

ESG テクニカルレビュー

InterSystems IRIS Data Platform: 分析とビジネス上の洞察のための効率的な統合データプラットフォーム

日付: 2019年10月 作成者: 上級IT検証アナリスト、Kerry Dolan

要約

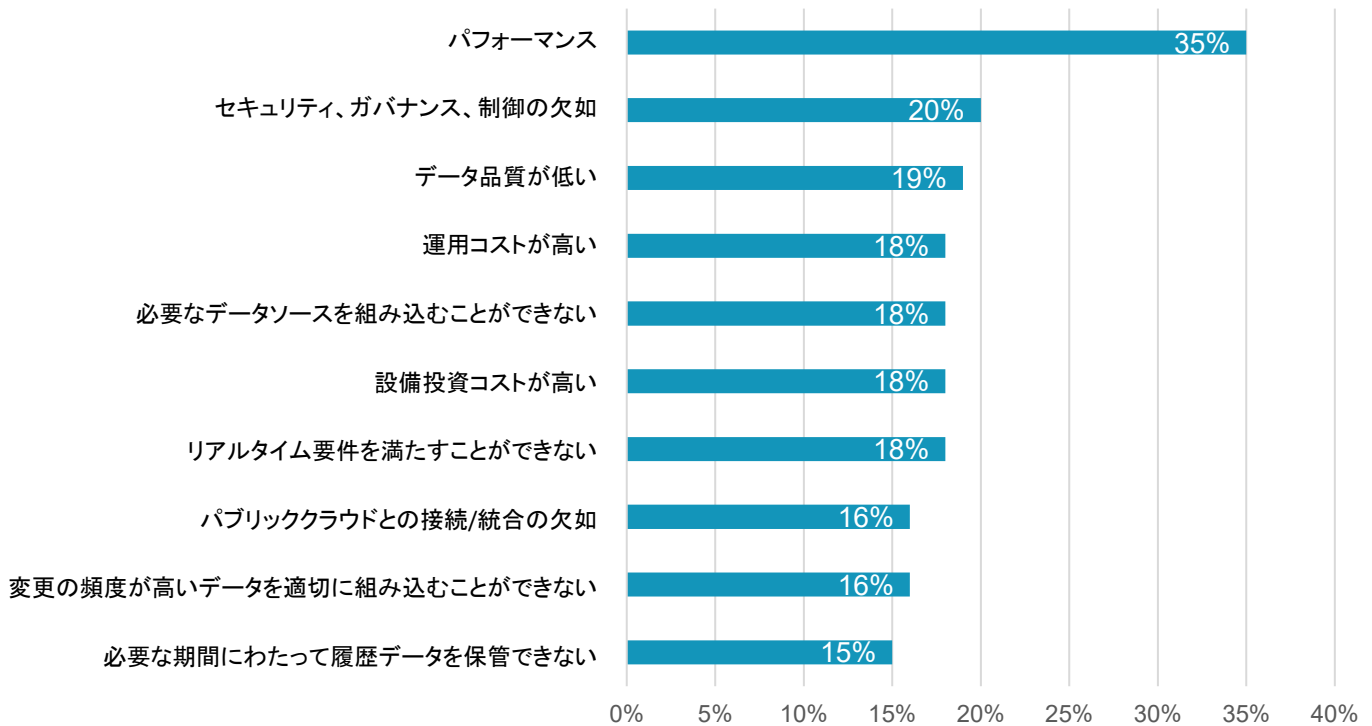
このESGテクニカルレビューレポートは、InterSystems IRIS Data Platformのパフォーマンステストを当社が検証した結果を記載したものです。テストは、3つの一般的なデータベースを比較して、概念実証環境(Proof of Concept)で実施しました。このテストは、金融サービス業界における一般的なクエリ実行のパフォーマンスの検証に重点を置いたものです。

課題

トランザクションと分析のいずれのビジネスプロセスも、データによって遂行されます。顧客およびパートナーとのトランザクションによって、収益および利益が決まります。一方、データ分析によって(そのデータが顧客やサプライヤから得られたものであれ、センサやマシンから得られたものであれ)重要な洞察が得られ、ビジネス上の意思決定のための情報を提供します。リアルタイムデータをビジネス上の意思決定を下すための知識に変換する能力は、成功と失敗の分かれ目となる可能性があります。データ管理テクノロジーは、分析のための強力な基盤を提供する必要があります。また、簡単な配備、高い費用効率、容易な拡張に加え、何よりも高いパフォーマンスが要求されます。遅延が生じると、クエリへの回答とビジネス変革が妨げられます。ESGがIT部門の意思決定者にデータ管理上の課題について尋ねると、最も多い回答はパフォーマンスでした。第2位以降の回答には、セキュリティと制御の問題、データ品質、コスト、リアルタイム要件への対応がありました。¹

図 1. 上位 10 のデータ管理の課題

データ管理/データウェアハウスソリューションで、
どのような課題に直面しましたか？(回答者の比率、回答者数=310、複数回答可)



出典: Enterprise Strategy Group

¹出典: ESG マスタサーベイ結果、「[The State of Data Analytics](#)」、2019年8月

このESGテクニカルレビューは、インターシステムズ社によって委託され、ESGからのライセンスに基づいて配布されます。

ほとんどの組織では、複数のデータベースを利用しており、そのためインフラストラクチャコストと人件費の両方が増大しています。組織は、リアルタイム、履歴、自然言語の処理とビジネスインテリジェンスを含むすべてのトランザクションと分析の負荷に対応するために必要なリソースを減らすことができれば、コストを削減できます。しかし、組織は、データの耐久性、拡張性、管理の容易性と共に、高いパフォーマンスを実現しなければなりません。

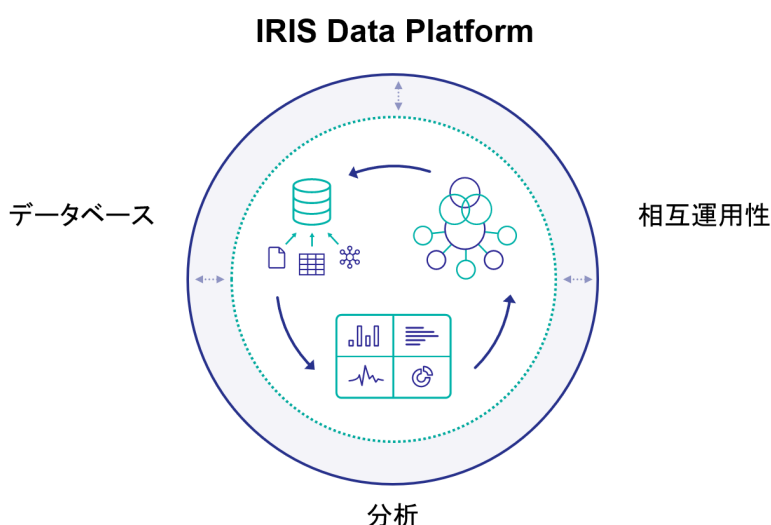
ソリューション: InterSystems IRIS Data Platform

InterSystems IRIS Data Platform は、複数のトランザクション処理と分析機能を提供し、インテリジェントなデータ集約型アプリケーションの構築において組織を支援するよう設計されました。InterSystems IRIS は、単一のソリューションで強力なデータ管理、分析、相互運用性を組み合わせた統合リアルタイムプラットフォームです。これは、組み込みのセキュリティ、実証済みの信頼性、サポートを提供する一方で、クラウド環境、オンプレミス環境、ハイブリッド環境で多様なワークロードと大規模なデータ集約型アプリケーションを処理し、最適化するため設計されました。

InterSystems IRIS は、すべての機能を備えた単一のプラットフォームで、単一のデータベースで、今日のビジネスの多様なニーズに対応します。これにより、組織は多様なシステムごとのソリューションを購入して管理し、複数のデータベースに格納してデータをマッピングする必要がありません。InterSystems IRIS には、以下の機能が含まれています。

- マルチワークロード:** InterSystems IRIS では、同じエンジンで大規模なトランザクションと分析の両方を処理します。これにより、分析のためのトランザクションデータの移動と変換が不要となり、複数のデータベースの配備による複雑性が解消します。InterSystems IRIS では、非リアルタイムデータの分析とトランザクションデータのリアルタイム分析を同時に実行する一方で、トランザクションデータにインデックスを付けることができます。これにより、ビジネスへのより早い対応が可能であり、よりよい洞察が得られます。
- マルチモデル:** InterSystems IRIS は、従来のサイロ化したデータベースアプローチではなく、同じデータを多様に表現できる単一のデータベースを備えています。これは、リレーショナル、キーバリュー、ドキュメント、オブジェクト、スキーマレスな多様な NoSQL データなど、多数のデータモデルをサポートしています。
- 相互運用性:** この包括的な統合およびオープンな分析プラットフォームは、効率的なデータコーディネーションとサービスオーケストレーションを実現し、データソースへの接続とビジネスプロセスを簡素化します。これには、サービスの特定、公開、オーケストレーション、収益化のためのライフサイクル全体の API 管理機能も含まれています。
- 分析:** InterSystems IRIS は、リアルタイム分析とバッチ分析機能(分散 SQL 処理、ビジネスインテリジェンス、予測モデリングランタイムエンジンと、電子メールやソーシャルメディアなどのソースからの非構造化データの自然言語処理を含む)を提供し、さらに Apache Spark などのビッグデータ処理フレームワークと連携が可能です。
- エンタープライズレベルのセキュリティ:** Kerberos, LDAP, KMIP、ロールベースのアクセス制御、送信中および保存時のデータの暗号化によって、きわめて堅牢なセキュリティが提供されます。

図 2. InterSystems IRIS Data Platform



出典: インターシステムズ社

InterSystems IRIS のアーキテクチャ

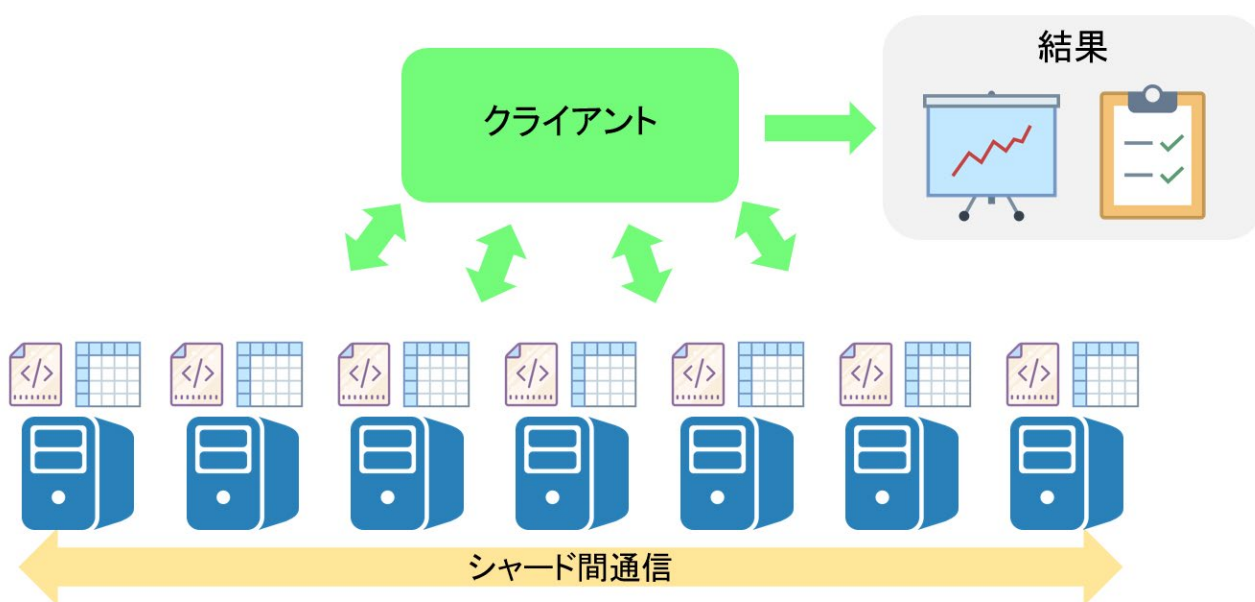
InterSystems IRIS では、トランザクションデータを高速に取り込み、格納する一方で、バッチ、リアルタイム、ACID 準拠データの分析ワークロードを、同時に処理できます。これにより、リアルタイムデータを処理する分析プラットフォームに移動することで生じる遅延がなくなります。また、並列 SQL 処理での垂直方向の拡張性(スケールアップ)を備え、水平方向の拡張性(スケールアウト)に対しては他社と異なるアプローチを採用しています。

大量のデータセットと複雑なクエリに対して、InterSystems IRIS は、水平方向の拡張性のために特別なアプローチを取っています。このアプローチでは、分散クエリ処理のためにコモディティサーバを利用し、インテリジェントなシャーディングと分散キャッシングを活用しています。データベースのシャーディングによって、ワークロードとテーブルを複数のノードに分散させることができます。シャーディングはアプリケーションに対して透過的であり、コードの変更が必要な場合でも、ごくわずかで済みます。

InterSystems IRIS ではシャードマスタを利用し、並列実行のためにアプリケーションコードをシャードにプッシュし、各シャードの結果を集約して、最終結果を返します(図 3 を参照)。各シャードは、他のデータすべての位置を認識しています。シャードは、クエリに必要なデータが別のシャードにある場合、必要なデータのみを取得するために、別のシャードに直接アクセスできます。他のソリューションでは、テーブル全体を他のノードにブロードキャストする必要があり、待ち時間とタイムアウトが増大します。このため、InterSystems IRIS は、シャードデータを複雑に結合するクエリの場合でも、高いパフォーマンスを達成できます。

シャーディングによって互いに素なパーティションが生成されるため、各サーバのキャッシュは独立しています。したがって、サーバをスケールアウトすることで、クラスタのメモリが直線的に増加します。このため、適切に規模を決定した場合、InterSystems IRIS の配備は、インメモリデータベースに近いレベルのパフォーマンスを達成できます。ただし、頻繁に使用するデータはキャッシュを使用し、すべてのデータをメモリに格納する必要はありません。また「あまり使用されない」データは最適化された形式で恒久的にディスクに保存されるため、高速なランダムアクセスを実現します。

図 3.水平方向の拡張性(スケールアウト)



出典: Enterprise Strategy Group

ESG による検証

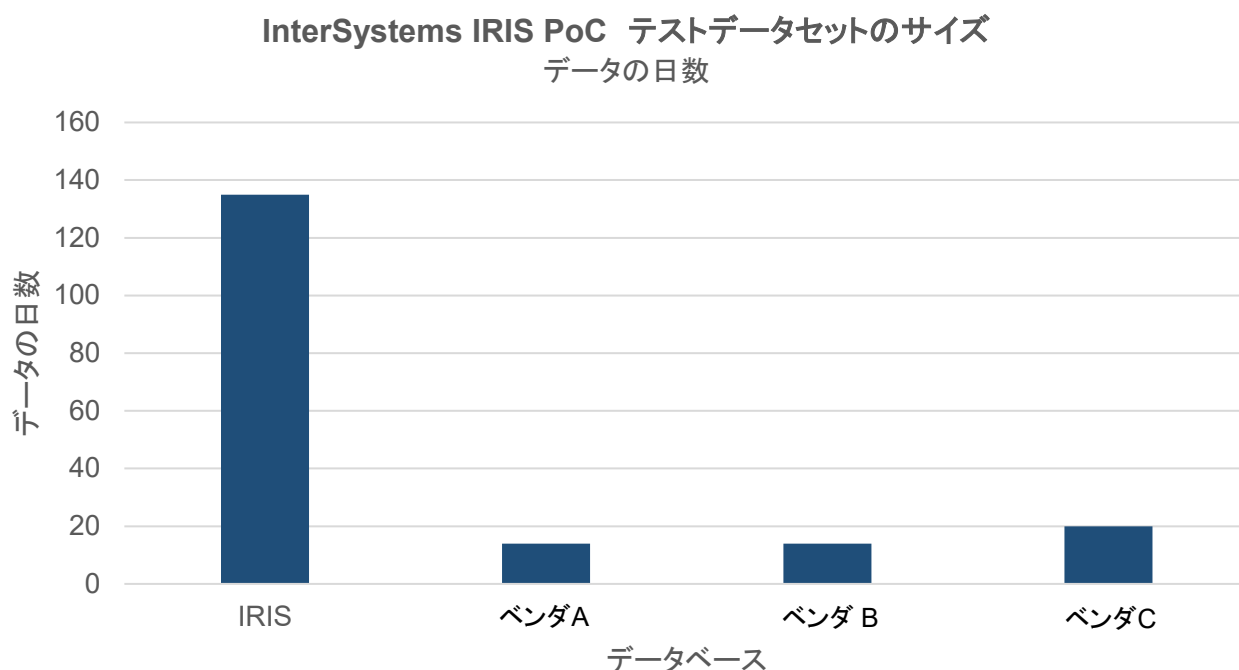
ESG は、金融業界のある企業によってデータベースアプリケーションの概念実証環境でのテスト結果を監査しました。テストでは、数千の金融商品を利用して、InterSystems IRIS と金融業界で一般に使用される他の複数のデータベースによる分析クエリのパフォーマンスを重点的に検証しました。使用したアプリケーションは企業全体の複数の基幹事業部門と接続されており、アルゴリズムによる取引アプリは、SQL クエリを厳しいパフォーマンスの制約内で正常に実行することが要求されます。その一方で、企業全体で数百人のユーザがクエリ結果に対して随時アクセスする必要もあります。ミリ秒単位のパフォーマンスが非常に重要です。要求されるパフォーマンス SLA (Service Level Agreement) 内にクエリを完了できない場合、アルゴリズムによる取引は実行されず、企業はその取引戦略を正常に実施できません。

テスト対象には、InterSystems IRIS と 3 つの他の一般的なデータベース配備を含めました。表 1 は各構成を示すもので、コンピューティングインフラストラクチャ (VM、コア、RAM など) が比較的小規模な InterSystems IRIS の構成で、競合ソリューションよりもデータに対する多くのクエリが実行されたことを示しています (表 1 と図 4 に示すように日数、GB、行数を測定しました)。

表 1. テストベッドの構成

データベース	構成	データの日数	データセットサイズ (GB)	データセットサイズ (行数)
InterSystems IRIS	4 VM: 8 コア、96GB の RAM、400GB のローカルストレージ	135	320	200,000,000
ベンダ A	8 VM: 16 コア、256GB の RAM、400GB のローカルストレージ	14	33	20,740,000
ベンダ B	8 VM: 16 コア、256GB の RAM、400GB のローカルストレージ	14	33	20,740,000
ベンダ C	それぞれ 24 コア、256GB の RAM を備えた 3 台の専用物理サーバおよび、2TB の共有 SAN ストレージ	20	47	29,620,000

図 4. データセットサイズの比較

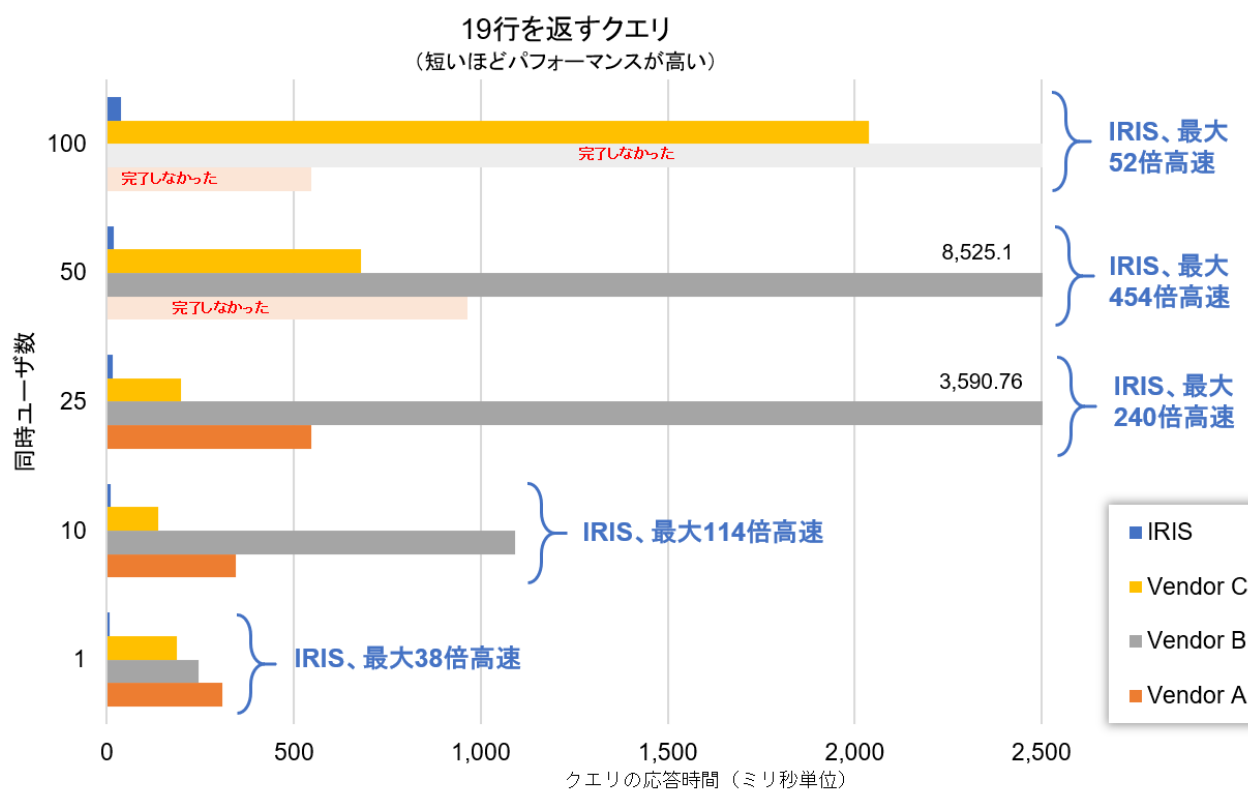


出典: Enterprise Strategy Group

5 つのクエリでテストを実行しました。すべてのクエリは、5~6 個のテーブルの複雑な JOIN (結合) であり、WHERE 句に 8~10 個の述語が含まれます。結果セットには 19~6,504 行が含まれ、各行には数百のフィールドがあります。同時ユーザ数を 1、10、25、50、100 と増やして、クエリを実行するたびに応答時間を測定しました。

図 5 には、19 行を返すクエリ 1 の結果を示します。パフォーマンスは、同時ユーザ数を増やしたときに、各クエリの完了までに必要な時間 (ミリ秒単位) を測定しました。

図 5.同時ユーザ数を増やしたときのクエリ 1 のパフォーマンス

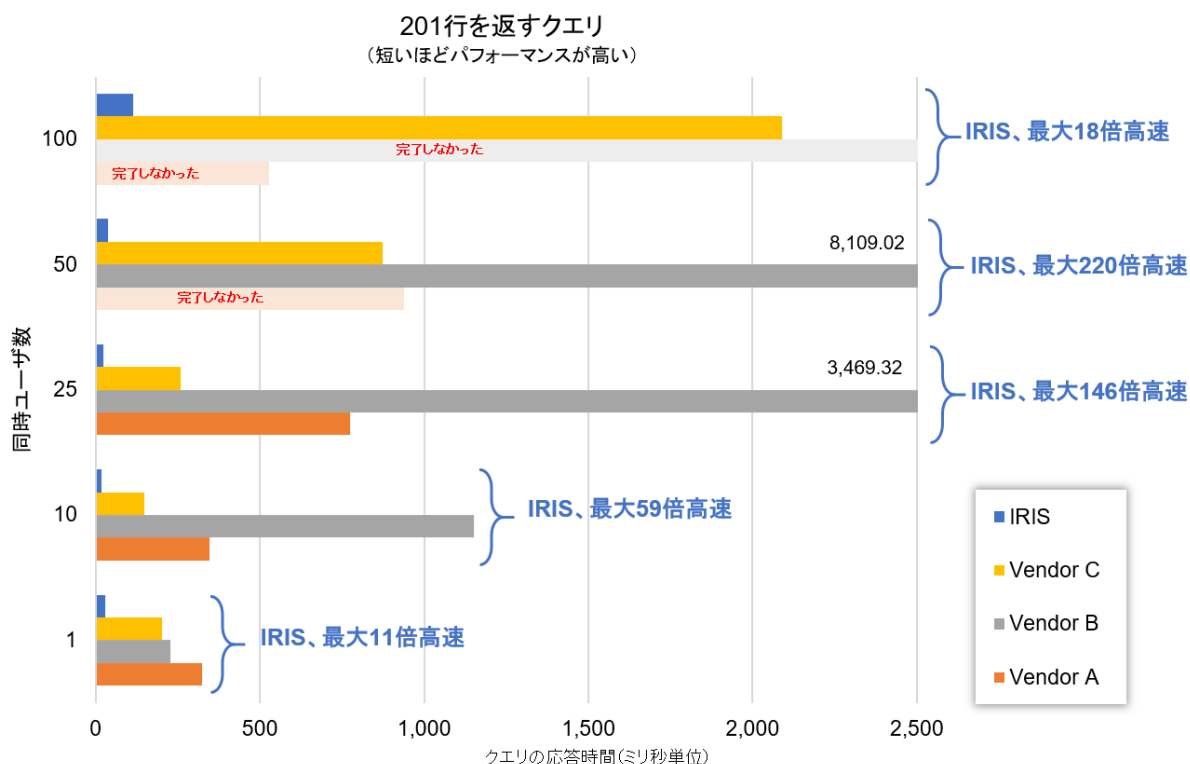


注:

- InterSystems IRIS は、他のベンダよりも小規模なインフラストラクチャで動作し、5~8 倍多いデータにアクセスしました。
- 一部のソリューションは大規模なクエリを完了することができず、その有用性が限られていました。このような結果は、注意書きと薄い色の棒グラフで示しています。

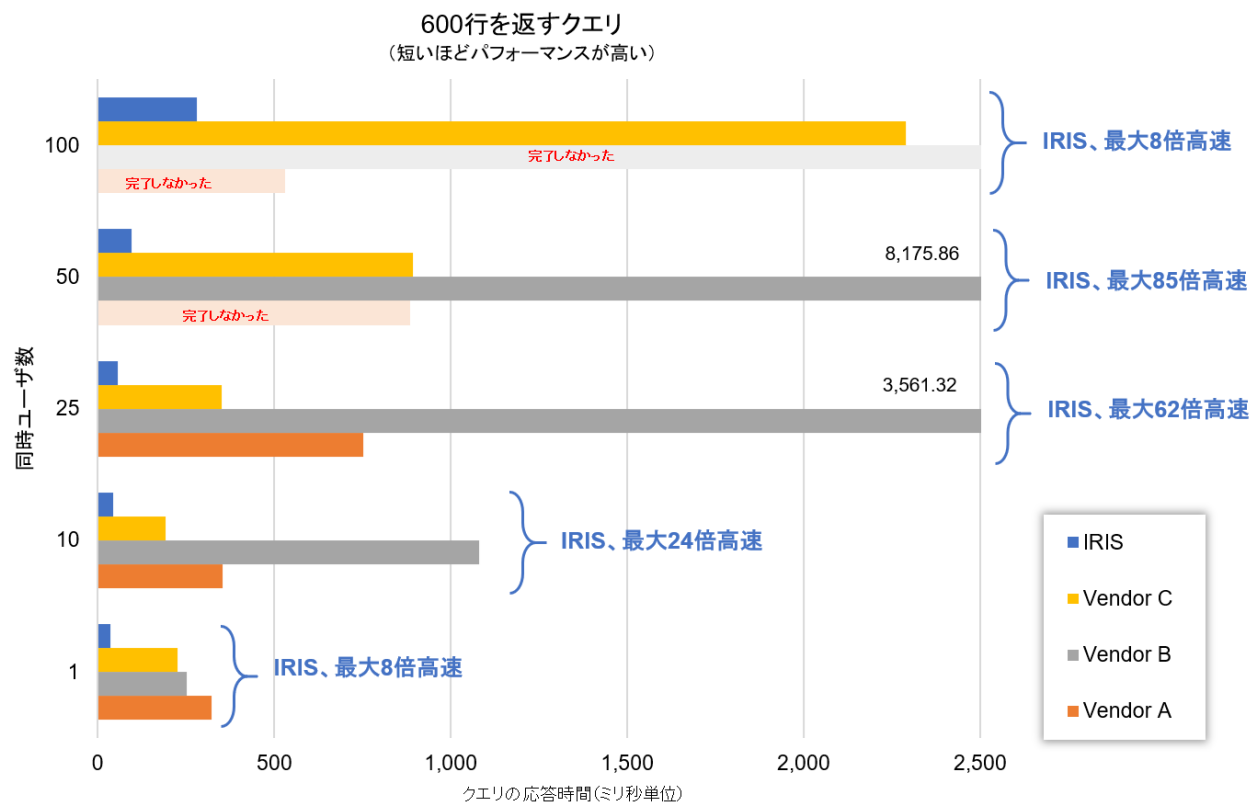
クエリ 2~4 でも同様の結果が示され、InterSystems IRIS が一貫して他のベンダより高速でした。クエリ 2 では 201 行が返され、クエリ 3 では 600 行が返され、クエリ 4 では 984 行が返されました。図 6、7、8 に、これらの結果を示します。

図 6. 同時ユーザ数を増やしたときのクエリ 2 のパフォーマンス



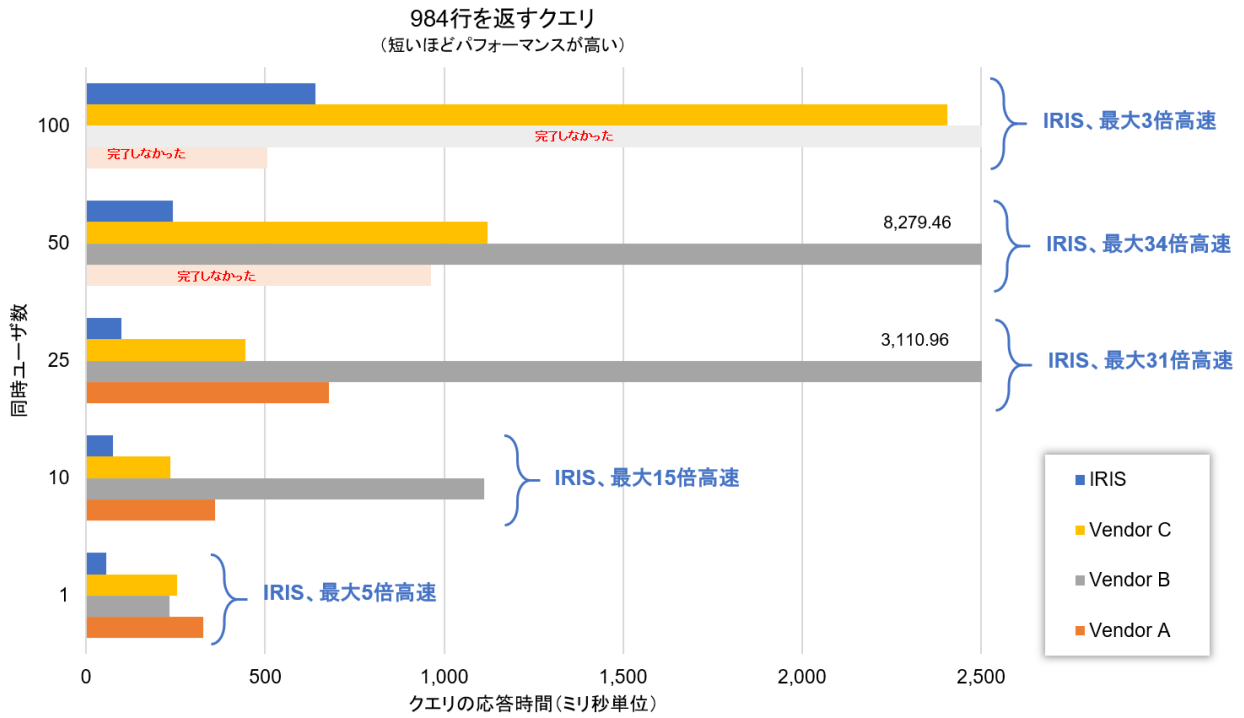
出典: Enterprise Strategy Group

図 7. 同時ユーザ数を増やしたときのクエリ 3 のパフォーマンス



出典: Enterprise Strategy Group

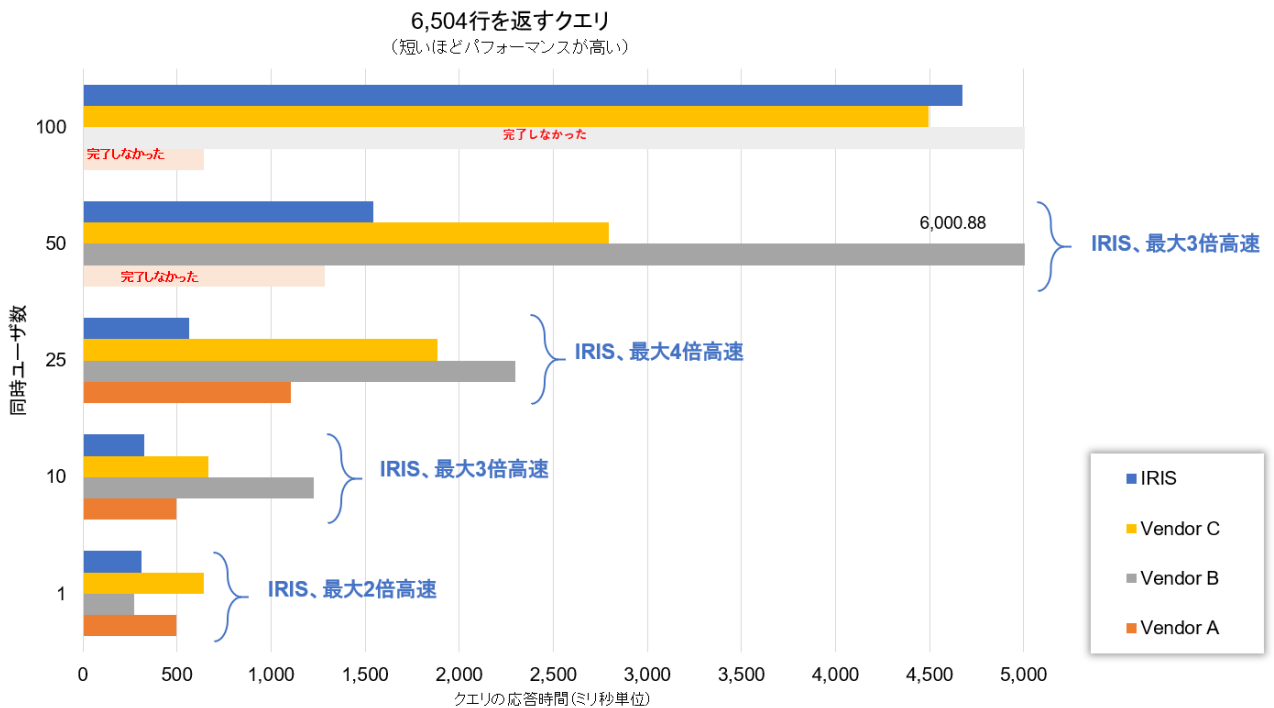
図 8. 同時ユーザ数を増やしたときのクエリ 4 のパフォーマンス



出典: Enterprise Strategy Group

図 9 には、6,504 行を返すクエリ 5 の結果を示します。この複雑度が高いクエリでは、同時ユーザ数が増えるにつれ、パフォーマンスへの影響がさらに大きくなりました。

図 9. 同時ユーザ数を増やしたときのクエリ 5 のパフォーマンス



出典: Enterprise Strategy Group

注:

- InterSystems IRIS は、他のベンダより小規模なインフラストラクチャで動作し、より多くの日数のデータにアクセスしました。
- InterSystems IRIS のパフォーマンスは、1 ユーザの 312 ミリ秒から、100 ユーザの 4,673 ミリ秒の範囲であり、処理時間が 14 倍に増大しました。
- ベンダ C は、100 ユーザで InterSystems IRIS より少し高速であり、InterSystems IRIS の 4,673 ミリ秒に対して 4,492 ミリ秒を達成しました。

i **これが重要である理由**

ESG の調査によると、組織は、重要なビジネス目標（運営効率の改善、経営コストの削減、ビジネス上の意思決定/製品の品質/予測精度の改善、製品を市場に投入するまでの期間の短縮）を達成するために、データ分析を重視しています。²つまり、組織は、ビジネスのあらゆる部分を改善するために分析を利用しており、正確かつ適時な意思決定を下すために迅速な回答を必要としています。アクセスできるデータが多く、クエリによって高速に結果が返されるほど、意思決定が改善されます。

ESG は、InterSystems IRIS がトランザクションと分析のワークロードに要求される堅牢なデータ管理機能とパフォーマンスを提供することを確認しました。IRIS は、他のベンダよりも小規模なインフラストラクチャで動作し、より多くのデータにアクセスする一方で、他のテスト対象データベースよりも一貫して高速に実行されました。InterSystems IRIS では、すべてのクエリタイプが高速なパフォーマンスで処理され、拡張が容易で、必要なインフラストラクチャの規模が比較的小さく、いかなる障害も生じませんでした。アクセスするデータを増やし、結果の提供を迅速化することにより、InterSystems IRIS はビジネス上の洞察を改善できます。

より深い洞察

組織は、情報の活用によって、競争上の優位性を獲得します。金融、医療、製造、通信、そして他の大半の業界の企業は、データ分析がよりよい意思決定に結びつくことを理解しています。しかし、連携の取れていないインフラストラクチャをセットアップして、トランザクションプラットフォームと分析プラットフォーム間でデータを移動すると、手間と時間がかかり、洞察が遅れ、コストが増大することがあります。組織は顧客、サプライヤ、市場の動向を深く、迅速に把握できるほど、経営上の意思決定を改善できます。課題は、できるだけ多くのデータから、できるだけ高速に洞察を得て、その洞察に基づいてすばやく行動することです。組織は、柔軟かつ拡張が容易で、費用効率とパフォーマンスの高いプラットフォームでリアルタイム分析を実行できれば、リアルタイムに変化を起こすことができます。

InterSystems IRIS は、組織がデータ駆動型のインテリジェンスとリアルタイムビジネスプロセスを組み合わせることによって意思決定を改善する目的で設計された、費用効率の高い統合プラットフォームです。また、InterSystems IRIS は、多様なトランザクションデータベースと分析データベースの必要性をなくすことによって、リアルタイムデータの分析クエリを高速化する革新的なナレッジプラットフォームです。組織は、データの移動、マッピング、変換を行う必要がなく、同一のエンジンを使用して、トランザクションワークロードと分析ワークロードを実行できます。データは同時に複数の方法で表示でき、システムごとのソリューションを複数配備する必要性がなくなり、パフォーマンスがより高速になります。

ESG は、複数の一般的なデータベースソリューションと比較して、InterSystems IRIS のパフォーマンスを検証しました。その結果、InterSystems IRIS では、他のソリューションよりも多くのデータと、小規模なインフラストラクチャを利用して、より高速なパフォーマンスが得られることが実証されました。したがって、組織は、より高速に、低コストで洞察を深めることができます。より深い洞察により、組織は、収益の増加、顧客体験の改善、製品やサービスのカスタマイズ、コンプライアンスの改善、リスクの低減、経営の効率化、コストの削減を実現できます。データ中心の洞察のために強力なデータ管理プラットフォームを探している組織には、InterSystems IRIS の評価を行うことを推奨します。

²出典: 同上

すべての商標は、それぞれの企業の財産です。本書に含まれる情報は、Enterprise Strategy Group(ESG)が信頼できるとみなした出典から入手されましたが、ESGによる保証の対象となりません。本書には、修正されることがある ESG の見解が含まれる場合があります。本書の著作権は、Enterprise Strategy Group, Inc.が保有しています。ハードコピー形式、電子形式、またはその他のいずれによるかにかかわらず、Enterprise Strategy Group, Inc.から明示的な同意を得ずに、本書を受け取る権限のない者に対して本書の全部または一部を複製または再配布することは、米国著作権法の違反であり、民事損害賠償訴訟の対象となり、該当する場合は刑事罰が科されます。ご不明な点がある場合は、ESG Client Relations (508.482.0188)までお問い合わせください。

ESG 検証レポートの目的は、すべてのタイプおよび規模の企業向けの情報技術ソリューションについて、IT 専門家に情報を提供することです。ESG 検証レポートは、購買の意思決定を下す前に実施すべき評価プロセスに代わるものではなく、新たに生まれている技術に関して洞察を提供することを意図しています。当社の目的は、IT ソリューションの比較的有效な特徴および機能を調査し、実際の顧客の課題を解決するためにソリューションをどのように利用できるかを示し、改善の必要な領域を特定することです。ESG 検証チームエキスパートの第三者的な観点は、本稼働環境でこれらの製品を利用している顧客のインタビューと、当社独自の実践テストに基づいています。



© 2019 by The Enterprise Strategy Group, Inc. All Rights Reserved.



www.esg-global.com



contact@esg-global.com



P.508.482.0188