

ソリューション概要

データセンター
インテル® FPGA によるデータベース・アクセラレーション



rENIACにおけるインテル® FPGAを活用した 大規模トランザクション・データベースの高速化

AIおよびデータ主導のビジネスに向けた Apache Cassandra® NoSQL データベースの変革

概要

「データ・トラフィックの
パフォーマンスは、
優れたユーザー体験に要求される
応答時間と拡張性を満たすための
極めて重要な要素です。
rENIAC は、Distributed Data
Engine テクノロジーを使用して
Cassandra® データベースの
アクセラレーションを行うことで、
世界中のソリューション・
プロバイダーや
企業のクリティカルな
NoSQL ワークロードを
高速化しています」

— rENIAC, CEO
Prasanna Sundararajan 氏

Apache Cassandra® NoSQL データベースは、IoT や不正検出からパーソナライゼーション、金融サービスまで、現代社会を支えるデータ負荷の高いユースケースに広く使用されています。Cassandra® は数多くのエンタープライズ・クラスの要件を満たしますが、トランザクションや AI のアプリケーションを処理するときに、限界に達します。rENIAC は、Cassandra® クライアントとデータベース・ノードの間に革新的な「中間」レイヤーを提供します。Data Engine と rENIAC ソフトウェアから成るこのソリューションでは、インテリジェントなキャッシュによって、ストレージとネットワークとの距離がより近くなります。インテル® アーキテクチャー上での動作とインテル® FPGA によるアクセラレーションの採用により、rENIAC は、要求の厳しいワークロードやアプリケーションに対して卓越したパフォーマンス、高スループット、低レイテンシーを実現し、AI の導入を容易にします。

課題

パフォーマンスの面では一切妥協することなく急増するデータを管理することが、多くの企業にとって差し迫った課題となっています。一般的な Cassandra® NoSQL データベースでは、複数の汎用サーバーにわたってデータベース・トラフィックを極めて効率良く処理し、高い書き込みスループットの提供とノードの追加による線形の拡張を実現できます。しかし、巨大なデータベース、AI モデルの学習処理やニューラル・ネットワーク推論、ビデオや画像処理などの高帯域幅データ、ミリ秒速度のトランザクションによってワークロードの要求がますます厳しくなると、多くのインフラストラクチャーは限界に達し、スループット、パフォーマンス、レイテンシーに悪影響が生じます。ほとんどのデータセンターでは、他のビッグデータの課題と同様に、さらに多くの DRAM メモリーを各 Cassandra® ノードに追加することで対応してきました。当然、1 ノード当たりの追加可能なメモリー容量には上限があり、この方法では、クラスターのメモリーを拡張するためだけに、さらに多くのノードの追加が必要となってしまいます。つまり、こうした方法が利用できるのは、各ノードに大量の DRAM を追加できる余裕がある場合に限られるということです。

rENIAC

ソリューション

rENIACは、インテル® FPGAでアクセラレーションを行うコスト効率の高いプラグ・アンド・プレイ・ソフトウェア・ソリューションを用いて、Cassandra® データベースの容量を最大化します。このソリューションは、新たなノードを追加することなく、高スループットの実稼動環境にて予測可能な低レイテンシーを提供し、そしてビジョン・プロセッシング・ユニット (VPU) の管理を簡素化します。インテル® CPU上で動作するとき、このソリューションは、データに基づくユースケースを幅広く実現し、既存のインフラストラクチャーを活用してAIを容易に導入できることを目的として設計されています。

このソリューションは、分散テクノロジーを使用してビッグデータ・ワークロードを高速化するもので、ソフトウェアの変更なしにApache Cassandra® 上で大幅なレイテンシーの低減とパフォーマンスの向上を実現します。rENIACの革新的な手法が、データストア・アクセラレーションを新たなレベルのデータベースと効率向上を可能にするサービスとして提供可能にします。

rENIACのソフトウェアは、インテル® FPGA、インテル® CPU、およびメモリーを使用して、トランザクションを高速化してデータエンジンとして機能します。rENIAC Data Engineは、高速データアクセスに最適化されたインテル® FPGAベースのインテル® FPGA プログラマブル・アクセラレーション・カード (インテル® FPGA PAC) を使用します。さらに、低レイテンシーと高スループットを維持し、オンプレミスでもパブリック・クラウドでも導入でき、Cassandra® セキュリティー・プロトコルに準拠しています。Cassandra® NoSQLの場合、これは、Cassandra® クライアントとCassandra® データベース・ノードの間に透過的なレイヤーを配置します。データをローカルストレージにキャッシュし、データをローカルストレージから供給するか、またはデータがローカルストレージで入手できない場合はバックエンド・データベースからフェッチすることによって、クエリーに応答します。アプリケーション・レイヤーは、rENIACのData Engineの上位に構築されます。Data Engineはデータ中心のワークロード向けのアクセラレーション・ソリューションを構築するためのテクノロジー・プラットフォームとなります。

アクセラレーションの向上

インテル® FPGAにより
rENIAC ソリューションが実現した高速化

6倍以上¹

レイテンシーの低減

25倍以上

99 パーセンタイルで予測可能な
低レイテンシー¹

TCO の改善

TCO の大幅な低減：
サーバー数の削減

最大 50%¹

Apache Cassandra® NoSQL パフォーマンスの最適化

インテル® FPGAは、ネットワーク・ベースのトランザクションの最適化に役立ちます。rENIACの基本アーキテクチャーは分散型であり、インテル® FPGAは、分散トランザクションの高速化に使用されています。Alibaba® Cloud F1 インスタンスなどのコンピューティング・インスタンスを、ソフトウェアやファームウェアを介してネットワーク・インスタンスに変えることで、大幅なネットワーク・アクセラレーションを可能にします。これは、負荷の高いeコマース環境やトランザクション環境、さらにはオンデマンドのビジネスモデルをサポートできるだけの容量と速度を備えた新たなレベルのデータベースを可能にします。業界標準のインテル® アーキテクチャー (インテル® FPGA、インテル® FPGA PAC、インテル® CPU など) が、rENIACソリューションの高水準のパフォーマンス、拡張性、信頼性を実現します。

一般的なソフトウェア・スタックでは、あらゆる処理がCPU内で実行されます。パケットが入ってくると、CPU上で動作するアプリケーションがそのパケットを検出し、CPU上で処理します。一方、rENIACでは、インテル® FPGAをCPUに相当するものとして使用します。これが、カーネルとCPUをバイパスしているため、データがネットワークのより近くに留まることになり、インテル® FPGAがハードウェアからの要求にすぐに応えることができます。その結果、パフォーマンスは6倍以上向上します。

業界標準として広く普及している既存のインテル® アーキテクチャーをrENIACのData Engineソフトウェアとともに活用することにより、次のようなメリットを容易に獲得することができます。

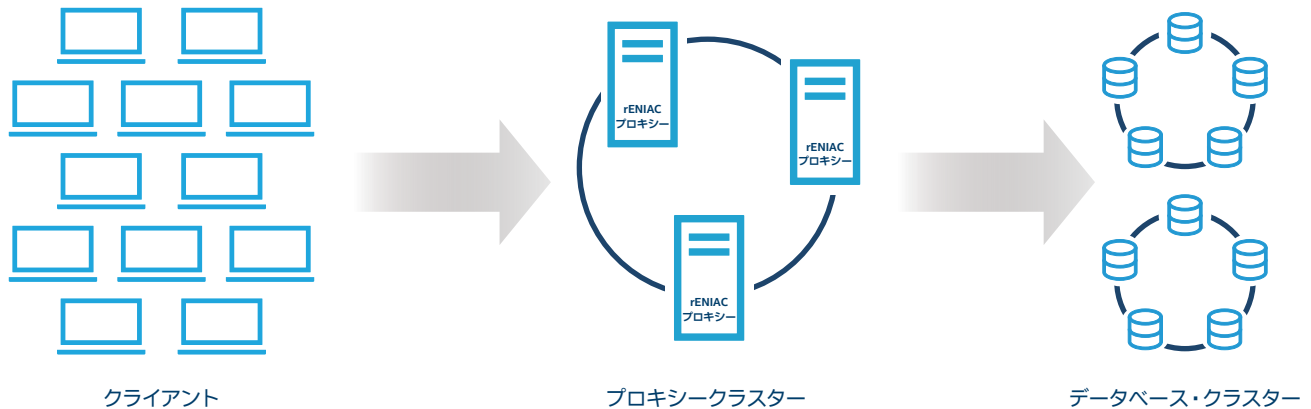
- **導入の容易さ**
インストールおよび利用が非常に容易で、ソフトウェアの変更が不要
- **エンタープライズ対応**
クラスタリング機能が高可用性と冗長性を提供
- **運用コストの低減**
サーバー数の削減とTCOの低減
- **将来に対応したテクノロジー**
インテル® CPU、FPGA、SSDを使用する、AI、ビッグデータ、IoT向けの拡張性に優れたインフラストラクチャー

「rENIACは、サーバー・インフラストラクチャーやデータベース・クラスター向けのデータ・ワークロードのアクセラレーションを行うために独自の手法を採用しています。rENIACのソフトウェアは、代替の実装に比べて最大6倍のスループット向上を提供できます。¹ さらに、同社のソリューションはコード変更が不要なので、ユーザーはシームレスに移行して、インテル® FPGAと高度なメモリー・テクノロジーを容易に利用することができます」

— インテル、プログラマブル・ソリューションズ
事業本部
データセンターおよびコミュニケーション部門
副社長 兼 事業本部長
John Sakamoto

データセンター・アーキテクチャーの再考

rENIACとインテルが提供するソリューションは、データセンター・アーキテクトやデータ・サイエンティストに対しても新たな可能性をもたらしています。DataStaxの共同設立者であるMatt Pfeil氏は次のように語っています。「FPGAがますます主流になるにつれて、Cassandra* ユーザーは要求の厳しいワークロードにおいてもパフォーマンスのさらなる向上が期待できるため、既存のシステムをアップグレードするソリューションを検討しています。rENIACは、そのようなコミュニティに確かなイノベーションを提供します」



rENIAC Data Engine の使用によるデータベースのアクセラレーション

ユースケースの例

デジタルサイト上で関連性のある広告を表示するには、自動的にブラウジング・パターンを識別し、購買層、地域などのユーザー・プロフィール・データを関連付ける必要があります。Cassandra* は、ブラウジング・パターンの追跡やユーザー・プロフィールの保管に広く利用されています。デジタルメディア企業や広告プロバイダーは一般的な KPI として、関連性の高いコンテンツを 10 ミリ秒以内に提供する必要があります。つまりこれは、Cassandra* が約 5 ミリ秒以内に適切なコンテンツを引き出せなければ、顧客とのつながりが失われることを意味します。Cassandra* を rENIAC とともに使用することで、デジタルメディア企業は、Cassandra* からできる限り低いレイテンシーで情報を引き出し、最小限のノード数でクエリー数を最大化することが可能になります。

オンデマンドで大規模なデータやリッチメディアをストリーミングする企業や、e コマース、不正検出、金融サービスなどは、一般的な Cassandra* の利用シーンにおける rENIAC アクセラレーションのメリットが発揮される例のごく一部に過ぎません。

Apache Cassandra*

Apache Cassandra* は、無料かつオープンソースの分散ワイド・カラム・ストアの NoSQL データベース管理システムで、多数の汎用サーバーにわたって大量のデータを処理するように設計されており、単一障害点もない高い可用性を提供します。

<http://cassandra.apache.org/> (英語)



rENIACとインテル® FPGA

Field Programmable Gate Array (FPGA) には、データセンターにおけるパフォーマンスや分析の問題を解決するという役割が求められ、その期待は高まり続けています。FPGA は厳しいハイパフォーマンス要件を満たすために不可欠なテクノロジーですが、既存のアプリケーションを変更せずにそのメリットを手に入れるには適切なソフトウェアが欠かせません。

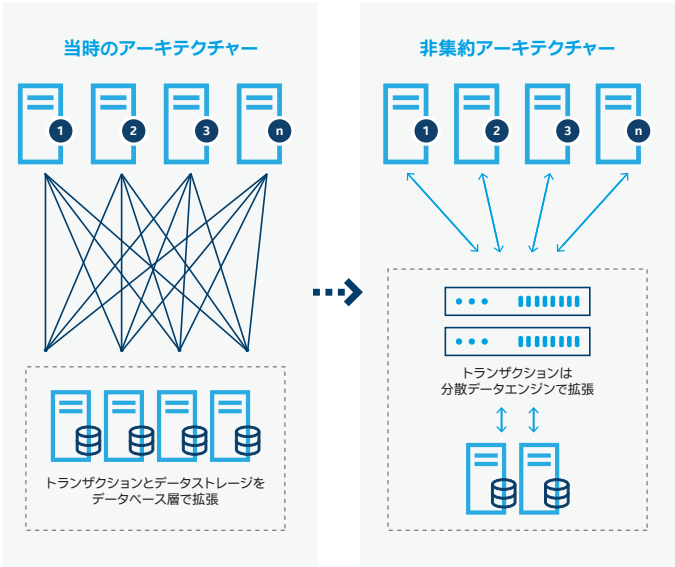
rENIAC は、アプリケーション・ソフトウェアを変更することなくデータとトラフィックを高速化するために、従来の CPU データフローを無効にする (つまりバイパスする) ソフトウェア・ソリューションを提供することで、インテル® FPGA プラットフォームの能力と柔軟性を活用します。その結果、FPGA のデザインやプログラミングに関する知識がなくても、クリティカルな NoSQL ワークロードを実行するプライベート・クラウドやパブリック・クラウドで予測可能なワークロード・パフォーマンスが得られます。

導入事例：デジタルメディア向けのコンテンツ・パーソナライゼーション

ある大手のオンライン・デジタル・メディア企業は、コンテンツのパーソナライズに注力しており、ユーザーのパーソナライゼーション・データをNoSQL上でCassandra*に保管していました。データベース・トランザクションのパフォーマンスについては、75パーセンタイルのサービスレベル・アグリーメント（SLA）でしか満たすことができず、収益を回収できませんでした。さらに、現在の事業拡大ペースを維持するためには、データベース・サーバーの台数を前年比で30%増やさなければならぬという状況にも直面することとなりました。

当時採用していたソリューションでは、トランザクション用とデータストレージ用の両方のCassandra*ノードを拡張せざるを得なかったため、SLAは向上することなく、インフラストラクチャーの設置面積とコストのみが大幅に増加してしまいました。

rENIACのDistributed Data Engineをネットワーク・サービスとして導入することで、トランザクションとデータストレージの拡張を分離することができました。その結果、トランザクションのスループットは8倍に増加し、レイテンシーは1/10に短縮され(99パーセンタイルで5～8ミリ秒のSLA)、コストを低減しながら、速度と収益を向上できました。¹



	当時のインフラストラクチャー	rENIAC 導入後	結果
データベース・サーバー数	150	18	1/8に削減
ノード当たりのデータベース・クエリー数	3,000	25,000	50%のコスト削減 ¹
クエリー当たりのレイテンシー(SLA)	75パーセンタイル：7～8 ms 98パーセンタイル：60 ms	99パーセンタイル：5～8 ms	レイテンシーを1/10に削減 収益を向上
ソフトウェア API	Cassandra* 2.x	Cassandra* 2.x	ソフトウェア変更なし

仕組みの概要

そのCassandra* NoSQL データベースは、書き込み操作を効率的に処理するように設計されました。そのため、実現できる読み取りレイテンシーと、ノードあたりに保管できるデータ量との間には常にトレードオフがあります。ノード当たりのデータ量が増加すると、演算およびI/Oの需要が増加するため、書き込みパイプラインがシステムに与えるプレッシャーが増加し、その結果、読み取りレイテンシーが低下します。さらに、その他の要因が組み合い、全体的なシステム・パフォーマンスに影響を与えます。例えば、ネットワークのレイテンシーとスループットは、Cassandra* クラスターの効率的な拡張に直接影響を及ぼします。同様に、Cassandra* クラスターのスループットは、SATA SSD または NVMe* ドライブなどのフラッシュメモリーへのアクセス効率に影響されます。エンジンの圧縮パフォーマンスも読み取りレイテンシーに影響を与えます。

データはCassandra* クラスターにキャッシュされます。rENIACではデータのコピーを行わないので、このソリューションはCassandra* に使用されるものと同じセキュリティ標準に従います。

rENIACのData Engineは、インテル® FPGAと汎用サーバーまたはクラウド・インスタンス (Alibaba* Cloud F1 インスタンスなど) を使用して、NoSQL データベースや検索などのデータ中心のワークロード向けのアクセラレーション・ソリューションを構築するテクノロジー・プラットフォームを提供します。rENIACのDistributed Data Engineは、データレイヤーとアプリケーション・レイヤーの切り離しを実現すると同時に、あらゆるボトルネックを解消するI/O アクセラレー

ターとして機能します。ストレージクラスのメモリーを低レイテンシーのネットワーク・スタックに直接接続するというこのエンジン独自のアーキテクチャーにより、大量のトランザクションが処理されます。

分散データベース向けのrENIAC NoSQL Data Engineは、クライアントコードやデータベースを一切変更することなく、最小限のコンフィグレーションだけで利用できるよう設計されています。rENIAC NoSQL Data Engine ノードは、コンフィグレーションされたポート上で、受信クエリーを聞き取ります。読み取りクエリーに対して、エンジンはクエリーの構文解析を行い、ローカルストレージ内でデータを探します。そして、データが見つかった場合は、結果をクライアントに返します。見つからなかった場合は、データベース・クラスターからデータを取得し、コピーをローカルストレージに格納して、結果をクライアントに返します。

挿入、更新、削除の操作の場合、Data Engineはクエリーをデータベース・クラスターに転送します。データベースがクエリーの処理に成功すると、エンジンがその応答をクライアントに転送し、データのコピーを更新します。

Data Engineは、パフォーマンスの把握に役立つメトリックとログメッセージを生成します。メトリックとログはエンジンからメトリックサービスに送信され、メトリックサービスがコンソールウィンドウやブラウザーに表示し、そのデータをディスクに格納します。

rENIAC Data Engineの主要な3つのコンポーネント

1. データエンジン: データ・ワークロード向けに完全かつ透過的なアプリケーション・レイヤーを実装します。
2. ストレージエンジン: カラム、ドキュメント、グラフなどの主要なデータモデルと、インテル® 3D XPoint™ メモリー、NVMe® SSD、SATA SSDなどの最新フラッシュメモリー・テクノロジーへのインターフェイスをサポートします。
3. ネットワーク・エンジン: ほとんどのネットワーク・プロトコル (TCP/IP、DHCP、ARP、ICMP など) をサポートし、障害の起こりやすいデータセンター環境内のLinux® 上の既存のデータ・ワークロードとの相互運用性を実現します。

アプリケーションの観点から見た rENIAC の Data Engine の主な特長

- CQL 互換 (アプリケーションの変更が不要)
- ライトスルー・キャッシュ
- 自動的なデータ一貫性
- 読み取りと書き込みのパイプラインの分離により、ミリ秒単位の読み取り SLA を実現
- 数百ものデータ・エンジン・ノードへの拡張性 (ゴシッププロトコルの使用)



インテル® Arria® 10 FPGA

インテル® Arria® 10 FPGA は、IoT アプリケーションにおけるディープラーニングの分野でパワフルな汎用アクセラレーターとして機能します。インテル®

Arria® 10 FPGA は、一般公開されている OpenCores® デザインを使用した場合、高速のコア性能と最大 20% の fmax 向上を実現しています。² インテル® Arria® 10 FPGA & SoC は、前世代の FPGA & SoC に比べて最大 40% の消費電力低減も実現しており、最大 1.5 Tera Floating-point Operations Per Second (TFLOPS) の速度を実現する業界唯一のハード浮動小数点デジタル信号処理 (DSP) ブロックを搭載しています。² インテル® FPGA は、高速処理、ハイパフォーマンス HD ビデオ解析、確定的な低レイテンシーに加えて、さまざまなタイプの画像センサーとのインターフェイスを実現する柔軟性も提供します。さらに、FPGA を採用した製品は、導入後のアップデートが可能であり、インフラストラクチャーが AI の最新のテクノロジーを常に利用できるようにします。

OEM 各社とソリューション・プロバイダーに対するインテル® FPGA とインテル® FPGA PAC の利点

柔軟性	• FPGA の機能は、デバイスの電源を入れるたびに変更することができます。設計エンジニアが変更や試行を行うときは、新しいコンフィギュレーション・ファイルをデバイスにダウンロードするだけです。 • FPGA に対する変更は頻繁に行うことができ、PC ボードの変更が必要なくコストもかかりません。
市場投入期間の短縮	• FPGA デバイスは汎用品として入手できます (一方、ASIC の製造サイクルには数カ月を要します)。 • FPGA の柔軟性により、OEM 各社は、その設計が正しく機能しテストが完了した時点で、直ちにシステムを出荷し、ソリューションを素早く市場に投入することができます。 • FPGA によって CPU のオフロードと高速化が実現されることで、システム全体の処理速度が向上します。
集積化	• 簡素化されたホワイトボックス統合。 • FPGA 内部の機能が増えるほど、基板上のデバイス数は減ります。そして、デバイス故障の可能性が低減する分、信頼性は向上します。 • 現在の FPGA には、オンダイ・プロセッサ、28 Gbps (またはそれ以上) のトランシーバー I/O、RAM ブロック、DSP エンジンなどが含まれています。
総保有コスト (TCO)	• FPGA なら、プロトタイプ・システムを出荷してフィールド試験を行うことができるだけでなく、大量生産に入る前でも、迅速な変更が可能であるため、リスクを低減できます。 • 長期にわたるデバイス供給 (15 年以上) をサポートしています。 • 基板上の電子デバイスの 1 つが生産終了 (EOL) になっても、OEM 生産機器の再設計や再認定のコストがかかりません。 • ASIC は同等の FPGA に比べて単価は低いものの、その構築には、開発費 (NRE)、高価なソフトウェア・ツール、専門の設計チーム、および長い製造サイクルが必要となります。

まとめ

rENIACのDistributed Data Engineは、データ・アプリケーションを実行するときの非効率性を解消し、スケールアウト型のアーキテクチャーでTCOを低減しながらパフォーマンスを大幅に向上します。既存のソフトウェア・スタックに対するソフトウェア変更が不要であるにもかかわらず、rENIACはビジネス継続性からパフォーマンス監視まで、さまざまな基幹業務のニーズをサポートします。高性能なIntel®アーキテクチャー上で実行し、さらにIntel®FPGAが提供する大幅なアクセラレーションを利用することで、rENIACはCassandra®NoSQLデータベースの高速化を提供し、データ主導のトランザクション社会を実現します。

rENIACとIntelのサポートにより、企業はパフォーマンスで妥協することなく、爆発的に増え続けるデータを管理することが可能になります。

詳細情報

Intel®FPGAについては、<http://www.intel.co.jp/fpga/>を参照してください。

AI向けのIntelのイノベーションについては、<http://www.intel.co.jp/ai/>を参照してください。

rENIACの詳細については、<http://www.reniac.com/> (英語)を参照いただくか、info@reniac.comまでお問い合わせください。

Apache Cassandra*の詳細については、<http://cassandra.apache.org/> (英語)を参照してください。

rENIAC, Inc.について

rENIAC (米国カリフォルニア州マウンテンビュー) は、パブリック・クラウド、ハイブリッド、およびオンプレミスのデータセンターにおけるクリティカルなワークロードに対して、既存のアプリケーションに対するソフトウェアの変更を行うことなく、レイテンシーの低減とスループットの向上を実現します。rENIACのDistributed Data Engineは、データベース、ネットワーク、ストレージのソリューションに役立つよう設計されており、より多くのCPUリソースをビジネス価値の創造向けに解放することができます。

<http://www.reniac.com/> (英語)

rENIAC



ハードウェア	rENIAC プロキシサーバー	Cassandra* クライアント	Cassandra* サーバー
プロセッサ	8コア (Intel® Xeon® Silver 4109T プロセッサ @2 GHz)	40コア (2 x Intel® Xeon® Gold 6148 プロセッサ @2.40 GHz)	32コア (2 x Intel® Xeon® Gold 6130 プロセッサ @2.10 GHz)
メモリー	64 GB DDR3	128 GB RDIMM, 2666MT/s, デュアルランク	192 GB RDIMM, 2666MT/s, デュアルランク
NVMe*/ストレージ	187 GB Intel® Optane™メモリー	750 GB NVMe*	224 GB NVMe*
オペレーティング・システム	CentOS* 7.6	CentOS* 7.6	CentOS* 7.6
ソフトウェア	rENIAC FPGA Data Engine	Apache Cassandra* v3.11	Apache Cassandra* v3.11

Cassandra* ワークロードのセットアップ: Apache Cassandra* v3.11, パーティション: 500万, パーティション4KB当たりのサイズ: 5カラム, カラム当たり780バイト

Cassandra* Stress コマンド: `./cassandra-stress read n=1000000 no-warmup -node 11.22.101.106 -port native=8001 -schemakeyspace=4kapache310 -col n=FIXED\{5\} size=FIXED\{780\} -pop dist=uniform\{1..5000000\} -rate threads=1`

使用したベンチマーク・ツール: Cassandra* Stress テスト (業界標準のテスト): http://cassandra.apache.org/doc/4.0/tools/cassandra_stress.html (英語)

² <http://www.intel.co.jp/content/www/jp/ja/products/programmable/fpga/arria-10.html>

性能の測定結果は2018年12月時点のテストに基づいています。また、現在公開中のすべてのセキュリティ・アップデートが適用されているとは限りません。詳細については、公開されている構成情報を参照してください。絶対的なセキュリティを提供できる製品はありません。

性能に関するテストに使用されるソフトウェアとワークロードは、性能がIntel®マイクロプロセッサ用に最適化されていることがあります。SYSmark*やMobileMark*などの性能テストは、特定のコンピューター・システム、コンポーネント、ソフトウェア、操作、機能に基づいて行われたものです。結果はこれらの要因によって異なります。製品の購入を検討される場合は、他の製品と組み合わせた場合の本製品の性能など、ほかの情報や性能テストも参考にして、パフォーマンスを総合的に評価することをお勧めします。詳細については、<http://www.intel.com/benchmarks/> (英語)を参照してください。

Intel®テクノロジーの機能と利点はシステム構成によって異なり、対応するハードウェアやソフトウェア、またはサービスの有効化が必要となる場合があります。実際の性能はシステム構成によって異なります。絶対的なセキュリティを提供できるコンピューター・システムはありません。詳細については、各システムメーカーまたは販売店にお問い合わせいただくか、<http://www.intel.co.jp/iot/>を参照してください。

記載されているコスト削減シナリオは、指定の状況と構成で、特定のIntel®プロセッサ搭載製品が今後のコストに及ぼす影響と、その製品によって実現される可能性のあるコスト削減の例を示すことを目的としています。状況はさまざまであると考えられます。Intelは、いかなるコストもコスト削減も保証いたしません。

Intelはサードパーティーのデータについて管理や監査を行っていません。本資料の内容をレビューし、ほかの情報も参考にして、参照しているデータが正確かどうかを確認してください。

Intel、インテル、Intelロゴ、Intel Insideロゴ、3D XPoint、Arria、Intel Optane、Xeonは、アメリカ合衆国および/またはその他の国におけるIntel Corporationまたはその子会社の商標です。

* その他の社名、製品名などは、一般に各社の表示、商標または登録商標です。