



メディアアラート

2018年12月11日

AWS、新たな Amazon EC2 インスタンスと ネットワーク拡張を発表

マイクロサービスやウェブサーバーなどのスケールアウト型ワークロードのコストを最大 45%削減する、Armベースのカスタム AWS Graviton プロセッサを搭載した新しい A1 インスタンス

機械学習や高性能コンピュートなど、スケールアウト型で分散型のワークロード向けに高スループットを提供する、新たな P3 および C5 インスタンスの 100 Gbps ネットワーキング機能

数万のコア越しに緊密に結合した高性能コンピューティング・アプリケーションを拡張する、新しい遅延最適化 Elastic Fabric Adapter

インターネットトラフィックをインテリジェントにルーティングして
地理的に分散されたアプリケーションの可用性と性能を高める AWS Global Accelerator

(シアトル、2018年11月27日発表) [Amazon.com, Inc.](https://www.amazon.com) (NASDAQ:AMZN) の関連会社である Amazon Web Services, Inc. (AWS) は本日、AWS re:Invent にて、Amazon Elastic Compute Cloud (Amazon EC2) 向けの 3 つの新たなインスタンスを発表しました。

- ・A1 インスタンス: スケールアウト型ワークロード向けにカスタム設計された AWS Graviton プロセッサを搭載したインスタンス
- ・P3dn GPU インスタンス: ディストリビューション型の機械学習および高性能コンピューティング・アプリケーションに最適なインスタンス
- ・C5n インスタンス: 最先端のコンピューティング、集中したワークロード向けに拡張したネットワーク帯域幅を提供するインスタンス

コンピューティングに最適化された新たな P3dn GPU および C5n インスタンスは、100 Gbps のネットワーク・スループットを提供し、高性能コンピューティング (HPC)、機械学習トレーニング、データ分析などの分散型ワークロードのスケールアウトを可能にします。

また、AWS は 2 つの新たなネットワークサービスの提供開始を発表しました。

- ・Elastic Fabric Adapter (EFA): AWS の柔軟性と拡張性を備えたオンプレミス HPC クラスターの性能を提供する、Amazon EC2 インスタンス向けネットワークアダプター
- ・AWS Global Accelerator: AWS のグローバルバックボーンとエッジロケーションを活用し、1 つ以上の AWS リージョンで稼働するアプリケーションの可用性と性能を改善するフルマネージドサービス

新たな Amazon EC2 インスタンスおよびネットワーク機能に関しては、以下の URL を参照ください。

<https://aws.amazon.com/ec2/> もしくは <https://aws.amazon.com/products/networking/>.

AWS のコンピュートサービス担当バイスプレジデントのマット・ガーマン (Matt Garman) は「お客様の要望で最も多いものに、基本的なワークロードのコストを下げる方法とクラウドにおいて要求の厳しいスケールアウト型の高性能なコンピュートや機械学習のワークロードの効率的な稼働方法の 2 つが挙げられます。今回、A1 インスタンスを追加したことで、コンテナ化されたマイクロサービスのような分散型アプリケーションをコスト効率良く運用できる選択肢をお客様に提供できます。A1 インスタンスは、過去 10 年に渡りハイパースケールのクラウドプラットフォームを開発してきた当社の専門知識を活用し、当社が新しくカスタム設計した AWS Graviton プロセッサ (ARM 命令セット) を搭載しています。当社の新しい P3dn および C5n インスタンスは、スケールアウト型で分散型のワークロード向けに、100 Gbps のネットワーク性能を提供し、分散した機械学習トレーニングや高性能コンピュートの処理速度を高めることができます。これらの新しいインスタンスにより、すでに業界で最も高性能でコスト効率に優れたコンピューティングプラットフォームを、新しいワークロードのニーズを満たせるように拡張することができます」と述べています。

スケールアウト型ワークロードのコストを大幅に削減する、カスタム設計の AWS Graviton プロセッサを搭載した新たな A1 インスタンス

今日、多くのお客様が益々クラウド上で様々なワークロードを稼働するようになり、性能を妥協せずにコストを削減できるソリューションを求めようになってきました。汎用プロセッサは多くのワークロードに優れた価値を引き続き提供しますが、コンテナ化されたマイクロサーバーや x86 命令セットに依存しないウェブティアアプリケーションのような新しいスケール型ワークロードの場合は、アプリケーションの演算負荷を共有できる、より小型で最新の 64 ビット Arm プロセッサで稼働するほうがコストと性能の両面で恩恵を得ることができます。本日発表された A1 インスタンスは、システムデザインとクラウドインフラストラクチャにおける AWS の広範囲に及ぶ専門知識を活用してカスタム設計されたプロセッサ (Graviton) を搭載し、初めて Arm プロセッサがクラウド上で利用できるようになります。A1 インスタンスでスケール型ワークロードを稼働すると、その他の Amazon EC2 汎用インスタンスに比べて最大 45% のコスト削減を実現できます。A1 インスタンスは、Amazon Linux 2、Red Hat、Ubuntu を含む複数の Linux ディストリビューション、および Amazon Elastic Container Service (Amazon ECS) や Amazon Elastic Container Service for Kubernetes (EKS) を含むコンテナサービスによりサポートされています。

100 Gbp のネットワーク性能を提供し、多くの演算を必要とするワークロードの処理を早める新たな P3dn および C5n インスタンス

今日、多くのお客様が、クラウド上で最もコンピューティンシブなワークロードの処理に、コンピューティングに最適化された AWS の C5 インスタンスと GPU 搭載 P3 インスタンスを使用し始めています。機械学習トレーニングから、計算流体力学や気象シミュレーションなどの HPC アプリケーション、ビデオエンコーディングに至るまで、こうした様々なワークロードは強力なプロセッサと高速なネットワーク機能から恩恵を得られます。AWS は、安全で拡張性があり、かつ柔軟な 100 Gbp のネットワーク性能を提供する最初のクラウドプロバイダーとなり、お客様は標準ドライバーとプロトコルを用いて、このサービスを HPC だけでなく、分析や機械学習、ビッグデータ、データレイクのワークロードに使用することができます。

・**P3dn インスタンス (来週提供予定)**: クラウド上で最も強力な機械学習トレーニング向けの GPU インスタンスです。P3 インスタンスにより、すでにお客様は機械学習モデルのトレーニング時間を数日から数時間に短縮していますが、新しい大規模な P3dn インスタンスの 100 Gbps のネットワーク性能により、複数の GPU インスタンスに機械学習ワークロードを分散させることで、トレーニング時間をさらに 1 時間未満に短縮することができます。2017 年 10 月に P3 インスタンスを発表して以来、機械学習の採用が広まり、それを活用して、より複雑なアプリケーションに取り組むようになったことで、高性能コンピューティングに対するお客様のニーズも高まってきました。新しい P3dn インスタンスは、従来の P3 インスタンスに比べて 4 倍高いネットワークスループットを提供し、最大 100 Gbps のネットワークスループット、高速な NVMe インスタンスストレージ、96 個の vCPUs を備えるカスタム Intel CPU、AVX512 命令および複数の 32GB メモリー搭載 NVIDIA Tesla V100 GPU のサポートを提供します。これにより、機械学習モデルのトレーニング性能を複数のインスタンス越しに直線的に拡張でき、前処理時間を短縮し、データ転送のボトルネックを解消しながら、自社の機械学習モデルの品質を向上させることができます。

HPC ワークロードの AWS への移行を支援する低遅延の Elastic Fabric Adapter (EFA)

従来、お客様はサイズが固定されたオンプレミスの HPC システムを使用していました。HPC システムは高コストで多大な資本投資が必要なため、企業は調達ができず、実現に時間がかかり、結果として生産性が下がり、検証も限定的になり、重要なプロジェクトの実施に遅れが出ていました。さらに、次回のインフラストラクチャー更新時まで、現在使用可能なハードウェア性能とテクノロジーに縛られ、新しい方法を採用する代わりに、既存のインフラストラクチャーでアプリケーションを稼働せざるを得ない状況でした。本日プレビュー版が提供された EFA は、HPC アプリケーションを拡張するのに不可欠なインスタンス内の通信性能を向上させ、AWS の柔軟性と拡張性といった付加価値と共に、オンプレミス HPC クラスター並みの性能をクラウド上で提供します。EFA は MPI (Message Passing Interface) に統合されているため、一切の変更を加えることなく、HPC アプリケーションを数万の CPU コアに拡張できます。EFA は Amazon EC2 P3dn および C5n インスタンス上で利用可能で、2019 年には新たな Amazon EC2 インスタンスタイプとして提供されます。これにより、事前の計画なしに自社のワークロードに最適なコンピューティング設定をオンデマンドで選択できる選択肢がさらに広がることになります。

グローバルに分散したアプリケーションの可用性と性能を改善する AWS Global Accelerator

お客様がより大規模で地理的に異なるユーザーを対象とするようになるにつれ、システムの運用において可用性と性能を高める必要が出てきました。これらのお客様は、高性能を得るために、複数の AWS リージョンにアプリケーションを展開していますが、これはユーザーを最適で健全なアプリケーション・エンドポイントに振

り分ける必要があることを意味します。また、アプリケーションの障害や性能テスト、トラフィックの上昇に応じて、アプリケーションのエンドポイントを各 IP アドレスとともに定期的に拡張、縮小する必要があります。アプリケーションを拡張すると、今度はそのアプリケーションに接続する各クライアントをアップデート(通常、ドメインネームサービスで行われる)しなければならず、結果として、各クライアントは新しいエンドポイントを見つけるのに多くの時間を要していました。本日提供開始の AWS Global Accelerator により、アプリケーションの可用性と性能は改善し、インターネットトラフィックをユーザーから複数の AWS リージョンで稼働するアプリケーションエンドポイントに向かわせることができます。AWS Global Accelerator は、AWS の広範で高可用性、ふくそうのないグローバルネットワークバックボーンとエッジロケーションを使用し、インターネットトラフィックをユーザーからアプリケーションエンドポイントに向かわせる一方、クライアントは地理的な位置、アプリケーションの健全さ、カスタム設定可能なルーティング規則に基づいて最適なアプリケーションエンドポイントに振り分けられます。また、AWS Global Accelerator は、アプリケーション毎に異なり、変化することのない静的な Anycast IP アドレス群を割り当てることで、アプリケーションを拡張した時にクライアントをアップデートする必要がなくなります。アプリケーションエンドポイントは継続的に管理され、問題が生じた時は AWS Global Accelerator がクライアントの設定を変更することなく、該当するクライアントを健全なエンドポイントに向かわせることもできます。

アマゾン ウェブ サービスについて

アマゾン ウェブ サービス(AWS)は 12 年の間に、世界で最も包括的かつ幅広く採用されたクラウドプラットフォームになっています。AWS は、米国、オーストラリア、ブラジル、カナダ、中国、フランス、ドイツ、インド、アイルランド、日本、韓国、シンガポールおよび英国の 19 の AWS リージョン、1 つのローカルリージョンと 57 のアベイラビリティゾーン(AZ)で、コンピューティング、ストレージ、データベース、ネットワーク、アナリティクス、マシンラーニング、人工知能(AI)、モノのインターネット(IoT)、モバイル、セキュリティ、ハイブリッド、仮想現実および拡張現実(VR/AR)、メディア、アプリケーションディベロプメント、デプロイメントおよびマネージメントに関する 125 種類以上の十分な機能を有するサービスを提供しています。AWS のサービスは、アジリティを高めながら同時にコストを削減できるインフラエンジンとして、急速に成長しているスタートアップや大手企業、および有数の政府機関を含む数百万以上のアクティブなお客様から信頼を獲得しています。AWS の詳細については以下の URL をご参照ください。<https://aws.amazon.com/>

Amazon.com について

Amazon は 4 つの理念を指針としています。競合他社ではなくお客様を起点にすること、創造への情熱、優れた運営へのこだわり、そして長期的な発想です。カスタマーレビュー、1-Click 注文、パーソナライズされたおすすめ商品機能、Amazon プライム、フルフィルメント by Amazon(FBA)、アマゾン ウェブ サービス(AWS)、Kindle ダイレクト・パブリッシング、Kindle、Fire タブレット、Fire TV、Amazon Echo、Alexa などは、Amazon が先駆けて提供している商品やサービスです。

報道関係からのお問い合わせ先
アマゾン ウェブ サービス ジャパン株式会社
広報担当 河村
Email: awsjp-pr@amazon.com

#####