

AWS ビッグデータ関連サービス 最新アップデート

2016 年 11 月 18 日

アマゾンウェブサービスジャパン株式会社
ソリューションアーキテクト | 内海英一郎



スピーカー 紹介



アマゾンウェブサービス
ジャパン株式会社

ソリューションアーキテクト

♥ Amazon Kinesis

♥ Java

♥ LMAX Disruptor

アジェンダ

以下のサービスについて 2016/7 以降のアップデートをご紹介します

- Amazon EMR
- Amazon S3
- Amazon Redshift
- Amazon Kinesis
- Amazon Machine Learning
- Amazon Elasticsearch Service
- Amazon DynamoDB
- Amazon QuickSight



アップデートをいち早くキャッチするために

以下の情報をチェックしましょう

- What's New at AWS
 - <https://aws.amazon.com/new/>
- フォーラムのリリースノート
 - <https://aws.amazon.com/releases/notes>
- AWS Blog
 - <https://aws.amazon.com/blogs/aws/>
- AWS Big Data Blog
 - <https://aws.amazon.com/blogs/big-data/>
- AWS 公式 Twitter アカウント
 - <https://twitter.com/awscloud>
- 各サービスのドキュメント

Amazon EMR

Amazon EMR アップデート

- 2016/08/02 – Amazon EMR release 5.0 now available: Apache Spark 2.0, Hive 2.1, enhanced debugging, and more
 - <https://aws.amazon.com/about-aws/whats-new/2016/08/amazon-emr-release-5-0-now-available-apache-spark-2-0-hive-2-1-enhanced-debugging-and-more/>
 - Spark 2.0, Hive 2.1, Presto 1.150, Tez 0.8.4, Zeppelin 0.6.1 (Snapshot), Hue 3.10
 - EMR がステップの失敗の根本原因を識別できる場合、詳細情報を提供

Amazon EMR アップデート

- 2016/09/21 – Amazon EMR now supports data encryption for Apache Spark, Tez, and Hadoop MapReduce
 - <https://aws.amazon.com/about-aws/whats-new/2016/09/amazon-emr-now-supports-data-encryption-for-apache-spark-tez-and-hadoop-mapreduce/>
 - EMR 4.8.0 以降
 - HDFS, EMRFS on S3 で At-rest 暗号化をサポート
 - EMR – S3 間, EMR クラスターのノード間通信で In-transit 暗号化をサポート

Amazon EMR アップデート

- 2016/11/03 – Now run real-time stream processing at scale with Apache Flink on Amazon EMR
 - <https://aws.amazon.com/about-aws/whats-new/2016/11/now-run-real-time-stream-processing-at-scale-with-apache-flink-on-amazon-emr/>
 - EMR 5.1.0
 - Flink 1.1.3, Zeppelin 0.6.2, HBase 1.2.3

Amazon EMR アップデート

- 2016/11/16 – Amazon EMR now supports fine-grained access control using cluster tags
 - <https://aws.amazon.com/about-aws/whats-new/2016/11/amazon-emr-now-supports-fine-grained-access-control-using-cluster-tags/>
 - クラスタに付与されたタグを条件とするアクセス制御が IAM ポリシー内で記述可能に

Amazon S3

Amazon S3 アップデート

- 2016/08/11 – Amazon S3 support for IPv6
 - <https://aws.amazon.com/releases/1150571418807500>

Amazon Redshift

Amazon Redshift アップデート

- 2016/09/15 – Amazon Redshift now supports Enhanced VPC Routing
 - <https://aws.amazon.com/about-aws/whats-new/2016/09/amazon-redshift-now-supports-enhanced-vpc-routing/>
 - COPY および UNLOAD のトラフィックを VPC 経由で転送可能に

Amazon Redshift アップデート

- 2016/09/30 – Amazon Redshift introduces new data type to support time zones in time stamps
 - <https://aws.amazon.com/about-aws/whats-new/2016/09/amazon-redshift-introduces-new-data-type-to-support-time-zones-in-time-stamps/>
 - TIMESTAMPTZ データタイプをサポート
 - Redshift 内部で UTC に変換されて保存

Amazon Redshift アップデート

- 2016/11/10 – Amazon Redshift adds connection limits for users and for databases and support for UTF-8 multibyte characters in object names
 - <https://aws.amazon.com/releases/3157984964781295>
 - ユーザーが同時にオープンすることができるデータベース接続の上限値が設定可能に
 - テーブル名, カラム名, その他データベースオブジェクト名に UTF-8 文字が利用可能に

Amazon Redshift アップデート

- 2016/11/11 – Amazon Redshift announces new data compression, connection management, and data loading features
 - <https://aws.amazon.com/about-aws/whats-new/2016/11/amazon-redshift-announces-new-data-compression-connection-management-and-data-loading-features/>
 - ANALYZE COMPRESSION がストレージ容量の削減率を報告
 - CREATE TABLE AS が自動圧縮をサポート
 - ソートキー順でのデータロードの改善

Amazon Kinesis

Amazon Kinesis アップデート

- 2016/08/11 – Introducing Amazon Kinesis Analytics
 - <https://aws.amazon.com/about-aws/whats-new/2016/08/introducing-amazon-kinesis-analytics/>
 - Kinesis Streams または Kinesis Firehose のストリーミングデータを SQL でリアルタイムに分析
 - 米国東部（バージニア北部）／米国西部（オレゴン）／欧州（アイルランド）リージョンで利用可能

Amazon Kinesis アップデート

- 2016/09/16 – Amazon Kinesis announces IAM role support in Kinesis Agent
 - <https://aws.amazon.com/about-aws/whats-new/2016/09/amazon-kinesis-announces-iam-role-support-in-kinesis-agent/>

Amazon Kinesis アップデート

- 2016/11/07 – Test drive Amazon Kinesis Firehose using built-in sample data
 - <https://aws.amazon.com/about-aws/whats-new/2016/11/test-drive-amazon-kinesis-firehose-using-built-in-sample-data/>
 - 配信ストリームの設定確認に株価のサンプルデータが利用可能に

Amazon Kinesis アップデート

- 2016/11/11 – Amazon Kinesis Streams console now supports configuring data retention and shard level metrics
 - <https://aws.amazon.com/about-aws/whats-new/2016/11/amazon-kinesis-streams-console-now-supports-configuring-data-retention-and-shard-level-metrics/>

Amazon Kinesis アップデート

- 2016/11/15 – Automate Amazon Kinesis Streams scaling and shard limit monitoring using new APIs
 - <https://aws.amazon.com/about-aws/whats-new/2016/11/amazon-kinesis-streams-scaling-and-shard-limit-monitoring-using-new-apis/>
 - Kinesis Streams に UpdateShardCount API が追加
 - ターゲットのシャード数を指定してスケールアップ／ダウンが可能に
 - 24 時間あたりの呼び出し回数やターゲットのシャード数に制限あり

Amazon Machine Learning

Amazon Machine Learning アップデート

- 2016/07/20 – Now Easily Delete Multiple Objects through the Amazon ML Console
 - <https://aws.amazon.com/about-aws/whats-new/2016/07/now-easily-delete-multiple-objects-through-the-amazon-ml-console/>
 - 複数のデータソース, モデル, 評価を選択して削除可能に

Amazon Machine Learning アップデート

- 2016/07/26 – Now Access CloudWatch Metrics directly from the Amazon ML Console
 - <https://aws.amazon.com/about-aws/whats-new/2016/07/now-access-cloudwatch-metrics-directly-from-the-amazon-ml-console/>

Amazon Elasticsearch Service

Amazon Elasticsearch Service アップデート

- 2016/07/27 – Amazon Elasticsearch Service now supports Elasticsearch 2.3
 - <https://aws.amazon.com/about-aws/whats-new/2016/07/amazon-elasticsearch-service-now-supports-elasticsearch-2-3/>
 - パイプライン集計, geo-point フィールドの改善, データの圧縮サポート
 - バージョン 1.5 と比較してパフォーマンス, メモリー管理, 耐障害性, セキュリティが改善

Amazon DynamoDB

Amazon DynamoDB アップデート

- 2016/08/15 – Announcing General Availability of aws-record
 - <https://aws.amazon.com/about-aws/whats-new/2016/08/aws-record-general-availability/>
 - <https://github.com/aws/aws-sdk-ruby-record>
 - Ruby 向けの Amazon DynamoDB 抽象化レイヤー

Amazon DynamoDB アップデート

- 2016/09/23 – The Amazon EMR-DynamoDB Connector for Apache Hive and Apache Spark is now open-source
 - <https://aws.amazon.com/about-aws/whats-new/2016/09/the-amazon-emr-dynamodb-connector-for-apache-hive-and-apache-spark-is-now-open-source/>
 - <https://github.com/awslabs/emr-dynamodb-connector>

Amazon QuickSight

Amazon QuickSight アップデート

- 2016/11/15 – Amazon QuickSight Now Generally Available
 - <https://aws.amazon.com/about-aws/whats-new/2016/11/amazon-quicksight-now-generally-available/>
 - Excel ファイル, SQL Server, MySQL, PostgreSQL, Salesforce, RDS, Redshift, S3 等のデータソースに対応
 - 超高速・並列・インメモリーの計算エンジン “SPICE” を搭載
 - Web ブラウザーと iOS モバイルアプリから利用可能
 - 2 か月間のフリートライアルおよび SPICE 無料枠あり
 - 米国東部（バージニア北部）／米国西部（オレゴン）／欧州（アイルランド）リージョンで利用可能

その他サービス

その他サービス アップデート

- 2016/08/18 – AWS Cost and Usage Report Data is Now Easy to Upload Directly into Amazon Redshift and Amazon QuickSight
 - <https://aws.amazon.com/about-aws/whats-new/2016/08/aws-cost-and-usage-report-data-is-now-easy-to-upload-directly-into-amazon-redshift-and-amazon-quicksight/>

その他サービス アップデート

- 2016/08/25 – New AWS Public Data Set - SpaceNet on AWS
 - <https://aws.amazon.com/about-aws/whats-new/2016/08/new-aws-public-data-set-spacenet-on-aws/>

その他サービス アップデート

- 2016/10/07 – Announcing the Hadoop Distributed File System Support for AWS Import/Export Snowball
 - <https://aws.amazon.com/about-aws/whats-new/2016/10/announcing-the-hadoop-distributed-file-system-support-for-aws-import-export-snowball/>

