

Vertica Advanced Analytics Platform

次世代ビッグデータ解析プラットフォーム

データシート

今日の大量データの処理

今日のデータインフラストラクチャーには、あらゆる場所からデータが入ってきます。CRMやERPといったビジネスシステム、マシン生成データの収集に用いられるセンサー、ツイートなどのソーシャルメディアのデータ、Webログやデータストリーム、ガスや電力の系統、モバイルネットワークなど、多様なデータソースがあります。このようにさまざまな場所から集められたデータの保存と管理には、多くの企業で低コストのストレージがばらばらに使用されており、高度なデータ分析を大規模に実施しようとする場合、複雑さとサイロ化が障害となります。

データの量、速度、多様性を管理するために、新しい革新的なビッグデータ分析プラットフォームが登場しています。それらは主にデータの規模と複雑性への対処を目的としていますが、中にはすべてのデータに一律な方法でアクセスするのが難しいものや、ごく限られたインフラストラクチャーオプションしか持たないものもあります。データ主導型企業のニーズに応えられる高度な分析プラットフォームとは、すべてのデータにアクセスでき、インフラストラクチャーに関する制限のないものだけです。データのマネタイゼーションから顧客の維持、コンプライアンス、トラフィックの最適化に至るまで、ビッグデータ分析プラットフォームを利用する企業は、小売、医療、通信、エネルギーなど、さまざまな産業の力学を変えつつあります。

ビッグデータ分析プラットフォームの主要なテクノロジー要件

そこで、現在と将来のデータの課題に応えるために、データ分析ソリューションに何が必要かを考えてみましょう。次の点について考えます。

大量のデータを統一的方法で分析: 多くの企業では、無制限の規模のデータ分析と、適切なときに適切な場所にデータを保存できる能力が求められています。現時点では、データのスケールはテラバイトやペタバイト単位です。近い将来は、エクサバイトも視野に入ってくるでしょう。

ビジネスの期待に応える: ユーザーは、結果を待つことを望みません。使用するソリューションは、高いスケーラビリティを備え、サービスレベルアグリーメント (SLA) とビジネスニーズを満たしながら、最低限の総所有コスト (TCO) で迅速に投資回収 (ROI) を実現できる必要があります。

一般的なツールの使用: ビッグデータ分析に抽出、変換、ロード型 (ETL) ツールやSQLベースのビジュアライゼーションを使用している場合、分析プラットフォームは、信頼性の高い強力なSQLを提供するとともに、すべてのツールとの組み合わせで動作することが検証されている必要があります。主要なベンダーのツールだけに対応しているものや、単独のインフラストラクチャープラットフォームに制限されるものは適しません。

Vertica Advanced Analytics Platform – 制限や妥協を排除

データベース分野の伝説的指導者であるマイケル・ストーンブレイカーが構想したVertica Advanced Analytics Platformは、最初のコード作成時からビッグデータ分析専用で構築されています。なぜでしょうか。それは、データウェアハウス、実証されていないオープンソースソリューション、そして従来どおりのビジネスの慣行が、テクノロジーの制約となり、ビジネスに苦しい妥協を強いているからです。Vertica Advanced Analytics Platformは、速度、スケーラビリティ、単純さ、オープン性を核として意識的に設計されており、分散型の圧縮されたカラム型アーキテクチャーを通じて分析ワークロードを処理するように構築されています。Vertica Advanced Analytics Platformは、驚異的な高速性 (クエリの実行が50 ~ 1,000倍高速)、エクサバイトのスケール (サーバー 1台あたりのデータ容量が10 ~ 30倍)、オープン性、単純さ (あらゆるビジネスインテリジェンス <BI> ツール、ETLツール、Hadoopなどが使用可能) と、従来のデータウェアハウスソリューションよりもはるかに低いコストを特長とし、実証されていないオープンソースソリューションよりも短期間で市場投入が可能です。

Verticaのパワーを生み出すテクノロジー

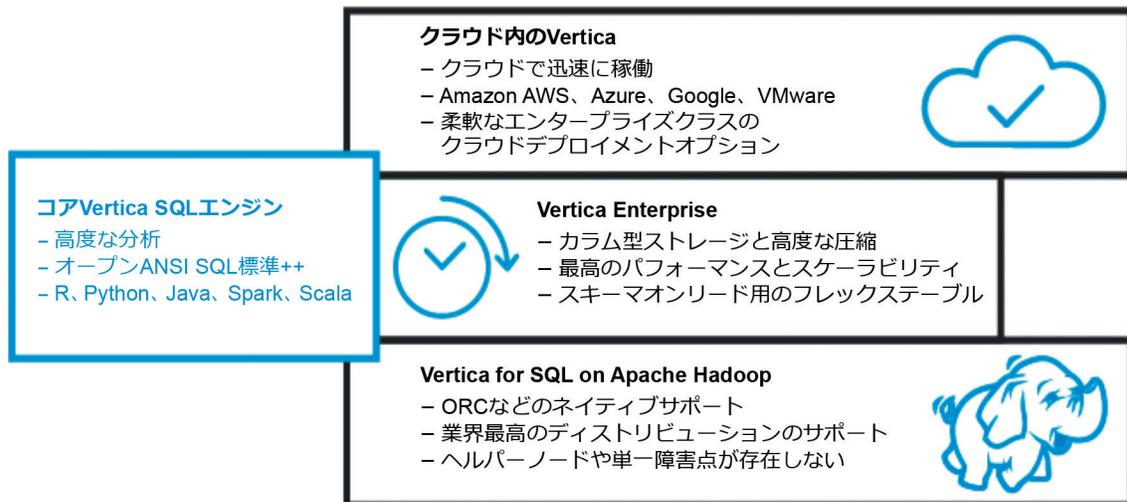
Verticaは、ビッグデータ分析の課題を解決するために一から設計されています。超並列処理システムにより、エクサバイトの規模に対応し、業界で最も条件の厳しいユースケースで実証されています。カラムストアとデータのアグレッシブな圧縮を利用することで、ビッグデータを超高速で分析できます。古い時代に構築された旧式の行ストアテクノロジーに比べて、クエリ時間を数時間から数分、あるいは数分から数秒に短縮できます。さらに、Verticaには、グラフ分析、三角形カウント、モンテカルロシミュレーション、地理空間分析といった高度なSQLベースの分析が組み込まれています。すなわち、フル機能の分析データベースシステムです。そして、これらすべての機能を、きわめて条件の厳しいユースケースでVerticaに直接読み込まれた「ホットデータ」に適用することも、HDFSやAmazon S3といった外部の場所にあるデータに適用することもできるので、データの移動なしに統一された分析表示を実現できます。

Verticaのインデータベース機械学習機能を使用すれば、ビッグデータを活用しながら、予測分析のプロセスをシンプル化、高速化することで、意思決定を改善し、競争上の立場を強化し、情報を入手するまでの時間を短縮できます。データの準備から展開まで、Verticaは機械学習のプロセス全体をサポートします。また、モデルを複数のVerticaクラスターに展開できるので、Verticaを分析エンジンとして埋め込んだソリューションも実現できます。

Verticaのすべてのリリースは、ビジュアライゼーションツールとETLツールに関して検証され、テストされています。一般的なSQLと、Java Database Connectivity (JDBC)/Open Database Connectivity (ODBC) がサポートされます。これにより、ユーザーは、これらのテクノロジーに関する何年分もの投資とトレーニングを節約できます。一般的なすべてのSQLプログラミングツールや言語がシームレスに動作するからです。TableauやMicroStrategyなどの最先端のBIおよびビジュアライゼーションツールや、Informatica、Talend、Pentahoなどの一般的なETLツールが緊密に統合されています。

Vertica Advanced Analytics Platformの核となるのは、今日の分析ワークロードの処理に特化して設計されたカラム指向のリレーショナルデータベースです。古い時代に少量のデータの処理を目的として設計された市販やオープンソースの行ストア型ソリューションと異なり、Vertica Advanced Analytics Platformには以下のような特長があります。

- 完備された高度なSQLベースの分析機能による強力なSQL分析
- クラスター方式のビッグデータ格納による、高いパフォーマンスのクエリと分析
- 高度な圧縮により、同等のデータ分析ソリューションに比べて少ないハードウェアとストレージで動作
- ワークロードの増加に容易に対応できる柔軟性とスケーラビリティ
- クエリの際のロードスルーブットと同時実行の向上



- RおよびPythonによる内蔵予測分析と、インデータベース機械学習アルゴリズム
- オーバーヘッドとチューニングのためのデータベース管理者 (DBA) による介入が少ない

Verticaは、最高のスケーラビリティを持つ大規模なビッグデータ分析を実現します。独自設計のメモリ-ディスクバランス型分散圧縮カラムストア方式により、今日のデータ分析ワークロードの処理を古い技術よりもはるかに高速に行えます。

デプロイメントモデルと利用モデルの幅広い選択

Verticaは、オンプレミスでも、Hadoop上でも、クラウドでも利用でき、統一された1つの分析エンジンにより、実証済みのビッグデータ分析を比類のない速度とスケールで実現します。インフラストラクチャーの制約とは、将来にわたって無縁です。

- Verticaは、自社のデータセンターにオンプレミスで展開することも、プライベートクラウドに展開することも、Amazon、Azure、Google、またはVMware®クラウドなどのパブリッククラウドに展開することもできます。その方法は簡単です。Vertica Enterpriseライセンスを取得して、同じ統一されたエンジンをクラウドに展開するだけです。追加の容量が必要で、オンプレミスハードウェアを準備する時間がない場合には、これは魅力的なオプションです。
- 統一されたVerticaエンジンを、Apache Hadoop®にネイティブに展開できます。Vertica SQL on Apache Hadoop®は、データの探索とSQL分析を高速化し、組織で使用されているHadoopディストリビューションでネイティブに実行できます。

クラウド — Verticaソフトウェアは、Amazon、Azure、Google、VMwareクラウドでの実行に最適化され、あらかじめ構成されています。Verticaを使用すれば、展開、セルフプロビジョニング、BIおよびETLソフトウェアツールとの統合を迅速に行う俊敏性と拡張性が得られます。小規模から出発してビジネスの成長とともに拡大できる柔軟性により、データウェアハウスをクラウドやオンプレミスにシームレスに移行し、また元に戻すことができます。このレベルの俊敏性があれば、妥協の必要はありません。

オンプレミス — Vertica Advanced Analytics Platformは、「シェアードナッシング」型の分散データベースであり、コスト効果の高い市販サーバーのクラスターで動作します。パフォーマンスを強化するには、クラスターにサーバーを追加するだけで済みます。Verticaはグリッドアーキテクチャーを採っているため、多数のCPUやSANを備えた「ビッグアイアン」サーバーを必要とする従来のデータベースに比べて、ハードウェアと規模拡大のコストを大幅に (70 ~ 90%) 削減できます。

また、クラスター化により、クエリやロードをクラスター内の複数のノードで並列処理することで、スループットを高め、パフォーマンスを向上させることができます。

Hadoop — Hadoopと組み合わせて使用する場合、Vertica for SQL on Apache HadoopはHadoopクラスターに直接インストールされます。強力なデータ分析機能のセットを利用することで、各プラットフォームを単独で使用する場合よりもはるかに多くの成果が得られます。クエリにヘルパーノードを使用しないため、単一障害点が存在しません。ORC、Parquet、AvroなどのネイティブHadoopファイル形式の読み込みと、Parquetの書き込みも可能です。Vertica SQLエンジンをHadoopクラスターにインストールすることで、高度で包括的なSQL on Hadoop機能を利用でき、TPC-DSクエリが100%変更なしで動作し、すべてのHadoopディストリビューションが使用できます。

ビッグデータのマネタイゼーションと活用

Vertica Advanced Analytics Platformを使用すれば、夢物語を飛び越えて、ビッグデータの価値を実現できます。以下に紹介するのは、Verticaを利用して、最も戦略性の高い資産であるデータの活用で成功した企業の例です。

- **Intuit** — 何十億件ものトランザクションを処理して、TurboTaxで納税準備をしている数百万人のユーザーに、個々のユーザーに合わせた回答を迅速に返します。
- **Conservation International** — 科学者が8,600万件のレコードを対象に調査地と種の比較分析をほぼリアルタイムで行い、気候、人間、土地利用の影響を評価するために役立っています。
- **Cerner** — 救急救命室 (EMR) 臨床医の効率を分析することで患者の治療を改善し、敗血症警告モデルに基づいて現在までに500人の命を救いました。
- **Anritsu** — 従来のEDWに代わってVerticaを新しい埋め込み分析データベースとして採用し、これまで理論上の存在だった予測分析ソリューションを実現することで、わずか4か月で投資を回収し、351%のROIを達成しました。
- **Guess** — モバイルデバイスを通じて重要な日次店舗レポートを作成することで、売上を正確に把握し、商品の供給と配分を改善し、顧客の購買行動に関する貴重な情報を入手しています。

コンセプトを現実へと変換するVerticaのパワーをお試しください。Vertica次世代ハイパフォーマンスSQL分析エンジンは、さまざまなニーズに応えられるように、オンプレミス、クラウド、Hadoopの3つの統合形式で提供されています。お客様のニーズはそれぞれ異なるので、分析データベースもそれぞれ異なる必要があります。www.vertica.com/tryで今すぐVerticaをお試しいただけます。

詳細情報

www.vertica.com

詳細情報: www.vertica.com

© 2017 Micro Focus. All rights reserved. Micro FocusおよびMicro Focusロゴは、英国、米国、およびその他の国におけるMicro Focusまたはその子会社または関連会社の商標または登録商標です。その他の商標については、それぞれの所有者が権利を有しています。
