



データグラビティはどのように分析環境をクラウドに誘引しているのか

# 目次

はじめに .....	3
ニュートンとデータの関係 .....	4
データグラビティ .....	7
分析における引力とは .....	8
速度 .....	9
クラウドに至る道 .....	10
データ、引力、そしてクラウド .....	11
ハイブリッド型データ分析が持つさまざまな側面 .....	12
移行を導く Tableau .....	14
Tableau およびその他のリソースについて .....	16

## はじめに

かつてデータとアプリケーションは、組織自体の施設でローカルにホスティングされるものでした。しかし、クラウドコンピューティング革命でこの当たり前の状況は変わりました。データがクラウドへと移行されつつある今、データグラビティが組織でのソフトウェアと分析の実行方法に変化をもたらしています。

パンデミックと経済危機はクラウドの採用を加速させる要因にしかならず、組織はデータドリブンなデジタル変革における重要なコンポーネントとしてクラウドを追求しています。データ分析およびクラウドテクノロジーは、効率の向上、最適化、コストの削減から顧客サービスの向上まで、さまざまな優れたメリットをもたらします。それらは、特殊なビジネス課題が存在している今、企業がアジャイル性と回復力を高めるのに役立ちます。

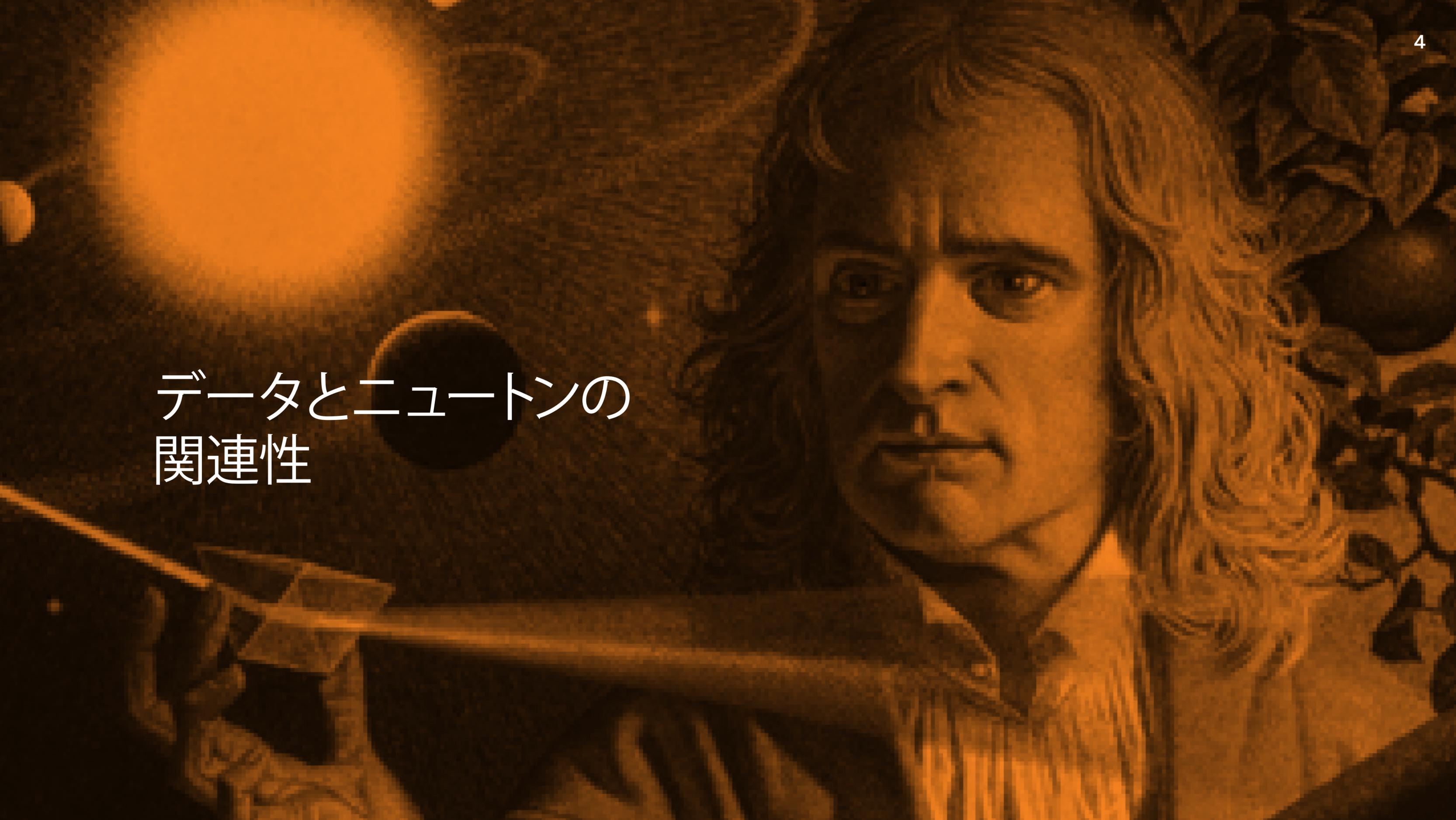
「パンデミックにより、クラウドの価値提案が実証されました。組織にとって、オンデマンドで拡張可能なクラウドモデルを使用してコスト効率とビジネスの継続性を達成できる能力が、デジタルビジネス変革計画を迅速に進めるための推進力となっています。パブリッククラウドサービスの使用の増加により、クラウドの採用はこれまで以上に「ニューノーマル」となっています」

— Gartner 社リサーチ担当バイスプレジデント、Sid Nag 氏

出典: Gartner プレスリリース

**Gartner、2021年に世界のパブリッククラウドエンドユーザーの支出が18%増加すると予測 (英語)**

2020年11月17日

A portrait of Isaac Newton, rendered in a dark, monochromatic style with a strong orange-brown tint. The image is a close-up of his face, looking slightly to the right. A large, solid black circle is superimposed on the left side of his face, partially obscuring his eye and cheek. The background is dark and textured, suggesting a historical setting.

# データとニュートンの 関連性

「宇宙の2体は互いに引き寄せ合う。その力はそれぞれの質量の積に正比例し、物体間の距離の2乗に反比例する」

— ニュートンの万有引力の法則

データ、アプリケーション、サービスはいずれもそれ自体の「引力(グラビティ)」を持っていますが、「質量」が最も大きいのはデータです。そのため、他のテクノロジーはデータに引き寄せられます。つまりデータがクラウドにあれば、データグラビティにより、他のアプリケーションやサービスもクラウドに引き寄せられます。そのため、データの取り込み、保存、分析、共有を迅速かつ簡単にできるようにするためにクラウドコンピューティングに依存する組織が増えれば増えるほど、データグラビティの力はより劇的になり、より大きな影響が生み出されます。

組織の多くはつい最近まで、データの保存とアプリケーションの実行をすべてオンプレミスで行っていました。しかし、クラウドコンピューティング革命でこの当たり前の状況は変わりました。そして、この革命によって、15世紀のニュートンによる法則が再度注目されることになりました。

データとアプリケーションはいずれオンプレミスからクラウドインフラに移行すると予想していた、草創期からのクラウド技術者である Dave McCrory 氏は、ニュートンの万有引力の法則が物体以外にも適用される可能性があることに気づきました。そしてこれを「データグラビティ」と名づけました。McCrory 氏はデータグラビティの概念に合わせて、万有引力の公式を次のように書き換えました。

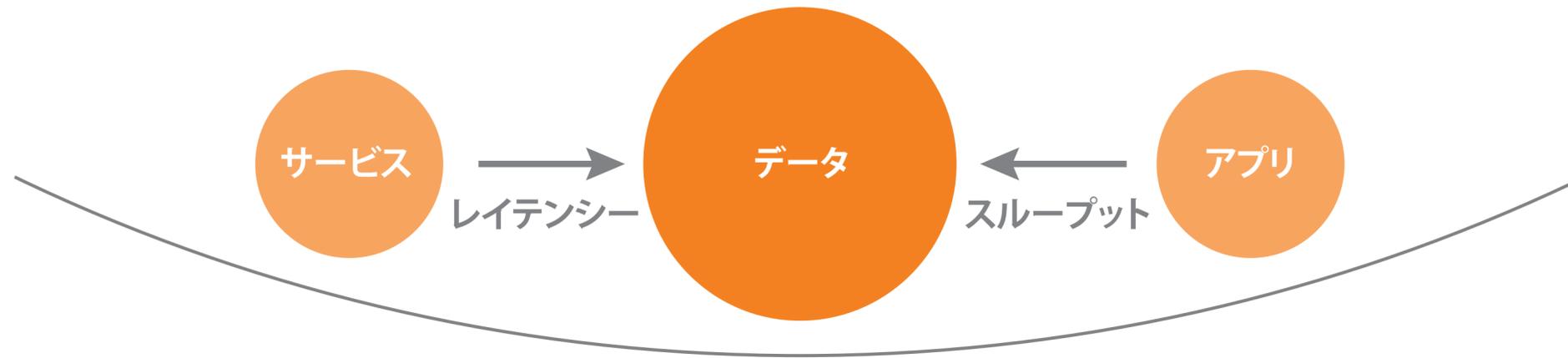
$$\frac{\left( \begin{array}{c} \text{データの} \\ \text{質量} \end{array} \times \begin{array}{c} \text{アプリケーションの} \\ \text{質量} \end{array} \right) \times \begin{array}{c} \text{1秒あたりの} \\ \text{リクエスト数} \end{array}}{\left( \begin{array}{c} \text{レイテンシー} \\ \text{(秒数)} \end{array} + \left( \begin{array}{c} \text{リクエストの} \\ \text{平均サイズ (MB)} \end{array} / \begin{array}{c} \text{1秒当たりの} \\ \text{帯域幅 (MB)} \end{array} \right) \right)^2}$$

簡単に言えば、データ、アプリケーション、サービスは、すべてそれぞれの引力を持っています。しかし、その中でデータが最も重要な存在であり、したがって最も強い引力を持っています。

ニュートンの頭にリンゴが落ちたのは、リンゴより地球の質量の方が大きく、引力として知られる現象でリンゴを引き寄せたからです。データも同様に、ソフトウェアアプリケーションやサービスを含む、その他のテクノロジーを引き寄せます。

## データグラビティ

データは、保存されている場所にアプリケーションとサービスを引き寄せる傾向を持っています。



\*出典: Dave McCorry 氏、[データグラビティ — クラウド内 \(英語\)](#)

## 分析における引力とは

**レイテンシー:** 処理を行う、または結果を出すのにかかる時間。

**スループット:** 単位時間あたりに処理を実行できる、または結果を出すことができる回数。

分析環境を考えるうえで、この2つの要素はなぜ重要なのでしょうか。

**レイテンシーとスループット**は、データグラビティを生み出す力です。ニュートンのリンゴとは異なり、データに周囲の物体を引き寄せる物理的な質量はありません。分析プロセスではその代わりに、レイテンシーとスループットが推進力として働きます。

分析環境に当てはめると、レイテンシーとは、クエリがソフトウェアアプリケーションからデータベースに送られ、戻ってくるまでの待ち時間のことです。

同様にスループットとは、ある一定の時間に、ソフトウェアアプリケーションがデータベースに対してクエリを実行できる回数のことです。

## 速度

適切な場所を選ぶことにより、速度を達成できます。

データは、人がそれを使って質問に答えを出せるときに初めて役に立ちます。そのようなインパクトを得るためには、タイムリーにデータにアクセスできる必要があります。クエリに何時間もかかるようでは、分析フローが中断され、有意義なアクションにつながるインサイトの取得が損なわれます。逆に、レイテンシーを低減しスループットを向上させればクエリはより短時間で返され、分析結果や答えもより短時間で得られるようになります。

システムのレイテンシーとスループットを左右する「変数」は数多くありますが、「定数」となるのが場所です。実体(データ、アプリケーション、サービス)が互いに近ければ、レイテンシーは低くスループットは高くなります。

この概念に基づくと、データを燃料とするアプリケーションやサービスに対して、データは引力を働かせます。したがってデータの場所は、速度を求めるうえで、データの収集、保存、分析に利用するアプリケーションの場所を決定するための重要な判断材料となります。

## クラウドに至る道

データハイウェイでは、行き先の標識はクラウドを指しています。クラウドインフラのスケラビリティと信頼性を活用しようと、ますます多くの企業がデータと分析をクラウドに移行しつつあります。

「パブリッククラウドの使用の増加により、あらゆる規模の組織でクラウドに対する支出が押し上げられています。パブリッククラウドの支出は、特に大規模な組織の間で IT 予算の重要項目となっています」

出典: [Flexera 2020 クラウドの現状 \(英語\)](#)

しかし、クラウドに至る道は1本ではありません。

データのみをオンプレミスからクラウドに移行する組織もあれば、インフラをクラウドプラットフォームに移行する組織もあります。多くの場合、組織はその両方を同時に行っています。さらには、設立当初からクラウドを利用し、Web アプリケーションとクラウドネイティブのデータのみで業務を行う組織も増えてきています。

## データ、引カ、そしてクラウド

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。

IDC の調査によると、テクノロジーの意思決定者のほぼ 50% が、パンデミックの結果としてクラウドコンピューティングの需要が中程度または大幅に増加すると予想しています。

「多くの CIO は、クラウドへの移行に対する投資が、パンデミックの際に報われたと語っています。簡単にスケールアップまたはスケールダウンできたからです」

— IDC 社最高リサーチ責任者、Meredith Whalen氏

出典: IDC、[COVID-19 が経営幹部にもたらした新しい優先事項](#) (英語)、2020 年 5 月

ビジネスインテリジェンスとクラウドコンピューティングテクノロジーは、ビジネスを効果的に支援するうえで重要であることが証明されています。そのアジリティは、デジタル変革において、多様で拡大し続けるデータセットの管理に役立ちます。動的で不安定な環境に直面しながらも、分析エクスペリエンスを高速に保ち、ビジネスをサポートするには、クラウドおよびエンタープライズ IT アーキテクチャ全体でサポートする分析環境の導入を計画する際に、データグラビティを考慮する必要があります。

しかし、現実にはクラウドに至る道が組織によって異なっているため、それは言葉で言うほど簡単なことではありません。サービスとアプリケーションをデータの近くに維持するには、環境の移行と進化に応じてプラットフォームを変更する必要がないように、ハイブリッド型のモデルをサポートする分析ツールとビジュアライゼーションツールが必要です。

## ハイブリッド型データ分析が持つさまざまな側面

分析環境のハイブリッド型モデルには、データがある場所にソフトウェアを導入するか、あるいはデータ戦略の進化に応じて導入環境を調整していくか、という選択肢があります。

分析インフラの計画策定では、データグラビティは単なる1要素に過ぎません。他にも、データアプリケーションの現在の利用状況や将来の計画という要素を検討する必要があります。

データを移行して、長期間使用されてきたワークフローを変更することは困難な作業となる可能性があるため、分析アプリケーションは移行過程のあらゆる段階でユーザーをサポートするものでなければなりません。

さらに何より重要なのは、システム管理者からデータを分析するエンドユーザーまでの全ユーザーに対し、そのニーズを考慮しなければならないという点です。オンプレミスのハードウェアとソフトウェアを管理する専門チームはありますか？あるいはビジネスユーザーは、完全に管理されたクラウドアプリケーションに容易にアクセスできることを求めていますか？

ハイブリッド型モデルは、データ、インフラ、アプリケーションに対し、オンプレミスおよびクラウドテクノロジーを両方サポートしています。そのため、分析環境を導入する場所をすぐに選べる柔軟性が提供されるとともに、環境の移行と変更の際にユーザーをサポートします。

データと分析のアプリケーションには、オンプレミスからクラウド上の完全ホスティングまで、さまざまな形があります。そして、それぞれにメリットがあります。

**オンプレミス** — データベースおよび分析アプリケーションは、組織が独自の施設でホスティングし管理します。つまり組織が、必要なハードウェアのプロビジョニングと、将来の需要に合わせたハードウェアの拡張に責任を負うということです。また、組織はソフトウェアの管理と保守も積極的に行います。

**IaaS および PaaS** — 独自のハードウェアを購入する代わりに、AWS、GCP、Microsoft Azure などのパブリッククラウドベンダーからインフラを借り、データベースや分析環境をクラウドに導入することができます。多くの場合、サービスとしてのインフラ (IaaS) やサービスとしてのプラットフォーム (PaaS) の利用には、コスト、スケーラビリティ、柔軟性の面でメリットがあります。

**St. Mary's Bank** は、データと分析をオンプレミスからクラウドに移行することで、約 4 万件のデータエラーを解決し、自動レポートによって週あたり約 15 時間を節約しています。

[St. Mary's Bank のストーリー全文を見る](#)

**完全ホスティング型の SaaS** — Web ベースの分析も SaaS として提供できます。これは、ハードウェアやソフトウェアの保守を考慮する必要がないことを意味します。

忘れてはならないのは、データが最大の質量を持っていることです。データの移動には、時間的にもリソース的にもコストがかかります。分析環境のハイブリッド型モデルでは、保存先のデータベースやホスティング元のインフラにかかわらず、データに接続することができます。

**Specialized Bicycle Components** 社は、インフラストラクチャを排除して、認証とセキュリティのアップグレードを可能にし、BI を五大洲に拡張しました。

[Specialized Bicycle Components 社のストーリー全文を読む](#)

## 移行を導く Tableau

組織が事業でクラウドを活用する方法は数多くあります。また、クラウドに至る道もさまざまです。アプリケーションは、現在と将来の両方のニーズを満たすものを選ばなければなりません。

Tableau によるクラウドへのアプローチは、お客様に選択を任せるシンプルなものです。お客様が分析環境の導入の方法と場所を選択します。そして、データがある場所にかかわらず、あらゆるデータを分析するように選択します。完全ホスティング型の SaaS ソリューションでも、クラウドプラットフォームまたはオンプレミスでの認可ソフトウェアの導入でも、Tableau なら思い通りに分析環境を導入して管理することができます。

クラウドに至る道のどの段階でも、Tableau が支援します。

### Tableau Online

Tableau Online は、クラウドでホスティングされたセルフサービス分析環境です。セキュアかつスケーラブルであり、サーバーを管理する必要はありません。

### Tableau Server

オンプレミスでもパブリッククラウドでも、Tableau Server を利用すると、データの価値を組織全体で活用できます。

### Tableau Desktop

ビジュアル分析の「ゴールドスタンダード」と呼ばれる Tableau Desktop は、直感的なインターフェイスによる制限のないデータ探索を実現します。

### Tableau Prep

Tableau Prep では、データの結合、形式変換、クリーニングのほか、データ準備フローの自動化も視覚的に直接行うことができます。

「Tableau Online は、組織のさまざまな側面を1つにまとめ上げる機会をもたらしてくれます。また、世界各地やさまざまな部署にいる Red Hat 社員に簡単に提供できるため、導入が加速され、コミュニティが形成されて盛り上がりも生み出されました」

— Red Hat 社ビジネスインテリジェンス・分析部門シニアマネージャー、Wes Gelpi 氏

[ストーリー全文を読む](#)

## Tableau について

Tableau は、人々と組織がさらにデータドリブンになるのを支援する、エンタープライズ対応の完全な統合ビジュアル分析プラットフォームです。オンプレミスでもクラウドでも、また Windows でも Linux でも、Tableau はテクノロジーへの既存の投資を生かし、お客様のデータ環境の変化と成長に合わせた規模の拡張が可能です。お客様の最も貴重な資産であるデータと人材の力を解き放ちます。

## その他のリソース

### リソースハブ: データと分析をクラウドで

分析の最新化、クラウドへの移行計画、クラウド投資の最適化の準備など、クラウドに至る道のどの段階でも役立つリソースをご用意しています。

#### [ハブにアクセス](#)

### クラウドに移行する準備ができた組織のために

クラウドでの成功に向けて計画しましょう。クラウドへの移行に向けて組織が尋ねるべき最も重要な質問を見つけてください。

#### [ホワイトペーパーをダウンロード](#)

### 分析機能をクラウドに移行すべきかどうかの判断

Tableau Server (オンプレミスまたはパブリッククラウドでの導入) または Tableau Online (完全ホスティング型 SaaS 分析) のどちらを選択するかについて、重要な考慮事項をご確認ください。

#### [ブログ記事を読む](#)