



# スタートアップ向け： データコスト最適化 ガイド

スタートアップが AWS のデータサービスを使用して  
スケールアップとコスト削減を実現する方法



# 目次

データコスト最適化の概要 .....	3
Well-Architected フレームワークの活用 .....	4
継続的なプロセスとしてのコスト最適化 .....	6
データ戦略のコストを最適化する 6 つの方法 .....	7
1. 情報インフラストラクチャのコストを大幅に最適化する .....	8
2. ストレージデータをコンピューティングデータから分離する ...	9
3. 予測可能なワークロードのキャパシティを計画して プロビジョニングする.....	10
4. 予測不可能なワークロードにはオンデマンド インスタンスキャパシティを使用する .....	11
5. ストレージレイヤーを一元化してデータの重複を避ける .....	12
6. 最大 100,000 USD の AWS Activate クレジットを 活用する .....	13
データコストの最適化に成功しているスタートアップの紹介 .....	14
次のステップ .....	20

# データコスト最適化の概要

スタートアップにとって、コストの最適化は最重要課題です。さまざまな手法によって有意義なコスト削減を達成できますが、スタートアップにおけるコスト最適化の取り組みは、その成長段階によって異なります。多くのスタートアップは製品開発に集中しており、他の時間を追加機能の開発に割くか、コスト管理に割くかの選択を迫られることがあります。アカウント構造を再編成するかコスト分析パイプラインを構築するかを選択です。いずれにしても、少ない労力で大きな効果が得られるアーキテクチャの変更に優先的に取り組むと、企業の勢いを維持できます。

スタートアップのビジネス意思決定者がスケーラブルなデータ戦略によって継続的に支出を最適化してコストコントロールを実現するうえで、Amazon Web Services (AWS) が役立ちます。包括的データ戦略により、スタートアップがどの段階からでもデータジャーニーに取り組めるよう支援します。

AWS では幅広いサービスと料金オプションを用意しており、ビジネスに必要なキャパシティを確保しながら柔軟にコストを管理できます。AWS の包括的データ戦略により、スタートアップはスケーラブルなデータレイクの迅速な構築、目的に応じた幅広いデータサービスの利用、統合データアクセス、セキュリティ、ガバナンスによるコンプライアンスの確保、パフォーマンスを犠牲にしない低コストでのシステム拡張などを実現し、極めて迅速かつ俊敏に、より優れた意思決定ができるようになります。

この日本語ガイドでは、コストをコントロールしながら最適なデータ戦略を実現するための、簡単に実践できるコスト最適化方法を紹介します。自社のお客様に真の価値をもたらす機能の開発に立ち戻り、以降で紹介するインサイトを活用して、長期的なコスト管理戦略でランウェイを伸ばしましょう。



# Well-Architected フレームワークの活用

競合する優先事項に追われる日々を送っているスタートアップ創業者の皆さんは、基盤となるインフラストラクチャや想定外のコストを気にせず、イノベーションやプロダクトマーケットフィットに時間を使いたいとお考えでしょう。

余計なことに時間を取られず、イノベーションの推進に必要な自由とコントロールを得るためには、スタートアップのアーキテクチャの枠組み、特にデータ分析アプリケーションのワークロードについて、適切な基盤を構築することが重要です。Well-Architected フレームワークを使用しない場合、製品開発におけるコストと遅延が増大し、スタートアップの利益と成長に悪影響を及ぼす可能性があります。

AWS クラウドで分析アプリケーションワークロードを設計、デプロイ、および構築する方法を理解することで、コストを最適化し、スタートアップを新たな高みへとスケールアップさせることができます。データ分析ワークロードに Well-Architected フレームワークを使用すると、データインフラストラクチャのコスト管理のための長期的な戦略を策定し、ビジネスのランウェイを伸ばすことができます。

Well-Architected フレームワークにおけるコスト最適化の柱は、最初の実証設計から本番ワークロードの継続的な運用まで、コストを最適化するためにシステムのライフサイクル全体にわたって改良と改善を継続的に行うプロセスです。このプロセスは、長年にわたる継続的なものです。

適切なソリューションと料金モデルを選択し、ビジネス成果の達成とコストの最小化を可能にする、コストを意識したシステムを構築できます。長期的なコスト最適化を実施するには、削除または縮小可能なデータ、インフラストラクチャリソース、分析ジョブを特定する必要があります。

# Well-Architected フレームワークの活用 (続き)



## 運用上の優秀性

プロセスやイベントのサポートにおいて、ビジネス価値と継続的な改善を実現するために、本番ワークロードを運用、管理、モニタリングする。



## セキュリティ

リスク評価と軽減策により、価値を実現しながら、情報、システム、アセットを保護する。



## 信頼性

動的なリソース管理でインフラストラクチャ、電源、システムの障害からワークロードを自動復旧し、運用上のしきい値を満たす。



## パフォーマンス効率

コンピューティングリソースを効率的に使って、最大限のパフォーマンスでワークロードを提供するためにオンデマンドの変更をサポートし、SLA を満たす。



## コスト最適化

ベストプラクティスやビジネスニーズに影響を与えることなく、不必要なコストを回避、排除したり、リソースを費用対効果の高いものに置き換える。



## サステナビリティ

より持続可能な未来を推進するために、ワークロードの設計、アーキテクチャ、および実装を評価して、エネルギー消費を削減して効率を改善できるようにする。

[Well-Architected フレームワークの詳細はこちら >](#)



## 継続的なプロセスとしての コスト最適化

コスト最適化は、最初概念実証の設計から本番ワークロードの継続的な運用まで、システムのライフサイクル全体にわたって改良と改善を継続的に行うプロセスです。このプロセスは、長年にわたり続くものです。そのため、自社に適したソリューションと料金モデルを選択することが非常に重要です。コストを最小化しながらビジネス目標を達成できる、コストを意識したシステムを構築する必要があります。長期的なコスト最適化を実施するには、削除または縮小可能なデータ、インフラストラクチャリソース、分析ジョブを特定する必要があります。

個々のデータ処理ステップやパイプライン分岐で、分析ワークフローのコストを判断することは重要です。分析ワークフローコストをこのような細かいレベルで把握することで、開発のためにエンジニアリングリソースをどこに集中させるかを定めることができますという利点があります。また、分析ポートフォリオ全体に対する投資対効果 (ROI) の計算にも役立ちます。

常に最もコスト効率の良いワークロードを実現するために、定期的にワークロードを見直し、新しいサービス、機能、コンポーネントを実装する機会を見つけましょう。分析ワークロードでは、ユーザー数が増え続け、データ量も指数関数的に増加するのが一般的です。未使用のデータ、インフラストラクチャ、ETL (抽出、変換、ロード) ジョブなどの未使用リソースを特定し、削除するための標準プロセスを組織全体で導入する必要があります。

# データ戦略のコストを最適化する 6 つの方法

自社のデータ分析ワークロードに適したコンピューティングおよびストレージソリューションをどのように選択すればいいでしょうか？ 初期設計の選択はコストに大きく影響する可能性があります。定常状態やスパイク性など、ワークロードのリソース要件を理解したうえで、要件を満たすソリューションとツールを選択してください。コスト最適化の機会を増やすため、過剰なプロビジョニングは避けましょう。

これらの目標を念頭に置いて、データ戦略のコスト最適化に役立つベストプラクティスを 6 つ紹介します。

- 1 情報インフラストラクチャのコストを大幅に最適化する >
- 2 ストレージデータをコンピューティングデータから分離する >
- 3 予測可能なワークロードのキャパシティを計画してプロビジョニングする >
- 4 予測不可能なワークロードにはオンデマンドインスタンスキャパシティを使用する >
- 5 データの重複を避ける >
- 6 AWS Activate クレジットを活用してデータ戦略を策定する >

# 1: 情報インフラストラクチャのコストを大幅に最適化する

スタートアップが迅速なスケールアップを実現できるということは、迅速な意思決定、急速に変化するビジネスニーズへの俊敏な対応、そして顧客の需要への対応ができることを意味します。処理するデータの量と速度の上昇に追従し、データ駆動型の意思決定をリアルタイムで行うには、スケーラビリティとパフォーマンスに優れたデータサービスを使用してデータを保存し、クエリを実行する必要があります。AWS は、データの保存、クエリ、活用を行うためのスケーラビリティとパフォーマンスに優れたサービスを提供しています。また、最適なコストパフォーマンスを実現するために、継続的に最適化を行っています。AWS では、大規模なユースケースに最適なコストパフォーマンスを実現するためのデータベースを提供しています。

**Amazon Redshift** は、他のクラウドデータウェアハウスと比較してコストパフォーマンスが最大で 3 倍高い、最も高速で経済的なクラウド型データウェアハウスです。AWS のインフラストラクチャのイノベーションは、データサービス全体にわたってお客様にメリットをもたらします。データサービスの多くは、**AWS Graviton** インスタンスでワークロードが実行されていて、ほぼコードを変更する必要なくコストパフォーマンスを向上できます。これらのサービスには、最大 50% 優れたコストパフォーマンスを誇る **Amazon Relational Database Service** (Amazon RDS)、最大 30% のコスト削減、最大 15% の Spark のパフォーマンス向上、最大 20% の **Amazon Aurora** のパフォーマンス向上を実現する **Amazon EMR** などがあります。**Amazon Elastic Compute Cloud** (Amazon EC2) インスタンスは **AWS Nitro System** を搭載しており、同じ CPU で動作する他の主要クラウドプロバイダーと比較して、ワークロードによっては 15% 以上高いスループット性能を実現できるため、データベースにアクセスするアプリケーションをより高速に実行でき、分析に必要なインサイトを得るまでの時間を短縮できます。

## 最大 50%

Amazon RDS でのコストパフォーマンス向上

## 最大 30%

Amazon EMR でのコスト削減

## 最大 20%

Amazon Aurora のパフォーマンス向上

## 2: ストレージデータをコンピューティングデータから分離する

一般的に、データアセットは年々飛躍的に増加します。しかし、スタートアップのコンピューティングのニーズは、同じペースで増加しないかもしれません。ストレージデータをコンピューティングデータから分離すると、ストレージとコンピューティングのコストを別々に管理できます。これにより、コストを最適化するさまざまな機能を実装して、コストを最小化することができます。そのためには、以下を行う必要があります。

### 必要に応じてデータを圧縮する

ほとんどの分析サービスが圧縮データにアクセスできます。データを圧縮して列指向形式で保管することで、クエリの速度向上やクエリトランザクションのコスト削減が期待できます。

### データストアとストレージクラスを選択する際にデータのアクセス頻度を考慮する

階層型ストレージによってデータセットを使用パターンに応じて異なるストレージ階層に移せるので、ストレージコストを最適化できます。[Amazon Simple Storage Service \(Amazon S3\)](#) は、さまざまなストレージクラスを提供し、経時的にコストを最適化することができます。S3 標準クラスのデータを S3 標準 - 低頻度アクセス (または、場合によっては Amazon S3 1 ゾーン - 低頻度アクセス) や、Amazon S3 Glacier に移行するコスト削減ライフサイクルプランを検討します。

### スポットインスタンスなどの低コストなコンピューティングリソースを適宜利用する

ストレージデータをコンピューティングデータから切り離すことで、ステータスなロジックの実装が可能になりますが、さらなるコスト削減のために [Amazon EC2](#) スポットインスタンスの利用を検討してください。スポットインスタンスは、開発や品質保証 (QA) アプリケーションのコストを削減するための優れたオプションでもあります。

### コンピューティングをデータの近くにデプロイしてデータ転送コストを削減する

AWS では、AWS からインターネットに転送されるデータや AWS Direct Connect を介して転送されるデータ、AWS リージョン間で転送されるデータ、アベイラビリティゾーン間で転送されるデータに対して、データ転送料が発生します。AWS へのデータ転送とアベイラビリティゾーン内でのデータ転送は無料です。

### Amazon S3 Select と Amazon S3 Glacier Select を使用してデータ取得を削減する

Amazon S3 Select と Amazon S3 Glacier Select により、アプリケーションではシンプルな SQL 式を使用してオブジェクトからデータのサブセットのみを取得できます。



### 3: 予測可能なワークロードのキャパシティを計画してプロビジョニングする

明確に定義されたデータワークロードの場合、平均的な使用パターンに基づいて事前にキャパシティを計画することで、リソースの使用率を向上させ、過剰なプロビジョニングを回避できます。急増するワークロードに対しては、ユーザーとワークロードの需要に対応する自動スケーリングを設定します。そのためには、以下を行います。

#### ワークロードのパターンと成長率に基づいて適切なインスタンスタイプを選択する

ワークロードのパフォーマンス要件を満たす CPU、メモリ、ネットワーキングなどのリソースのニーズを検討します。過剰なプロビジョニングを避けるために、適切なインスタンスタイプを選択しましょう。最適化された Amazon EC2 インスタンスでは、最適なパフォーマンスとインフラストラクチャコストでワークロードを実行できます。例えば、成長率が低い場合は、選択するインスタンスのサイズを小さくすると、より細かい増分変更が可能になります。

#### 平均または中規模のワークロード使用量に基づいてサイズ適正化をデプロイする

サイズ適正化は、可能な限り低いコストで、インスタンスのタイプやサイズを、ワークロードのパフォーマンスとキャパシティの要件に合わせるプロセスです。また、デプロイ済みのインスタンスを確認して、キャパシティや他の要件を犠牲にせず、コスト削減につながるダウンサイジングを行う機会を見出すプロセスでもあります。

#### 自動スケーリング機能を使用してピーク需要に対応する

多くの分析システムコンポーネントは、需要に応じて動的にスケーリングでき、過剰なプロビジョニングを抑えることができます。可能であれば、スケールアップではなくスケールアウトする分析システムを設計してください。

## 4: 予測不可能なワークロードにはオンデマンドインスタンスキャパシティを使用する

サーバーレスソリューションの料金は、処理されたデータ量または使用されたコンピューティングリソース量で決まりますが、この料金はアクティブに実行されているワークロードがある場合にのみ請求されます。長時間稼働するインフラストラクチャを使用しながらアイドル状態のリソースに料金を支払う必要がない場合とこのアプローチを比べてください。

### アドホック SQL ワークロードに Amazon Athena を使用する

**Amazon Athena** は、標準 SQL を使用して簡単に Amazon S3 のデータを直接分析できるサーバーレスクエリサービスです。Amazon Athena では、実行したクエリに対してのみ料金が発生します。クエリごとにスキャンしたデータ量に基づいて料金が請求されます。

### 実行頻度の低い ETL ジョブには Amazon EMR ではなく AWS Glue を使用する

**AWS Glue** は、データの分類、クリーニング、リッチ化、さまざまなデータストアやデータストリーム間での確実なデータ移動をシンプルかつコスト効率よく実現する、フルマネージド ETL サービスです。AWS Glue では、ETL プロセスで使用されたリソースに対してのみ料金が発生します。一方、**Amazon EMR** は通常、半永続的なデータストレージを必要とする実行頻度の高いジョブに利用します。

### 一時的なワークロードまたは短期的な開発とテストのニーズにはオンデマンドのリソースを使用する

実験またはテスト目的のデータ処理を低コストで行うには、EC2 スポットインスタンス上の Amazon EMR のようなオンデマンドリソースを選択してください。分析ワークロードの初期計画段階およびテスト段階でオンデマンドリソースを使用することで、さまざまなアーキテクチャでイテレーションを行ってから、Well-Architected 分析プロジェクトのすべての要素をジョブに合わせて最適化できます。これには、適切なインスタンスタイプを選択することや、分析パイプラインの実行に必要な時間を満たせるようにサーバーレスアーキテクチャを使用することなどが含まれます。



## 5: ストレージレイヤーを一元化してデータの重複を避ける

コスト意識の高いチームは、さまざまなレベルでデータを共有して、幅広く奥深いインサイトを  
得るだけでなく、複雑さと支出を最小限に抑えたいと考えています。例えば、常に更新データを  
ロードして変換する中央データウェアハウスと、ダッシュボードアプリケーション、アドホック  
クエリ、データサイエンスなど、さまざまなワークロードに対応するビジネスインテリジェンス  
(BI) および分析クラスターで、データを共有する必要があるとします。しかし、あるシステム (プ  
ロデューサー) からデータをアンロードし、別のシステム (コンシューマー) にコピーすること  
は、特にコンシューマーの数が増えるほど、コストがかかり、遅延が発生する可能性があります。  
このアプローチでは、各コンシューマーに関連するデータのサブセットを提供するために個別の  
ETL ジョブを構築し、維持する必要があります。ビジネスクリティカルなデータの使用状況を定  
期的にモニタリングし、コンプライアンスを確保するためのセキュリティ対策も必要であり、複  
雑さはさらに増していきます。分析ワークロードは、可能な限り、データを重複させることなく  
共有できる技術ソリューションを採用する必要があります。スタートアップでこのシフトを進め  
る場合は、以下を行ってください。

### テナント間でデータを共有するための中央ストレージレイヤーを実装する

- データレイクアーキテクチャなどの一元化されたストレージレイヤーを使用すると、さまざま  
なチームが同じデータリソースを利用できます。
- データレイクのある部分にアクセスする必要があるチームが、そのデータの関連サブセット  
のみを利用できるように、セキュリティの分離を適用できます。
- AWS Lake Formation のような市場で実績のあるソリューションを使用して、安全なデータ  
レイクを構築します。AWS のデータレイクは、優れたコストパフォーマンスで、データのサ  
イロを解消し、エンドツーエンドのデータインサイトの最大化を支援します。AWS では数  
十万ものデータレイクが実行されています。

## 6: 最大 100,000 USD の AWS Activate クレジットを活用する\*

スタートアップの創業者は、AWS Activate プログラムに参加し、いつでもそのメリットを享受することができます。スタートアップのステージに応じて、最大 100,000 USD の AWS Activate クレジット\*のほか、最大 800,000 USD 相当のサードパーティサービス、製品、メンバーシップの割引または無料利用を含む Activate 限定特典を受けることができます。

さらに、クレジットの発行が承認されると、以下のようなお得な特典を利用できます。

- **限定オファー**：AWS Activate クレジットをお持ちであれば、**Notion**、**Miro**、**HubSpot**、**Stripe** などの信頼できる組織が提供する限定オファーも利用できるようになります。割引、無料試用期間、クレジット、その他スタートアップに適した特典はすべて、スタートアップの成長と成功の後押しを目的としたものです。
- **コスト最適化ツール**：AWS Activate クレジットをお持ちであれば、AWS Activate コンソールでコストとクレジットの概要を確認し、コストを最適化することができます。この概要では、残高をリアルタイムで確認でき、クレジットの有効期限が迫っている場合はアラートが表示され、コストを最適化するためのヒントが提供されるため、最大限にコストを活用することができます。このステップは急がずに、自身のペースで進めてください。データアーキテクチャを構築する準備ができたなら、そのコストを相殺するためのクレジットを申請できます。

[AWS Activate の詳細はこちら >](#)

\*AWS クレジットはすべて米ドルで提供されており、AWS 販促クレジット利用規約の対象となります。

# データコストの最適化に 成功しているスタートアップ の紹介

エンドツーエンドのデータ戦略で AWS のコスト最適化の柱を活用してコストを削減した、さまざまなステージのスタートアップの事例

# TextNow が AWS PrivateLink によってデータ転送料金を 93% 削減

## 課題

2009 年に 2 人の大学生が立ち上げたモバイルアプリ TextNow は、通信技術を通じてユーザーを無料または限りなく無料に近い形でつなげることを使命としています。2020 年 3 月に新型コロナウイルス感染症によるパンデミックが発生したとき、企業はデジタル広告の支出を削減し始めました。TextNow のソフトウェアテクニカルマネージャーの Andres Ivanov 氏は次のように述べています。「広告事業は、当社の主な収益の柱です。利益を回復するための計画が必要でした」。

## AWS のソリューション

同社は AWS のコストを削減する方法を模索し始め、**AWS PrivateLink** サービスレディパートナーの **Datadog** を利用することでデータ転送料金を削減する機会を得ました。AWS PrivateLink を使用すると、企業はパブリックインターネットにデータを公開することなく、AWS またはオンプレミスでホストされているサービスと仮想プライベートクラウドとの間にプライベート接続を確立できます。

## 成果

プライベートネットワークを介してトラフィックを保護することで、TextNow はデータセキュリティを強化しながらデータ転送料金を大幅に削減することができました。例えば、AWS の米国東部リージョンにおけるギガバイトあたりのデータ転送コストが 0.09 USD から 0.01 USD になりました。

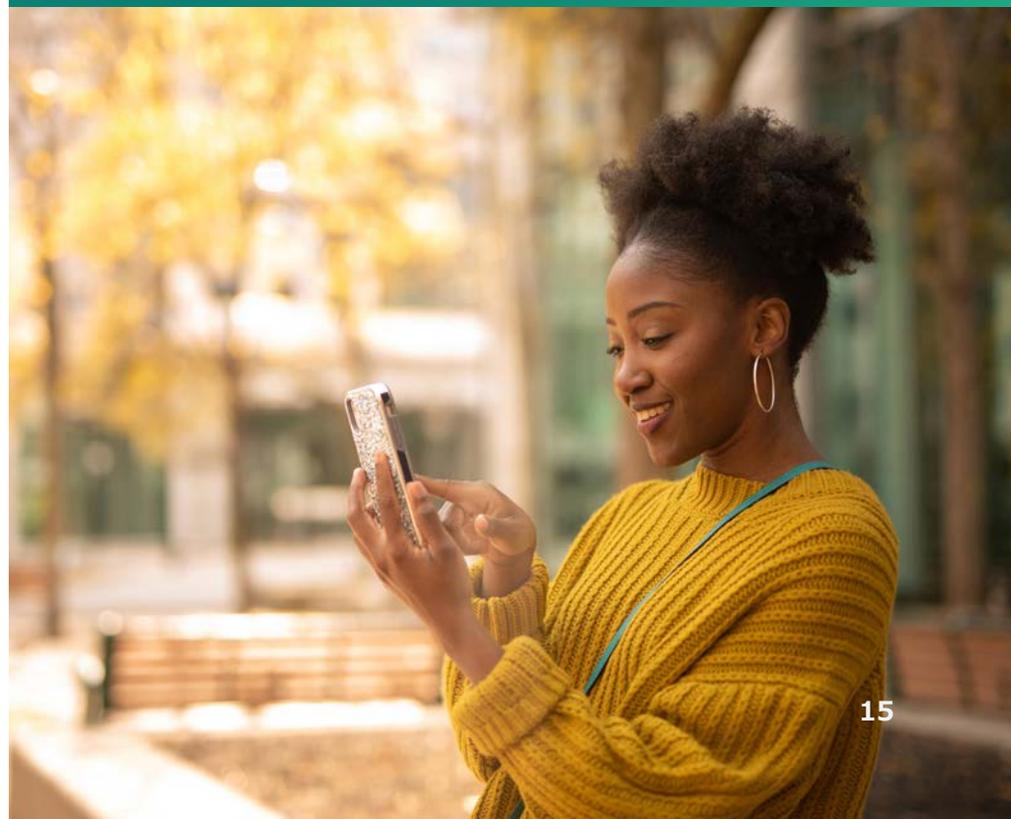
[全文を読む >](#)



# textnow

「AWS PrivateLink を使用することで、データ転送料金を 93% 削減できました。すばらしい成果であり、この新しいソリューションで取り組みを進める重要な原動力となっています」

TextNow、ソフトウェアテクニカルマネージャー、  
Andres Ivanov 氏



# DCI がクラウドコストを 27% 削減し、AWS で長期的成長を実現

## 課題

Digital Commerce Intelligence (DCI) は、市場と競争に関するインサイトを提供し、顧客がデータ駆動型の戦略を策定できるようサポートしています。DCI が成長するにつれ、当時利用していたクラウドサービスプロバイダーでは必要な柔軟性が得られなくなり、結果として予測不可能なコンピューティングコストとデータベースコストが生じていました。同社は投資家から、以前のクラウドプロバイダーで生じていたような請求の問題を回避するために、AWS への移行を提案されました。

## AWS のソリューション

同社は AWS スタートアッププログラムを使用してわずか 6 か月間で AWS に移行しました。スタートアップに無料のツールやリソースを提供する **AWS Activate** に参加したことで、同社のビジネスは迅速に進むようになりました。アカウントチームと AWS サポートエンジニアのガイドを受けることで、DCI はデータ収集ツール、SQL Server、メッセージングキュー、Kubernetes クラスタ、イメージレジストリ、コンピューティングを AWS に移行できました。

## 成果

これまでに、AWS への移行によって IT の月間コストを 27% 削減しています。DCI は、アルゴリズムを実行して結果を顧客に提供するために大量のデータを取り込んで処理する必要があるため、こうしたコスト削減は重要です。提供された結果から、DCI の顧客は、ビジネスをよりインテリジェントに進めるために必要な市場インサイトを得ることができます。

[全文を読む >](#)



「コスト削減により、人材や製品開発に資金を回して、ビジネスの競争力を高めることができます」

Digital Commerce Intelligence、CTO、Konstantinos Kitsaras 氏



# SkyTouch Technology、ビジネスクリティカルなアプリケーションをAWS クラウドに移行することで俊敏性とコストメリットを確保

## 課題

Choice Hotels の一部門である SkyTouch Technology は、業界で最も広く利用されているクラウドベースのホテル管理システムを提供しています。ホテル業界は、新型コロナウイルス感染症のパンデミックにより大きな打撃を受けました。客室稼働率が 30% 以上低下する中、SkyTouch の最大の課題の 1 つは、いかに運営コストを削減するかということでした。

## AWS のソリューション

SkyTouch は、重要なビジネスアプリケーションを AWS に移行するという大きな仕事を成し遂げました。SkyTouch は、顧客に対するダウンタイムの影響を最小限に抑えるだけでなく、システム、顧客、サードパーティ間で数多くの統合を計画する必要がありました。**AWS Migration Acceleration Program (MAP) for Windows** によって、SkyTouch はベストプラクティス、ツール、インセンティブを利用して移行目標をより迅速に達成することができました。

## 成果

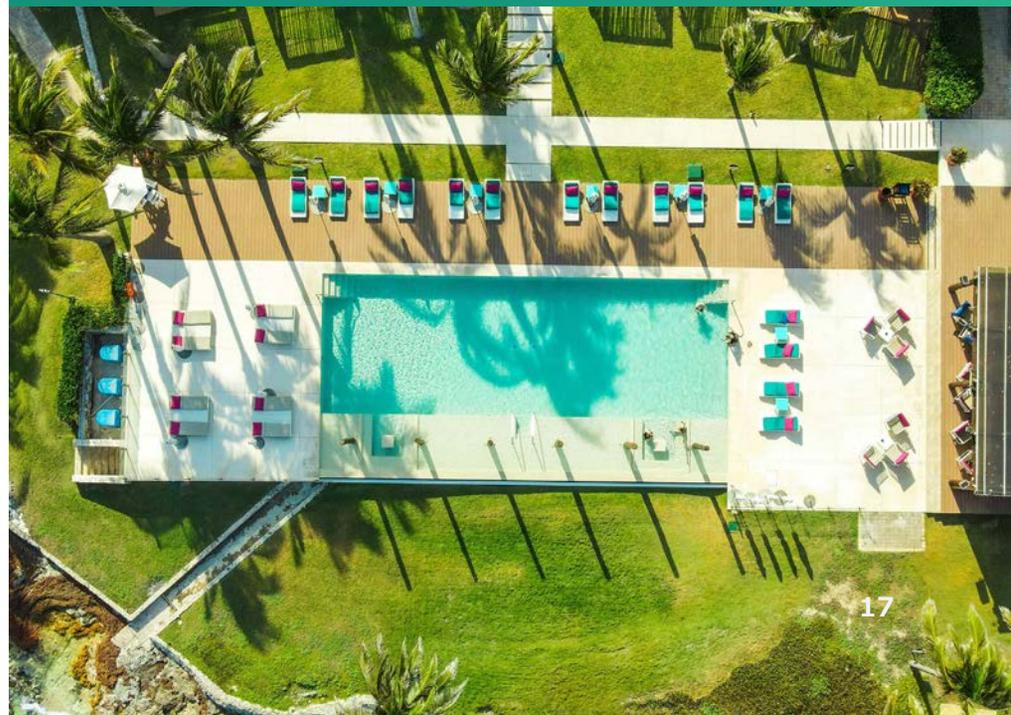
2018 年に AWS と提携して以降、SkyTouch のホテル管理システムのダウンタイムは前年比で 77% 削減されています。また、AWS Auto Scaling の導入や、AWS Savings Plans、AWS リザーブドインスタンスの活用により、インフラストラクチャコストの最適化にも成功しています。これらの取り組みにより、オンデマンド料金と比較してデータコストを 35% 削減しました。

[全文を読む \(英語\) >](#)



「当社の需要は大きな起伏があります。簡単にスケールアップ、スケールダウンできるインフラストラクチャをプロビジョニングすることで、需要に応じたキャパシティを確保することができます。この柔軟性により、インフラストラクチャリソースのサイズ適正化にかかるオーバーヘッドを大幅に削減することができます」

SkyTouch Technology、クラウドプラットフォームエンジニアリング担当ディレクター、Steve Garvey 氏



# Majelan が Amazon Athena を使用してユーザーの行動の把握とコストコントロールを実現

## 課題

2019年にリリースされた Majelan は、50か国の1,800万以上のポッドキャストに無料でアクセスできるアプリです。このインテリジェントなアプリケーションは、コンテンツがどのように聴かれているかに関するきめ細かいデータを収集し、レコメンデーションや番組改善を行うことを目的としています。当初、Majelan は Amazon Elasticsearch Service と自社サーバーを使用して分析を行っていました。しかし、ユーザー数やプラットフォーム利用が増えるにつれ、データの保存や分析の面で、これらのサービスの限界に直面するようになりました。

## AWS のソリューション

Majelan には、大量のデータを分析できるハイパフォーマンスなソリューションが必要でした。その技術的なニーズを AWS と Google Cloud Platform (GCP) の双方に相談し、最終的に AWS と **Amazon Athena** を選択しました。

## 成果

Amazon Athena は、大量のデータに対して SQL クエリを実行し、クエリ時間を10～15分から数秒に短縮することを可能にします。また、このソリューションにより、Majelan のコストコントロールも大幅に向上しました。「計算とデータ処理を行わない場合、Athena で発生するのは比較的低コストのストレージ料金だけです」と Majelan のデータディレクターの Thomas Fillon 氏は述べています。

[全文を読む >](#)



「S3 ファイルをクロールして Athena に構造化ビューを提供するインデックス作成ロボットを使用すると、そのサービスでリクエストごとに2億件のイベントを読み取る必要がなくなります。Athena のコストはクエリを実行するために読み取られるデータ量とリンクしているため、これは重要なことです。コストを制限するには、すべてのデータを使用しないインテリジェントなクエリを作成すること、例えば、日付、ユーザー、使用タイプでファイルを分割することで実現できます」

Majelan、データディレクター、Thomas Fillon 氏



# Apify がウェブインサイトを強化し、AWS でクラウドコストを 25% 削減

## 課題

2015 年に設立されたプラハのスタートアップである Apify は、ウェブスクレイピングと自動化のプラットフォームとオープンソースツールのセットを開発し、大量のウェブデータの収集と分析、ウェブ処理の自動化によって企業の業務改善を支援しています。Apify はスタートアップとして、信頼性とコスト効果に優れ、顧客ニーズの拡大に応じて容易にスケールできるインフラストラクチャを必要としていました。

## AWS のソリューション

Apify は、スタートアップに無料でツールやリソースを提供するプログラムである **AWS Activate** を通じて AWS の利用を開始しました。Apify のアプローチでは、顧客がウェブデータに確実にアクセスして分析できるように、大量の情報を高速処理する必要があります。**Amazon EC2** と、クラウドまたはオンプレミスで Kubernetes アプリケーションを実行およびスケールするマネージドコンテナサービスである **Amazon Elastic Kubernetes Service** (Amazon EKS) を使用して、毎月 2,000 万件のウェブ自動化ジョブを処理しています。

## 成果

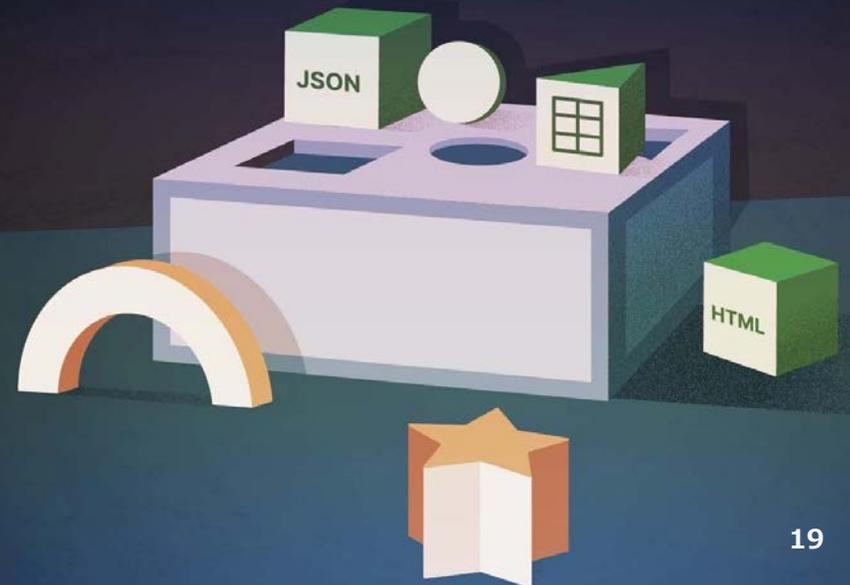
AWS でサービスを構築することで同社は急成長を遂げ、現在では世界 179 か国、1,000 組織にサービスを提供しています。また、月に 1,000 テラバイトのデータを処理する規模にまでスケールし、耐障害性の高いワークロードを最大 90% の割引価格で実行する Amazon EC2 スポットインスタンスを使用して、クラウドの総コストを 25% 削減しています。

[全文を読む >](#)



「AWS を選択した主な理由は、AWS Activate を通じて受けたクレジット、トレーニング、そしてコストコントロールのサポートにあります。このようなサポートが初期段階での成功に寄与したことは間違いありません」

Apify, CTO, Marek Trunkat 氏



# 次のステップ

データ戦略のコストの最適化はスタートアップにとって最優先事項であり、その最適化によって、チームは最も時間をかける必要がある業務に、継続的に時間を費やせるようになります。

AWS を利用することで、コストをコントロールし、継続的にデータの支出を最適化しながら、ニーズに合ったスケーラブルなモダンアプリケーションを構築できます。AWS では、コストを効果的に管理し、必要なパフォーマンスとキャパシティを維持するための柔軟性を備えた幅広いサービスと価格オプションを提供しており、スタートアップが最大限のコスト削減を達成できるよう支援します。

**AWS の利用を開始するには、AWS Data for Startups  
をご覧ください >**