST05-BP02 ワークロードのすべてのコンポーネントを分析する \(p. 53\)](#)
- [COST05-BP03 各コンポーネントの詳細な分析を実行する \(p. 54\)](#)
- [COST05-BP04 コスト効率の高いライセンスを提供するソフトウェアを選択する \(p. 56\)](#)
- [COST05-BP04 組織の優先順位に従ってコストが最適化されるようにこのワークロードのコンポーネントを選択する \(p. 57\)](#)
- [COST05-BP06 異なる使用量について経時的なコスト分析を実行する \(p. 59\)](#)

COST05-BP01 組織のコスト要件を特定する

チームメンバーと協力して、コストの最適化とこのワークロードのその他の柱とのバランス (パフォーマンスや信頼性など) を定義します。

このベストプラクティスが確立されていない場合のリスクレベル: 高

実装のガイダンス

ほとんどの組織では、情報技術 (IT) 部門には複数の小さなチームがあり、抱える議題や注力分野は、所属メンバーの専門分野とスキルを反映してそれぞれに異なります。組織全体の目標、優先事項、目的を把握し、各部門や各プロジェクトがそうした目標にどのように貢献しているかを理解する必要があります。人員、設備、技術、資材、外部サービスなど、すべての重要なリソースを分類しておくことは、組織の目標を達成し、包括的な予算計画を立てるうえで不可欠です。こうした体系的なアプローチでコストを特定し、把握することが、組織にとって現実的で健全なコスト計画を立てるための基本です。

ワークロードのサービスを選択する場合は、組織の優先順位を理解することが重要です。コスト最適化と、AWS Well-Architected フレームワークの他の柱 (パフォーマンスや信頼性など) とのバランスを図ります。このプロセスは、組織の目標、市況、業務のダイナミクスの変化を反映するために、体系的かつ定期的に実施する必要があります。十分にコスト最適化されたワークロードとは組織の要件に最も適合するソ

リユースであって、必ずしも最低コストのソリューションとは限りません。製品、ビジネス、技術、財務など、組織内のすべてのチームと会合し、情報を収集します。競合する利益または代替アプローチ間のトレードオフの影響を評価し、重点領域を決定するか、一連のアクションを選択する際に十分な情報に基づいて意思決定を下せるようにします。

たとえば、新しい機能の市場投入までの時間を短縮することは、コストの最適化よりも重視されることがあります。または、非リレーショナルデータ用にリレーショナルデータベースを選択すれば、データ型に合わせて最適化されたデータベースに移行してアプリケーションを更新するよりも、システムの移行が簡素化されます。

実装手順

- 組織のコスト要件を特定する: 製品管理、アプリケーション所有者、開発および運用チーム、管理、財務のメンバーなど、組織のチームメンバーとミーティングを行います。このワークロードとそのコンポーネントに対して、Well-Architected の柱に優先順位を付けます。柱を順番に並べたリストを作成してください。また、それぞれの柱に重みを付け、その柱が他の柱よりどの程度重視されているかや、2 つの柱の重点度がどの程度類似しているかを示すことができます。
- 技術的負債に対処し、文書化する: ワークロードのレビュー中に、技術的負債に対処します。バックログ項目を文書化して、後日、リファクタリングやリアーキテクティングで最適化を進めることを目標に、ワークロードを再検討します。生じたトレードオフを他のステークホルダーに明確に伝えることが大切です。

リソース

関連するベストプラクティス:

- [REL11-BP07 可用性の目標と稼働時間のサービスレベルアグリーメント \(SLA\) を満たす製品を設計する](#)
- [OPS01-BP06 トレードオフを評価する](#)

関連するドキュメント:

- [AWS 総保有コスト \(TCO\) 計算ツール](#)
- [Amazon S3 ストレージクラス](#)
- [クラウド製品](#)

COST05-BP02 ワークロードのすべてのコンポーネントを分析する

現在のサイズやコストに関係なく、すべてのワークロードが分析されることを確認します。見直しを行う際には、現在のコストや予想コストなどの潜在的利益を織り込む必要があります。

このベストプラクティスが確立されていない場合のリスクレベル: 高

実装のガイダンス

組織にビジネス価値をもたらすべく設計されたワークロードコンポーネントには、さまざまなサービスが含まれる場合があります。コンポーネントごとに、ビジネスニーズに対応する特定の AWS クラウド サービスを選択できます。何が選択されるかは、それらのサービスに関する知識や使用経験などの要因によって違ってきます。

組織の要件 (「[COST05-BP01 組織のコスト要件を特定する](#)」を参照) を特定したら、ワークロードのすべてのコンポーネントを徹底的に分析します。現時点と今後予測されるコストとサイズを考慮して、各コンポーネントを分析します。分析のコストを、ワークロードのライフサイクル全体で削減が見込まれる額と

比較検討してください。対象となるワークロードのコンポーネントをすべて分析するための労力が、そのコンポーネントの最適化により見込まれる節約額や改善に見合っていないかもしれません。例えば、提案されたリソースのコストが月額 10 USD で、予測負荷が月額 15 USD を超えない場合に、コストを 50% (月額 5 USD) 削減するために 1 日分の労力を費やすようでは、システムの寿命全体にわたって得られると考えられる利益を超えることになるかもしれません。データに基づくより高速でより効率的な予測を使用すると、このコンポーネントの全体的な成果を最善のものにできます。

ワークロードは時間の経過とともに変化する可能性があり、ワークロードのアーキテクチャや使用方法が変化すると、適切だったサービスの組み合わせが最適ではなくなってしまうことがあります。サービスの選択に関する分析には、現在および将来のワークロードの状態と使用量レベルが組み込まれる必要があります。将来のワークロードの状態や使用量に合わせてサービスを運用すると、今後の変更に必要な労力を軽減または削除できることになり、全体的なコストを削減できます。例えば、最初は Amazon EMR Serverless を使用してみるのが適切かもしれません。ただし、そのサービスの使用量が増えてきたら、Amazon EC2 の Amazon EMR に移行することで、ワークロードの該当コンポーネントのコストを削減できる可能性があります。

現在の属性に関係なく、すべてのワークロードコンポーネントを戦略的に見直すことで、時間の経過に伴い、目に見えて機能を強化し、コストを削減できる可能性があります。このレビュープロセスに費やす労力は、それに見合う効果が得られるか慎重に検討した上で、決める必要があります。

[AWS Cost Explorer](#) および [AWS Cost and Usage Report \(CUR\)](#) では、PoC (概念実証) または実行中の環境のコストを分析できます。また、[AWS Pricing Calculator](#) を使用して、ワークロードのコストを見積もることもできます。

実装手順

- ワークロードコンポーネントをリスト化する: ワークロードのコンポーネントのリストを作成します。これは、各コンポーネントが分析されたことを確認するための検証として使用されます。費やされる労力は、組織の優先順位によって定義されたワークロードの重要性を反映している必要があります。リソースをグループ化すると、機能的に効率が向上します (例えば、複数のデータベースがある場合は本番データベースストレージ)。
- コンポーネントリストに優先順位を付ける: コンポーネントリストを取得して、労力をかける順で優先順位付けをします。これは通常、コンポーネントのコストが最も高価なものから最も安価なものへ、または組織の優先順位で定義されている重要度の順に並べられます。
- 分析を実行する: リストの各コンポーネントについて、使用可能なオプションとサービスを確認し、組織の優先順位に最適なオプションを選択します。

リソース

関連するドキュメント:

- [AWS Pricing Calculator](#)
- [AWS Cost Explorer](#)
- [Amazon S3 ストレージクラス](#)
- [クラウド製品](#)

COST05-BP03 各コンポーネントの詳細な分析を実行する

各コンポーネントの、組織にかかる全体的なコストを調べます。運用および管理のコスト、特にクラウドプロバイダーが提供するマネージドサービスを使用するコストを考慮して、総保有コストを計算します。レビューを行う際には、潜在的利益 (分析に費やされた時間がコンポーネントのコストに比例しているなど) を織り込む必要があります。

このベストプラクティスが確立されていない場合のリスクレベル: 高

実装のガイダンス

時間短縮を検討して、チームが技術的な負債の返済、イノベーション、付加価値機能、ビジネスの差別化要素の構築に集中できるようにします。例えば、データベースをできるだけ迅速にオンプレミス環境からクラウドにリフトアンドシフト (リホストともいいます) して、後で最適化する必要がある場合です。AWS でマネージドサービスを利用して、ライセンスコストの排除または削減を模索することには、時間をかけるだけの価値があります。AWS のマネージドサービスによって、OS のパッチ適用やアップグレードなど、サービス維持に伴う運用上および管理上の負担が軽減されるため、イノベーションとビジネスに集中できます。

マネージドサービスはクラウド規模で運用されるため、トランザクションまたはサービス単位でコストを削減できます。アプリケーションのコアアーキテクチャを変更せずに、具体的なメリットを生み出すための潜在的最適化作業を行うことができます。例えば、[Amazon Relational Database Service \(Amazon RDS\)](#) などの database-as-a-service プラットフォームに移行するか、[AWS Elastic Beanstalk](#) などのフルマネージド型のプラットフォームにアプリケーションを移行することで、データベースインスタンスの管理に費やす時間を削減できます。

通常、マネージドサービスは、十分なキャパシティを確保するために設定できる属性を備えています。この属性を設定およびモニタリングして、余剰キャパシティを最小限に抑え、パフォーマンスを最大化する必要があります。AWS Management Console や AWS API および SDK を使用して AWS Managed Services の属性を変更し、需要の変化に合わせてリソースのニーズを調整できます。例えば、Amazon EMR クラスターまたは Amazon Redshift クラスターのノード数を増減して、規模をスケールアウトまたはスケールインできます。

また、AWS リソースの複数のインスタンスを圧縮して、高密度に使用することもできます。例えば、単一の Amazon Relational Database Service (Amazon RDS) データベースインスタンスで、複数の小さなデータベースをプロビジョニングできます。使用量が増えたら、スナップショットや復元プロセスを使用して、そのデータベースの 1 つを専用 Amazon RDS データベースインスタンスに移行できます。

マネージドサービスでワークロードをプロビジョニングする際は、サービスキャパシティの調整要件を理解する必要があります。主な要件としては、時間、労力、通常のワークロードオペレーションへの影響などが一般には考えられます。プロビジョニングされたリソースでは変更が発生するまでの時間が許容され、このために必要なオーバーヘッドをプロビジョニングする必要があります。サービス変更に必要な継続的労力は、システムに統合する API と SDK や Amazon CloudWatch などのモニタリングツールを使用することで、実質ゼロまで減らすことができます。

[Amazon RDS](#)、[Amazon Redshift](#)、[Amazon ElastiCache](#) は、マネージドデータベースサービスを提供します。[Amazon Athena](#)、[Amazon EMR](#)、[Amazon OpenSearch Service](#) はマネージド分析サービスを提供します。

[AMS](#) は、エンタープライズのお客様やパートナーに代わって AWS インフラストラクチャを運用するサービスです。コンプライアンスに準拠したセキュアな環境で、ワークロードをデプロイできます。AMS では、エンタープライズクラウド運用モデルとオートメーションを使用して、組織の要件を満たし、クラウド移行を高速化し、継続的な管理コストを削減できます。

実装手順

- 詳細な分析を実行する: コンポーネントリストを使用して、各コンポーネントを優先度が高いものから処理します。優先度がより高く、より多くのコストがかかるコンポーネントについては、追加の分析を実行し、利用可能なすべてのオプションとその長期的な影響を評価します。優先度の低いコンポーネントの場合、使用状況の変化によってコンポーネントの優先度が変更するかどうかを評価し、かける労力の適切性の分析を実行します。
- 管理されているリソースと管理されていないリソースを比較する: 自社で管理するリソースの運用コストを考え、それを AWS が管理するリソースと比較します。例えば、Amazon EC2 インスタンスで実行しているデータベースをレビューし、Amazon RDS (AWS マネージドサービス) というオプションと比較します。または、Amazon EMR と、Amazon EC2 で Apache Spark を実行する場合を比較します。セルフマネージドワークロードから AWS のフルマネージドワークロードに移行する際は、オプションを慎

重に研究してください。最も重要な 3 つの考慮事項は、使用する [マネージドサービスの種類](#)、[データ移行](#) に使用するプロセス、[AWS 責任共有モデルの理解](#) です。

リソース

関連するドキュメント:

- [AWS 総保有コスト \(TCO\) 計算ツール](#)
- [Amazon S3 ストレージクラス](#)
- [AWS クラウド 製品](#)
- [AWS 責任共有モデル](#)

関連動画:

- [Why move to a managed database?](#) (マネージドデータベースに移行する理由)
- [What is Amazon EMR and how can I use it for processing data?](#) (EMR の概要とこれを使用したデータの処理方法)

関連する例:

- [Why move to a managed database?](#) (マネージドデータベースに移行する理由)
- [Consolidate data from identical SQL Server databases into a single Amazon RDS for SQL Server database using AWS DMS](#) (AWS を使用してさまざまな SQL サーバーデータベースから単一の Amazon RDS for SQL Server データベースにデータを統一する)
- [Deliver data at scale to Amazon Managed Streaming for Apache Kafka \(Amazon MSK\)](#) (Amazon Managed Streaming for Apache Kafka (Amazon MSK) に大規模にデータを配信する)
- [Migrate an ASP.NET web application to AWS Elastic Beanstalk](#) (ASP.NET ウェブアプリケーションを AWS Elastic Beanstalk に移行する)

COST05-BP04 コスト効率の高いライセンスを提供するソフトウェアを選択する

オープンソースソフトウェアは、ワークロードに多大なコストをもたらすソフトウェアライセンスコストを排除することができます。ライセンスされたソフトウェアが必要な場合は、CPU などの任意の属性に結びついたライセンスは避け、出力または結果に結びついたライセンスを探します。これらのライセンスのコストは、提供するメリットに応じてより密にスケールされます。

このベストプラクティスが確立されていない場合のリスクレベル: 低

実装のガイダンス

「オープンソース」は、ソフトウェア開発の文脈で生まれた用語であり、ソフトウェアが特定の無料配布基準に準拠しているという意味です。オープンソースソフトウェアは、誰でも検査、変更、拡張できるソースコードで構成されています。組織はビジネス要件、エンジニアのスキル、予測される使用状況、その他の技術的な依存関係を踏まえて、ライセンスコストを最小限に抑えるため、AWS でオープンソースソフトウェアを使用することを検討できます。つまり、[オープンソースソフトウェア](#)を使用することで、ソフトウェアライセンスのコストを削減できます。オープンソースソフトウェアへの変更は、ワークロードサイズが拡大するにつれ、ワークロードコストに大きな影響を与える可能性があります。

ライセンスを取得したソフトウェアで得られる効果を、ワークロードの最適化にかかる総コストに照らして測定してください。ライセンス変更とその変更がワークロードコストに与える影響をモデリングしま

す。あるベンダーがデータベースライセンスのコストを変更したなら、それがワークロードの全体的な効率にどのような影響を与えるかを調査します。ベンダーの過去の価格アナウンスを検討して、ベンダー製品全体のライセンス変更の傾向を検討してください。ライセンスコストは、ハードウェアごとにスケールするライセンス (CPU バウンドライセンス) など、スループットや使用量とは関係なくスケールされる場合があります。こうしたライセンスは、それに伴う成果が見られないままコストが急増する可能性があるため、避けてください。

例えば、Linux オペレーティングシステムを搭載した Amazon EC2 インスタンスを us-east-1 で運用する場合、Windows で実行する別の Amazon EC2 インスタンスを運用する場合と比較して、約 45% のコスト削減になります。

[AWS Pricing Calculator](#) では、Amazon RDS インスタンスや各種データベースエンジンなど、ライセンスオプションが異なるさまざまなリソースのコストを総合的に比較できます。さらに、AWS Cost Explorer では、既存のワークロード (特にライセンスがさまざまに異なるワークロード) のコストを有益な視点で検討できます。ライセンス管理については、[AWS License Manager](#) でソフトウェアライセンスの管理と処理を合理化できます。お客様は、AWS クラウド でお好みのオープンソースソフトウェアをデプロイして運用できます。

実装手順

- ライセンスオプションを分析する: 利用可能なソフトウェアのライセンス条項を確認します。必要な機能を備えたオープンソースバージョンを探し、ライセンスされたソフトウェアの利点がコストを上回っているかどうかを調べます。好条件があれば、ソフトウェアのコストに見合う利点が得られます。
- ソフトウェアプロバイダーを分析する: ベンダーによる料金履歴またはライセンスの改定履歴を確認します。特定のベンダーのハードウェアまたはプラットフォームで実行することについての懲罰的な条件など、結果に見合わない変更を調べます。また、監査の実行方法や課される可能性のある罰則についても確認します。

リソース

関連するドキュメント:

- [Open Source at AWS](#)
- [AWS 総保有コスト \(TCO\) 計算ツール](#)
- [Amazon S3 ストレージクラス](#)
- [クラウド製品](#)

関連する例:

- [オープンソースに関するブログ](#)
- [AWS Open Source Blogs](#)
- [最適化とライセンス評価](#)

COST05-BP04 組織の優先順位に従ってコストが最適化されるようにこのワークロードのコンポーネントを選択する

ワークロードのすべてのコンポーネントを選択したときのコストを考慮します。これには、アプリケーションレベルのサービスとマネージドサービス、またはサーバーレス、コンテナ、イベント駆動型アーキテクチャを使用して、全体のコストを削減することが含まれます。オープンソースソフトウェアやライセンス料金がつかからないソフトウェア、または代替品を使用して、支出を最小限に抑えます。

このベストプラクティスが確立されていない場合のリスクレベル: 中

実装のガイダンス

すべてのコンポーネントを選択する際は、サービスのコストとオプションを考慮します。これには、[Amazon Relational Database Service \(Amazon RDS\)](#)、[Amazon DynamoDB](#)、[Amazon Simple Notification Service \(Amazon SNS\)](#)、[Amazon Simple Email Service \(Amazon SES\)](#) などのアプリケーションレベルのサービスとマネージドサービスを使用して、組織の全体的なコストを削減することが含まれます。

コンピューティングには [AWS Lambda](#) などのサーバーレスやコンテナを使用し、静的ウェブサイトには [Amazon Simple Storage Service \(Amazon S3\)](#) を使用します。可能であればアプリケーションをコンテナ化し、[Amazon Elastic Container Service \(Amazon ECS\)](#) や [Amazon Elastic Kubernetes Service \(Amazon EKS\)](#) などの AWS マネージドコンテナサービスを使用します。

オープンソースソフトウェア、またはライセンス料金のないソフトウェア (コンピューティングワークロード用の Amazon Linux、データベースを Amazon Aurora に移行するなど) を使用して、ライセンスコストを最小限に抑えます。

[AWS Lambda](#)、[Amazon Simple Queue Service \(Amazon SQS\)](#)、[Amazon Simple Notification Service \(Amazon SNS\)](#)、[Amazon Simple Email Service \(Amazon SES\)](#) などのサーバーレスまたはアプリケーションレベルのサービスを使用できます。これらのサービスではリソースを管理する必要がなく、コード実行、キューサービス、メッセージ配信の機能を利用できます。もう 1 つの利点は、使用量に応じてパフォーマンスとコストをスケールインするため、コスト配分とコストの帰属が効率的になることです。

サーバーレスサービスでは、[イベント駆動型アーキテクチャ](#)を使用することもできます。イベント駆動型アーキテクチャはプッシュベースであるため、イベントはルーターで発生してもオンデマンドで取得されます。この方法では、イベントをチェックするために定期的にポーリングする費用が発生しません。つまり、ネットワーク帯域幅の消費を抑え、CPU 使用率は低く、アイドルなフリートキャパシティは少なくなり、SSL/TLS ハンドシェイクも減ります。

サーバーレスの詳細については、[Well-Architected Serverless Application lens サーバーレスアプリケーションレンズのホワイトペーパー](#)を参照してください。

実装手順

- 各サービスを選択してコストを最適化する: 優先順位リストと分析を使用して、組織の優先順位に最も合致する各オプションを選択します。需要に合わせてキャパシティを増やすのではなく、より低いコストでより優れたパフォーマンスを得られる可能性がある他のオプションを検討します。例えば、AWS 上のデータベースに対する予想されるトラフィックをレビューし、インスタンスサイズを増やすか、Amazon ElastiCache サービス (Redis または Memcached) を使用してデータベースにキャッシュメカニズムを提供するかを検討します。
- イベント駆動型アーキテクチャを評価する: サーバーレスアーキテクチャを使用すると、分散マイクロサービスベースのアプリケーション向けにイベント駆動型アーキテクチャを構築することもできます。これを利用すると、スケーラブルで回復性が高く、迅速かつコスト効果の高いソリューションを構築できます。

リソース

関連するドキュメント:

- [AWS 総保有コスト \(TCO\) 計算ツール](#)
- [AWS でのサーバーレス](#)
- [EDA とは](#)
- [Amazon S3 ストレージクラス](#)

- [クラウド製品](#)
- [Amazon ElastiCache for Redis](#)

関連する例:

- [Getting started with event-driven architecture \(イベント駆動型アーキテクチャで開始する\)](#)
- [イベント駆動型アーキテクチャ](#)
- [How Statsig runs 100x more cost-effectively using Amazon ElastiCache for Redis](#)
- [AWS Lambda 関数を使用するためのベストプラクティス](#)

COST05-BP06 異なる使用量について経時的なコスト分析を実行する

ワークロードは時間の経過とともに変化することがあります。それぞれのサービスまたは機能のコスト効率は、使用レベルによって異なります。各コンポーネントについて予想使用量に基づく経時的な分析を行うことで、ワークロードのコスト効率性がそのライフタイム全体にわたって維持されます。

このベストプラクティスが確立されていない場合のリスクレベル: ミディアム

実装のガイダンス

AWS で新しいサービスや機能がリリースされると、ワークロードに最適なサービスが変化する可能性があります。求められる労力は、潜在的な利点が反映されたものである必要があります。ワークロードレビューの頻度は、組織の要件によって異なります。ワークロードにかなりのコストがかかっている場合、新しいサービスの運用が早いほどコスト削減が最大になるため、レビュー頻度が高い方が有利です。レビューのトリガーとしては、使用パターンの変化も挙げられます。使用量が大幅に変化した場合は、別のサービスを使った方がよい場合もあります。

データを AWS クラウドに移動する必要がある場合、AWS が提供するバリエーション豊かな製品やパートナーツールを選択して、データセットを移行できます。データセットは、ファイル、データベース、マシンイメージ、ブロックボリューム、あるいはテープバックアップであっても構いません。例えば、大量のデータを AWS に対して入出力する場合や、エッジでデータを処理する場合、AWS の目的別デバイスのいずれかを使用して、コスト効果が高い方法でペタバイト規模のデータをオフラインで移動できます。別の例としては、より速いデータ転送速度が必要な場合、VPN よりも、ビジネスに必要な安定した接続性能を提供する直接接続サービスの方が安価な場合があります。

さまざまな使用状況において繰り返したコスト分析を基にして、スケーリングアクティビティをレビューします。結果を分析して、複数のインスタンスタイプと購入オプションを使用したインスタンスの追加に合わせてスケーリングポリシーを調整できるかを確認します。設定をレビューして、最小限を削減してもユーザーリクエストを処理できる (ただしより小さなフリートサイズで) かを確認し、予想される高需要を満たすためにリソースを追加します。

組織の関係者と話し合っ、さまざまな使用状況について繰り返しコスト分析を実行し、[AWS Cost Explorer](#) の予測機能を使用して、サービス変更によって発生する可能性のある影響を予測します。AWS Budgets、CloudWatch 請求アラーム、AWS Cost Anomaly Detection を使用して使用状況レベルのトリガーをモニタリングし、最もコスト効果が高いサービスをなるべく迅速に特定して実装します。

実装手順

- 予測された使用パターンを定義する: マーケティングや製品所有者などの組織と協力して、ワークロードに対して期待および予測される使用パターンを文書化します。これまでと今後両方のコストと使用量の増加についてビジネス上の関係者と話し合い、増加がビジネス要件に沿ったものであることを確認します。自社の AWS リソースを使用するユーザーが増える日、週、月を特定します。そのタイミングで既

存のリソースのキャパシティを増やすか追加サービスを導入して、コストを削減しパフォーマンスを向上させる必要があります。

- 予測された使用量に基づくコスト分析を実行する: 定義された使用パターンを使用して、これらの各ポイントで分析を実行します。分析作業は、潜在的な結果を反映する必要があります。例えば、使用量の変化が大きい場合は、コストと変化を確認するために詳細な分析を実行する必要があります。つまり、コストが増えていれば、ビジネスにおける使用量も同様に増えているはずです。

リソース

関連するドキュメント:

- [AWS 総保有コスト \(TCO\) 計算ツール](#)
- [Amazon S3 ストレージクラス](#)
- [クラウド製品](#)
- [Amazon EC2 Auto Scaling](#)
- [クラウドデータの移行](#)
- [AWS Snow Family](#)

関連動画:

- [AWS OpsHub for Snow Family](#)

正しいリソースタイプ、リソースサイズ、リソース数を選択する

最適なリソースタイプ、リソースサイズ、リソース数を選択することで、最低限のリソースコストで技術要件を満たすことができます。適切なサイズ変更アクティビティは、ワークロードのすべてのリソース、各リソースのすべての属性、および適切なサイジング操作に関連する作業が考慮されます。適切なサイズ変更は、使用パターンや、AWS の値下げや 新しい AWS リソースタイプなどの外部要因の変化により、反復プロセスになることがあります。適切なサイズ設定の作業コストが、ワークロードの寿命全体にわたって可能と考えられる削減額を上回る場合は、1 回だけ設定することも可能です。

AWS には以下のような多様なアプローチがあります。

ベストプラクティス

- [COST06-BP01 コストモデリングを実行する \(p. 60\)](#)
- [COST06-BP02 データに基づいてリソースタイプ、リソースサイズ、リソース数を選択する \(p. 62\)](#)
- [COST06-BP03 メトリクスに基づいて自動的にリソースタイプ、リソースサイズ、リソース数を選択する \(p. 63\)](#)

COST06-BP01 コストモデリングを実行する

組織の要件 (ビジネスニーズや既存のコミットメントなど) を特定して、ワークロードとその各コンポーネントのコストモデリング (全体コスト) を実行します。さまざまな予測負荷におけるワークロードのベンチマークアクティビティを実行し、コストを比較します。モデリングの際には、潜在的な利点を織り込む必要があります。例えば、費やされた時間がコンポーネントのコストに比例しているなどです。

このベストプラクティスが確立されていない場合のリスクレベル: 高

実装のガイダンス

ワークロードと各コンポーネントのコストモデリングを実行してリソース間のバランスを把握し、特定のパフォーマンスレベルに応じてワークロード内の各リソースの適切なサイズを見つけます。コストの考慮事項を理解すると、計画されたワークロードのデプロイにおける価値実現の成果評価時に、組織のビジネスケースや意思決定プロセスがわかります。

さまざまな予測負荷におけるワークロードのベンチマークアクティビティを実行し、コストを比較します。モデリングの際には、費やした時間がコンポーネントのコストまたは予想される削減額に比例しているといった潜在的な利点を織り込む必要があります。このプロセスのベストプラクティスについては、[AWS Well-Architected フレームワークのパフォーマンス効率の柱に関するレビューセクション](#)を参照してください。

例えば、コンピューティングリソースからなるワークロードのコストモデリングを作成する場合、[AWS Compute Optimizer](#) がワークロード実行におけるコストのモデリングに役立ちます。使用履歴に基づき、コンピューティングリソースの正しいサイズ設定に関するレコメンデーションを提供します。CloudWatch エージェントが Amazon EC2 インスタンスにデプロイされていることを確認し、AWS Compute Optimizer 内でより正確な推奨事項を得ることができるメモリメトリクスを収集します。リスクレベルに応じて複数のレコメンデーションを作成できる機械学習が使われている無料サービスであるため、コンピューティングリソースにとって理想的なデータソースです。

他のサービスやワークロードコンポーネントのサイズを適正化するために、カスタムログをデータソースとして使用できる [サービスが複数](#) あります。例えば、[AWS Trusted Advisor](#)、[Amazon CloudWatch](#)、[Amazon CloudWatch Logs](#) などです。AWS Trusted Advisor はリソースをチェックして使用率が低いリソースにフラグを立てるため、リソースのサイズを適正化しコストモデリングを作成するのに役立ちます。

コストモデリングのデータとメトリクスに関する推奨事項は以下の通りです。

- モニタリングはユーザーエクスペリエンスを正確に反映する必要があります。対象期間に適切な間隔を選択して、平均の変わりや最大値や 99 パーセンタイル値をじっくり見極めます。
- すべてのワークロードのサイクルをカバーするために必要な分析期間の適切な間隔を選択します。たとえば、分析を 2 週間間隔で実行する場合、1 か月サイクルで使用率が高くても見逃す場合があり、過小プロビジョニングにつながる可能性があります。
- 既存のコミットメント、他のワークロード用に選択された料金モデル、イノベーションを迅速化しコアビジネスバリューに集中する能力を考慮し、計画されたワークロードに対して適切な AWS サービスを選択します。

実装手順

- リソースのコストモデリングを実行する: ワークロードまたは概念実証を、テストする特定のリソースタイプとサイズを持つ別のアカウントにデプロイします。テストデータを使用してワークロードを実行し、出力結果のほか、テスト実行時のコストデータを記録します。その後、ワークロードを再デプロイするか、リソースタイプとサイズを変更して、テストをもう一度実行します。コストモデリングの際には、これらのリソースで使用する可能性のある製品のライセンス料と、これらのリソースをデプロイおよび管理する推定運用 (作業またはエンジニア) コストを含めます。一定期間 (時間単位、日次、月次、年次、三年次) のコストモデリングを考慮します。

リソース

関連するドキュメント:

- [AWS Auto Scaling](#)
- [Identifying Opportunities to Right Size](#) (サイズ適正化の機会を特定する)
- [Amazon CloudWatch の機能](#)

- [コスト最適化: Amazon EC2 サイズの適正化](#)
- [AWS Compute Optimizer](#)
- [AWS 料金見積りツール](#)

関連する例:

- [Perform a Data-Driven Cost Modelling](#) (データ駆動型のコストモデリングを実行する)
- [計画した AWS リソース設定のコストを見積もる](#)
- [適切な AWS ツールの選択](#)

COST06-BP02 データに基づいてリソースタイプ、リソースサイズ、リソース数を選択する

ワークロードとリソースの特性に関するデータに基づいて、リソースのサイズやタイプを選択します。例えば、コンピューティング、メモリ、スループット、書き込み頻度などです。この選択は通常、以前の(オンプレミス)バージョンのワークロード、ドキュメント、ワークロードに関する他の情報ソースを用いて行います。

このベストプラクティスが確立されていない場合のリスクレベル: 中

実装のガイダンス

Amazon EC2 では、さまざまなユースケースに合わせて、CPU、メモリ、ストレージ、ネットワークキャパシティのレベルが異なる、幅広いインスタンスタイプを選択肢として用意しています。インスタンスタイプごとに CPU、メモリ、ストレージ、ネットワーク機能の組み合わせが異なるため、プロジェクトに適したリソースの組み合わせを柔軟に選択できます。どのインスタンスタイプにもサイズが複数用意されており、ワークロードの需要に基づいてリソースを調整できます。必要なインスタンスタイプを判断するには、インスタンスで実行する予定のアプリケーションまたはソフトウェアのシステム要件に関する詳細情報を収集する必要があります。これらの詳細には、次の内容を含める必要があります。

- オペレーティングシステム
- CPU コア数
- GPU コア
- システムメモリ (RAM) の容量
- ストレージタイプとスペース
- ネットワーク帯域幅要件

コンピューティング要件の目的と必要なインスタンスを明らかにしたうえで、さまざまな Amazon EC2 インスタンスファミリーを検討してください。次のインスタンスタイプファミリーが提供されています。

- 汎用
- コンピューティング最適化
- メモリ最適化
- ストレージ最適化
- 高速コンピューティング
- HPC 最適化

各 Amazon EC2 インスタンスファミリーで達成可能な具体的な目的とユースケースの詳細については、「[AWS インスタンスタイプ](#)」を参照してください。

お客様のニーズに最適な特定のインスタンスファミリーとインスタンスタイプを選択するには、システム要件の収集が不可欠です。インスタンスタイプ名は、ファミリー名とインスタンスサイズで構成されます。例えば、t2.micro インスタンスは T2 ファミリーに属するマイクロサイズです。

ワークロードとリソースの特性 (例えば、コンピューティング、メモリ、スループット、書き込み頻度) に基づいて、リソースのサイズやタイプを選択します。この選択は通常、コストモデリング、以前のバージョンのワークロード (オンプレミスバージョンなど)、ドキュメント、ワークロードに関する他の情報ソース (ホワイトペーパー、公開ソリューション) を用いて行います。AWS 料金見積りツールやコスト管理ツールを使用すれば、十分な判断材料を基にインスタンスのタイプ、サイズ、構成を決定できます。

実装手順

- データに基づいてリソースを選択する: コストモデリングのデータを用いて、予想されるワークロードの使用レベルを選定し、特定されたリソースタイプとサイズを選択します。コストモデリングデータに基づいて、インスタンスに求められるデータ転送速度を考慮しつつ、仮想 CPU の数、総メモリ (GiB)、ローカルインスタンスストアボリューム (GB)、Amazon EBS ボリューム、ネットワークパフォーマンスレベルを決定します。常に詳細な分析と正確なデータに裏付けられた選択を行い、パフォーマンスの最適化とコスト管理の効率化を両立させましょう。

リソース

関連するドキュメント:

- [AWS インスタンスタイプ](#)
- [AWS Auto Scaling](#)
- [Amazon CloudWatch の機能](#)
- [EC2 Right Sizing によるコスト最適化](#)

関連動画:

- [Selecting the right Amazon EC2 instance for your workloads](#)
- [Right size your service](#)

関連する例:

- [It just got easier to discover and compare Amazon EC2 instance types](#)

COST06-BP03 メトリクスに基づいて自動的にリソースタイプ、リソースサイズ、リソース数を選択する

現在実行しているワークロードからのメトリクスを用いて、コストを最適化する適切なサイズやタイプを選択します。コンピューティング、ストレージ、データ、ネットワーキングなどのサービスに対して、適切なスループット、サイジング、ストレージのプロビジョニングを行います。これは、自動スケーリングなどのフィードバックループまたはワークロードのカスタムコードで行うことができます。

このベストプラクティスが確立されていない場合のリスクレベル: 低

実装のガイダンス

ワークロード内に、実行中のワークロードのアクティブなメトリクスを使用してそのワークロードを変更するフィードバックループを作成します。[AWS Auto Scaling](#) などのマネージドサービスを使用して、適切なサイジング操作を実行できます。AWS では、[API](#) や [SDK](#) のほか、最小限の労力でリソースを変更す

るための機能も提供しています。Amazon EC2 インスタンスの停止と起動のワークロードをプログラムして、インスタンスサイズやインスタンスタイプを変更できます。これにより、適切なサイジングによる利点が得られるだけでなく、変更に必要なほぼすべての運用コストを削減することもできます。

AWS サービスの中には、[Amazon Simple Storage Service Intelligent-Tiering](#) のように、タイプやサイズを自動的に選択する機能が組み込まれているものがあります。Amazon S3 Intelligent-Tiering では、使用パターンに基づいて、高頻度アクセスと低頻度アクセスの 2 つのアクセスティア間でデータが自動的に移動します。

実装手順

- ワークロードのメトリクスを設定して可観測性を高める: ワークロードの主要なメトリクスを取得します。これらのメトリクスは、ワークロード出力などのカスタマーエクスペリエンスに関する示唆を提供し、CPU やメモリの使用状況などのリソースのタイプとサイズの違いに合わせて調整されます。コンピューティングリソースの場合、パフォーマンスデータを分析して Amazon EC2 インスタンスのサイズを適切に設定します。アイドル状態のインスタンスと利用率の低いインスタンスを特定します。注目すべき重要なメトリクスは、CPU 使用率とメモリ使用率です (たとえば、[AWS Compute Optimizer およびメモリ使用率の有効化による適切なサイジング](#)で説明したように、90% の時間で 40% の CPU 使用率)。4 週間の最大 CPU 使用率およびメモリ使用率が 40% 未満のインスタンスを特定します。これらのインスタンスは、コスト削減のために適切なサイズを設定する必要があります。Amazon S3 などのストレージリソースについては、[Amazon S3 ストレージレンズ](#)を使用することで、バケットレベルでさまざまなカテゴリの 28 のメトリクスを確認でき、デフォルトで 14 日間の履歴データをダッシュボードで確認することが可能です。Amazon S3 ストレージレンズのダッシュボードは、要約、コスト最適化、またはイベントごとにフィルタリングして特定のメトリクスを分析できます。
- 適切なサイジングのレコメンデーションを表示する: AWS Compute Optimizer の適切なサイジングのレコメンデーションとコスト管理コンソールの Amazon EC2 の適切なサイジングツールを使用するか、リソースの適切なサイジングを行う AWS Trusted Advisor を確認してワークロードを調整します。さまざまなリソースの適切なサイジングを行う場合、[適切なツール](#)を使用し、Amazon EC2 インスタンス、AWS ストレージクラス、または Amazon RDS インスタンスタイプのいずれであっても、[適切なサイジングのガイドライン](#)に従うことが重要です。ストレージリソースには、Amazon S3 ストレージレンズを使用できます。これにより、オブジェクトストレージの使用状況、アクティビティの傾向を可視化し、コストを最適化してデータ保護のベストプラクティスを適用するための実用的な推奨事項を作成できます。[Amazon S3](#) ストレージレンズが組織全体のメトリクスの分析から得た状況に応じた推奨事項を使用して、ストレージを最適化するための手順をすぐに実行できます。
- メトリクスに基づいて自動的にリソースタイプとサイズを選択する: ワークロードメトリクスを使用して、ワークロードリソースを手動または自動で選択します。コンピューティングリソースの場合、AWS Auto Scaling を設定したり、アプリケーション内でコードを実装したりすると、頻繁な変更が要求される場合に必要となる労力を減らすことができるほか、手動プロセスより早く変更を実装できる可能性もあります。1 つの Auto Scaling グループ内でオンデマンドインスタンスとスポットインスタンスのフリートを起動し、自動的にスケールできます。スポットインスタンスの利用による割引に加え、リザーブドインスタンスや Savings Plans を利用して、通常のオンデマンドインスタンスの価格から割引を受けることができます。これらの要素をすべて組み合わせることで、Amazon EC2 インスタンスのコスト削減を最適化し、アプリケーションに必要なスケールとパフォーマンスを決定できます。また、[Auto Scaling グループ \(ASG\) で属性ベースのインスタンスタイプ選択 \(ABS\)](#) 戦略を使用すると、vCPU、メモリ、ストレージなどの属性セットとしてインスタンスの要件を表現できます。新しい世代のインスタンスタイプがリリースされると自動的に使用し、さらに Amazon EC2 スポットインスタンスでより広い範囲のキャパシティにアクセスできます。Amazon EC2 フリートと Amazon EC2 Auto Scaling が指定した属性に適合するインスタンスを選択して起動するため、手動でインスタンスタイプを選択する必要がなくなります。ストレージリソースには、[Amazon S3 Intelligent Tiering](#) と [Amazon EFS 低頻度アクセス機能](#)を利用できます。これらの機能のおかげで、データアクセスのパターンが変化しても、パフォーマンスへの影響や運用上のオーバーヘッドなしに、ストレージコストを自動的に削減するストレージクラスを自動選択できるようになります。

リソース

関連するドキュメント:

- [AWS Auto Scaling](#)
- [AWS の適切なサイジング](#)
- [AWS Compute Optimizer](#)
- [Amazon CloudWatch の機能](#)
- [CloudWatch の準備](#)
- [CloudWatch カスタムメトリクスの発行](#)
- [Amazon EC2 Auto Scaling の開始方法](#)
- [Amazon S3 ストレージレンズ](#)
- [Amazon S3 Intelligent-Tiering](#)
- [Amazon EFS 低頻度アクセス](#)
- [SDK を使用して Amazon EC2 インスタンスを起動する](#)

関連動画:

- [サービスの適切なサイズ設定](#)

関連する例:

- [Amazon EC2 フリート向け Auto Scaling 用の属性ベースのインスタンスタイプの選択](#)
- [スケジュールされたスケーリングを使用したコストに対する Amazon Elastic Container Service の最適化](#)
- [Amazon EC2 Auto Scaling で予測スケーリング](#)
- [Optimize Costs and Gain Visibility into Usage with Amazon S3 Storage Lens](#) (Amazon S3 ストレージレンズを使用してコストを最適化し、使用状況を可視化する)
- [Well-Architected ラボ: 適切なサイジングのレコメンデーション \(レベル 100\)](#)
- [Well-Architected ラボ: AWS Compute Optimizer および メモリ使用率の有効化によるサイズ適正化 \(レベル 200\)](#)

最適な料金モデルを選択する

ベストプラクティス

- [COST07-BP01 料金モデルの分析を実行する \(p. 65\)](#)
- [COST07-BP02 コストに基づいてリージョンを選択する \(p. 67\)](#)
- [COST07-BP03 費用対効果の高い条件を提供するサードパーティーの契約を選択する \(p. 69\)](#)
- [COST07-BP04 このワークロードのすべてのコンポーネントに対して料金モデルを実装する \(p. 70\)](#)
- [COST07-BP05 管理アカウントレベルで料金モデル分析を実行する \(p. 72\)](#)

COST07-BP01 料金モデルの分析を実行する

ワークロードの各コンポーネントを分析します。コンポーネントとリソースが長期間実行されるか (コミットメント割引)、動的および短期実行 (スポットまたはオンデマンド) とするかを決定します。コスト管理ツールのレコメンデーションを使用して、ワークロードに対して分析を行います。これらのレコメンデーションにはビジネスルールを適用して、高いリターンを実現します。

このベストプラクティスが確立されていない場合のリスクレベル: 高

実装のガイドンス

AWS には複数の[料金モデル](#)があり、組織のニーズと製品に合った最も費用対効果の高い方法でリソース料金を支払うことができます。チームと協力して最適な料金モデルを決定します。可用性にもとづいて決定すると、複数のオプションを組み合わせた料金モデルになることもよくあります

オンデマンドインスタンスを利用すると、コンピューティング容量またはデータベース容量の料金を、長期コミットメントや前払い金なしで、実行するインスタンスによって 1 時間単位または秒単位 (最小 60 秒) で支払うことができます。

Savings Plans は、1 年または 3 年の期間で一定の使用量 (1 時間あたりのドルで計測) をコミットする代わりに、Amazon EC2、Lambda、AWS Fargate (Fargate) を低価格で利用できる柔軟な料金モデルです。

スポットインスタンスは Amazon EC2 の料金の仕組みであり、予備のコンピューティング容量を割引価格の時間単価 (オンデマンド価格の最大 90% オフ) でリクエストでき、前払い金は不要です。

リザーブドインスタンスを利用すると、容量に対して前払いすることで、最大 75% の割引を受けることができます。詳細については、[予約を利用したコストの最適化](#)を参照してください。

本稼働、品質、開発の各環境に関連するリソースに、Savings Plan を含めることもできます。あるいは、サンドボックスリソースは必要なときのみ稼働するため、そういう環境のリソースにオンデマンドモデルを選択することもできます。Amazon [スポットインスタンス](#)を使用して Amazon EC2 のコストを削減したり、[コンピューティング Savings Plans](#)を使用して Amazon EC2、Fargate、Lambda のコストを削減したりします。[AWS Cost Explorer](#) レコメンデーションツールは、Savings Plans を利用したコミットメント割引の機会を提供します。

過去に Amazon EC2 の[リザーブドインスタンス](#)を購入したり、組織内で既にコスト配分を実践したりしている場合は、当面は Amazon EC2 リザーブドインスタンスを引き続き使用できます。ただし、より柔軟にコストを節約できる仕組みとして、いずれは Savings Plans を使用する戦略に則ることをお勧めします。AWS Cost Management の Savings Plans (SP) レコメンデーションを更新して、いつでも新しい Savings Plans レコメンデーションを作成できます。リザーブドインスタンス (RI) を使用して、Amazon RDS、Amazon Redshift、Amazon ElastiCache、Amazon OpenSearch Service のコストを削減します。Savings Plans とリザーブドインスタンスには、全額前払い、一部前払い、前払いなしの 3 つのオプションがあります。AWS Cost Explorer RI および SP 購入レコメンデーションで提供されたレコメンデーションを使用します。

スポットのワークロードを実行する機会を見つけるには、使用量全体の 1 時間ごとのビューを使用して、定期的に生じる使用量や伸縮性の変化を探します。スポットインスタンスは、耐障害性と柔軟性が高いさまざまなアプリケーションに使用できます。これには、ステートレスウェブサーバー、API エンドポイント、ビッグデータアプリケーションや分析アプリケーション、コンテナ化されたワークロード、CI/CD、その他柔軟性の高いワークロードなどがあります。

Amazon EC2 および Amazon RDS インスタンスを、使用していないとき (就業後や週末) にオフにできるか、分析します。このアプローチによって、24 時間 365 日使用する場合と比較して 70% 以上のコスト削減になります。特定の時間にのみ使用できるようにする必要がある Amazon Redshift クラスターがある場合は、そのクラスターを一時停止して、後で再開できます。Amazon Redshift クラスターや Amazon EC2 および Amazon RDS インスタンスが停止すると、コンピューティングに対する請求が停止され、ストレージ料金のみが適用されます。

[オンデマンドキャパシティ予約](#) (ODCR) は料金割引ではないことに注意してください。キャパシティ予約は、リザーブドキャパシティでインスタンスを実行しているかどうかにかかわらず、同等のオンデマンド料金で課金されます。キャパシティ予約は、実行する予定のリソースに対して十分な容量を提供する必要がある場合に、検討します。ODCR は、必要がなくなればキャンセルできるため、長期コミットメントと結びつける必要はありませんが、Savings Plans またはリザーブドインスタンスが提供する割引のメリットを受けることもできます。

実装手順

- ワークロードの伸縮性を分析する: Cost Explorer の時間単位の粒度またはカスタムダッシュボードを使用して、ワークロードの伸縮性を分析します。実行中のインスタンス数の定期的な変更を調べます。短期間のインスタンスはスポットインスタンスまたはスポットフリースの候補です。
 - [Well-Architected Lab: Cost Explorer](#) (Well-Architected ラボ: コストの可視化)
 - [Well-Architected Lab: Cost Visualization](#) (Well-Architected ラボ: コストの可視化)
- 既存の料金契約を見直す: 長期にわたる必要性について、現在の契約やコミットメントを見直します。現在締結しているものと、それらのコミットメントをどの程度使用しているかを分析します。既存の契約による割引やエンタープライズ契約を活用します。[エンタープライズ契約](#)では、お客様のニーズに最適になるように契約を調整するオプションがあります。長期コミットメントの場合、リザーブド料金割引、リザーブドインスタンス、特定のインスタンスタイプ、インスタンスファミリー、AWS リージョン、アベイラビリティゾーンにおける Savings Plans を検討します。
- コミットメント割引分析を実行する: アカウントで Cost Explorer を使用して、Savings Plans とリザーブドインスタンスのレコメンデーションを確認します。必要な割引を適用し、リスクを認識した上で、正しいレコメンデーションを実装していることを確認するには、[Well-Architected ラボ](#)に従ってください。

リソース

関連するドキュメント:

- [リザーブドインスタンスのレコメンデーションへのアクセス](#)
- [インスタンス購入オプション](#)
- [AWS エンタープライズ](#)

関連動画:

- [Save up to 90% and run production workloads on Spot \(最大 90% 節約し、Spot で本番環境のワークロードを稼働\)](#)

関連する例:

- [Well-Architected Lab: Cost Explorer](#) (Well-Architected ラボ: コストの可視化)
- [Well-Architected Lab: Cost Visualization](#) (Well-Architected ラボ: コストの可視化)
- [Well-Architected Lab: Pricing Models](#) (Well-Architected ラボ: 料金モデル)

COST07-BP02 コストに基づいてリージョンを選択する

リソースの料金は各リージョンで異なる場合があります。リージョンによるコストの差異を特定し、レイテンシー、データレジデンシーおよびデータ主権に関する要件を満たす場合にのみ、よりコストの高いリージョンにデプロイします。リージョンコストを織り込むことで、このワークロードに対して支払う料金の合計を最低限に抑えることができます。

このベストプラクティスを活用しない場合のリスクレベル: 中

実装のガイダンス

それらの [AWS クラウド インフラストラクチャ](#) はグローバルであり、[世界中の複数の場所でホストされており](#)、AWS リージョン、アベイラビリティゾーン、Local Zones、AWS Outposts、Wavelength Zones を中心に構築されています。リージョンとは世界中の物理的な場所であり、各リージョンは、AWS

が複数のアベイラビリティゾーンを設置している地理的に離れた地域です。各リージョン内の複数の独立した場所であるアベイラビリティゾーンは、1つ以上の独立したデータセンターで構成されます。各データセンターは、冗長性のある電源、ネットワーク、接続を備えています。

各 AWS リージョン は現地マーケットの条件内で運用されており、土地代、回線代、電気代、税金などのコストが異なるため、リソース料金は各リージョンで異なります。世界的に最小料金で稼働できるように、ソリューションのコンポーネントまたは全体を運用する特定のリージョンを選択します。ここでは [AWS 計算ツール](#) を使用して、ロケーションタイプ (リージョン、Wavelength Zone、ローカルゾーン) とリージョンごとにサービスを検索し、さまざまなリージョンでワークロードのコストを見積もります。

ソリューションを設計する際、ユーザーに近いコンピューティングリソースの場所を探して、レイテンシー低下とデータ主権の強化を図ることが推奨されます。ビジネス、データプライバシー、パフォーマンス、セキュリティの要件に基づいて、地理的場所を選択します。エンドユーザーが世界中にいるアプリケーションの場合は、複数の場所を使用します。

データプライバシー、セキュリティ、ビジネス要件に義務がない場合は、AWS のサービスの料金がより低いリージョンを使用してワークロードをデプロイします。例えば、デフォルトのリージョンが ap-southeast-2 (シドニー) であり、他のリージョンを使用するにあたっての制約 (データプライバシー、セキュリティなど) がない場合、重要ではない (開発とテスト) Amazon EC2 インスタンスを north-east-1 (バージニア北部) リージョンにデプロイすると、コストを抑えることができます。

	コンプライアンス	レイテンシー	コスト	サービス/機能
リージョン 1	✓	15 MS	\$\$	✓
リージョン 2	✓	20 MS	\$\$\$	X
リージョン 3	✓	80 MS	\$	✓
リージョン 4	✓	15 MS	\$\$	✓
リージョン 5	✓	20 MS	\$\$\$	X
リージョン 6	✓	15 MS	\$	✓
リージョン 7	✓	80 MS	\$	✓
リージョン 8	✓	15 MS	\$	X

リージョン機能マトリックステーブル

前述のマトリックステーブルから、他のリージョンに比べてレイテンシーが低く、サービスが利用可能で、コストが最も低いリージョンであるため、このシナリオではリージョン 4 が最適なオプションであることがわかります。

実装手順

- AWS リージョン の料金を確認する: 現在のリージョンのワークロードコストを分析します。サービスおよび使用タイプ別の最も高いコストから、利用可能な他のリージョンのコストを計算します。予測される費用削減効果がコンポーネントまたはワークロードの移動コストを上回っている場合は、新しいリージョンに移行します。

- 複数のリージョンにデプロイする場合の要件を確認する: ビジネス要件と義務 (データプライバシー、セキュリティ、パフォーマンス) を分析して、複数リージョンを使用すべきでない制約があるかどうかを確認します。単一リージョンを使用するよう制限する義務がない場合は、複数のリージョンを使用します。
- 必要なデータ転送を分析する: リージョンを選択する際は、データ転送コストを考慮します。データは顧客に近く、またリソースに近いところに置いてください。データ転送が最小限でデータの流れがいい、よりコストの低い AWS リージョンを選択します。データ転送のビジネス要件によって、[Amazon CloudFront](#)、[AWS PrivateLink](#)、[AWS Direct Connect](#)、[AWS Virtual Private Network](#) を使用して、ネットワークコストの削減、パフォーマンスの向上、セキュリティの強化を実現できます。

リソース

関連するドキュメント:

- [リザーブドインスタンスのレコメンデーションへのアクセス](#)
- [Amazon EC2 料金](#)
- [インスタンス購入オプション](#)
- [リージョン別の表](#)

関連動画:

- [Save up to 90% and run production workloads on Spot](#)

関連する例:

- [Overview of Data Transfer Costs for Common Architectures \(一般的なアーキテクチャでのデータ転送コストの概要\)](#)
- [Cost Considerations for Global Deployments \(グローバルデプロイにおけるコストの考慮事項\)](#)
- [What to Consider when Selecting a Region for your Workloads \(ワークロードに応じたリージョンを選択する際の注意点\)](#)
- [Well-Architected ラボ: Restrict service usage by Region \(Level 200\) \(リージョンごとにサービスの仕様を制限する \(レベル 200\)\)](#)

COST07-BP03 費用対効果の高い条件を提供するサードパーティーの契約を選択する

コスト効率に優れた契約と条件により、これらのサービスのコストが、提供されるメリットに見合ったものとなります。組織に追加のメリットを提供するときに、それに合わせてスケーリングする契約と料金を選択します。

このベストプラクティスが確立されていない場合のリスクレベル: 中

実装のガイダンス

クラウド環境のコスト管理に役立つさまざまな製品が流通しています。こうした製品は、ターゲットとなる顧客の要件に応じて、コストガバナンスやコスト可視化を重視したものや、コスト最適化を重視したものなど、機能面で違いが見られる場合があります。効果的なコスト最適化とガバナンスの鍵を握る要因の1つは、料金モデルが適正で、必要な機能が揃った適切なツールを使用することです。製品ごとに料金モデルは異なります。月額請求総額の一定割合を支払うものや、実際に削減したコストの一部 (一定割合) を支払うものがありますが、必要な分だけ支払う従量課金制が理想的です。

クラウドでサードパーティーのソリューションやサービスを利用する場合は、期待する成果に合わせて料金体系を選ぶことが重要です。料金は、コスト最適化の結果とサービスの価値に合わせてスケールする必要があります。例えば、実際に削減できたコストの一部を支払う成功報酬型のソフトウェアの場合、削減率が上がるほど(成果)、請求額も高くなります。支出負担が増えるに従い、支払う金額も増えるライセンス契約は、コスト最適化の点で必ずしも最適解とは限りません。ただし、請求書のあらゆる項目で優遇を受けられる場合、こうした変動料金が妥当なケースもあるかもしれません。

例えば、Amazon EC2 のレコメンデーションを提供し、請求総額の一定割合を課金するソリューションでは、優遇なしの他のサービスを利用すると、割高になる場合があります。もう 1 つの例は、管理対象となるリソースのコストを一定の割合で支払うマネージドサービスです。インスタンスサイズが大きくなっても必ずしも管理の負担が増えるわけではありませんが、請求額は高くなります。こうしたサービス料金設定に、コスト最適化のプログラムや、効率を向上するサービス機能が含まれていることを確認してください。

顧客は市場に流通しているこれらの製品を比較的高度である、または使いやすいと受け止めるかもしれません。こうした製品のコストを考慮し、長期的に見てコスト最適化の可能性があるかどうかを考える必要があります。

実装手順

- サードパーティーの契約と諸条件を分析する: サードパーティーの契約の料金を確認します。さまざまな使用レベルに対応したモデリングを行い、新しいサービスの使用や、ワークロードの増加による現在のサービスの増加など、新たなコストを考慮します。追加コストによってビジネスに必要なメリットが得られるかどうかを判断します。

リソース

関連するドキュメント:

- [リザーブドインスタンスのレコメンデーションへのアクセス](#)
- [インスタンス購入オプション](#)

関連動画:

- [Save up to 90% and run production workloads on Spot](#)

COST07-BP04 このワークロードのすべてのコンポーネントに対して料金モデルを実装する

永続的に実行されるリソースでは、Savings Plans やリザーブドインスタンスなどのリザーブドキャパシティを利用する必要があります。短期的な使用には、スポットインスタンスまたはスポットフリートを使用するように設定します。オンデマンドインスタンスは、中断することのできない、かつリザーブドキャパシティに対して長時間稼働しない短期ワークロードに対してのみ使用されます(リソースタイプに応じて、期間の 25% から 75%)。

このベストプラクティスが確立されていない場合のリスクレベル: 低

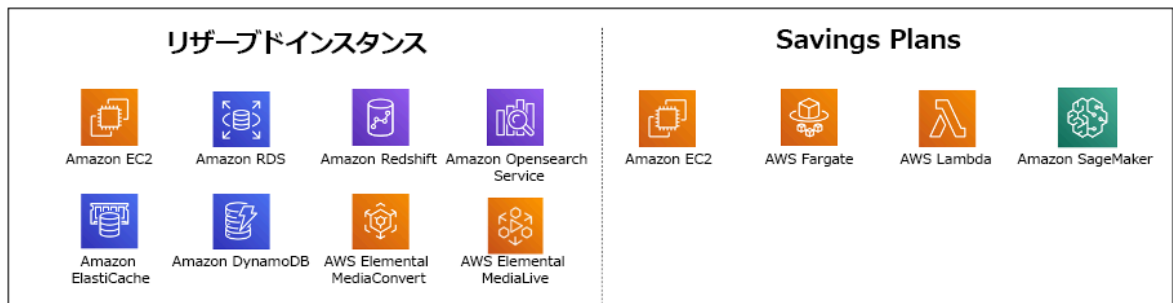
実装のガイダンス

コスト効率を上げるために、AWS は過去の使用状況に基づいて確約利用(コミットメント)のレコメンデーションをいくつか提示します。これらのレコメンデーションを参考にして、実際に節約できるコストと、そのコミットメントの活用法を理解できます。これらのサービスをオンデマンドまたはスポットで利用することも、一定期間の利用を確約してリザーブドインスタンス(RI)や Savings Plans(SP)でオンデマ

ンドコストを削減することもできます。ワークロードを最適化するには、各ワークロードコンポーネントや複数の AWS サービスだけでなく、該当するサービスのコミットメント割引、購入オプション、スポットインスタンスについても理解する必要があります。

ワークロードのコンポーネントの要件を検討し、これらのサービスのさまざまな料金モデルを理解しましょう。コンポーネントの可用性要件を定義します。ワークロードで関数を実行する複数の独立したリソースの有無、ワークロードの継続的に必要となる要件を確認します。デフォルトのオンデマンド料金モデルと他の適用可能なモデルを使用して、リソースのコストを比較します。リソースまたはワークロードコンポーネントで変更可能なものはすべて考慮します。

例えば、AWS のこのウェブアプリケーションアーキテクチャを検討してみましょう。このサンプルワークロードは、Amazon Route 53、AWS WAF、Amazon CloudFront、Amazon EC2 インスタンス、Amazon RDS インスタンス、ロードバランサー、Amazon S3 ストレージ、Amazon Elastic File System (Amazon EFS) など、複数の AWS サービスで構成されています。これらのサービスをそれぞれ見直し、さまざまな料金モデルでコストをどれくらい削減できるのかを確認する必要があります。RI または SP を利用できるものもあれば、オンデマンドでしか利用できないものもあります。次の図からわかるように、AWS の一部のサービスは利用を確約し、RI または SP を使用できます。



リザーブドインスタンスと Savings Plans で確約利用した AWS サービス

実装手順

- 料金モデルを実装する: 分析結果に基づいて、Savings Plans やリザーブドインスタンスを購入するか、スポットインスタンスを実装します。コミットメントの初回購入時には、リストの上位 5 件または 10 件のレコメンデーションを選択し、翌月または翌々月までの結果をモニタリングして分析します。このプロセスは AWS Cost Management Console が案内してくれます。コンソールから RI または SP のレコメンデーションを確認し、その内容 (タイプ、支払い、期間) をカスタマイズし、時間単位の確約利用料 (例えば、1 時間あたり 20 USD) を確認して、カートに追加します。割引は、対象となる使用量に自動的に適用されます。コミットメント割引で定期的に少量を購入します (例: 2 週間ごと、または 1 か月ごと)。中断可能またはステートレスなワークロードにスポットインスタンスを実装します。最後に、Amazon EC2 オンデマンドインスタンスを選択し、残りの要件にリソースを割り当てます。
- ワークロードレビューサイクル: ワークロードに対し、特に料金モデルのカバレッジを分析するレビューサイクルを実装します。ワークロードが必要なカバレッジを達成したら、部分的に (2 か月ごと)、または組織の使用状況の変化に応じて、追加のコミットメント割引を購入します。

リソース

関連するドキュメント:

- [Understanding your Savings Plans recommendations](#)
- [リザーブドインスタンスのレコメンデーションへのアクセス](#)
- [リザーブドインスタンスの購入方法](#)
- [インスタンス購入オプション](#)
- [スポットインスタンス](#)
- [Reservation models for other AWS services](#)

- [Savings Plans Supported Services](#)

関連動画:

- [Save up to 90% and run production workloads on Spot](#)

関連する例:

- [Savings Plans を購入する前に考慮すべきことは何ですか？](#)
- [Cost Explorer で使用率とコストを分析する方法を教えてください](#)

COST07-BP05 管理アカウントレベルで料金モデル分析を実行する

請求やコスト管理ツールをチェックして、コミットメントや予約を利用した推奨される割引を確認し、管理アカウントレベルで定期的に分析を実行します。

このベストプラクティスが確立されていない場合のリスクレベル: 低

実装のガイダンス

コストモデリングを定期的に行うと、複数のワークロードにまたがって最適化する機会が得られます。例えば、複数のワークロードでオンデマンドインスタンスを使用している場合、集計レベルでは変更リスクが低くなり、コミットメントベースの割引を運用すると全体的なコストが低くなります。2週間から1か月の定期的なサイクルで分析を実行することを推奨します。これにより、調整のための小口購入が可能になり、ワークロードやコンポーネントの変更に合わせて料金モデルの調整を続けることができます。

[AWS Cost Explorer](#) のレコメンデーションツールを使用して、管理アカウントでコミットメント割引を適用する機会を見つけます。管理アカウントレベルでのレコメンデーションは、リザーブドインスタンスまたは Savings Plans (SP) 割引共有が有効になっている AWS 組織内のすべてのアカウントにまたがる使用量を考慮して計算され、アカウント全体で節約を最大化できるコミットメントを推奨します。

管理アカウントレベルでの購入は、多くの場合、最大限の節約を目指して最適化されますが、特定の連結アカウントでの利用に最初に割引を適用したい場合など、連結アカウントレベルで SP を購入することを検討する状況もあります。メンバーアカウントのレコメンデーションは個別のアカウントレベルで計算され、それぞれのアカウントごとに節約を最大化します。アカウントが RI と SP の両方のコミットメントを所有している場合、それらは次の順序で適用されます。

ゾーン RI > 標準 RI > コンバーチブル RI > Instance Savings Plan > Compute Savings Plan

管理アカウントレベルで SP を購入した場合、割引率の高いものから低いものに基づいて節約が適用されます。管理アカウントレベルの SP は、すべての連結アカウントを調べて、割引が最も高いところで節約が適用されます。節約が適用される場所を制限したい場合は、連結アカウント単位で Savings Plan を購入すると、そのアカウントが対象となるコンピューティングサービスを実行しているときはいつでも、割引が最初に適用されます。アカウントが対象となるコンピューティングサービスを実行していない場合、割引は同じ管理アカウントにある他の連結アカウント間で分配されます。割引共有はデフォルトでオンになっていますが、必要に応じてオフにできます。

一括請求ファミリーでは、Savings Plans は最初に所有者アカウントの使用に適用され、次に他のアカウントの使用に適用されます。これは、共有を有効にしている場合にのみ発生します。Savings Plans は、最も高い節約率に最初に適用されます。節約率が等しい使用量が複数ある場合は、Savings Plans のレー

トが最も低い最初の使用に Savings Plans が適用されます。Savings Plans は、残りの使用量がなくなるか、コミットメントが使い果たされるまで引き続き適用されます。残りの使用量はオンデマンド料金で請求されます。AWS Cost Management の Savings Plans レコメンデーションを更新して、いつでも新しい Savings Plans レコメンデーションを作成できます。

インスタンスの柔軟性を分析すると、レコメンデーションに沿ってコミットできます。コストモデリングを作成するにあたって、さまざまなリソースオプションの可能性を含めたワークロードの短期コストを分析し、AWS 料金モデルの分析を行い、それらをビジネス要件に合わせて、総保有コストと [コスト最適化](#) の機会を見出します。

実装手順

コミットメント割引分析を実行する: アカウントで Cost Explorer を使用して、Savings Plans とリザーブドインスタンスのレコメンデーションを確認します。Savings Plan のレコメンデーションを理解し、月次費用の見積もりと、月次節約額の見積もりを行っていることを確認します。リザーブドインスタンスまたは Savings Plans 割引共有が有効になっている AWS 組織内のすべてのアカウントにまたがる使用量を考慮して計算された管理アカウントレベルでのレコメンデーションをレビューし、アカウント全体で節約を最大化します。Well-Architected ラボに従って、必要な割引を適用し、リスクを認識したうえで、正しいレコメンデーションを実装していることを確認します。

リソース

関連するドキュメント:

- [AWS の料金の仕組みはどのようになっていますか?](#)
- [インスタンス購入オプション](#)
- [Saving Plan Overview \(Savings Plans の概要\)](#)
- [Saving Plan recommendations \(Savings Plans のレコメンデーション\)](#)
- [リザーブドインスタンスのレコメンデーションへのアクセス](#)
- [Understanding your Saving Plans recommendation](#)
- [How Savings Plans apply to your AWS usage](#)
- [一括請求を利用した Saving Plans](#)
- [共有リザーブドインスタンスと Savings Plans の割引の有効化](#)

関連動画:

- [Save up to 90% and run production workloads on Spot \(最大 90% 節約し、Spot で本番環境のワークロードを稼働\)](#)

関連する例:

- [AWS Well-Architected Lab: Pricing Models \(Level 200\)](#)
- [AWS Well-Architected Labs: Pricing Model Analysis \(Level 200\)](#)
- [Savings Plans を購入する前に考慮すべきことは何ですか?](#)
- [How can I use rolling Savings Plans to reduce commitment risk?](#)
- [When to Use Spot Instances \(スポットインスタンスの使用が適している場合\)](#)

データ転送を計画する

クラウドの利点は、マネージド型のネットワークサービスであることです。スイッチ、ルーター、その他の関連するネットワーク機器などのフリートの管理や運用は不要になります。クラウド内のネットワーキ

ングリソースは CPU とストレージと同じように消費され、同じように実際に使用した分だけを支払うこととなります。クラウドでコストを最適化するには、ネットワーキングリソースを効率的に使用する必要があります。

ベストプラクティス

- [COST08-BP01 データ転送モデリングを実行する \(p. 74\)](#)
- [COST08-BP02 データ転送コストを最適化するコンポーネントを選択する \(p. 75\)](#)
- [COST08-BP03 データ転送コストを削減するサービスを実装する \(p. 77\)](#)

COST08-BP01 データ転送モデリングを実行する

組織の要件を取りまとめ、ワークロードとその各コンポーネントのデータ転送モデリングを実行します。これにより、現在のデータ転送要件に対する最低コストを特定できます。

このベストプラクティスが確立されていない場合のリスクレベル: 高

実装のガイダンス

クラウドでソリューションを設計する際、習慣的にオンプレミスのデータセンターを使用してアーキテクチャを設計してしまったり、知識が不足していたりするせいで、データ転送料金を見落としがちです。AWS のデータ転送料金は、送信元、送信先、トラフィック量によって決まります。設計段階からこれらの料金を織り込めば、コスト削減につながる可能性があります。総保有コスト (TCO) を正確に見積もるには、ワークロードのどこでデータ転送が発生するかや転送のコスト、関連するメリットを把握することがきわめて重要です。これにより、十分な情報に基づいてアーキテクチャ設計上の変更や承諾の決定ができます。たとえば、アベイラビリティゾーン間でデータをレプリケートするマルチアベイラビリティゾーンを設定したとします。

ワークロードにおいてデータを転送するサービスコンポーネントをモデリングし、求められる信頼性と耐障害性を実現するために、これが許容されるコスト (両方のアベイラビリティゾーンのコンピューティングとストレージに支払うのと同様) であると判断します。さまざまな使用量のレベルでコストをモデリングします。ワークロード使用量は経時変化します。また、サービスの種類ごとに異なるレベルで費用対効果が向上する場合があります。

データ転送をモデリングする際には、取り込まれるデータの量とデータの転送元を考慮します。また、処理されるデータ量と、必要なストレージやコンピューティングのキャパシティについても検討してください。モデリング中は、ワークロードのアーキテクチャに即したネットワークのベストプラクティスに従い、見込まれるデータ転送コストを最適化してください。

AWS Pricing Calculator を使用して、特定の AWS サービスのコストの見積りと、予想されるデータ転送を確認できます。ワークロードを (テスト目的または実稼働前の環境で) 既に実行している場合は、[AWS Cost Explorer](#) または [AWS Cost and Usage Report \(CUR\)](#) を使用してデータ転送コストを把握し、モデリングします。PoC (概念実証) を設定するか、またはワークロードをテストして、現実的な条件でシミュレートされた負荷を用いてテストを実行します。ワークロードのさまざまな需要に応じてコストをモデリングできます。

実装手順

- 要件を特定する: 送信元と送信先の間で予定されているデータ転送の主な目標とビジネス要件は何ですか? 最終的にどのようなビジネス成果を期待していますか? ビジネス要件を収集し、期待される成果を定義します。
- 送信元と送信先を特定する: データ転送の送信元と送信先はどこですか (AWS リージョン 内の転送、AWS サービスへの転送、インターネットへの発信など)?
 - [AWS リージョン 内のデータ転送](#)
 - [AWS リージョン 間のデータ転送](#)

- [インターネットへのデータ転送](#)
- データ分類を特定する: 転送されるデータの分類はどうなっていますか? データの種類は? データの大きさは? データ転送の頻度は? 機密データですか?
- 使用する AWS サービスまたはツールを特定する: このデータ転送にはどの AWS サービスを使用しますか? 既にプロビジョニングされているサービスを別のワークロードに使用できますか?
- データ転送コストを計算する: [AWS の料金ページ](#)で、前もって作成したデータ転送モデルを使用して、ワークロードのデータ転送コストを計算します。ワークロードの使用量が増減した場合の、使用量別のデータ転送コストを計算します。ワークロードアーキテクチャに複数のオプションがある場合は、比較のために各オプションのコストを計算します。
- コストを結果にリンクする: 発生したデータ転送コストごとに、ワークロードで達成した結果を指定します。コンポーネント間の転送であればデカップリングのため、アベイラビリティゾーン間の転送であれば冗長性のためかもしれません。
- データ転送モデルを作成する: すべての情報を収集したら、複数のユースケースやさまざまなワークロードの基準となる、データ転送の概念モデルを作成します。

リソース

関連するドキュメント:

- [AWS caching solutions](#)
- [AWS の料金](#)
- [Amazon EC2 の料金](#)
- [Amazon VPC の料金](#)
- [Understanding data transfer charges](#)

関連動画:

- [Monitoring and Optimizing Your Data Transfer Costs](#)
- [S3 Transfer Acceleration](#)

関連する例:

- [Overview of Data Transfer Costs for Common Architectures](#)
- [AWS AWS 規範ガイド](#)

COST08-BP02 データ転送コストを最適化するコンポーネントを選択する

すべてのコンポーネントが選択され、データ転送コストを低減するようアーキテクチャが設計されます。これには、ワイドエリアネットワーク (WAN) 最適化やマルチアベイラビリティゾーン (AZ) 設定などのコンポーネントの使用が含まれます。

このベストプラクティスが確立されていない場合のリスクレベル: 中

実装のガイドンス

データ転送を念頭に置いたアーキテクチャでは、データ転送コストを最小限に抑えることができます。このアーキテクチャでは、コンテンツ配信ネットワークを使用してユーザーに近いデータを特定したり、お客様のプレミスと AWS をつなぐ専用ネットワーク接続が使用される場合があります。WAN の最適化やアプリケーションの最適化によって、コンポーネント間で転送されるデータ量を減らすこともできます。

AWS クラウド との間やその中でデータを転送する場合、データ転送を最適化する適切な AWS サービスを選択するために、さまざまなユースケース、データの性質、利用可能なネットワークリソースに基づいて転送先を把握することが不可欠です。AWS は、多様なデータ移行要件に対応する幅広いデータ転送サービスを提供しています。組織内のビジネスニーズに基づいて、適切な[データストレージ](#)と[データ転送](#)のオプションを選択してください。

ワークロードアーキテクチャを計画または確認するときは、次の点を考慮してください。

- AWS 内で VPC エンドポイントを使用する: VPC エンドポイントを使用すれば、VPC とサポート対象の AWS サービスとの間にプライベート接続を確立できます。これで、データ転送コストが発生する可能性のある公開インターネットの使用を回避できます。
- NAT ゲートウェイを使用する: [NAT ゲートウェイ](#)を配置して、プライベートサブネットのインスタンスがインターネットまたは VPC 外部のサービスに接続できるようにします。NAT ゲートウェイの背後にあるリソースのうち、最多量のトラフィックを送信しているリソースのアベイラビリティゾーンが、NAT ゲートウェイと同じかどうかを確認してください。違う場合は、そのリソースと同じアベイラビリティゾーンに新しい NAT ゲートウェイを設置し、AZ 間のデータ転送料金を削減します。
- AWS Direct Connect を使用する: AWS Direct Connect は、公開インターネットをバイパスし、オンプレミスネットワークと AWS との間にプライベート接続を直接確立します。インターネットを介して大量のデータを転送するよりも、この方がコスト効率が高く、堅実です。
- リージョン境界をまたいでデータを転送しない: AWS リージョン間 (あるリージョンから別のリージョンへの) データ転送には通常、料金がかかります。複数のリージョンをまたいで転送する場合は、慎重に検討したうえで決断してください。詳細については、「[複数リージョンのシナリオ](#)」を参照してください。
- データ転送を監視する: Amazon CloudWatch および [VPC フローログ](#)を使用して、データ転送とネットワークの使用量に関する詳細情報をキャプチャします。VPC 内でネットワークインターフェイスとの間を行き来するネットワークトラフィックについて、IP アドレスや範囲などのキャプチャされた情報を分析してください。
- ネットワーク使用状況を分析する: AWS Cost Explorer、CUDOS Dashboard、CloudWatch などの計測とレポートのツールを使用して、ワークロードのデータ転送コストを把握してください。

実装手順

- データ転送用にコンポーネントを選択する: データ転送モデリング (「[データ転送モデリングを実行する \(p. 74\)](#)」を参照) を使用して、データ転送コストが現時点で最も高い箇所や、ワークロードの利用状況が変化した場合に最も高くなると予測される箇所に着目します。データ転送の必要性を排除または削減 (またはコストを削減) する代替アーキテクチャや追加のコンポーネントを探します。

リソース

関連するベストプラクティス:

- [COST08-BP01 データ転送モデリングを実行する \(p. 74\)](#)
- [COST08-BP03 データ転送コストを削減するサービスを実装する \(p. 77\)](#)

関連するドキュメント:

- [クラウドデータの移行](#)
- [AWS caching solutions](#)
- [Deliver content faster with Amazon CloudFront](#)

関連する例:

- [Overview of Data Transfer Costs for Common Architectures](#)

- [AWS Network Optimization Tips](#)
- [Optimize performance and reduce costs for network analytics with VPC Flow Logs in Apache Parquet format](#)

COST08-BP03 データ転送コストを削減するサービスを実装する

データ転送コストを削減するサービスを実装します。例えば、エッジロケーションやコンテンツ配信ネットワーク (CDN) を使用してエンドユーザーにコンテンツを配信する、アプリケーションサーバーまたはデータベースの前にキャッシュレイヤーを構築する、クラウドへの接続に VPN ではなく専用ネットワーク接続を使用するなどです。

このベストプラクティスが確立されていない場合のリスクレベル: 中

実装のガイダンス

ネットワークデータ転送の使用量を最適化するのに役立つさまざまな AWS サービスがあります。ワークロードのコンポーネント、種類、クラウドアーキテクチャにもよりますが、これらのサービスはクラウド上でのトラフィックの圧縮、キャッシュ、共有、分散に役立ちます。

- [Amazon CloudFront](#) は、低レイテンシーかつ高速な転送速度でデータを転送するグローバルなコンテンツ配信ネットワークです。世界中のエッジロケーションでデータをキャッシュすることで、お客様のリソースの負荷を軽減します。CloudFront を使用してレイテンシーを最低限に抑え、世界中の多数のユーザーにコンテンツを配信するための管理労力を軽減できます。[Security Savings Bundle](#) を使用すると、経時的に使用量を増加させる計画がある場合に、CloudFront の使用量を最大 30% 節約できます。
- [AWS Direct Connect](#) では、AWS への専用ネットワーク接続を確立できます。このサービスによって、ネットワークコストの削減、帯域幅の増加、インターネット経由の接続よりも安定したネットワーク接続が可能になります。
- [AWS VPN](#) を使用すると、プライベートネットワークと AWS グローバルネットワークとの間に安全なプライベート接続を確立できます。シンプルな接続とフルマネージド型の伸縮自在なサービスは、小規模なオフィスやビジネスパートナーに最適です。
- [VPC エンドポイント](#) により、プライベートネットワークを利用して AWS のサービス間の接続が可能になり、パブリックデータ転送と [NAT ゲートウェイ](#) のコスト削減に利用できます。[ゲートウェイ VPC エンドポイント](#) では、時間単位の料金は発生せず、Amazon S3 と Amazon DynamoDB がサポートされます。[インターフェイス VPC エンドポイント](#) は [AWS PrivateLink](#) によって提供され、時間単位の料金と GB あたりの使用料がかかります。
- [NAT ゲートウェイ](#) にはスケーリングと管理機能が組み込まれており、スタンドアロンの NAT インスタンスとは対照的にコストを削減できます。NAT ゲートウェイはトラフィックの多いインスタンスと同じアベイラビリティゾーンに配置し、Amazon DynamoDB または S3 にアクセスが必要なインスタンスでは VPC エンドポイントを使用して、データ転送とデータ処理コストを削減することを検討してください。
- コンピューティングリソースを備えた [AWS Snow Family](#) デバイスを使用して、エッジでデータを収集および処理します。AWS Snow Family デバイス ([Snowcone](#)、[Snowball](#)、[Snowmobile](#)) を使用すると、ペタバイト規模のデータをコスト効率よくオフラインで AWS クラウドに移動できます。

実装手順

- サービスを実装する: データ転送モデリングを使用し、VPC フローログを確認して、サービスとワークロードタイプに基づいて適切な AWS ネットワークサービスを選択します。最大のコストと最大のボリュームフローがどこにあるかを調べます。AWS のサービスを確認し、転送を減らすか排除するサービス (特にネットワークとコンテンツ配信) があるかどうかを評価します。また、データへの繰り返しのアクセス、または大量のデータがあるキャッシュサービスを探します。

リソース

関連するドキュメント:

- [AWS Direct Connect](#)
- [AWS の製品](#)
- [AWS caching solutions](#)
- [Amazon CloudFront](#)
- [AWS Snow Family](#)
- [Amazon CloudFront Security Savings Bundle](#)

関連動画:

- [Monitoring and Optimizing Your Data Transfer Costs](#)
- [AWS コスト最適化シリーズ: CloudFront](#)
- [NAT ゲートウェイのデータ転送料金を削減するにはどうすればよいですか?](#)

関連する例:

- [How-to chargeback shared services: An AWS Transit Gateway example](#)
- [Understand AWS data transfer details in depth from cost and usage report using Athena query and QuickSight](#)
- [Overview of Data Transfer Costs for Common Architectures \(一般的なアーキテクチャでのデータ転送コストの概要\)](#)
- [Using AWS Cost Explorer to analyze data transfer costs](#)
- [Cost-Optimizing your AWS architectures by utilizing Amazon CloudFront features](#)
- [NAT ゲートウェイのデータ転送料金を削減するにはどうすればよいですか?](#)

需要を管理しリソースを供給する

クラウドに移行すると、お支払いは必要な分のみになります。必要な時にワークロードの需要に合わせたリソースを供給できるため、コストがかかる無駄なオーバプロビジョニングを排除できます。また、スロットル、バッファ、キューを使用して需要を変更すると、少ないリソースで需要を円滑に処理できます。

ジャストインタイム供給による経済的利点を得るには、リソース障害、高可用性、プロビジョニング時間を考慮したプロビジョニングの必要性とのバランスを保つことが必要です。需要が固定か変動的かに応じて、最小限の手間で環境を管理できるようにメトリクスと自動化の作成計画を立てます。スケーリング中でも同様のことが言えます。需要を変更するときは、ワークロードが許容できる最大遅延を把握しておく必要があります。

AWS では、需要の管理とリソースの供給にさまざまなアプローチを使用できます。以下の各セクションでは、こうしたアプローチのベストプラクティスを説明します。

ベストプラクティス

- [COST09-BP01 ワークロードの需要に関する分析を実行する \(p. 79\)](#)
- [COST09-BP01 需要を管理するためのバッファまたはスロットルを実装する \(p. 81\)](#)
- [COST09-BP03 リソースを動的に供給する \(p. 83\)](#)

COST09-BP01 ワークロードの需要に関する分析を実行する

ワークロードの需要を経時的に分析します。分析が季節的傾向を考慮し、ワークロードのライフタイム全体にわたる動作条件を正確に反映したものであることを確認します。分析を行う際には、費やされた時間がワークロードのコストに比例しているなどの潜在的利益を織り込む必要があります。

このベストプラクティスが確立されていない場合のリスクレベル: 高

実装のガイダンス

クラウドコンピューティングのワークロード需要を分析するには、クラウド環境で開始されるコンピューティングタスクのパターンと特性を理解する必要があります。この分析により、ユーザーはリソース割り当てを最適化し、コストを管理し、パフォーマンスが必要なレベルを満たしていることを確認することができます。

ワークロードの要件を把握します。組織の要件に、リクエストに対するワークロードの応答時間を含める必要があります。応答時間は、需要が管理されているかどうか、または需要を満たすためにリソースの供給を変更する必要があるかどうかを判断するために使用できます。

分析には、需要の予測可能性と再現性、需要の変化率、需要の変化量を含める必要があります。分析は、月末処理や休日のピークなどの時季的な変動が組み込まれるように、十分な期間にわたって実行します。

分析作業では、スケーリングの実装による潜在的な利点が反映されるようにします。コンポーネントの予想される合計コスト、ワークロードのライフタイムにおける使用量の増減およびコストの増減に注目します。

クラウドコンピューティングのワークロード需要分析を行う際に考慮すべき重要な点は、次のとおりです。

1. リソース使用率とパフォーマンスメトリクス: 時間の経過に伴う AWS リソースの使用状況を分析します。ピーク時とオフピーク時の使用パターンを特定して、リソースの割り当てとスケーリング戦略を最適化します。応答時間、レイテンシー、スループット、エラー率などのパフォーマンスメトリクスをモニタリングします。これらのメトリクスは、クラウドインフラストラクチャの全体的な状態と効率を評価するのに役立ちます。
2. ユーザーとアプリケーションのスケーリング動作: ユーザーの行動と、こうした行動がワークロードの需要にどのように影響するかを理解します。ユーザートラフィックのパターンを調べることは、コンテンツの配信とアプリケーションの応答性を向上させるうえで役立ちます。需要の増加に伴ってワークロードがどのようにスケールするかを分析します。ロードの変動に対応するために、自動スケーリングパラメータが適切かつ効果的に設定されているかどうかを判断します。
3. ワークロードタイプ: バッチ処理、リアルタイムデータ処理、ウェブアプリケーション、データベース、機械学習など、クラウドで実行されているさまざまなタイプのワークロードを特定します。ワークロードのタイプごとに、リソース要件とパフォーマンスプロファイルが異なる場合があります。
4. サービスレベルアグリーメント (SLA): 実際のパフォーマンスを SLA と比較して、コンプライアンスを確保し、改善が必要な分野を特定します。

[Amazon CloudWatch](#) を使用して、メトリクスの収集と追跡、ログファイルの監視、アラームの設定を行い、AWS リソースの変化に自動的に対応します。Amazon CloudWatch を使すると、リソース使用率、アプリケーションパフォーマンス、運用状態についてシステム全体の可視性を得ることもできます。

[AWS Trusted Advisor](#) では、ベストプラクティスに従ってリソースをプロビジョニングすることで、システムのパフォーマンスと信頼性を向上させ、セキュリティを強化し、コスト節減の機会を探ることができます。また、本番環境以外のインスタンスをオフにして、需要の増減に合わせて Amazon CloudWatch や Auto Scaling を使用することもできます。

最後に、[AWS Cost Explorer](#) または [Amazon QuickSight](#) を CUR ファイルまたはアプリケーションログとともに使用して、ワークロード需要の高度な分析を実行できます。

全体的には、包括的なワークロード需要分析により、組織はリソースのプロビジョニング、スケーリング、最適化について情報に基づいた意思決定が可能になり、パフォーマンス、コスト効率、ユーザー満足度の向上につながります。

実装手順

- 既存のワークロードデータを分析する: 既存のワークロード、以前のバージョンのワークロード、または予測された使用パターンのデータを分析します。Amazon CloudWatch、ログファイルとモニタリングデータを使用して、ワークロードの使用状況についてインサイトを得ます。ワークロードの全サイクルを分析し、月末や年末のイベントなどの季節的な変化のデータを収集します。分析に反映される労力は、ワークロードの特性を反映する必要があります。最大の労力は、需要に最も大きな変化がある価値の高いワークロードに割り当てられる必要があります。需要の変化が最小である低価値のワークロードには、最小の労力を割り当てる必要があります。
- 外部の影響を予測する: ワークロードの需要に影響を与え、または変化させる可能性のある組織全体からのチームメンバーとミーティングを行います。一般的なチームは販売、マーケティング、ビジネス開発です。当該メンバーと協力して、業務のサイクルや、ワークロードの需要を変化させるイベントがあるかどうかを把握します。このデータを使用してワークロードの需要を予測します。

リソース

関連するドキュメント:

- [Amazon CloudWatch](#)

- [AWS Trusted Advisor](#)
- [AWS X-Ray](#)
- [AWS Auto Scaling](#)
- [AWS Instance Scheduler](#)
- [Getting started with Amazon SQS](#) (Amazon SQS の開始方法)
- [AWS Cost Explorer](#)
- [Amazon QuickSight](#)

関連動画:

関連する例:

- [Monitor, Track and Analyze for cost optimization](#)
- [Searching and analyzing logs in CloudWatch](#)

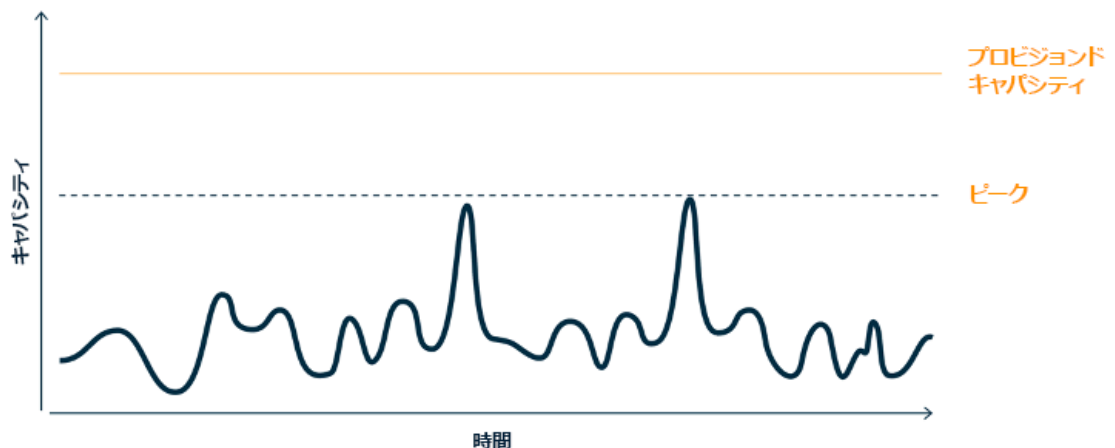
COST09-BP01 需要を管理するためのバッファまたはスロットルを実装する

バッファリングとスロットリングは、ワークロードの需要を修正し、ピークを滑らかにします。クライアントが再試行を実行するときにスロットリングを実行します。バッファリングは、リクエストを保存し、後日まで処理を延期するために実装します。スロットルとバッファが、クライアントが要求された時間内にレスポンスを受け取るように設計されていることを確認します。

このベストプラクティスが確立されていない場合のリスクレベル: 中

実装のガイダンス

クラウドコンピューティングでは、需要を管理し、ワークロードに必要なプロビジョンドキャパシティを削減するために、バッファまたはスロットリングの実装が不可欠です。パフォーマンスを最適化するには、ピークを含む総需要、リクエストの変化のペース、必要な応答時間を測定することが重要です。クライアントにリクエストの再送機能がある場合は、スロットリングの適用が現実的です。逆に、クライアントに再試行の機能がなければ、バッファソリューションの実装が理想的なアプローチです。バッファは、入ってくるリクエストの交通整理を行い、動作速度がさまざまに異なるアプリケーションとの通信を最適化します。



大容量のプロビジョンドキャパシティが必要な顕著なピークが 2 つある需要曲線

上の図に示す需要曲線を持つワークロードがあるとします。このワークロードには 2 つのピークがあり、これらのピークを処理するために、オレンジの線で示されるリソース容量がプロビジョニングされます。このワークロードで使用されるリソースとエネルギーは需要曲線の下領域ではなく、プロビジョンドキャパシティのラインの下領域で示されます。これら 2 つのピークを処理するには、プロビジョンドキャパシティが必要であるためです。ワークロードの需要曲線を平坦化することで、ワークロードに必要なプロビジョンドキャパシティを削減し、環境への影響を減らすことができます。ピークをならすには、スロットリングまたはバッファリングのソリューションの実装を検討してください。

理解を深めるために、スロットリングとバッファリングについて見ていきましょう。

スロットリング: 需要側に再試行の機能がある場合は、スロットリングを実装できます。スロットリングでは、その時点でリクエストを処理できない場合は、後で再試行する必要があることが需要側に通知されます。需要側は一定時間待ってから、リクエストを再試行します。スロットリングの運用には、リソースの最大量およびワークロードのコストを制限できるという利点があります。AWS では、[Amazon API Gateway](#) を使用してスロットリングを実装できます。

バッファベース: バッファベースのアプローチでは、プロデューサー (メッセージをキューに送信するコンポーネント)、コンシューマー (キューからメッセージを受信するコンポーネント)、キュー (メッセージをためておくキュー) を使用してメッセージを格納します。メッセージはコンシューマーによって読み取られ、処理されるため、コンシューマーのビジネス要件を満たせる動作速度でメッセージを実行できます。バッファを中心にした方法を採用することで、プロデューサーが送信したメッセージはキューまたはストリームに蓄えられ、コンシューマーがそれぞれの運用上の需要に応じたペースでアクセスできるようになります。

AWS でバッファリングアプローチを実装する際は、複数のサービスから選択できます。[Amazon Simple Queue Service \(Amazon SQS\)](#) は、コンシューマーが単独で個別のメッセージを読むことができるキューを提供するマネージドサービスです。[Amazon Kinesis](#) は、多数のコンシューマーが同じメッセージを読むことができるストリームを提供します。

バッファリングとスロットリングは、ワークロードの需要を変化させ、ピークを滑らかにします。クライアントがアクションを再試行する場合はスロットリングを使用し、リクエストを保留して後で処理する場合はバッファリングを使用します。バッファベースのアプローチを採用する場合は、必要な時間内にリクエストを処理するようにワークロードを設計し、作業の重複リクエストを処理できるようにします。全体的な需要、変化率、および要求される応答時間を分析して、必要なスロットルまたはバッファのサイズを適正化します。

実装手順

- クライアントの要件を分析する: クライアントのリクエストを分析して、クライアントが再試行を実行できるかどうかを判断します。再試行を実行できないクライアントの場合、バッファを実装する必要があります。全体的な需要、変化率、および要求される応答時間を分析して、必要なスロットルまたはバッファのサイズを決定します。
- バッファまたはスロットルを実装する: ワークロードにバッファまたはスロットルを実装します。Amazon Simple Queue Service (Amazon SQS) などのキューは、ワークロードコンポーネントにバッファを提供できます。Amazon API Gateway は、ワークロードコンポーネントのためにスロットリングを提供できます。

リソース

関連するベストプラクティス:

- [SUS02-BP06 需要曲線を平坦化するためにバッファリングまたはスロットリングを実装する](#)
- [REL05-BP02 リクエストのスロットル](#)

関連するドキュメント:

- [AWS Auto Scaling](#)
- [AWS Instance Scheduler](#)
- [Amazon API Gateway](#)
- [Amazon Simple Queue Service](#)
- [Getting started with Amazon SQS](#)
- [Amazon Kinesis](#)

関連動画:

- [Choosing the Right Messaging Service for Your Distributed App](#)

関連する例:

- [Managing and monitoring API throttling in your workloads](#)
- [Throttling a tiered, multi-tenant REST API at scale using API Gateway](#)
- [Enabling Tiering and Throttling in a Multi-Tenant Amazon EKS SaaS Solution Using Amazon API Gateway](#)
- [Application integration Using Queues and Messages](#)

COST09-BP03 リソースを動的に供給する

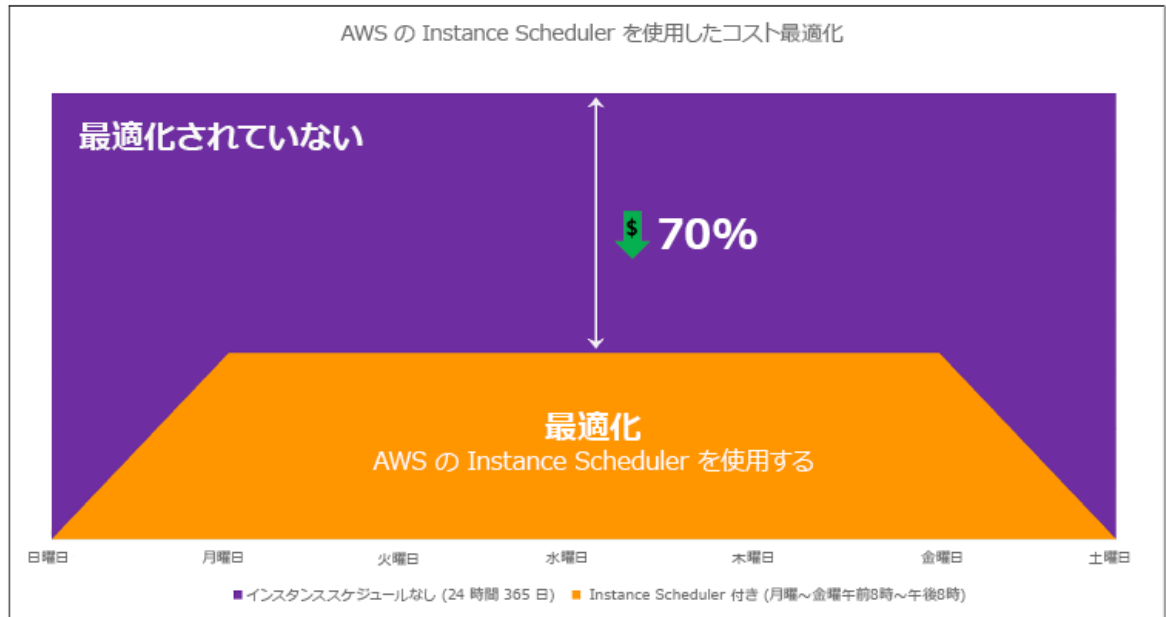
リソースを計画的なやり方でプロビジョニングします。これは、自動スケーリングなどの需要ベース、または需要が予測可能でリソースが時間に基づいて提供される時間ベースで行います。これらの手法を使用すると、過剰プロビジョニングやプロビジョニング不足を最小限に抑えることができます。

このベストプラクティスを活用しない場合のリスクレベル: 低

実装のガイダンス

AWS のお客様がアプリケーションに利用できるリソースを増やし、需要に合わせてリソースを供給する方法はいくつかあります。これらのオプションの 1 つは、Amazon Elastic Compute Cloud (Amazon EC2) および Amazon Relational Database Service (Amazon RDS) インスタンスの起動と停止を自動化する AWS インスタンススケジューラを使用することです。もう 1 つのオプションは AWS Auto Scaling を使用することです。この方法では、アプリケーションやサービスの需要に基づいてコンピューティングリソースを自動的にスケーリングできます。需要に応じてリソースを供給することで、使用したリソースに対してのみ支払いを行い、必要なときにリソースを起動してコストを削減し、必要でないときにリソースを終了することができます。

[AWS での Instance Scheduler](#) を使用すると、Amazon EC2 インスタンスや Amazon RDS インスタンスを決まった時間に停止および開始するように設定できます。これにより、例えばユーザーが毎朝 8 時に Amazon EC2 インスタンスにアクセスし、夜 6 時以降は必要としないなど、一貫した時間パターンがある同一リソースの需要に応えることができます。この解決方法では、リソースを使用しないときは停止し、必要なときに開始することで、運用コストを削減できます。



AWS Instance Scheduler によるコストの最適化。

また、AWS Systems Manager クイックセットアップを使用してシンプルなユーザーインターフェイス (UI) を使用して、アカウントやリージョン全体で Amazon EC2 インスタンスのスケジュールを簡単に設定できます。AWS Instance Scheduler を使用して Amazon EC2 または Amazon RDS インスタンスをスケジュールでき、既存のインスタンスを停止および起動できます。ただし、Auto Scaling グループ (ASG) の一部であるインスタンスや、Amazon Redshift や Amazon OpenSearch Service などのサービスを管理しているインスタンスは、停止および開始できません。Auto Scaling グループにはグループ内のインスタンスに対して独自のスケジュールがあり、これらのインスタンスが作成されます。

[AWS Auto Scaling](#) により、変化する需要に対応するためにキャパシティを調整して、最低限のコストで安定かつ予測可能なパフォーマンスを維持できます。これは、Amazon EC2 インスタンス、スポットフリート、Amazon ECS、Amazon DynamoDB、Amazon Aurora と統合されたアプリケーションのキャパシティを拡張するためのフルマネージド型の無料サービスです。Auto Scaling では、リソースの自動検出によってワークロード内の設定可能なリソースを検出できます。また、パフォーマンス、コスト、または両者のバランスを最適化するためのスケーリング戦略が組み込まれており、予測スケーリングによって定期的に発生する急増に対応することができます。

Auto Scaling グループをスケーリングするには、複数のスケーリングオプションを使用できます。

- 常に現在のインスタンスレベルを維持
- 手動スケーリング
- スケジュールに基づくスケーリング
- 需要に応じたスケーリング
- 予測スケーリングを使用する

Auto Scaling ポリシーは異なり、動的スケーリングポリシーとスケジュールスケーリングポリシーに分類できます。動的ポリシーには、手動または動的スケーリング、スケジュールスケーリングまたは予測スケーリングがあります。スケーリングポリシーは、動的スケーリング、スケジュールスケーリング、予測スケーリングに使用できます。また、[Amazon CloudWatch](#) のメトリクスとアラームを使用して、ワークロードのスケーリングイベントをトリガーできます。最新の機能や改善点にアクセスできる [起動テンプレート](#) を使用することをお勧めします。起動設定を使用する場合、すべての Auto Scaling 機能が利用できるわけではありません。例えば、スポットインスタンスとオンデマンドインスタンスの両方を起動する、

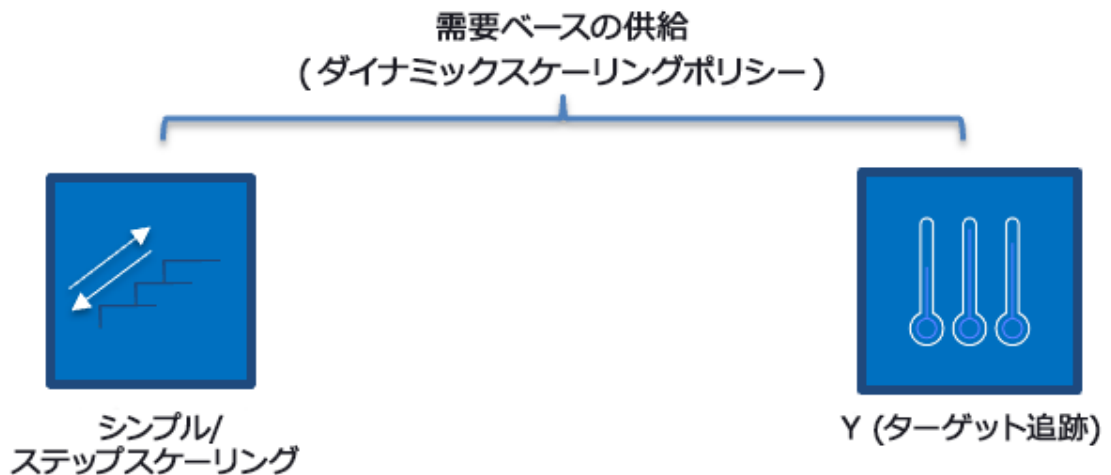
または複数のインスタンスタイプを指定する Auto Scaling グループを作成することはできません。これらの機能を設定するには、起動テンプレートを使用する必要があります。起動テンプレートを使用するときは、それぞれバージョンを作成することをお勧めします。起動テンプレートのバージョンニングにより、すべてのパラメータのサブセットを作成できます。その後、それを再利用して同じ起動テンプレートの他のバージョンを作成できます。

AWS Auto Scaling を使用するか、[AWS API または SDK を使用してコードにスケーリングを組み込むことができます](#)。これにより、環境を手動変更していた運用コストがなくなり、その結果、全体的なワークロードコストが削減され、変更をより迅速に実行できるようになります。またこれにより、いつでもワークロードのリソースを需要に合わせて調達できます。ベストプラクティスに従って組織に動的にリソースを供給するには、AWS クラウド の水平スケーリングおよび垂直スケーリングと、Amazon EC2 インスタンスで実行されるアプリケーションの特性を理解する必要があります。このベストプラクティスに従うには、クラウド財務管理チームとテクニカルチームが協働することをお勧めします。

[Elastic Load Balancing \(Elastic Load Balancing\)](#) は、複数のリソースに需要を分散することで、スケーリングを支援します。ASG と Elastic Load Balancing を使用する、トラフィックを最適にルーティングして受信リクエストを管理し、Auto Scaling グループ内の 1 つのインスタンスに負荷がかかりすぎないようにすることができます。リクエストは、キャパシティや使用率を考慮せずに、ラウンドロビン方式でターゲットグループのすべてのターゲットに分散されます。

一般的な Amazon EC2 メトリクスは、CPU 使用率、ネットワークスループット、Elastic Load Balancing で確認されたリクエストとレスポンスのレイテンシーなどの標準メトリクスです。可能な場合は、カスタマーエクスペリエンスの指標となるメトリクスを使用する必要があります。このメトリクスは一般には、ワークロード内のアプリケーションコードから生成されるカスタムメトリクスです。このドキュメントでは、需要を動的に満たす方法を詳しく説明するために、Auto Scaling を需要ベースの供給モデルと時間ベースの供給モデルの 2 つのカテゴリに分類し、それぞれについて詳しく説明します。

需要ベースの供給: クラウドの伸縮性を活用して、ほぼリアルタイムの需要状況に応じて、変化する需要に対応するリソースを供給できます。需要ベースの供給の場合、API やサービス機能を活用すると、アーキテクチャ内のクラウドリソースの量をプログラムで変更できます。これにより、アーキテクチャ内のコンポーネントの規模を変えたり、需要が急増したときにリソースの数を増加させてパフォーマンスを維持したり、需要が後退したときにキャパシティを減少させてコストを節減させたりできます。

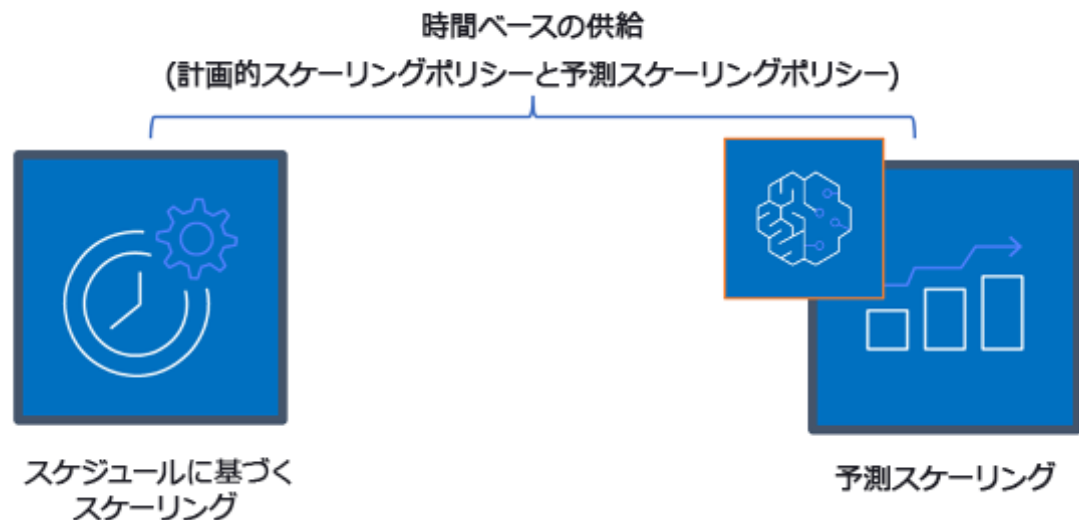


需要ベースの動的スケーリングポリシー

- シンプル/ステップスケーリング: メトリクスをモニタリングし、カスタマーが手動で定義したステップに従ってインスタンスを追加/削除します。
- ターゲット追跡: サーモスタットのような制御メカニズムで、インスタンスを自動的に追加または削除して、メトリクスをカスタマー定義の目標に維持します。

需要ベースのアプローチで設計する場合、主に 2 つの点を考慮する必要があります。第 1 に、新しいリソースをどれだけ早くプロビジョニングする必要があるかを理解することです。第 2 に、需要と供給の差異が変動することを理解することです。需要の変動ベースに対処できるようにしておくだけでなく、リソースの不具合にも備えておく必要があります。

時間ベースの供給: 時間ベースのアプローチでは、リソースのキャパシティを予測可能な需要、または時間ごとに明確に定義された需要に合わせます。このアプローチは、通常、リソースの使用率に依存せず、リソースが必要な特定の時間にそのリソースを確保します。また、起動手順、およびシステムや一貫性のチェックにより、遅延なくリソースを提供できます。時間ベースのアプローチでは、繁忙期に追加のリソースを投入したり、キャパシティを拡大したりできます。



時間ベースのスケールアップポリシー

スケジュールされたまたは予測される Auto Scaling を使用して、時間ベースのアプローチを実装できます。営業開始時など、特定の時間にワークロードをスケールアウトまたはスケールインするようにスケジュールできるため、ユーザーがアクセスしたときや需要が増加したときにリソースを利用可能にしておくことができます。予測スケールアップでは、パターンを使用してスケールアウトします。一方スケジュールに基づくスケールアップでは、事前に定義された時間を使用してスケールアウトします。属性ベースの [インスタンスタイプの選択 \(ABS\) 戦略](#) を Auto Scaling グループで使用することもできます。これにより、インスタンスの要件を vCPU、メモリ、ストレージなどの属性のセットとして表現できます。これにより、新しい世代のインスタンスタイプがリリースされると自動的に使用し、さらに Amazon EC2 スポットインスタンスでより広い範囲のキャパシティにアクセスできます。Amazon EC2 フリートと Amazon EC2 Auto Scaling が指定した属性に適合するインスタンスを選択して起動するため、手動でインスタンスタイプを選択する必要がなくなります。

また、[AWS API や SDK](#) および [AWS CloudFormation](#) を使用すると、必要に応じて自動的にプロビジョニングしたり、環境全体を削除したりできます。このアプローチは、所定の営業時間や一定期間にのみ実行される開発環境またはテスト環境に適しています。API を使用した環境内のリソースサイズのスケールアップ (垂直スケールアップ) にも対応しています。例えば、インスタンスのサイズやクラスを変更して、本番稼働ワークロードをスケールアップできます。これを行うには、インスタンスを停止・起動して、別のインスタンスのサイズやクラスを選択します。この手法は、使用中にサイズの拡大、パフォーマンス (IOPS) の調整、ボリュームタイプの変更が可能な Amazon EBS Elastic Volumes などのリソースにも適用できます。

時間ベースのアプローチを設計する際は、主に 2 つの点を考慮する必要があります。1 つ目は使用パターンの一貫性についてであり、第 2 に、パターンを変更した場合の影響です。予測精度は、ワークロード

をモニタリングし、ビジネスインテリジェンスを使用することで高めることができます。使用パターンに大幅な変更がある場合は、時間を調整して予測対象範囲に収まるようにします。

実装手順

- スケジュールされたスケーリングを設定: 需要の変化を予測できるため、時間ベースのスケーリングは適切な数のリソースを適時に提供できます。また、リソースの作成と設定が、需要の変化に対応するのに十分ではない場合にも役立ちます。ワークロード分析を活用して、AWS Auto Scaling を使用してスケジュールに基づくスケーリングを設定します。時間ベースのスケジューリングを設定するには、予測スケーリングまたはスケジュールに基づくスケーリングを使用し、予想される、または予測可能な負荷の変化に合わせて、事前に Auto Scaling グループの Amazon EC2 インスタンス数を増やすことができます。
- 予測スケーリングの設定: 予測スケーリングを使用すると、トラフィックフローの日次および週次パターンに先立って、Auto Scaling グループ内の Amazon EC2 インスタンス数を増やすことができます。定期的にトラフィックのスパイクがあり、アプリケーションの起動に時間がかかる場合は、予測スケーリングの使用を考慮すべきです。予測スケーリングを使用すると、見積もられた負荷の前にキャパシティを初期化できるため、性質上後手に回る動的スケーリング単体と比較して、より迅速にスケールできます。例えば、ユーザーが始業時間とともにワークロードの仕様を開始し、終業時間後は使用しない場合、予測スケーリングを使用すれば、始業時間前にキャパシティを追加できるため、トラフィックの変化に反応する動的スケーリングで生じる遅延を排除できます。
- 動的自動スケーリングの設定: アクティブなワークロードメトリクスに基づいてスケーリングを設定するには、Auto Scaling を使用してください。分析を使用して、正しいリソースレベルで起動するように Auto Scaling を設定し、ワークロードが要求された時間内にスケールすることを検証します。1 つの Auto Scaling グループ内でオンデマンドインスタンスとスポットインスタンスのフリートを起動し、自動的にスケールできます。スポットインスタンスの利用による割引に加え、リザーブドインスタンスや Savings Plans を利用して、通常のオンデマンドインスタンスの価格から割引を受けることができます。これらの要素をすべて組み合わせることで、Amazon EC2 インスタンスのコスト削減を最適化し、アプリケーションに必要なスケールとパフォーマンスを実現できます。

リソース

関連するドキュメント:

- [AWS Auto Scaling](#)
- [AWS での Instance Scheduler](#)
- Auto Scaling グループのサイズをスケールする
- [Getting Started With Amazon EC2 Auto Scaling](#)
- [Amazon SQS の開始方法](#)
- [Amazon EC2 Auto Scaling のスケジュールされたスケーリング](#)
- [Amazon EC2 Auto Scaling の予測スケーリング](#)

関連動画:

- [Auto Scaling のターゲット追跡スケーリングポリシー](#)
- [AWS での Instance Scheduler](#)

関連サンプル:

- [Amazon EC2 フリート向け Auto Scaling 用の属性ベースのインスタンスタイプの選択](#)
- [スケジュールされたスケーリングを使用したコストに対する Amazon Elastic Container Service の最適化](#)

- [Amazon EC2 Auto Scaling で予測スケーリング](#)
- [AWS CloudFormation で Instance Scheduler を使用して Amazon EC2 インスタンスをスケジュールするにはどうすればよいですか？](#)

継続的最適化

AWS では、新しいサービスを確認してワークロードに実装することで、経時的に最適化していきます。

AWS で新しいサービスと機能がリリースされた際は、ベストプラクティスとして、既存のアーキテクチャの決定事項を確認し、費用対効果を維持することが推奨されます。要件の変化に応じて、不要になったリソース、コンポーネント、ワークロードを積極的に排除します。経時的に最適化を行うには、次のベストプラクティスを考慮します。

ワークロードを長期的に最適化し、組織で [CFM](#) 文化を改善しながら、クラウドの運用コストを評価し、時間のかかるクラウド運用をレビューして、関連する AWS サービス、サードパーティー製品、またはカスタムツール ([AWS CLI](#) または [AWS SDK](#) など) を採用し、それらを自動化して人間の労力を減らします。

トピック

- [レビュープロセスを定義し、ワークロードを定期的に分析 \(p. 89\)](#)
- [オペレーションの自動化 \(p. 92\)](#)

レビュープロセスを定義し、ワークロードを定期的に分析

ベストプラクティス

- [COST10-BP01 ワークロードレビュープロセスを開発する \(p. 89\)](#)
- [COST10-BP02 このワークロードを定期的に見直し、分析する \(p. 91\)](#)

COST10-BP01 ワークロードレビュープロセスを開発する

ワークロードレビューの基準とプロセスを定義するプロセスを開発します。レビューを行う際には、潜在的利益を織り込む必要があります。例えば、コアワークロードや、請求の 10% 超に値するワークロードは四半期または 6 か月ごとにレビューし、10% 以下のワークロードは年に 1 回レビューするなどです。

このベストプラクティスが確立されていない場合のリスクレベル: 高

実装のガイダンス

ワークロードの費用対効果を最大にするには、ワークロードを定期的にレビューし、新しいサービス、機能、コンポーネントを実装する機会があるかどうかを把握する必要があります。全体的なコスト削減を達成するには、潜在的なコスト削減量に比例したプロセスを行う必要があります。例えば、支出全体の 50% を占めるワークロードは、支出全体の 5% を占めるワークロードよりも定期的かつ徹底的にレビューする必要があります。外部要因または変動性を考慮します。ワークロードにより特定の地域、特定の市場セグメントにサービスが提供されていて、その領域での変化が予測される場合、レビュー頻度を高くすることでコスト削減につながる可能性があります。レビューで考慮すべきもう 1 つの要因は、変更を運用する労力です。変更のテストおよび検証に多大なコストがかかる場合は、レビューの頻度を下げる必要があります。

古くなったレガシーコンポーネントやリソースには維持するための長期的なコストがかかることや、新しい機能を実装できないことを考慮します。テストと検証にかかる現在のコストが、提案されている利益を上回っている場合があります。しかし、ワークロードと現在のテクノロジーとのギャップが時間の経過とともに大きくなるにつれて、変更にかかるコストが増加し、結果として巨額のコストになることがあります。たとえば、新しいプログラミング言語に移行するときの費用対効果は現時点で低いとします。しかし、5年後には、その言語に精通した人材のコストが増加する可能性があります。ワークロードが増加すると、さらに大規模なシステムを新しい言語に移行することになり、結果的にこれまでよりもさらに多大な労力を要します。

ワークロードをコンポーネントに分割し、コンポーネントのコストを割り当て (コストの見積りで可)、各コンポーネントの横に要因 (労力や外部市場など) を一覧表示します。この指標を使用して、各ワークロードのレビュー頻度を決定します。たとえば、ウェブサーバーが高コストで、変更の労力が低く、外部要因が高い場合は、レビュー頻度が高くなります。中央データベースが中程度のコストで、変更の労力が高く、外部要因が低い場合は、レビューの頻度は中程度になります。

新しいサービス、設計パターン、リソースの種類、設定が利用できるようになった時点で、これらを評価するプロセスを定義し、ワークロードコストを最適化します。[パフォーマンスの柱のレビュー](#)や[信頼性の柱のレビュー](#)のプロセスと同様、最適化と改善作業を特定、検証、優先順位付けし、修正を公開してバックログに組み込みます。

実装手順

- レビュー頻度を定義する: ワークロードとそのコンポーネントを確認する頻度を定義します。継続的な改善とレビューの周期のための時間とリソースを割り当て、ワークロードの効率性と最適化を向上させます。これは要因の組み合わせであり、組織内のワークロード、またワークロード内のコンポーネントによって、異なる場合があります。一般的な要因には、収益またはブランドの観点から評価された組織にとっての重要性、ワークロードの実行にかかる総コスト (運用コストとリソースコストを含む)、ワークロードの複雑さ、変更の実装の容易性、ソフトウェアライセンス契約、ある変更がライセンス違反によるライセンス費用の重大な増加を生じさせるかどうかなどが含まれます。コンポーネントは、ウェブサーバーやデータベース、コンピューティングリソースやストレージリソースなど、機能的または技術的に定義できます。それに応じて要因のバランスをとり、ワークロードとそのコンポーネントのための期間を設定します。例えば、ワークロード全体は 18 か月ごとに、ウェブサーバーは 6 か月ごとに、データベースは 12 か月ごとに、コンピューティングおよび短期ストレージは 6 か月ごとに、長期ストレージは 12 か月ごとに、それぞれ確認することができます。
- レビューの十分性を定義する: ワークロードまたはワークロードコンポーネントのレビューに費やされる労力を定義します。レビュー頻度と同様に、これは複数の要因のバランスです。最も大きな利益をもたらす取り組みに集中できるように、改善の機会を定期的に評価し、優先順位を設定します。同時に、これらの活動に必要な作業量を見積もります。予想される結果が目標に達しておらず、作業コストがさらにかかる場合は、代わりの一連のアクションを使用して作業を繰り返します。レビュープロセスには、漸進的な継続的改善を可能にする時間とリソースを含める必要があります。例えば、データベースコンポーネントの分析に 1 週間、コンピューティングリソースの分析に 1 週間、ストレージのレビューに 4 時間を、それぞれ費やすように決めます。

リソース

関連するドキュメント:

- [AWS ニュースブログ](#)
- [クラウドコンピューティングのタイプ](#)
- [AWS の最新情報](#)

関連する例:

- [AWS サポートからのプロアクティブなサービス](#)
- [SAP ワークロードの定期的なワークロードレビュー](#)

COST10-BP02 このワークロードを定期的に見直し、分析する

既存のワークロードは、それぞれ定義されたプロセスに基づいて定期的に見直され、新しいサービスを導入できるか、既存のサービスを置き換えることができるか、またはワークロードをリアーキテクトできるかを確認します。

このベストプラクティスが確立されていない場合のリスクレベル: ミディアム

実装のガイダンス

AWS は定期的に新しい機能を追加しているため、最新のテクノロジーを利用して、より迅速に実験やイノベーションができます。[AWS の最新情報](#)には、AWS が行っている追加の詳細や、AWS のサービス、機能、リージョン拡大のお知らせの概要が、随時提供されています。発表されたリリースの詳細を確認して、既存のワークロードの見直しや分析にそれらを使用できます。新しい AWS のサービスと機能の利点を得るには、ワークロードでレビューを行い、必要に応じて新しいサービスや機能を実装する必要があります。つまり、場合によっては、ワークロードに使用している既存のサービスを置き換えたり、ワークロードをモダナイズして新しい AWS のサービスを導入したりする必要があるということです。例えば、ワークロードを見直して、メッセージングコンポーネントを Amazon Simple Email Service に置き換えることができます。これにより、すべての機能を低コストで提供しながら、インスタンスのフリートの運用と維持にかかるコストを削減できます。

ワークロードを分析して潜在的な機会を見出すには、新しいサービスだけではなく、ソリューション構築における新しい方法も考慮する必要があります。AWS の [This is My Architecture](#) の動画を確認して、他のお客様のアーキテクチャの設計、課題、解決策について学びます。[All-In series](#) を確認して、AWS のサービスの実際のアプリケーションやお客様事例を参照します。また、基本的なクラウドアーキテクチャパターンのベストプラクティスを説明、検証、詳説している [Back to Basics](#) 動画シリーズも視聴できます。他のソースとして [How to Build This](#) 動画もあります。これは、AWS のサービスを使用して、実用最小限の製品 (MVP) を実現するための素晴らしいアイデアを持つ人々を支援するように設計されています。確固たるアイデアを持った世界中の構築者が、経験豊富な AWS のソリューションアーキテクトからのアーキテクチャに関するガイダンスを得ることができます。最後に、[使用開始](#)のリソース資料を確認します。ステップバイステップのチュートリアルがあります。

レビュープロセスを実行する前に、合意されたレビュープロセスに従いながら、ワークロードにおけるビジネスの要件、特定のサービスまたはリージョンを使用するためのセキュリティおよびデータのプライバシー要件、パフォーマンス要件に従います。

実装手順

- ワークロードを定期的に見直す: 定義したプロセスを使用して、指定した頻度でレビューを実行します。各コンポーネントに適正な労力を費やしていることを確認します。このプロセスは、コスト最適化のためにサービスを選択した最初の設計プロセスに似ています。サービスとこのサービスがもたらすメリットを分析します。今回は、長期的なメリットだけでなく、変更を行うコストも考慮します。
- 新しいサービスの実装: 分析の結果、変更を実施する場合は、まずワークロードのベースラインを実行し、各アウトプットの現在のコストを把握します。変更を実施し、分析を実行して、各アウトプットの新しいコストを確認します。

リソース

関連するドキュメント:

- [AWS ニュースブログ](#)
- [AWS の最新情報](#)
- [AWS ドキュメント](#)

- [AWS の使用開始](#)
- [AWS の一般的なリソース](#)

関連動画:

- [AWS - This is My Architecture](#)
- [AWS - Back to Basics](#)
- [AWS - All-In series](#)
- [How to Build This](#)

オペレーションの自動化

ベストプラクティス

- [COST11-BP01 オペレーションのオートメーションを実行する \(p. 92\)](#)

COST11-BP01 オペレーションのオートメーションを実行する

クラウドでのオペレーションの労力コストを評価します。管理タスク、デプロイ、その他オペレーションにおけるオートメーションを使用した時間と労力の削減を定量化します。オペレーションの労力に必要な時間とコストを評価し、可能な部分は管理タスクを自動化して、人間による手作業を削減します。

このベストプラクティスが確立されていない場合のリスクレベル: 低

オペレーションを自動化すると、安定性とスケーラビリティが向上し、可視性、信頼性、柔軟性が上がります。また、人的リソースを解放しメトリクスを改善することで、コストを削減し、イノベーションを加速できます。ワークロードをデプロイ、管理、運用する際に安定した信頼性の高いエクスペリエンスを提供することで、手動タスクの頻度を減らし、効率を高め、企業にメリットをもたらします。インフラストラクチャのリソースを手動の運用タスクから解放し、それらをより価値の高いタスクやイノベーションに使用できるため、ビジネス成果が向上します。企業は、クラウドでワークロードを管理するための、実績がありテスト済みの方法を求めています。そのようなソリューションは安全、高速でありコスト効果が高く、最小限のリスクで最大限の信頼性を得られるものでなくてはなりません。

クラウドにおけるオペレーションコスト全体を確認して、必要な労力に基づきオペレーションに優先順位をつけることから始めます。例えば、クラウドに新しいリソースをデプロイするのにかかる時間、既存のものを最適化する変更にかかる時間、必要な構成を設定するのにかかる時間はどのくらいでしょうか。オペレーションと管理のコストを考慮に入れて、人間による操作の合計コストを確認します。管理タスクのオートメーションに優先順位を付けて、人間の手作業を減らします。レビューを行う際には、潜在的利益を織り込む必要があります。例えば、タスクを手動で実行する場合にかかる時間を自動で実行する場合と比較します。繰り返し実行する価値の高いアクティビティを優先します。通常、人的エラーを起こすリスクが高いアクティビティから自動化するのが良い方法です。このようなリスクは望ましくない追加の運用コスト (オペレーションチームの追加作業時間など) を発生させることが多いからです。

AWS のサービス、ツール、サードパーティ製製品を使用して、特定の要件に対してどの AWS オートメーションを実装しカスタマイズするかを選択します。次の表は、AWS のサービスを使用して管理とオペレーションを自動化することで実現できる主なオペレーション機能や性能の一部を示したものです。

- [AWS Audit Manager](#): AWS の使用状況を継続的に監査し、リスクとコンプライアンスの評価を簡易化します
- [AWS Backup](#): データ保護を一元的に管理および自動化します。
- [AWS Config](#): コンピューティングリソースを設定し、設定とリソースインベントリを査定、監査、評価します。

- [AWS CloudFormation](#): 高可用性リソースを Infrastructure as Code とともに起動します。
- [AWS CloudTrail](#): IT の変更管理、コンプライアンス、コントロール。
- [Amazon EventBridge](#): イベントをスケジュールし、AWS Lambda をトリガーしてアクションを実行します。
- [AWS Lambda](#): 繰り返し実行するプロセスを、イベントでトリガーしたり、Amazon EventBridge を使用して固定スケジュールで実行したりすることで、プロセスを自動化します。
- [AWS Systems Manager](#): ワークロードの開始と停止、オペレーションシステムへのパッチ適用、設定の自動化、管理の続行。
- [AWS Step Functions](#): ジョブをスケジュールし、ワークフローを自動化します。
- [AWS Service Catalog](#): コンプライアンスとコントロールを備えたテンプレート消費および Infrastructure as Code。

時間短縮を検討して、チームが技術的な負債の返済、イノベーション、付加価値機能に集中できるようにします。例えば、オンプレミス環境からクラウド環境に「リフトアンドシフト」式にできるだけ早急に移行して、最適化は後回しにする必要性が生じる場合があります。[Amazon Relational Database Service](#)、[Amazon EMR](#)、[Amazon WorkSpaces](#)、[Amazon SageMaker](#) など、ライセンスコストを排除または削減する AWS のマネージドサービスを利用して、コスト削減が可能なところを模索することには時間をかけるだけの価値があります。マネージドサービスによってサービス維持に伴う運用上および管理上の負担が軽減されるため、イノベーションに集中できます。さらに、マネージドサービスはクラウドのスケールで運用されるため、トランザクション単位またはサービス単位でコストを削減できます。

AWS の製品やサービスの使用と同時にオートメーションを導入したいが、組織にそのスキルがない場合、[AWS Managed Services \(AMS\)](#)、[AWS プロフェッショナルサービス](#)、または [AWS パートナー](#) に連絡して、オートメーションの導入を増やし、クラウドにおけるオペレーショナルエクセレンスを向上させます。

[AWS Managed Services \(AMS\)](#) は、エンタープライズのお客様やパートナーに代わって AWS インフラストラクチャを運用するサービスです。コンプライアンスに準拠したセキュアな環境で、ワークロードをデプロイできます。AMS では、エンタープライズクラウド運用モデルとオートメーションを使用して、組織の要件を満たし、クラウド移行を高速化し、継続的な管理コストを削減できます。

[AWS プロフェッショナルサービス](#) を利用すると、AWS を使用して目的のビジネス成果を達成し、オペレーションを自動化することもできます。AWS プロフェッショナルサービスは、エンタープライズクラウドコンピューティングの該当分野におけるお客様の業務をサポートする、グローバルで専門性の高いプラクティスを提供します。専門性の高いプラクティスは、ソリューション、テクノロジー、業界の対象分野にわたるベストプラクティス、フレームワーク、ツール、サービスを通じて、対象を絞ったガイダンスを提供します。このガイダンスを使用して、自動化された堅牢で俊敏な IT オペレーションや、クラウドセンタに最適化されたガバナンス機能をデプロイできます。

実装手順

- 構築は 1 回、デプロイは多数: AWS CloudFormation、AWS SDK、AWS Command Line Interface (AWS CLI) などの infrastructure-as-Code を使用して 1 回デプロイし、同じ環境やディザスタリカバリのシナリオで何度も使用します。デプロイ中にタグを付け、他のベストプラクティスで定義されている消費を追跡します。[AWS Launch Wizard](#) を使用して、多数の一般的なエンタープライズワークロードをデプロイする回数を削減します。AWS Launch Wizard は、AWS のベストプラクティスに従って、エンタープライズワークロードのサイズ決定、設定、デプロイにわたってガイドを提供します。また、[AWS Service Catalog](#) を使用することもできます。これを利用すると、Infrastructure-as-Code の承認を受けたテンプレートを作成および管理して AWS で使用できるため、承認を受けたセルフサービス型クラウドリソースを誰でも利用できます。
- オペレーションのオートメーション: ルーチンオペレーションを、人が介入せずに自動的に実行します。AWS のサービスやツールを使用すると、実装する AWS のオートメーションを選択し、固有の要件に合わせてカスタマイズできます。例えば、[EC2 Image Builder](#) を使用して仮想マシンやコンテナイメージを構築、テスト、デプロイし、AWS またはオンプレミスでできるようにします。目的のアクションが AWS のサービスで実現できない場合や、リソースのフィルタリングを伴うより複雑なアクションを必要とする場合は、[AWS CLI](#) または AWS SDK ツールを使用して、オペレーションを自動化

します。AWS CLI は、AWS コンソールを使用せず、スクリプト経由で、AWS のサービスのコントロールおよび管理プロセス全体を自動化する機能を提供します。AWS のサービスとやり取りする希望する AWS SDK を選択します。その他のコード例については、[AWS SDK Code examples repository](#) を参照してください。

リソース

関連するドキュメント:

- [Modernizing operations in the AWS クラウド](#) (AWS クラウドでのオペレーションのモダナイゼーション)
- [AWS Services for Automation](#) (オートメーションのための AWS のサービス)
- [AWS Systems Manager Automation](#)
- [AWS automations for SAP administration and operations](#) (SAP の管理とオペレーションのための AWS のオートメーション)
- [AWS Managed Services](#)
- [AWS プロフェッショナルサービス](#)
- [Infrastructure and automation](#) (インフラストラクチャとオートメーション)

関連する例:

- [Reinventing automated operations \(Part I\)](#) (オペレーション自動化の再発明 (パート I))
- [Reinventing automated operations \(Part II\)](#) (オペレーション自動化の再発明 (パート II))
- [AWS automations for SAP administration and operations](#) (SAP の管理とオペレーションのための AWS のオートメーション)
- [AWS Lambda による IT オートメーション](#)
- [AWS Code Examples Repository](#)
- [AWS Samples](#) (AWS のサンプル)

まとめ

コストの最適化とクラウドの財務管理は、継続的な取り組みです。財務チームやテクノロジーチームと定期的に協力し、アーキテクチャのアプローチをレビューし、コンポーネントの選択をアップデートしていく必要があります。

AWS が目指しているのは、お客様がコストの最小化と、弾力性、応答性、順応性に優れたデプロイを両立する後押しをすることです。デプロイコストを最適化するために、本書でご紹介したツール、テクニック、ベストプラクティスをぜひご活用ください。

寄稿者

本書の寄稿者は次のとおりです

- Amazon Web Services、Well-Architected コスト最適化の柱リード、Ben Mergen
- Amazon Web Services、コスト最適化部門ビジネス開発リード、Keith Jarrett
- Amazon Web Services、ビジネスデベロッパーマネージャー、Arthur Basbaum
- Amazon Web Services、コマーシャルアーキテクト、Jarman Hauser

その他の資料

詳細については、次の資料を参照してください。

- [AWS Well-Architected Framework](#)
- [AWS アーキテクチャセンター](#)

改訂履歴

このホワイトペーパーの更新に関する通知を受け取るには、RSS フィードをサブスクライブしてください。

変更	説明	日付
ベストプラクティスガイダンスの更新 (p. 98)	ベストプラクティスを更新し、柱全体における新しいガイダンスを追加。	October 3, 2023
ベストプラクティスガイダンスの更新 (p. 98)	ベストプラクティスを更新し、以下の領域に関する新しいガイダンスを追加。 ガバナンス 、 コストと使用量のモニタリング 、 最適な料金モデルを選択する 、 需要を管理しリソースを供給する 。	July 13, 2023
新しいフレームワークの更新 (p. 98)	規範ガイダンスを使用してベストプラクティスを更新、および新しいベストプラクティスを追加。質問 COST 11 とベストプラクティス COST11-BP01 を追加。	April 10, 2023
ホワイトペーパーの更新 (p. 98)	新しい実装ガイダンスを使用してベストプラクティスを更新。	December 15, 2022
ホワイトペーパーの更新 (p. 98)	ベストプラクティスに加筆し、改善計画を追加。	October 20, 2022
マイナーな更新 (p. 1)	イントロダクションに持続可能性の柱を追加。	December 2, 2021
マイナーな更新 (p. 98)	リンクを更新。	April 25, 2021
マイナーな更新 (p. 98)	リンクを更新。	March 10, 2021
新しいフレームワークの更新 (p. 98)	CFM、新しいサービス、Well-Architected との統合を追加。	July 8, 2020
ホワイトペーパーの更新 (p. 98)	AWS の変更内容を反映し、カスタマーレビューから学習した内容を追加。	July 1, 2018
ホワイトペーパーの更新 (p. 98)	AWS の変更内容を反映し、カスタマーレビューから学習した内容を追加。	November 1, 2017
初版発行 (p. 98)	コスト最適化の柱 - AWS Well-Architected フレームワークを公開。	November 1, 2016