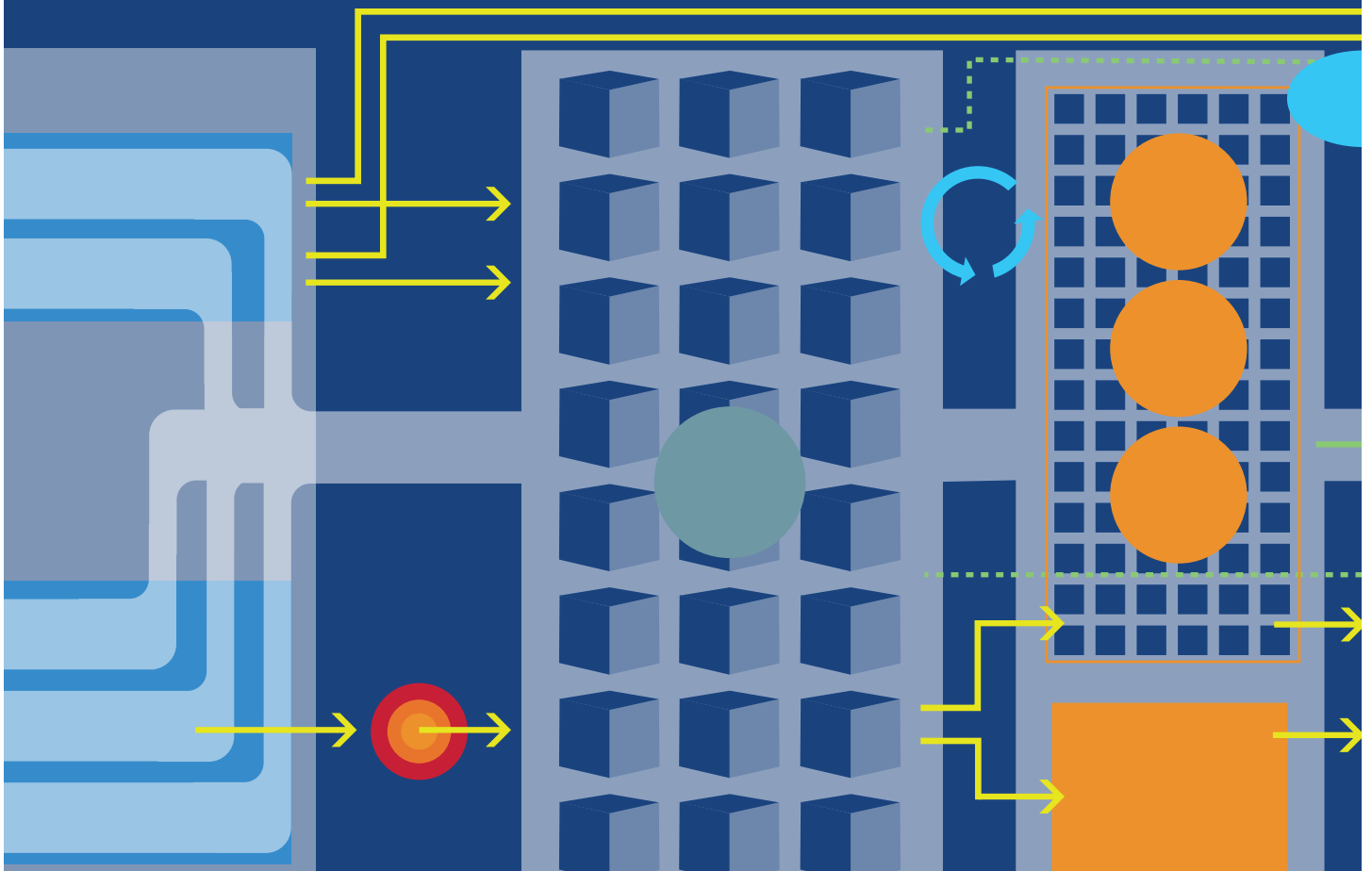




# 次世代のクラウド BI: Amazon EC2 上の Tableau Server



# 目次

はじめに.....	3
導入と拡張性.....	4
セキュリティとガバナンス.....	5
セルフサービス BI のユーザーエクスペリエンス.....	6
ケーススタディ .....	8
TripAdvisor .....	8
Netflix .....	9
最初の一步.....	10
Tableau およびその他のリソースについて .....	11

# はじめに

## 要約

Amazon Web Services (AWS) と Tableau Software で実現する包括的なビジネスインテリジェンスソリューションは、導入に時間がかからず、セキュリティ保護も容易で、あらゆるユーザーが利用できます。Tableau を利用することで、Amazon Redshift、Amazon Elastic MapReduce (EMR)、Amazon Aurora、Amazon Relational Database Service (Amazon RDS) などの AWS データサービスで管理しているデータの分析や共有を行うことができます。このホワイトペーパーでは、データの表示、理解、共有を進めるために、Amazon Elastic Compute Cloud (EC2) 上で Tableau Server を稼働させることをお勧めする理由を説明します。これにより、AWS 上でミッションクリティカルなビジュアル分析プラットフォームを十分に活用できるようになります。

## クラウドでの最先端のビジネスインテリジェンスプラクティスの必要性

デジタル変革と、増加を続ける「モノ」により、これまでにない量のデータが今までにないスピードで生成されています。データの使用例が増え続け、また意思決定のためにデータにアクセスする必要性が高まっている今、データを組織全体にとって有益なリソースにできる方法を見つけることは複雑な課題となっています。しかし、適切なツールをユーザーに提供できれば、データは非常に大きな資産となります。ますます多様になるデータストリーム全体からインサイトを取得しようと取り組む組織は、他者とは一線を画す競争力を得ることになります。

データ分析やビジュアライゼーションに使われるアプリケーションは、データそのものに引き寄せられます。そのため現在は、クラウドへの大規模な移行が進んでいます。Amazon Redshift、Amazon EMR、Amazon Aurora、Amazon RDS など、AWS のビッグデータおよびデータベースサービスを利用すると、どのような組織でもデータの収集や利用が可能となります。クラウドベースで大規模なビッグデータ処理を行っている組織でも、まだデータの分析にはあまり手を付けていない組織でも、ビジネス部門や IT 部門のユーザーがデータからパターンを視覚化し、分析してインサイトを得ることができるになれば大きなメリットが得られます。そのためには、Tableau のような最先端の BI ソリューションが役立ちます。

Amazon EC2 は、AWS クラウド上で高性能な仮想マシンのインスタンスを柔軟にコスト効率良く利用できるサービスで、Tableau Server を稼働させるのに理想的なプラットフォームです。Amazon EC2 を利用することで、ハードウェアの調達や展開の手間をかけることも膨大な先行投資も必要なく、BI の処理に必要なだけの仮想マシンインスタンスを利用することができます。オンプレミスで高可用性を実現し、リソース利用率を最適化するには多大な労力が必要になりますが、Amazon EC2 ではインスタンス数の増減や負荷分散などが自動的に行われるので、手作業の負荷も大幅に節約できます。

Amazon EC2 と Tableau を組み合わせれば、モダン BI を短期間で導入し、データの活用度を高めることができます。Amazon Redshift や Amazon EMR、Amazon Aurora、Amazon RDS を初めとするさまざまな AWS サービスで生成されたデータからインサイトを引き出し、発見した情報を社内で共有できれば、AWS の投資価値をさらに高めることができます。

# 導入と拡張性

## 規模に応じてアジャイル性を維持するための課題

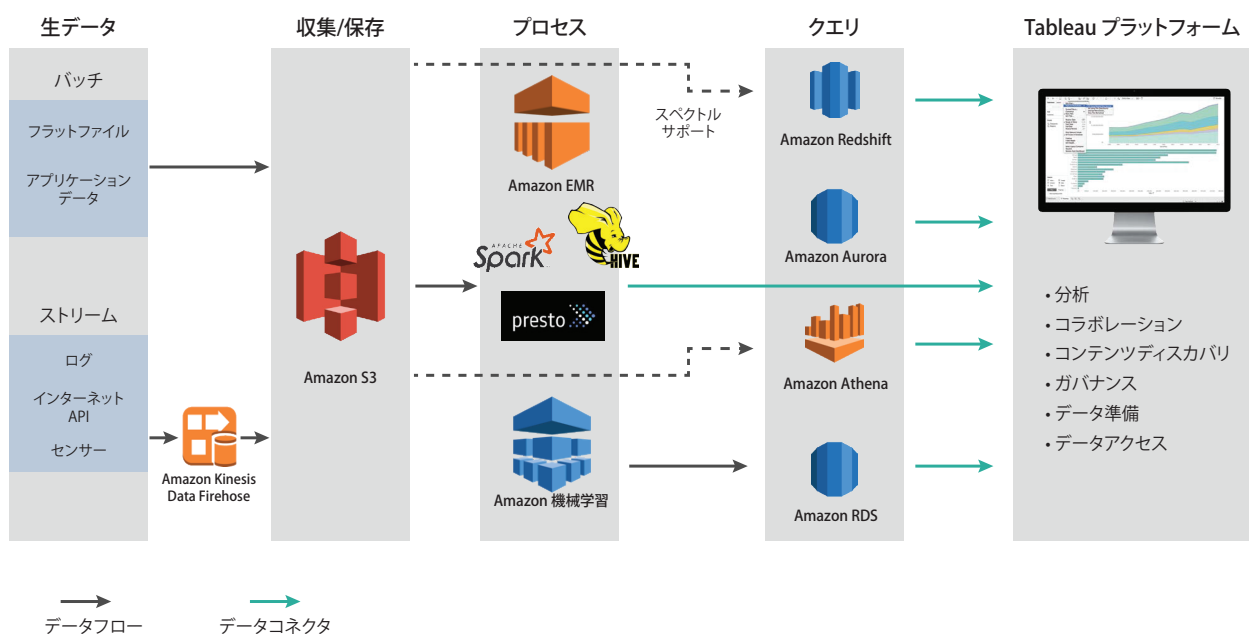
オンプレミス環境に従来型の BI 製品を取り入れる場合、導入も拡張も複雑で、かなりの時間と費用が必要になります。新たなデータソースに対応するには、その度に投資を行い、リードタイムを取る必要があります。また、新たな地域に広げていく場合にも、サポート対象の人数にかかわらず、追加投資が必要になります。

Tableau の一番の特長は直感的な UI です。これにより、どのようなスキルレベルのビジネスユーザーでもビジュアル分析を利用できます。さらに、基盤となるインフラストラクチャの管理やソフトウェアのインストール、維持、更新を担う IT 担当者にとっても、柔軟で扱いやすい製品となっています。ただし、そのスケーラビリティを十分に生かすために、モダン BI ではハードウェアレベルのアジャイル性が求められます。これはオンプレミス環境では難しい場合がありますが、Amazon EC2 ではそのような課題の多くを回避できます。

## クラウドの柔軟性

Amazon EC2 では、高パフォーマンスな仮想マシンのインスタンスを柔軟かつコスト効率良く、簡単に利用できることで、Tableau のワークロードを稼働させ、必要に応じて拡張や縮小を行うことも容易にできます。AWS ではさまざまな種類のインスタンスが揃っており、インスタンス数の増減や負荷分散を自動的に行うことができ、長期契約は不要で使用量に応じた価格設定となっています。オンプレミスでの展開とは異なり、ハードウェアの調達や実装が不要なので、Amazon EC2 ならば必要な容量を数分で用意することができます。容量の拡張が必要になっても、即座に対応することができます。Amazon EC2 で Tableau を稼働させると、Amazon Redshift や Amazon Aurora、Amazon Elastic MapReduce、Amazon Relational Database Service などの AWS 製品を初めとするさまざまな AWS のワークロードに加え、サードパーティのソリューションとも簡単に統合することができます。

## 「AWS 上の Tableau」プラットフォーム



## セキュリティとガバナンス

### データの整合性と下流での分析

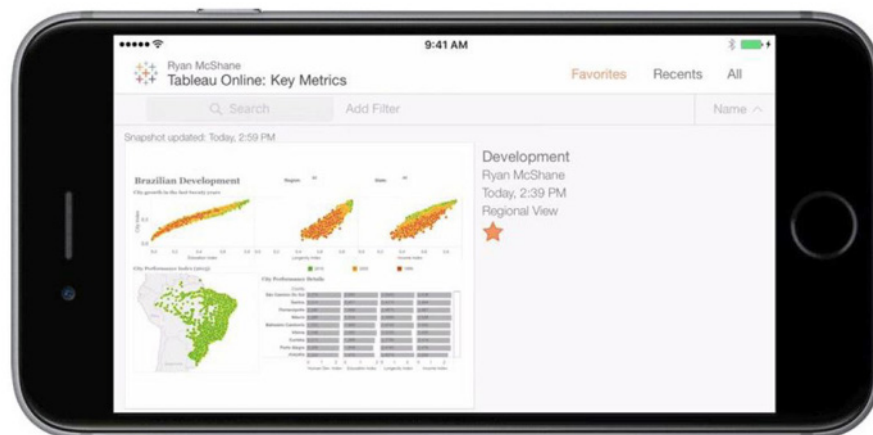
BI ソリューションを活用することになれば、通常は、大量のデータが一か所に集約され、大きな規模で共有されることとなります。そのデータの一部あるいはそのすべてが機密性の高いものです。社内で広く利用できるように、ユーザーにとって使いやすい BI ソリューションも数多く存在しますが、その多くは、企業のデータや知的財産をしっかりと保護できるようなセキュリティに欠けています。たとえば、表計算ソフト (スプレッドシート) を利用した下流での分析は、長年にわたって行われています。ただしスプレッドシートでは、データへのアクセス制御が、できたとしても非常に限られた形でしか行えません。共有されているすべてのユーザーが編集を加えることができる一方で、誰がいつどのような変更を加えたのかを管理者が確認できる監査証跡の機能は備わっていません。数式が壊れていたり、データに不備があっても、そのエラーがどこで起きたのかを確かめる手段は、現実的にほとんどありません。残念ながら旧式の BI プラクティスでは、セキュリティやガバナンス関連の脆弱性に関する例がいくつも挙げられます。

### 必要なセキュリティとガバナンス

Amazon が提供する世界トップクラスのクラウドセキュリティに加えて、Tableau Server はプロジェクトやワークブックに関する広範かつ詳細なセキュリティおよびガバナンス機能を提供します。強固なデータガバナンスの確立を支援するために、Tableau Data Server では、誰もが使えるように、データソースの定義と認証、および計算やデータフィールドの標準化を管理者が行えるようになっています。これによって、大規模で複雑なチーム間でデータモデルを共有するときに発生する可能性があるバージョン管理やセキュリティの問題を最小限に抑えることができます。さらに便利なことに、Tableau Data Server ではユーザーのローカルマシンにデータベースドライバーをインストールする必要もありません。

管理者は、パブリッシュする前に、フィルターや行レベルのセキュリティを設定することで、どのユーザーがどのデータを参照でき、どのユーザーがどのダッシュボードを利用できるかを指定できます。たとえば、地域担当マネージャーと売上レポートを共有する際に、西部担当のマネージャーには西部の売上だけを、東部担当のマネージャーには東部の売上だけを見せたい場合があります。そのような場合は行レベルのセキュリティを設定すれば、どちらのマネージャーにも同じビューを提供しながら、それぞれのマネージャーに担当地域のデータだけを見せることができます。Microsoft Active Directory やその他の認証プロセスとの統合により、データを参照できるユーザーやグループを制御するパーミッションの設定も簡単に行うことができます。稼働中のデータベースに接続するデータソースについては、パブリッシュされたビューをユーザーがクリックしたら、データベースへのアクセスに必要なログイン情報の入力を求めるかどうかを、管理者が設定することもできます。Tableau Server はバージョン管理機能も備えているので、必要に応じて、ビューを以前のバージョンに戻すのも簡単です。

Tableau Server にパブリッシュされたダッシュボードには、どのブラウザからでも安全にアクセスし、編集することができるため、より多くのチームメンバーが Web ブラウザやタブレットからデータを編集することで、データを掘り下げ、新しい質問の答えを得ることができます。作成したビューを、Salesforce や Microsoft SharePoint などの一般的な Web アプリケーションに埋め込むこともできます。さらに、Tableau Mobile では、モバイルデバイスからデータやダッシュボードにアクセスできるので、どこにいても作成や参照ができます。



## セルフサービス BI のユーザーエクスペリエンス ユーザーエクスペリエンスのジレンマ

このようなメリットを実現するうえで、多くの企業で大きな課題となっているのが、社内に展開した BI 製品がユーザーにあまり受け入れられないことです。通常の業務でデータを扱わない人たちにとって、ほとんどの BI 製品は不必要なまでに複雑で、インターフェイスはわかりづらく、質問の答えが得られるようになっていないので、使いたいと思えないのです。用意された BI 製品があまりにも複雑なものだと、自分のために作られているものとは感じられず、使う理由もないと考えてしまいがちです。製品のインターフェイスが扱いづらいため、ユーザーは自分の勘に頼るか、社内の他の人に頼ったりして、それまでのやり方を変えずに作業し続ける可能性が高くなります。

また、提供されている BI 製品ではデータの分析や視覚化の方法が限られていて、求めているタイプの情報にはあまり適していないと感じると、その製品を業務に取り入れていこうとは思わないものです。こうしたさまざまな条件を BI 製品がクリアしておらず、また、そこから得られるデータをユーザーが自分の業務で特に必要としていなければ、わざわざ手間暇をかけてまで使い方を習得するほどのメリットはないと思うのも無理のないことです。

## 誰もが利用できる画期的なテクノロジーと高速なビジュアル分析

Tableau は、思考しながら操作できるように設計されているので、さまざまなデータソースをドラッグしてシンプルな 1 つのキャンバスにドロップし、指定した変数に従ってデータを表示して、自然な感覚で瞬時にパターンを見つけ出すことができます。誰でも、浮かんだ疑問の答えを探すことや、傾向やチャンスを見つけ出すことができ、データを基に確信を持って意思決定を行えるようになります。Tableau Desktop は、単純なスプレッドシートからデータベースや複雑なビッグデータ環境まで、オンプレミスかクラウドかを問わず、ほぼあらゆるデータソースに対応しています。Tableau Desktop で作成したダッシュボードは Tableau Server で安全に共有できるので、得られた情報をユーザー自身がセルフサービスで、社内の他のユーザーやパートナーと共有することができ、データの用途の拡大とデータに基づく意思決定に役立ちます。

VizQL™ を初めとする特許取得済みのテクノロジーによって、クエリとビジュアライゼーションが 1 つのプロセスに統合されているため、ユーザーは自分の業務で求められる課題から注意をそらすことなく、データに関する質問の答えを見つけ出すことができます。Tableau Desktop を使用すれば、誰でもさまざまなデータソースに接続してビジュアライゼーションを作成できます。そこで新しい質問の答えを探し、データから傾向やチャンスを見出し、データに基づく意思決定を行うことができます。また、既存のデータから新しい計算を構築することや、リファレンスラインや予測の表示もでき、統計情報の概要を確認したり、トレンドや回帰、相関の分析を行うこともできます。Tableau Server を利用すると、インタラクティブなプレゼンテーションを作成することができます。プレゼンターの意図したストーリーは維持したまま、聞き手が自由にデータを探索できるので、さまざまなチームが共同で取り組む場合などにも効果を発揮します。

さらに、独自のダイナミックコード生成と最新の並行処理テクニックを使用する Tableau のデータエンジン、Hyper が、大規模または複雑なデータセットを迅速に分析できるよう支援します。これにより、抽出の作成とクエリの実行を迅速に行い、何十億ものデータ行のアドホック分析を行うことができます。

## ケーススタディ

### AWS 上の Tableau から得るインサイトでユーザー当たりの収益を増加させている TripAdvisor 社

世界トップの旅行サイトを運営する TripAdvisor 社では、AWS で Tableau を稼働させることにより、ここまでに紹介したさまざまなメリットを実現しています。TripAdvisor 社の収益最適化 (RevOps) 部門は、ユーザー当たりの収益を年間 15% 増加させることを目標とする部署です。この目標達成のために、会計四半期当たりおよそ 45 回の A/B テストを実施しています。

Tableau を使用する前、同社の収益最適化チームのプロダクトマネージャーは、各 A/B テストの終了後、その内容をまとめたメールを受信していました。しかし、そのメールの内容ではたいいてい、PM たちが知りたいことの半分程度にしか答えが出ておらず、自分で計算をする必要があり、本当に必要な情報が不足していました。

RevOps チームの PM は、答えを得るために必要な詳細情報を求めて、端末を起動し、一連の作成済みの Hive クエリを実行して、結果が出るまで待つ必要がありました。そして、出力されたデータを Excel に取り込み、さらに数時間かけてピボットテーブルや vlookup、さまざまな計算式を駆使して、求めている答えをはじき出します。このクエリでは答えが得られない場合には、ソフトウェアエンジニアにカスタムクエリの作成を依頼しなければならず、それを待っているとその日はもうぼぼつぶれてしまっていました。

最終的に、これではチームメンバーの役に立たないことを実感した同社は、BI 製品が必要であると認識しました。つまり、収益を増加させるためにサイトに関する新しいアイデアの生成をより迅速に反復すること、クエリを実行してより迅速な意思決定に必要なデータを収集すること、測定に関する自信を高めること、組織全体に結果を通知する方法を向上させること、各 A/B テストからより多くを学ぶことが可能な製品です。また、日報で PM が求めている答えをすべて網羅することで、エンジニアにかかっている負担を軽減し、展開したソリューションを簡単に拡張できるようにするための方法も必要となっていました。

TripAdvisor では、TripDNA という名前のソリューションを構築することを決定しました。これは、Amazon EC2 上で Tableau Server を使用して、Amazon Redshift のデータウェアハウスからデータを引き出すものです。3 時間足らずで展開できたこのソリューションによって、PM の業務が根底から変わりました。PM は自ら、テスト結果を分析できるようになり、エンジニアに頼らずに、テストの過程で発生した質問の答えを見つけられるようになっています。

Tableau と AWS を活用した TripDNA を利用するようになり、TripAdvisor では、ユーザー当たりの収益を初年度で 22%、2 年目の前半には 29% 増加させています。メールを受け取ってスプレッドシートにまとめ、あちこち参照してまわる必要もなくなり、エラーや手作業での計算を劇的に減らすことができました。経営陣にとっても、目標に向けたチームの進捗状況がわかりやすくなり、達成されつつある結果に対する信頼感も高まり、PM との連携もよりスムーズになっています。



## Netflix が Tableau と AWS を使用してクラウドでの分析を構築した方法

Netflix は、よく知られている通り、全インターネットトラフィックの 3 分の 1 以上をサポートするまで成長したため、それに見合うようにデータ機能を拡張する必要がありました。アマゾン ウェブ サービスおよび Tableau 上に構築された同社の大規模な分析プラットフォームは、スケーラブルで柔軟なビジネスインテリジェンスをクラウドに構築したいと考えている組織にとって非常に優れた青写真となります。

その会員数は 8,600 万人からさらに増え続けており、レポート作成プラットフォームマネージャーの Albert Wong 氏は、分析に関するすべてを規模に合わせて構築する必要があります。前述の同社のデータプラットフォームは、Amazon S3 に取り込まれるイベントデータおよび運用データに基づいて構築された、複雑で洗練されたシステムです。データは最終用途に合わせて Amazon Redshift または一連の NoSQL データプロセッサに送信され、その後、Tableau データ抽出に集約されます。このデータレイク/データウェアハウス戦略により、Netflix は膨大な量のデータの収集と保存が可能になるだけでなく、分析および探索するデータの概要ビューを表示することもできます。もちろん、すべてのデータ接続や抽出は Amazon EC2 上の Tableau Server で完了します。

Netflix は Tableau Data Server を使用しているため、データソースの再利用や、広範なユーザー全体の管理が可能です。たとえば、Albert Wong 氏が開発した最も重要なダッシュボードの 1 つは、各国内の使用率と視聴パターンを示すものです。これらのダッシュボードにより、各国のマネージャーは視聴者向けの番組を簡単に管理できます。

ダッシュボードから恩恵を受けるユーザーは何十人もいますが、そのために使用しているのは 1 つのデータソースのみです。同氏はそのデータソースについてグループレベルおよび個人レベルの両方のパーミッションを設定し、各チームが適切な情報にアクセスできるようにしました。Data Server を使用することで、管理された 1 つのデータセットを使用してチームが複数の質問への回答を得ることが可能になっています。

## 最初の一步

プロビジョニングしている Amazon EC2 インスタンス上に Tableau Server を自社で導入する、Tableau Server 内の AWS CloudFormation テンプレートを使用して AWS Quick Start に導入する、あるいは AWS Marketplace Amazon Machine Image (AMI) を使用して AWS 上に Tableau Server を導入することができます。

## Tableau は AWS インフラストラクチャ上に導入可能

### Amazon

サーバー

ワークブック、接続または抽出

デスクトップ

### お客様が開始する方法

Tableau Server	AWS セルフデプロイ	AWS クイックスタート	AWS Marketplace AMI
実稼働準備完了	✓	✓	✗
アップグレード可能	✓	✓	✗
スケーラブル	✓	✓	✗
分散インストール	✓	✓	✗
Active Directory サポート	✓	✗	✗
14日間トライアルライセンス	✓	✓	✓
BYOL ライセンス	✓	✓	✓
サブスクリプションライセンス	✓	✓	✓

実稼働PoC/テスト/評価

AWS で Tableau を実行するうえで最も柔軟性とスケーラビリティが高いのは、Amazon EC2 のインスタンスを用意し、Tableau Server をダウンロードしてそのインスタンス上にインストールすることです。これにより組織は、非常に細かい管理が可能になり、ビジネスの要求に応じて自由に拡張できるようになります。パフォーマンスを重視する場合は、m4.4xlarge など、vCPU 数を 16 確保するようにします。

AWS Marketplace では、ワンクリックで Tableau Server インスタンスを導入できる機能が提供されています。BYOL (Bring-Your-Own-License) というライセンス持込型のインスタンスタイプでは、Tableau Server の無料トライアル版を利用でき、購入を決定したら、Tableau からライセンスを取得することができます。これにより、組織の AWS 環境に Tableau Server AMI を迅速に導入できます。

また Tableau は時間当たり 10、25、50 または 100 ライセンス単位で利用でき、このライセンス料金はそのまま AWS の請求に含まれます。これは AWS で Tableau を実行する最も柔軟性と拡張性の低い方法であり、概念実証のために最も使用されています。ライセンスの変更が必要な場合は、キャンセルして再度発注する必要があります。また、この方法ではクラスタリングは利用できないので、特に、複雑で要件の厳しいワークロードには向かない場合があります。それでも、AWS で Tableau の稼働を始めるには、最も時間がかからない、簡単な方法です。長期契約も不要で、AWS からの請求にすべてまとめられるので、手間がかからず非常に便利です。

# Tableau について

Tableau は、規模に応じた超高速セルフサービス分析を通じてお客様がデータを見て理解できるように支援する、全てがそろった使いやすいエンタープライズ対応のビジュアル BI プラットフォームです。Tableau と AWS サービスを統合することで、企業は、組織のデータとテクノロジーへの既存の投資を最大限に活用することができます。Tableau Server は Amazon のクラウドインフラストラクチャとシームレスに連携するため、すべてのアプリケーションを AWS に導入することを希望するお客様は、Tableau が提供する完全なソリューションを利用できます。最も貴重なアセットであるデータと人の力を解き放ちます。

## その他のリソース

[AWS への Tableau Server の導入に関するテクニカルガイド](#)

[AWS 上の Tableau Server に関するクイックガイド](#)

[AWS Marketplace から Tableau の使用を開始](#)

[Tableau とアマゾン ウェブ サービスの詳細情報](#)

[Tableau 製品の詳細情報](#)

[ホワイトペーパー: Tableau および Redshift のパフォーマンスの最適化](#)