



クラウドを活用したセンシング／ モニタリングなどデータ分析の実現

Amazon Web Services Japan K.K.
Takayuki Shimizu, Solutions Architect

2017. 11. 14

登壇者の紹介



清水 崇之

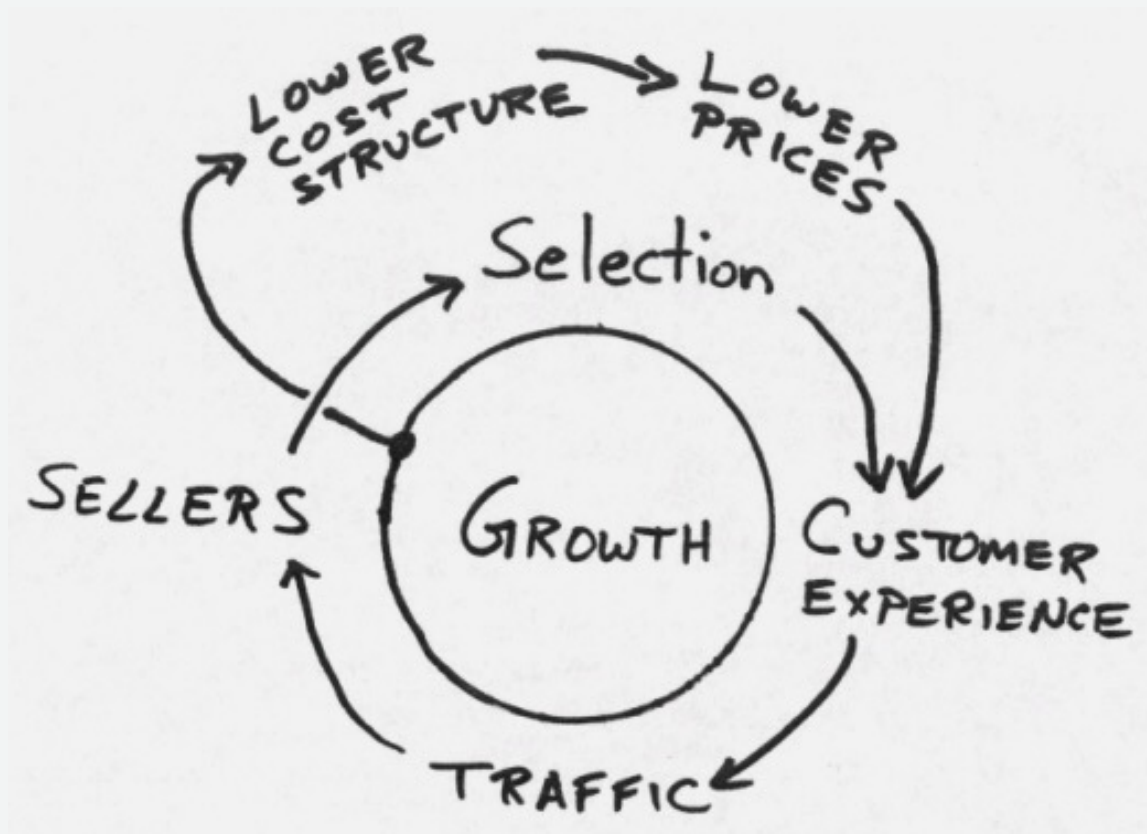
- ソリューションアーキテクト
- 西日本担当 @ 中之島オフィス
 - AWS導入時の技術サポート
 - イベント登壇など

Amazon と AWS のご紹介

Amazon のビジネス



規模の経済 × ITの活用 で薄利多売のビジネスモデル



Amazon Dash Button



Amazon Echo



ALEXA SKILLS KIT

新規機能の追加:

ASK は、音声を使ったアプリを簡単に構築ができます。



AWS上に構築
自動音声認識(ASR)
自然言語理解(NLU)
学習機能

ALEXA VOICE SERVICE

比類の無い配信:

AVSを使うことで、あなたのコンテンツをどこにでも配信できます。

amazon
echo
amazon
fireTV



Alexa, play Adele from Prime Music
(Prime Musicからアデルの曲をかけてください)

Alexa, ask Uber to request a ride
(Uberにタクシーを呼ぶよう頼んでください)

Amazon Go

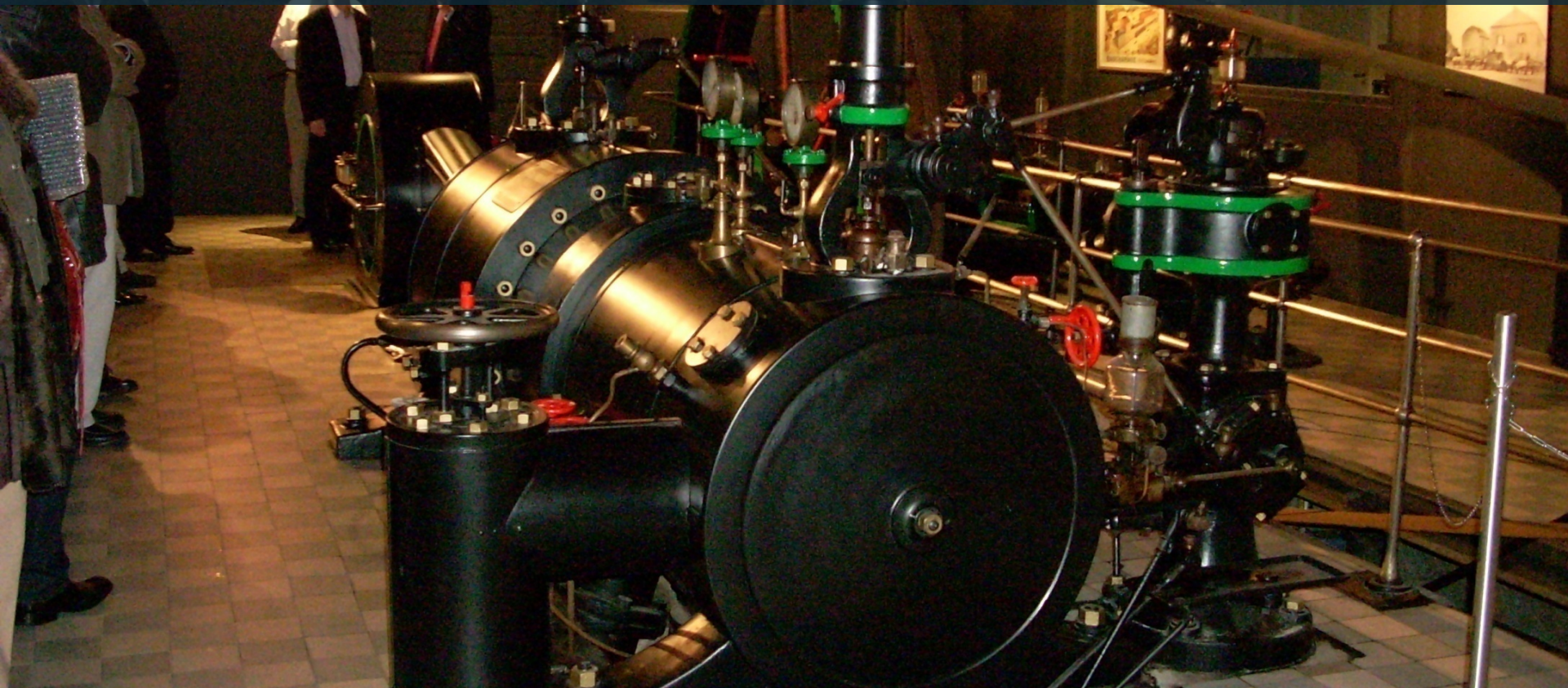


AWS (Amazon Web Service) とは Amazon の DNA を持ったクラウドサービス

クラウドとは？

クラウドのインパクトを
歴史上の出来事に例えると・・・

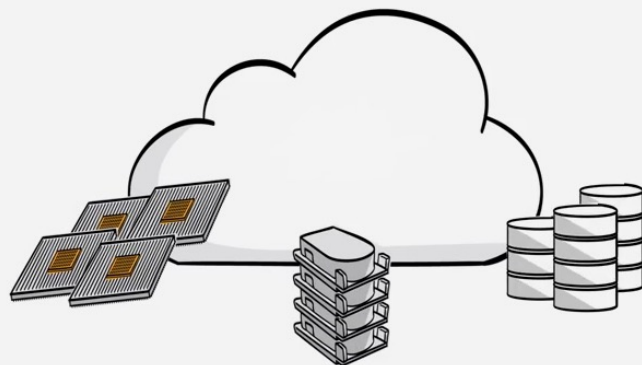
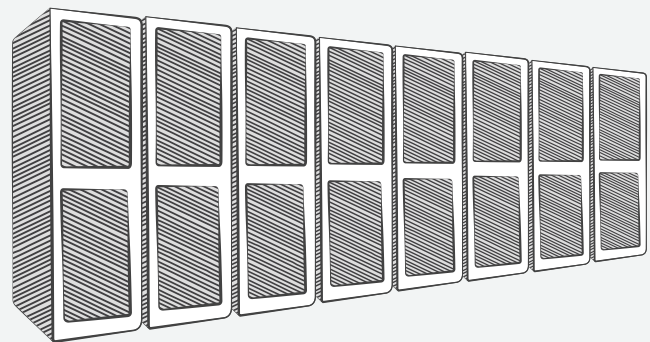
発電機所有が差別化要因だった時代の終焉



発電所＋送電網の出現がパラダイムシフトに



同様の変革が IT の世界にも



初期投資が発生

余剰・不足のリスク

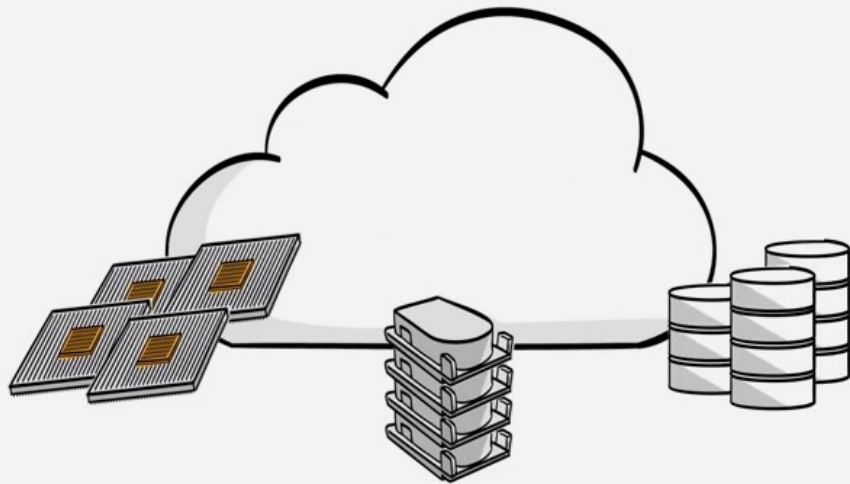
固定費

初期投資が不要

必要な分だけ利用可能

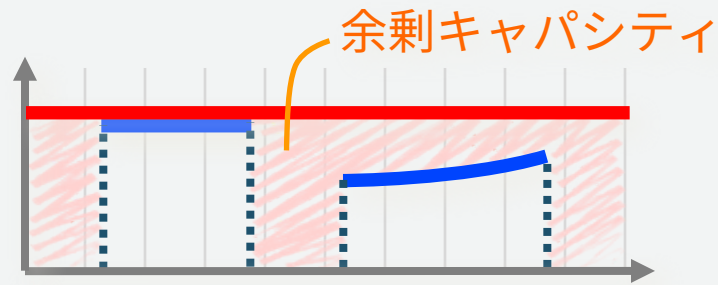
変動費

Amazon が考えるクラウドとは？

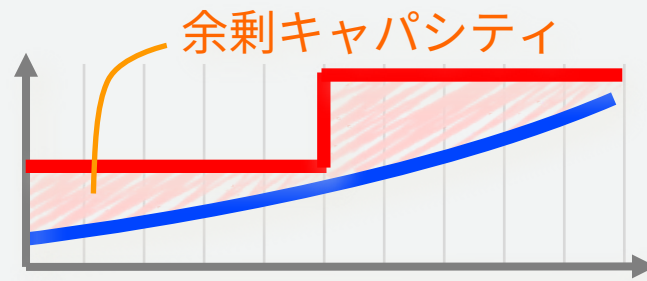


必要な時に、必要なだけ、低価格で
IT リソースを提供

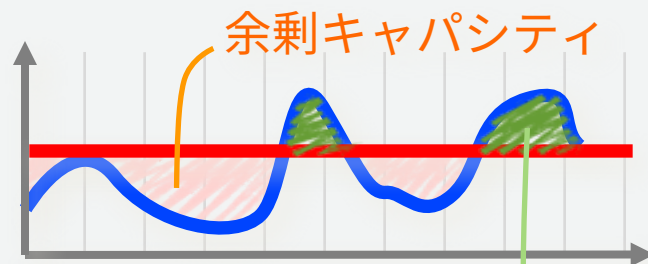
Before



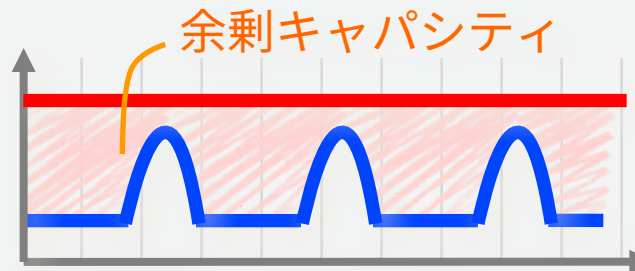
オンとオフ



急成長やM&A



予測できないピーク

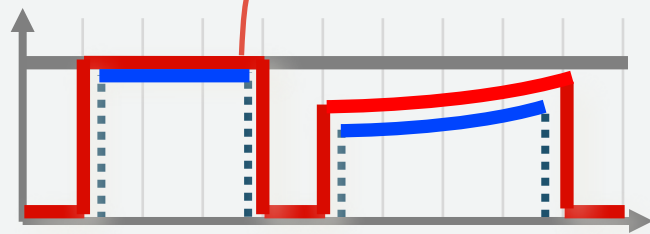


予測可能なピーク

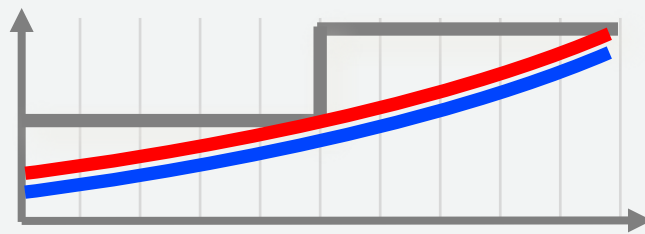
キャパシティ不足：機会損失

After

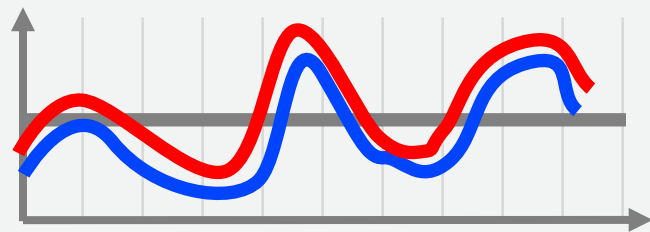
IT余剰と不足からの開放



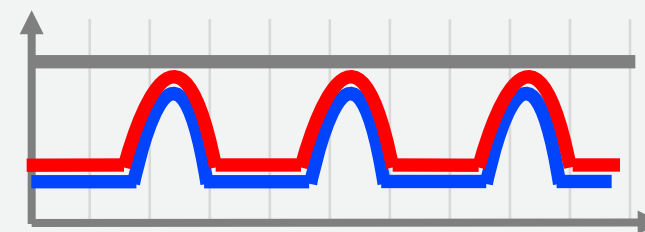
オンとオフ



急成長やM&A

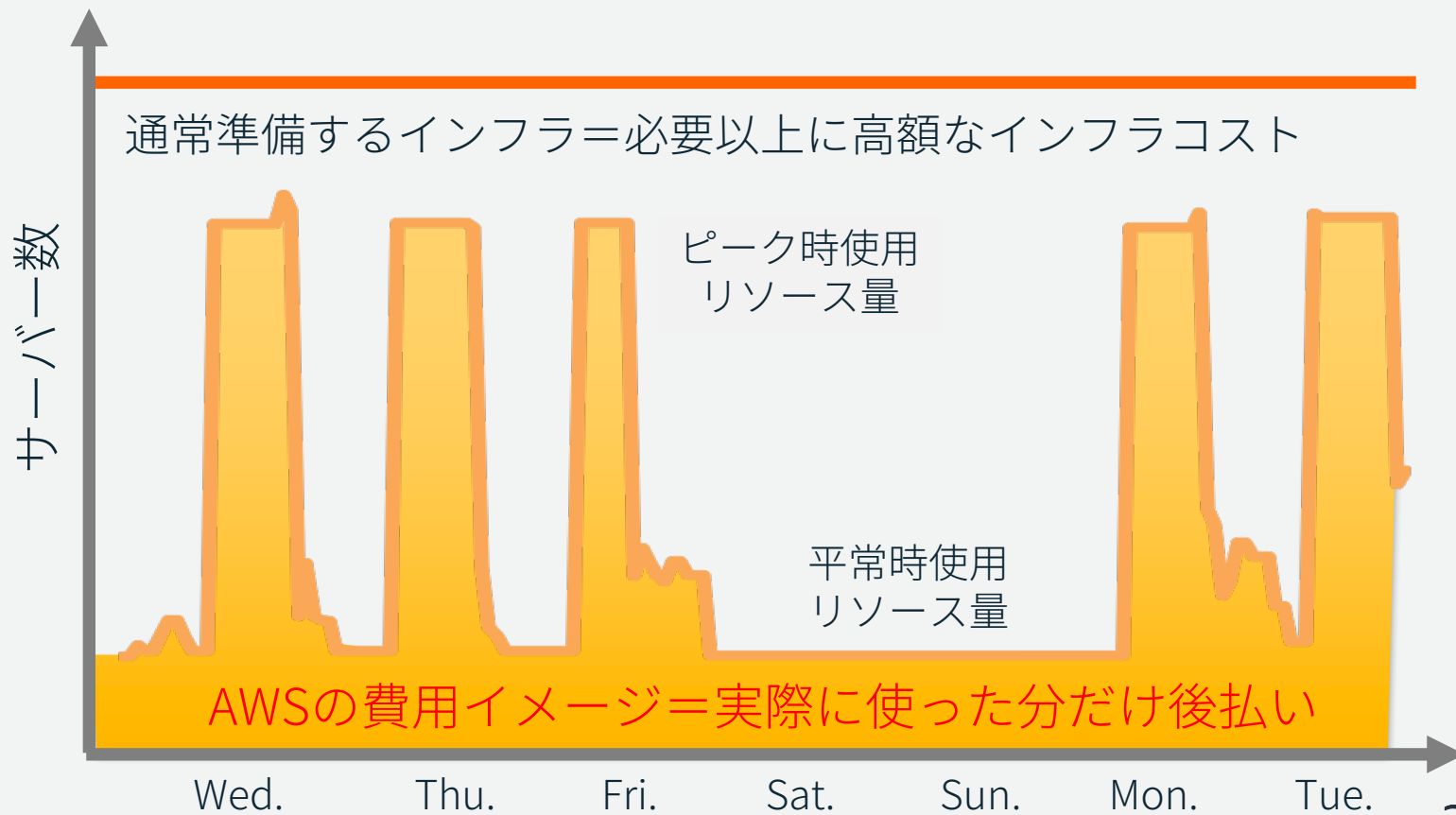


予測できないピーク



予測可能なピーク

初期投資0円 & 従量課金



継続的な値下げ

規模の拡大と
イノベーション

資本
投資

より多くの
顧客獲得

過去10年間で 60 回以上の値下げ

技術
投資

値下げ

効率
改善



すぐ用意できる

新しいインフラの構築は
複雑かつ遅くなりがち
障害対応・維持負担も高い



5分でインフラを用意
維持管理なし



なるべく多くの
ビジネスニーズに、
「できる」と答えられる



新しいデプロイ環境



新しいテスト環境



障害時の代替サーバー用意

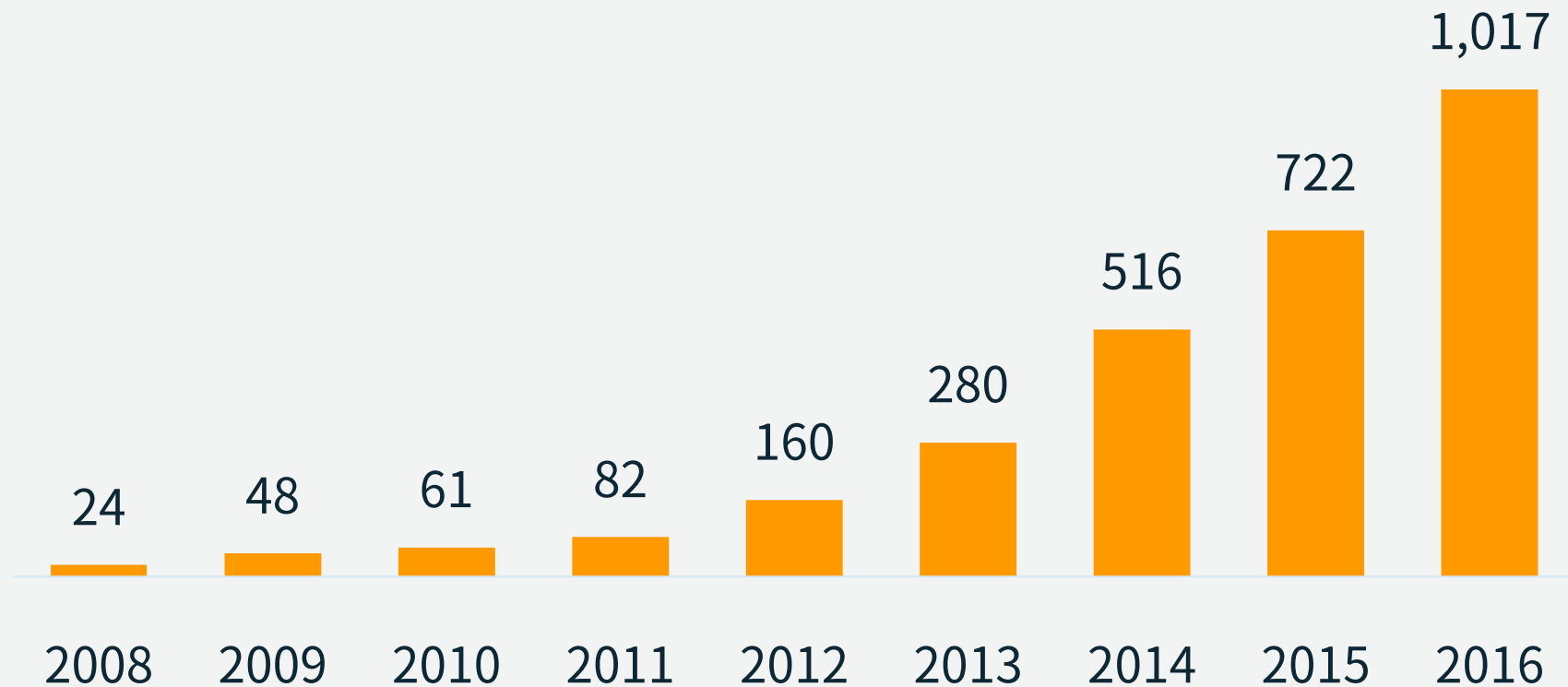


メディア掲載による
アクセス増対応



1,000台のサーバを
追加・削除

お客様のフィードバックにもとづく継続的なイノベーション



90 以上のクラウドサービスを提供



リージョンは世界中に展開

複数のアベイラビリティゾーンで構成されておりデータセンタ単位の冗長性を確保

16

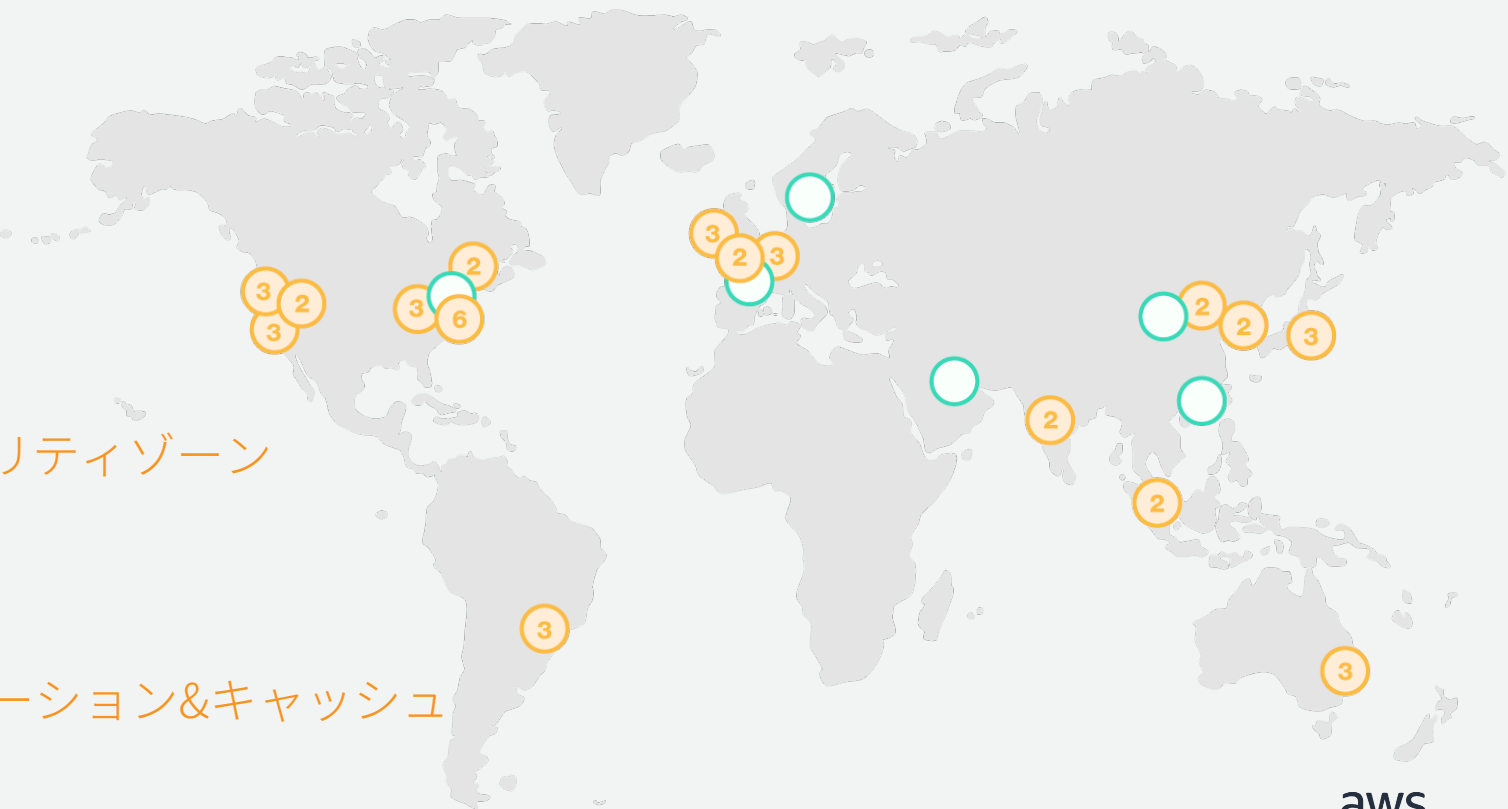
リージョン

44

アベイラビリティゾーン

100

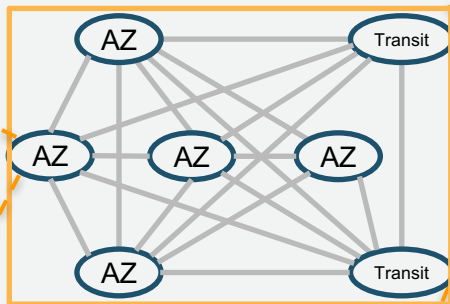
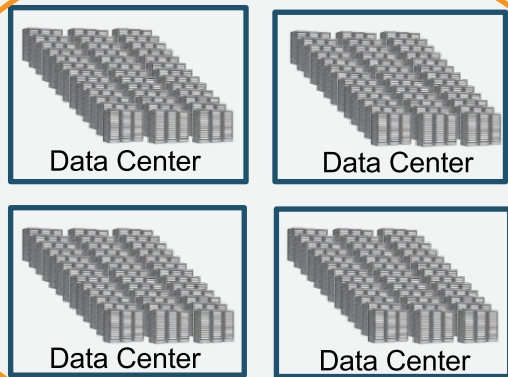
エッジロケーション&キャッシュ



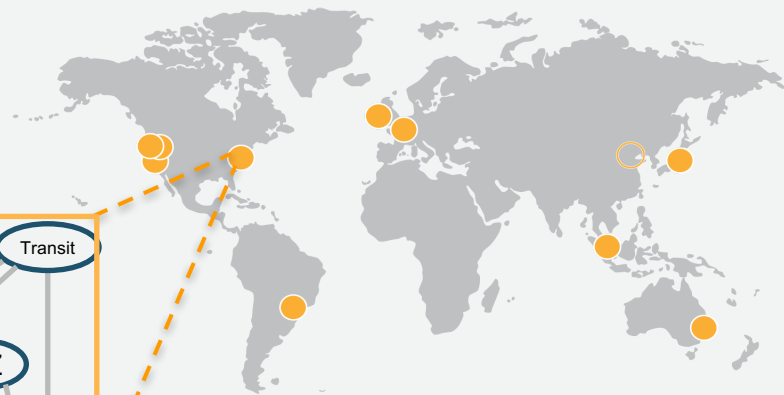
アベイラビリティゾーンとは？

データセンターレベルの冗長性により高い可用性を実現する

A data center hosts around 50K servers or more
Data centers are connected with private network links

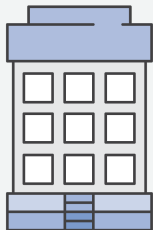


- 2 or more AZ in a region
- AZ consists of 1 or more DCs
 - A DC belongs to one AZ
- less than 1/4 ms latency between DCs
 - Low latency/high bandwidth network



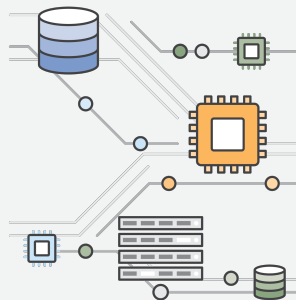
セキュリティ

物理 セキュリティ



秘匿性の高いデータセンター群
非常に限定されたデータセンターへ
の物理アクセス
全ての物理アクセスを記録・監査


論理 セキュリティ (ハードウェア・ソフトウェア・ネットワーク)



全てのシステム変更を管理・記録
段階的なアップデートとリリース
ストレージ廃棄プロセスの明確化
自動化されたモニタリングと監査

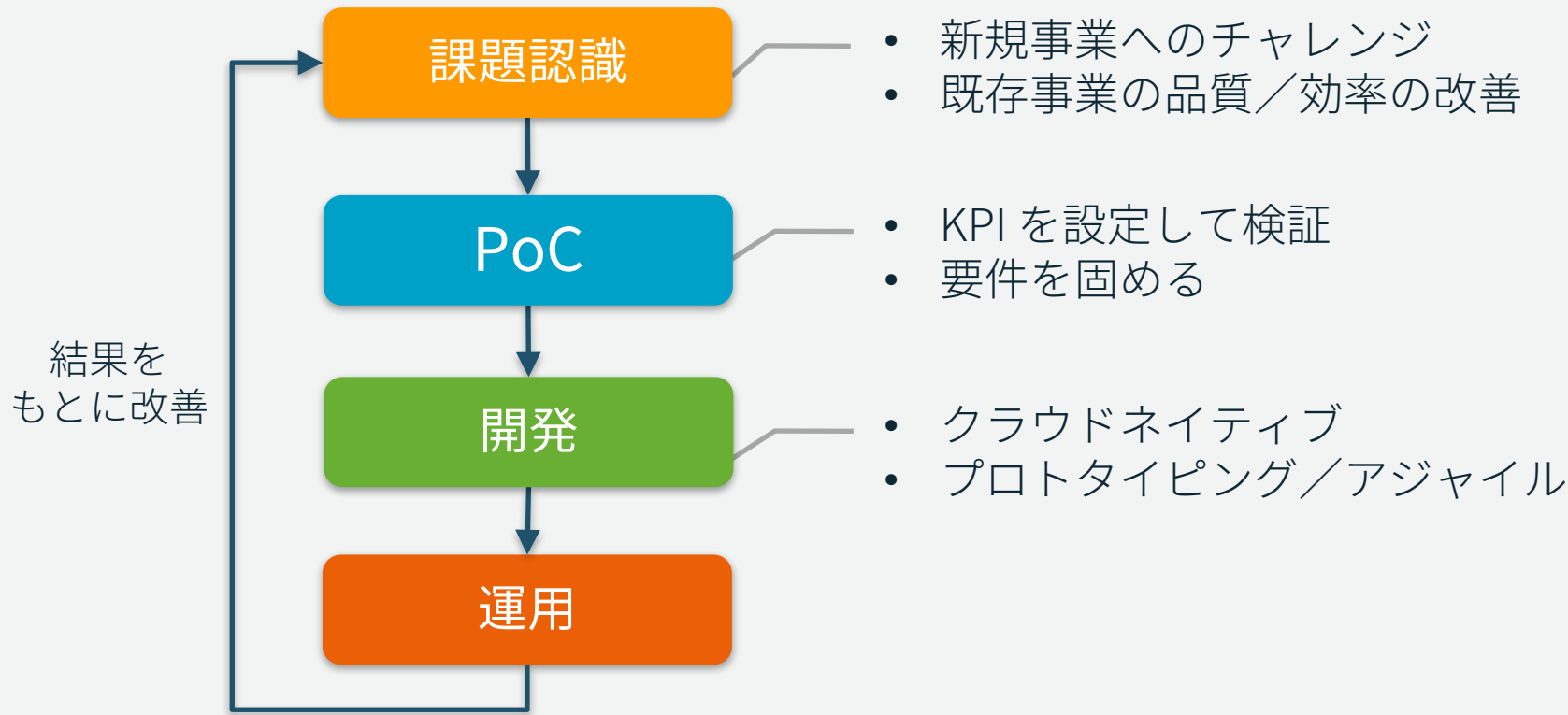
認定および認証



A close-up photograph of a child's hand, wearing a white long-sleeved shirt, placing a blue LEGO brick onto a structure. The structure is composed of various colored bricks: a tall red brick, a blue brick, a yellow brick, a green brick, and a white brick. The background is a light-colored, textured surface.

ビルディングブロックのように組み合わせることで
素早くアプリケーション構築が可能

クラウドならスモールスタートできる



事例紹介

参考資料

大阪ガス株式会社 様

https://media.amazonaws.com/jp/summit2016/1B_02.pdf

ダイードロリンコ株式会社 様

<https://d0.awsstatic.com/events/jp/2017/summit/slide/D4T5-6.pdf>

株式会社 NEXCO システムズ 様

<https://d1.awsstatic.com/events/jp/2017/summit/slide/D2T5-2.0fc6ea7430314376cc38195fdac71f9fb3be68af.pdf>

センシング／モニタリングなどデータ分析に
クラウドをどのように活用するか？

クラウドで構築するIoTプラットフォーム

求められる要件

モニタリング

位置情報管理・状態監視・
実績把握・動線把握

予防予知保全

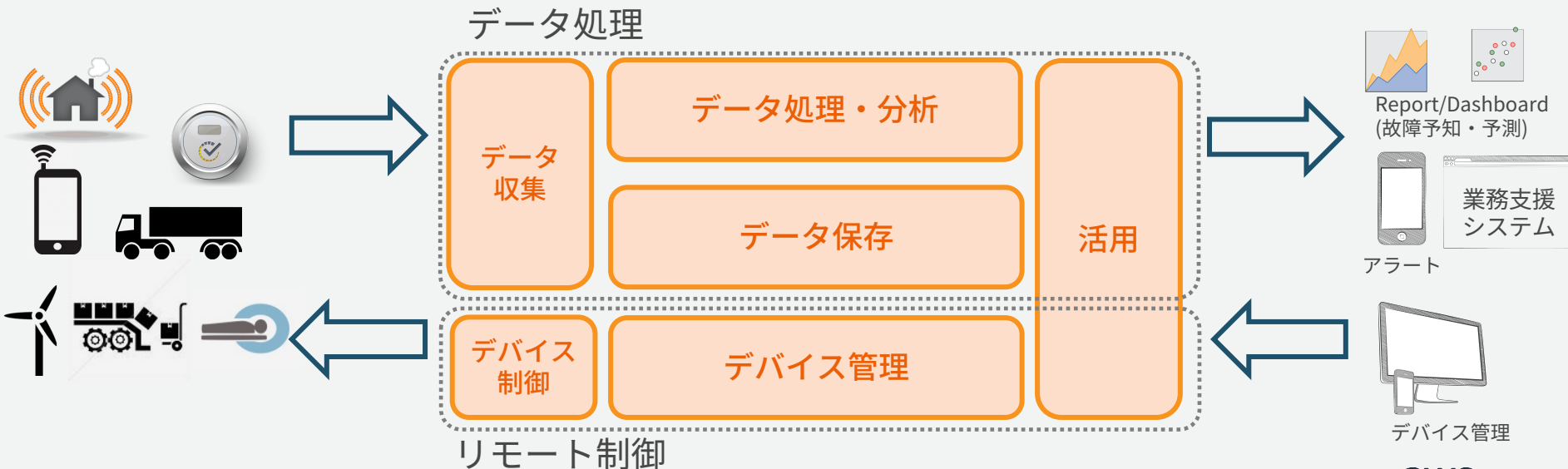
稼働実績・異常監視

作業効率・自動化

保守作業指示

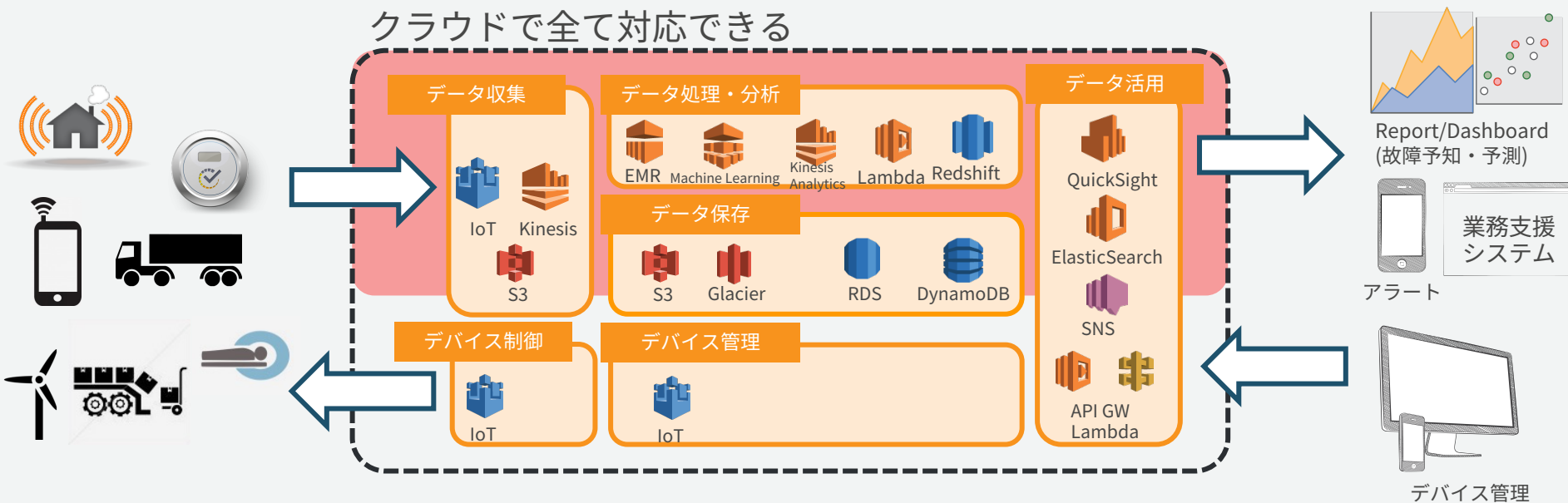
遠隔制御

機器運用・
ファームアップ



クラウドで構築するIoTプラットフォーム

ビルディングブロックのように組み合わせることで素早く構築が可能



データ処理の 4 つのステップ

収集

収集した
データを
リアルタイム
に基盤に転送

保存

データを
長期的に保存、
検索

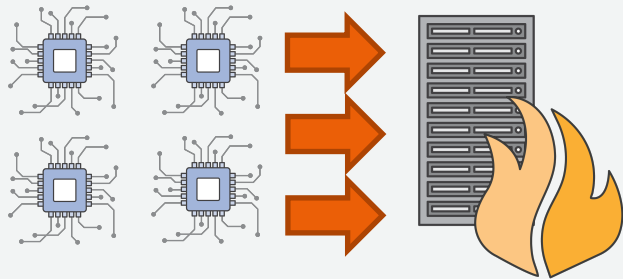
分析

大規模データ
を高速に分析
(解析)

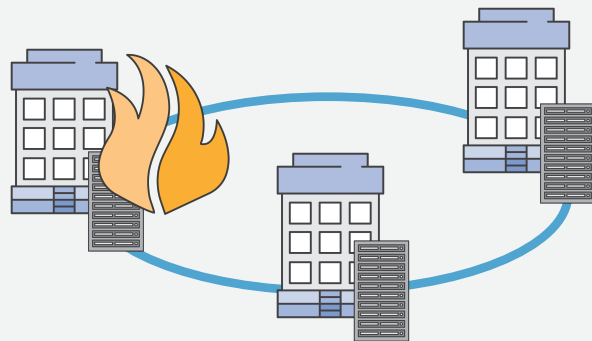
活用

分析結果を
人が参照
しやすい形で
提供

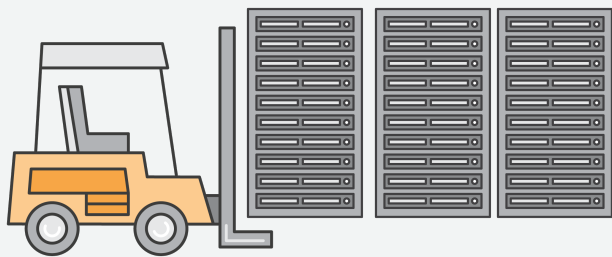
オンプレミスでのデータ収集の課題



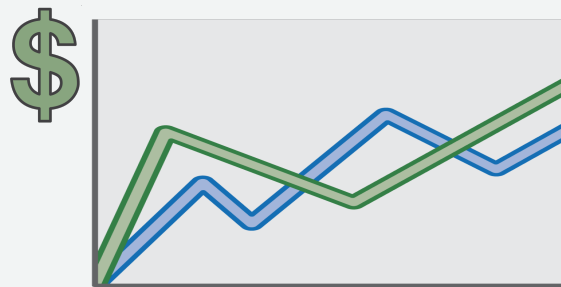
アクセス増加時の負荷分散



冗長化・可用性



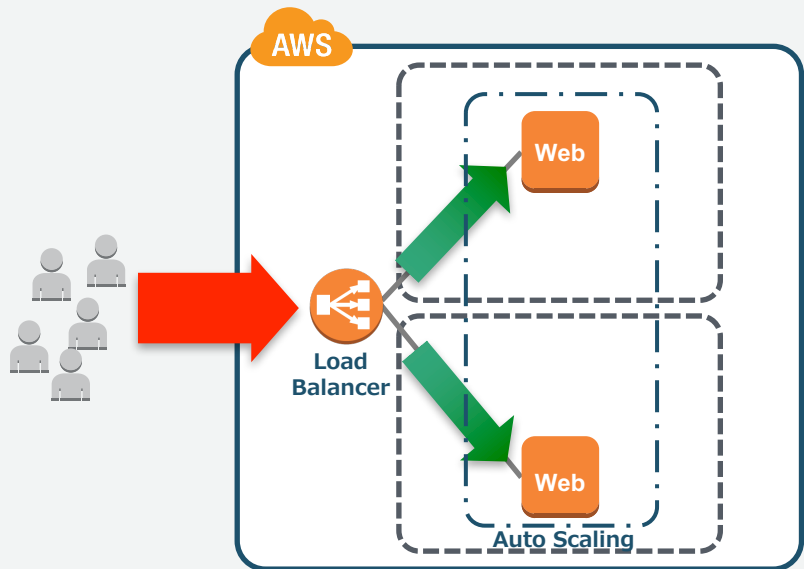
基盤の構築と運用



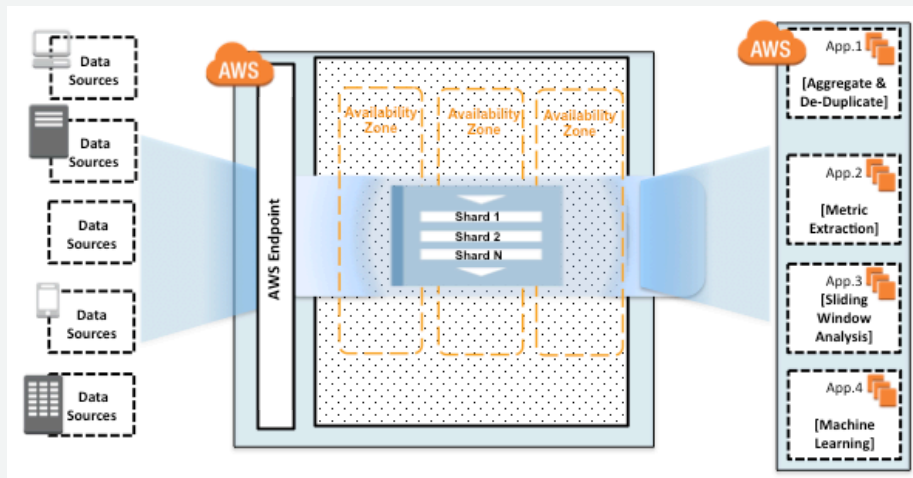
コストの増大

クラウドを利用したデータ収集

ロードバランシング
スケールアウト

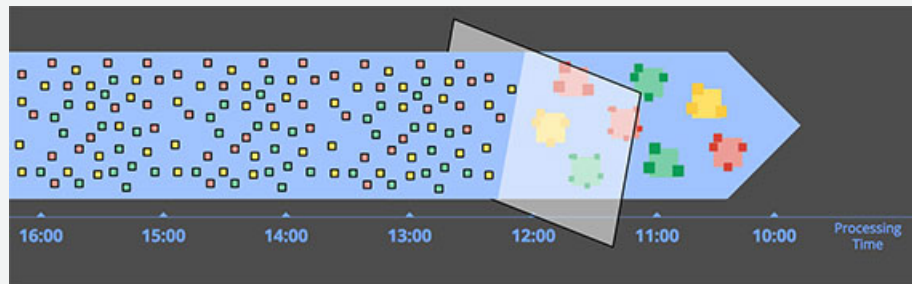
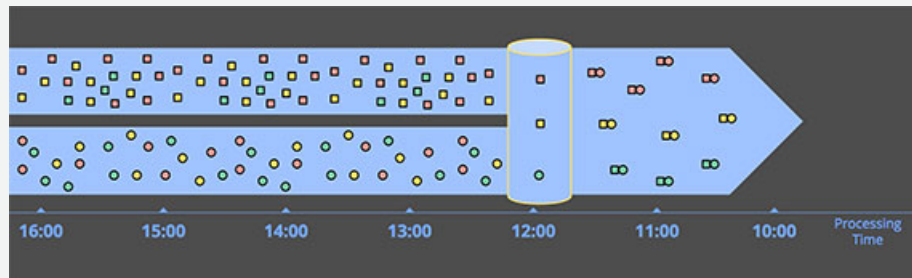
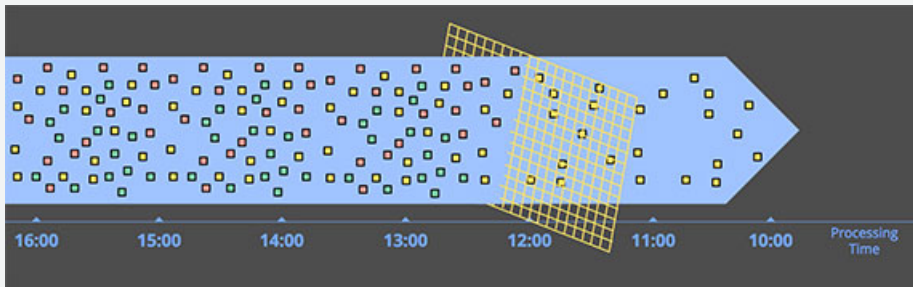


Kinesis などマネージドサービス
ストリーム処理



ストリーム処理

無限のデータを逐次処理して、永続的に処理をし続ける



引用元

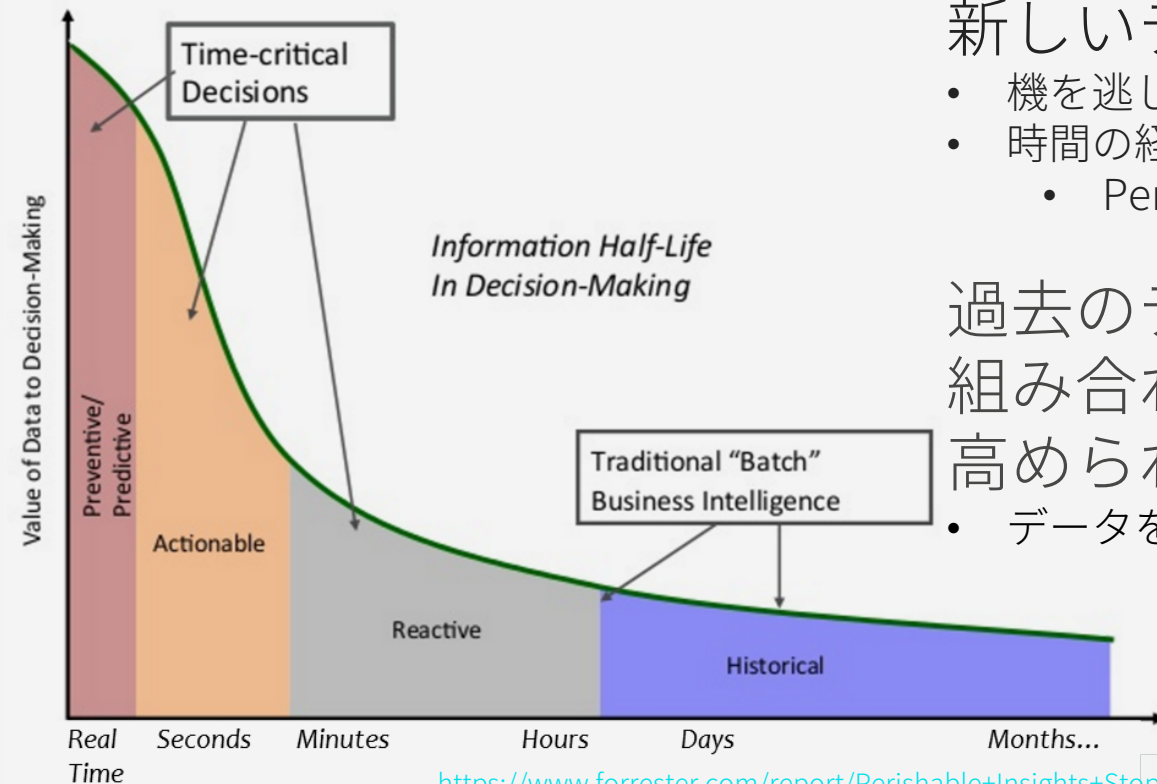
<https://www.oreilly.com/ideas/the-world-beyond-batch-streaming-101>

<https://www.oreilly.com/ideas/the-world-beyond-batch-streaming-102>

ストリーム処理の典型的なシナリオ

	STEP 1	STEP 2	STEP 3
シナリオ/ 業界別	取込-変換-ロードの加速	継続的なメトリクスの生成	機械学習や次のアクションにつながる洞察
広告 マーケティング	広告主、入札者情報の集約	カバレッジ、売り上げ、コンバージョンなどの広告メトリクス	ユーザーの広告エンゲージメント、入札エンジンの最適化
IoT	センサー、テレメトリデータの取込	運用のためのメトリクスやダッシュボード	デバイス運用の高機能化やアラートニング
ゲーム	オンラインデータ集約 例) トップ10プレイヤー	MMOGのライブダッシュボード	リーダーボードの生成やプレイヤースキルマッチ
個人向けWeb サービス	クリックストリーム分析	インプレッションやPVなどのメトリクス	レコメンドエンジンや積極的なユーザーケアプログラム
運用 セキュリティ	DevOpsツール、VPCフローログの取込	CloudWatch Logsをサブスクライブして、リアルタイムにログを分析する	異常検知

時間の経過とともに、データの価値は失われていく



新しいデータほど価値がある

- 機を逃してはならない
- 時間の経過とともに、洞察は陳腐化していく
 - Perishable Insights (Mike Gualtieri)

過去のデータと新しいデータを組み合わせれば、もっと価値を高められる

- データをつなぎ合わせる術を得る必要がある

データ処理の 4 つのステップ

収集

収集した
データを
リアルタイム
に基盤に転送

保存

データを
長期的に保存、
検索

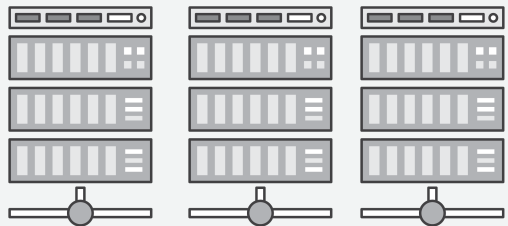
分析

大規模データ
を高速に分析
(解析)

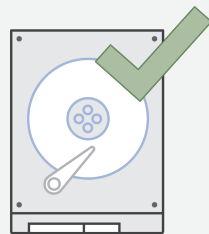
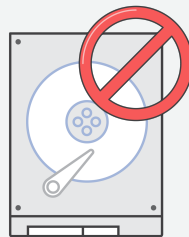
活用

分析結果を
人が参照
しやすい形で
提供

オンプレミスでのデータ保存の課題



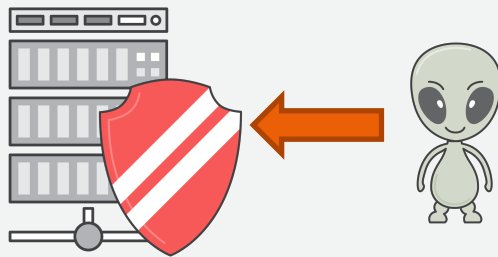
肥大化するデータ容量



堅牢性



分散ストレージの構築と運用



セキュリティ

クラウドを利用したデータ保存

- データを3つ以上のデータセンターに自動複製
- 99.999999999%の耐久性
- 1GBあたり月約3円・容量は無制限で保存した分だけ課金

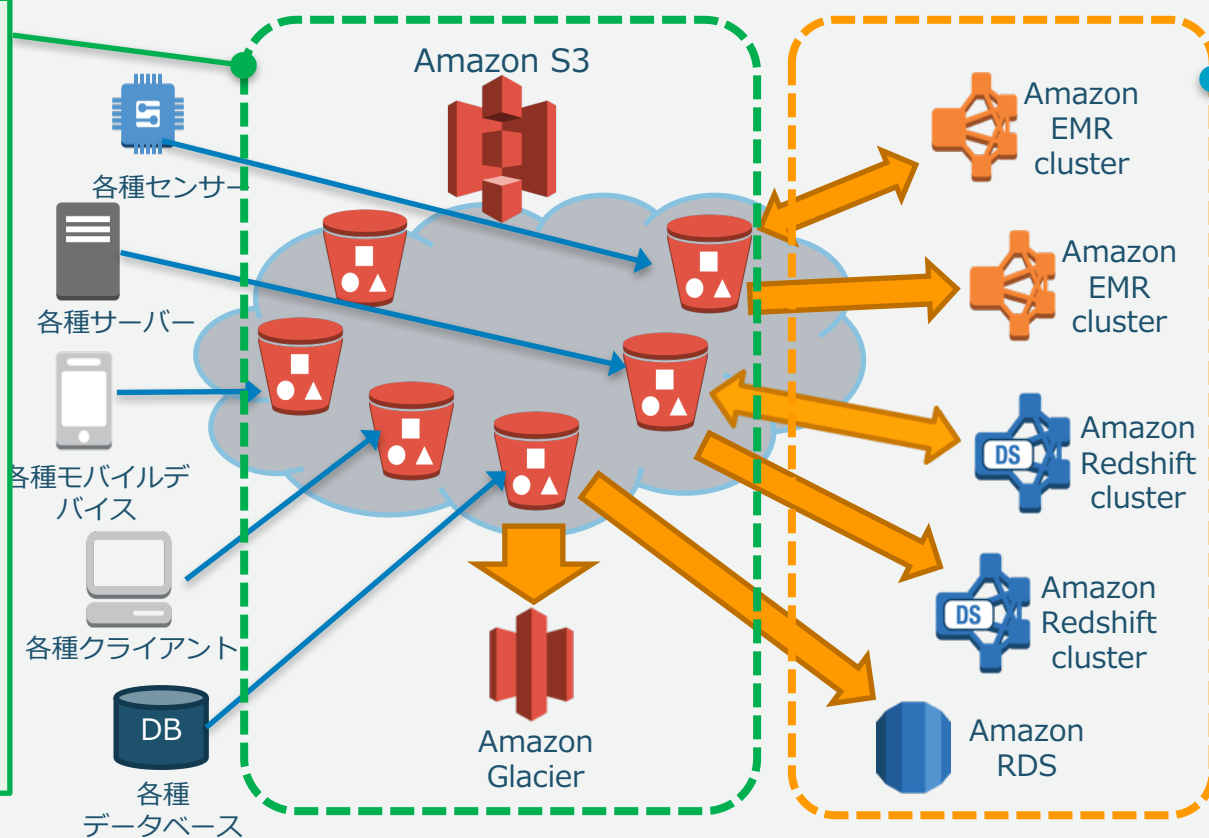


データレイクとしてのクラウドストレージ

恒久的な永続化の主体は、DWH や Hadoop (HDFS) 自体ではなく、S3 に一元化

必要な時に Redshift や EMR のクラスターを生成して、不要になれば削除するといったライトウェイトな運用が可能

S3にロードデータが保存されていれば分析用途別の後続での利用が容易に実現可能



従来のアーキテクチャでは、DWH や RDB、あるいは、HDFS 自体が恒久的なデータ蓄積・保存先となりがち

One Fact が様々な実体に複写されている状況

ストレージとデータ処理の結合度が高いため、ストレージ容量拡張や、CPUやメモリ容量拡張の際に手間が発生

データ処理の4つのステップ

収集

収集した
データを
リアルタイム
に基盤に転送

保存

データを
長期的に保存、
検索

分析

大規模データ
を高速に分析
(解析)

活用

分析結果を
人が参照
しやすい形で
提供

オンプレミスでのデータ分析の課題

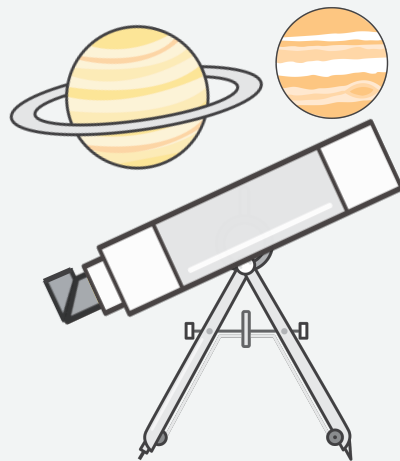
- データの特性によって分析方法は異なる
- それぞれの分析方法にあわせて基盤やツールが必要
→システムが複雑化、コストが肥大化



顕微鏡



虫めがね

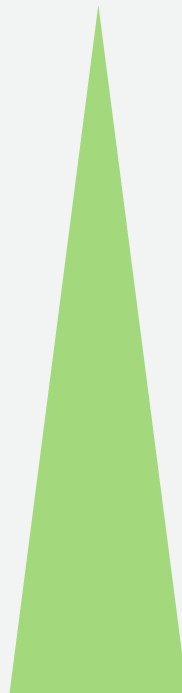


望遠鏡

データの特徴と分析方法

Hot Data	頻繁に読み書き メモリで処理 データサイズは小～中 短期間
Warm Data	頻繁ではない読み書き 常にメモリで処理しない データサイズは中～大 中期間
Cold Data	散発的な読み書き ディスクに保存 データサイズは大～ 長期間

サイズ



パフォーマンス



データの特徴と分析方法

Hot Data

リレーショナルデータベース
MySQL, PostgreSQL, Oracle, SQL Server

Warm Data

NoSQL データベース
memcached, redis, mongoDB, Hbase

Cold Data

データウェアハウス
DWH

Hadoop
MapReduce, Hive, Spark

ストリーム

大容量データストア
アーカイブ

データの特徴と分析方法

Hot Data

Warm Data

Cold Data

リレーショナルデータベース
MySQL, PostgreSQL, Oracle, SQL Server



Amazon
RDS

データウェアハウス
DWH



Amazon
Redshift

NoSQL データベース
memcached, redis, mongoDB, Hbase



Amazon
ElastiCache



Amazon
DynamoDB

Hadoop
MapReduce, Hive, Spark



Amazon
EMR



AWS Glue

ストリーム



Amazon
Kinesis

大容量データストア
アーカイブ



Amazon
S3



Amazon
Glacier



AWS Snowball*



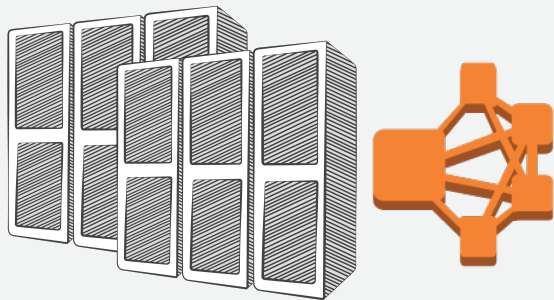
Amazon
Athena

Apache Hadoop (EMR)

分散処理基盤

分散処理に必要なリソースを簡単に調達/廃棄可能

- Master/Core/Task node
- EMRFS
- Bootstrap Action, Steps



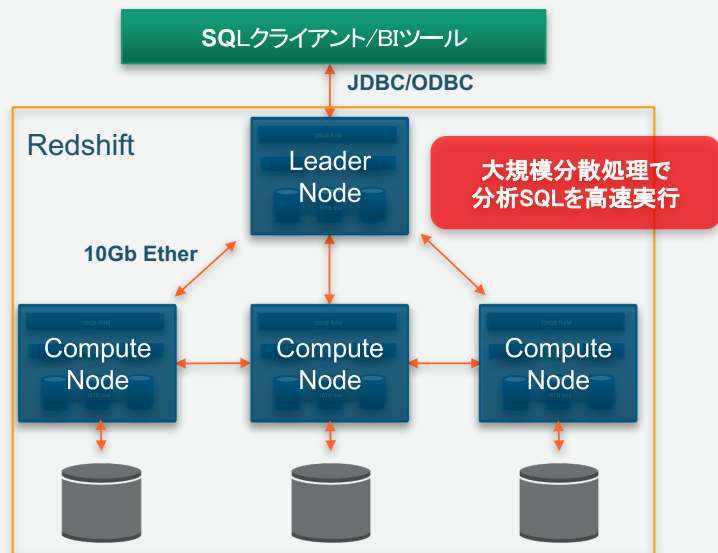
分散処理アプリケーション

複雑な設定無しに分散アプリケーションを利用可能

- Hadoop, YARN
- Hive, Pig, HBase, Spark..

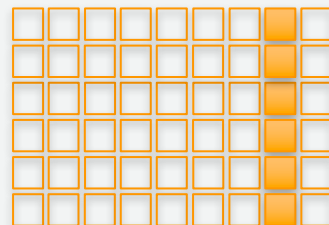


DWH データウェアハウス (Redshift)

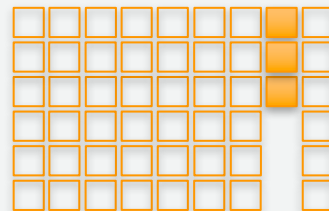


MPP : Massive Parallel Processing

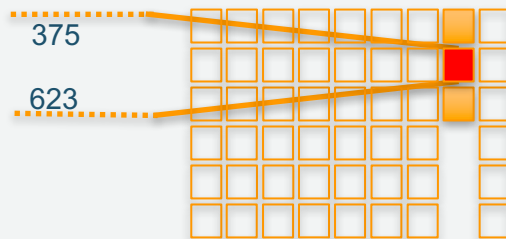
- 1つのタスクを複数のノードで分散して実行する仕組み
- Redshiftではリーダーノードがタスクをコンピューターノードに分散して実行
- ノードを追加する（スケールアウト）でパフォーマンス向上



列指向
カラムナー型
ストレージ



データ圧縮

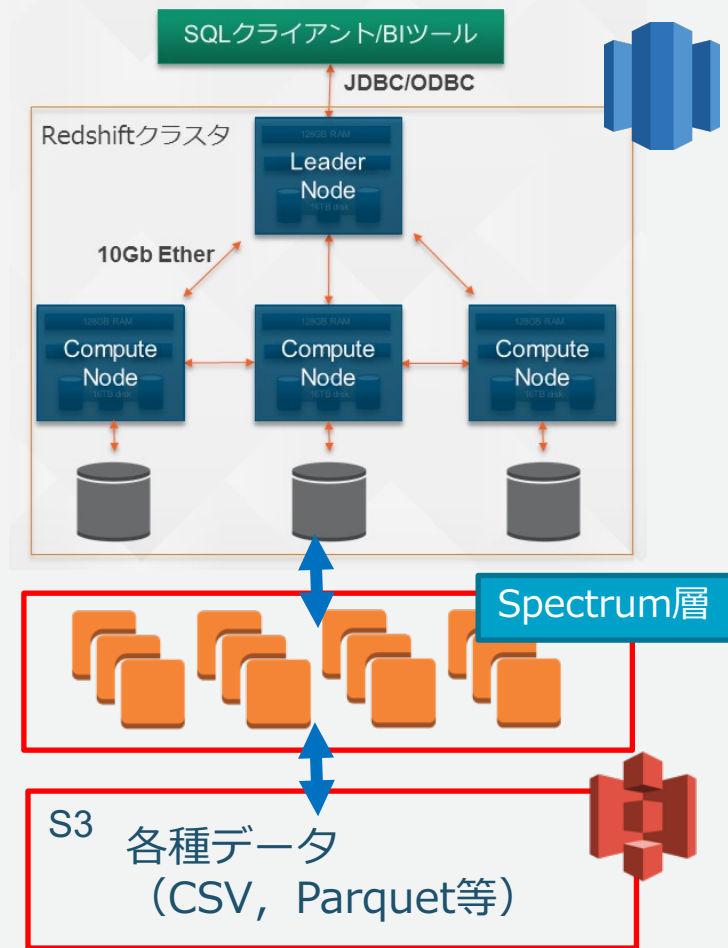


ゾーンマップ

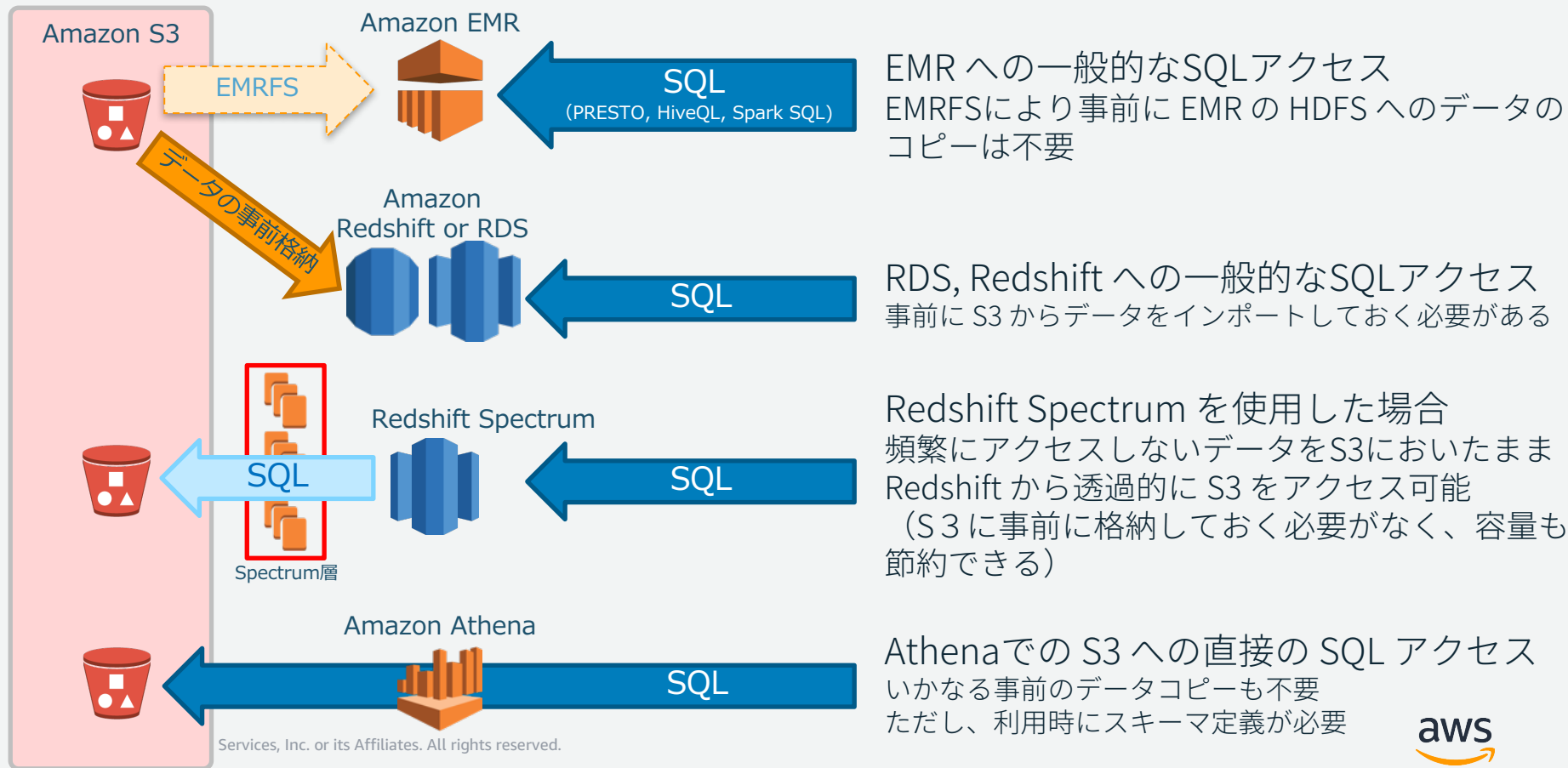
Redshift Spectrum

大規模スケールアウトの処理層（Spectrum層）を使い、S3上のデータに対してSQLを実行する機能を提供

- S3上に置いたファイルを外部テーブルとして定義し、Amazon Redshiftのクエリがそのまま活用できる
- データローディングにかかる手間、時間が不要
- 同じS3データセットに対し、色々なノードタイプ、サイズで作ったAmazon Redshiftクラスターからアクセス可能



アドホック・クエリーのバリエーション



EMR への一般的なSQLアクセス
EMRFSにより事前に EMR の HDFS へのデータの
コピーは不要

RDS, Redshift への一般的なSQLアクセス
事前に S3 からデータをインポートしておく必要がある

Redshift Spectrum を使用した場合
頻繁にアクセスしないデータをS3においたまま
Redshift から透過的に S3 をアクセス可能
(S3 に事前に格納しておく必要がなく、容量も
節約できる)

Athenaでの S3 への直接の SQL アクセス
いかなる事前のデータコピーも不要
ただし、利用時にスキーマ定義が必要

データ処理の 4 つのステップ

収集

収集した
データを
リアルタイム
に基盤に転送

保存

データを
長期的に保存、
検索

分析

大規模データ
を高速に分析
(解析)

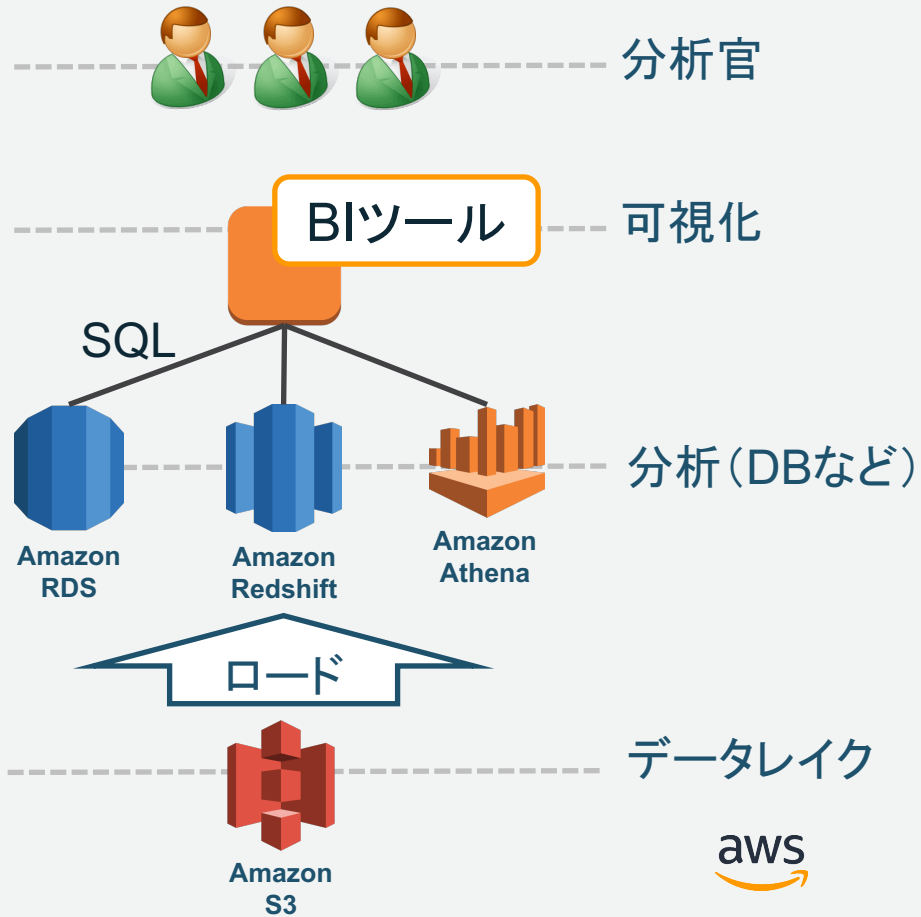
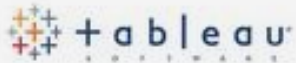
活用

分析結果を
人が参照
しやすい形で
提供

データ可視化は任意のツールが選択可能

一般的な BI ツール

多彩なパートナーソリューション・
OSSをEC2上で活用



クラウドベースの BI ツール (QuickSight)

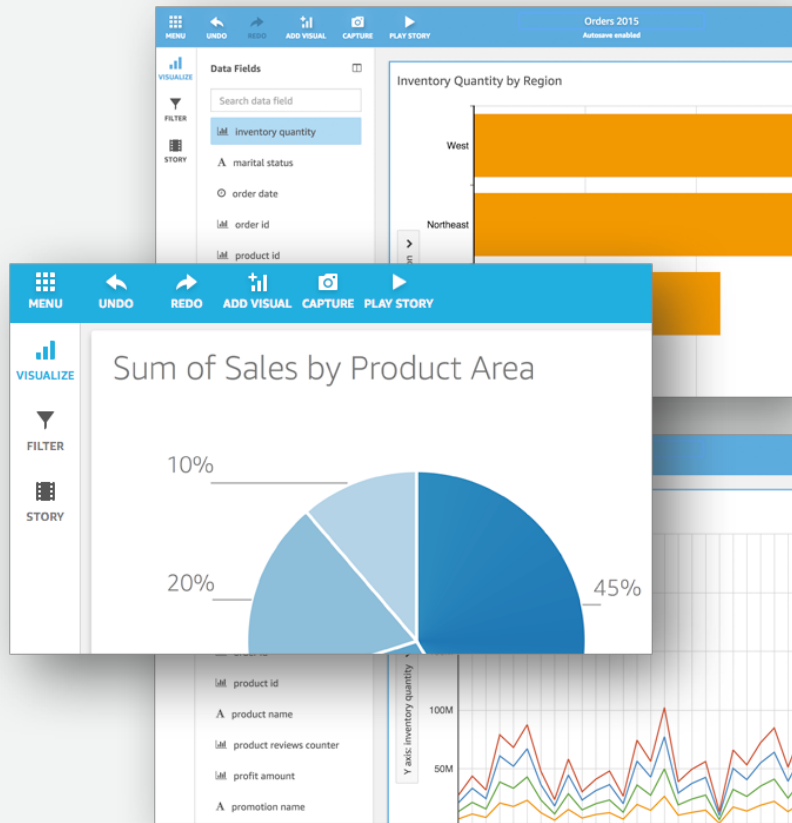
クラウドベースのBIツール

AWS内部のデータソースを自動探索し簡単に開始

直感的なビジュアル化機能(Auto Graph)を搭載

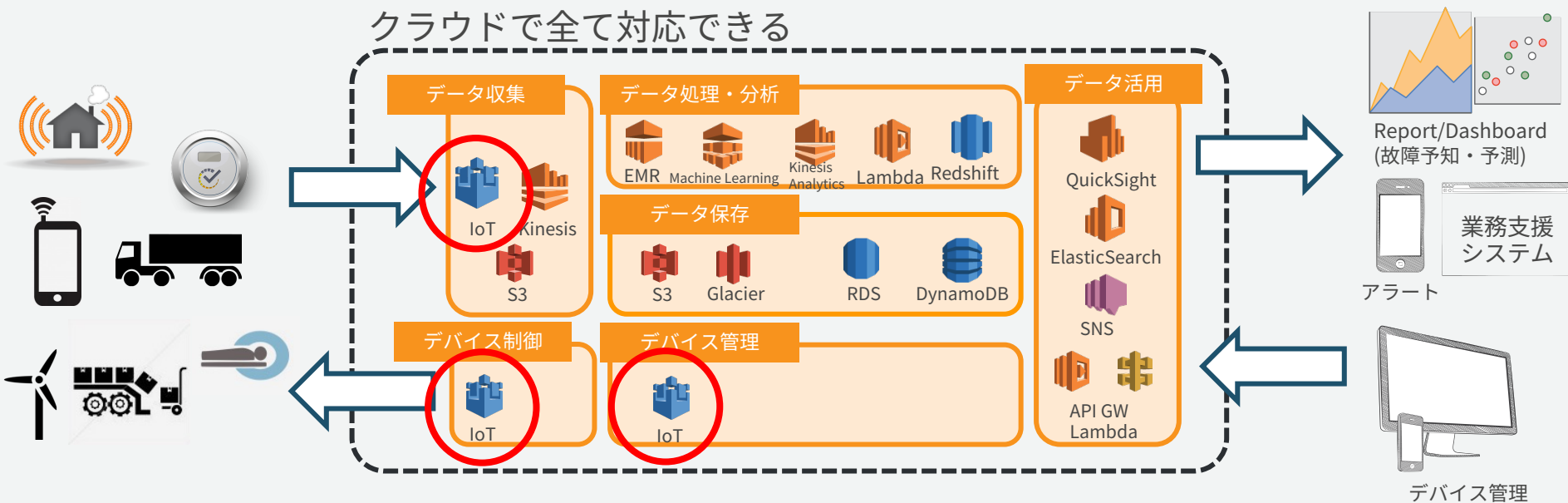
新規開発の SPICE エンジンを搭載し高速に動作

既存のパートナー製品との連携もサポート



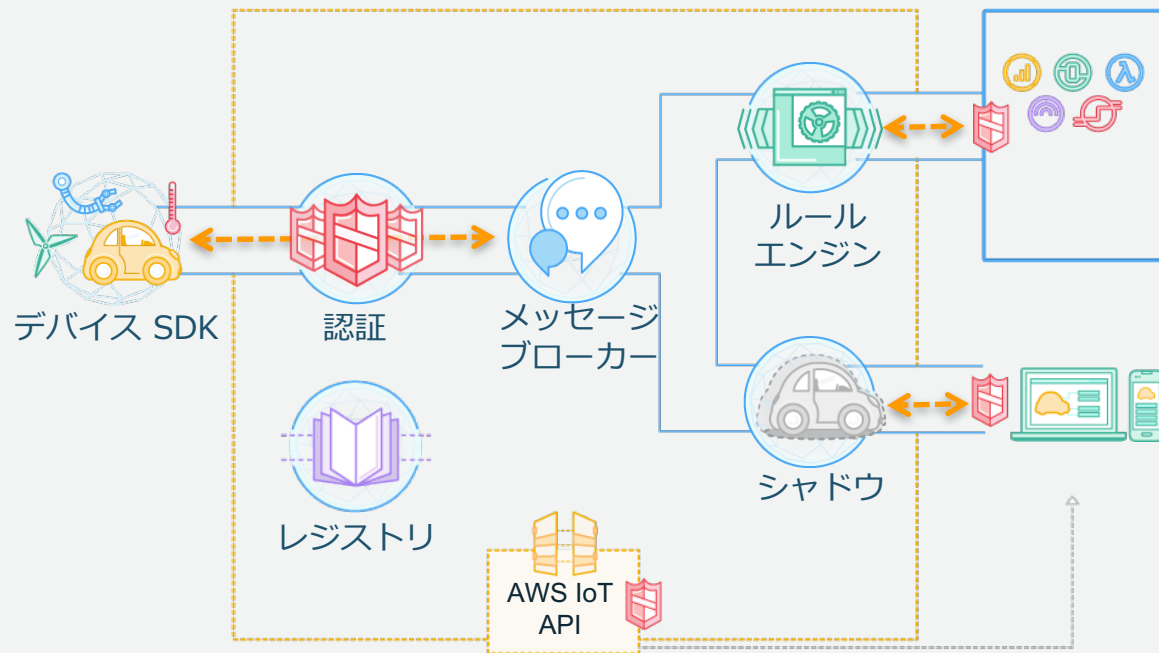
クラウドで構築するIoTプラットフォーム

ビルディングブロックのように組み合わせることで素早く構築が可能



AWS IoT

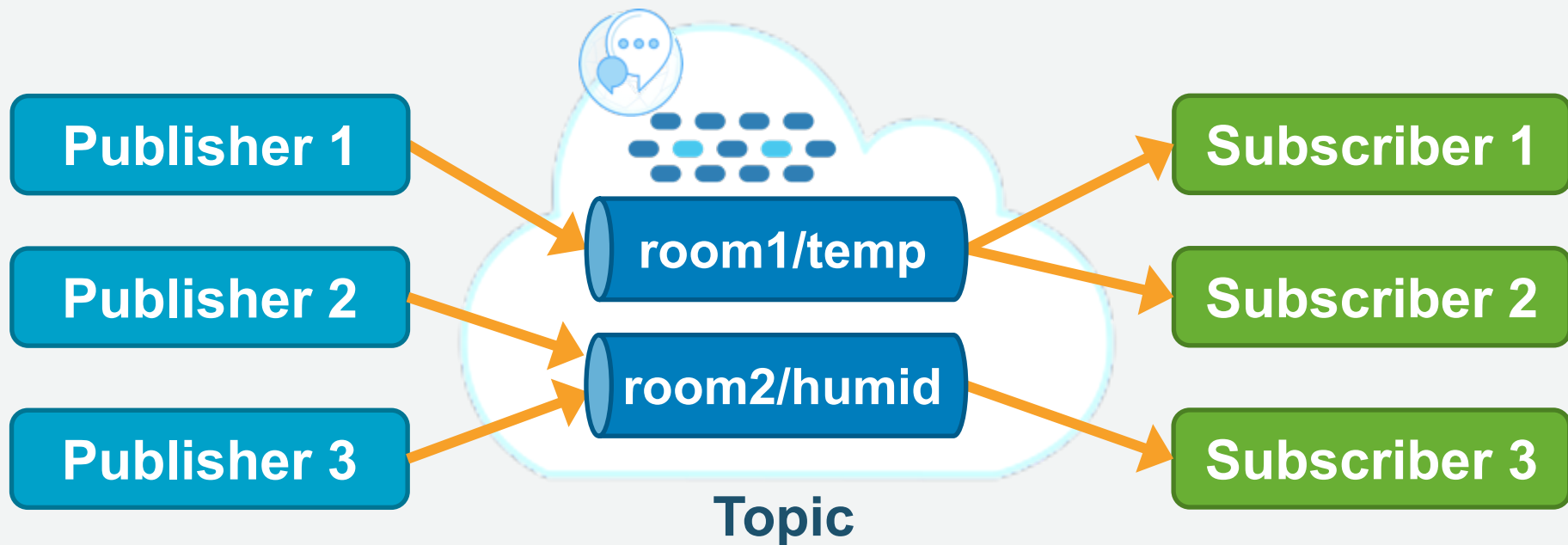
デバイスとクラウドの双方向コミュニケーションを実現する



- HTTPとMQTTに対応
- ハイランザクシオンに耐える高いスケーラビリティ
- ルールとアクションを定義可能
- 利用した分だけの支払い
- AWSサービスとのシームレスな連携

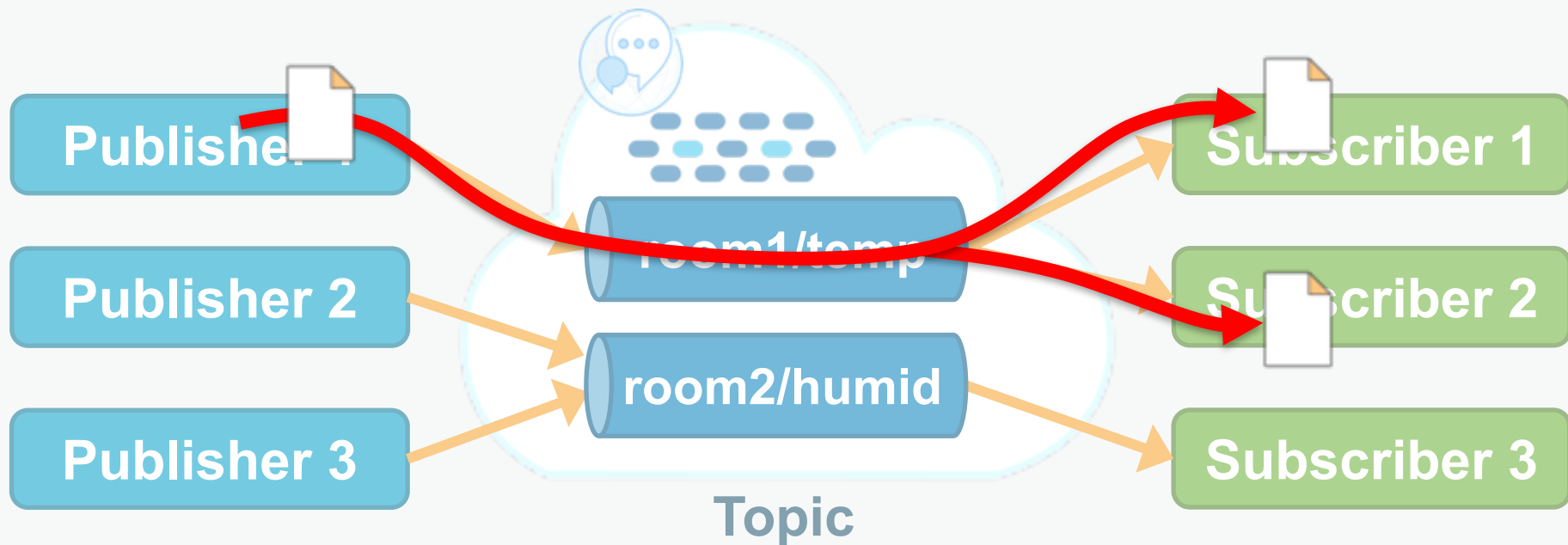
MQTT Pub-Sub メッセージ交換モデル

Topic を介してメッセージを双方向に送受信できる



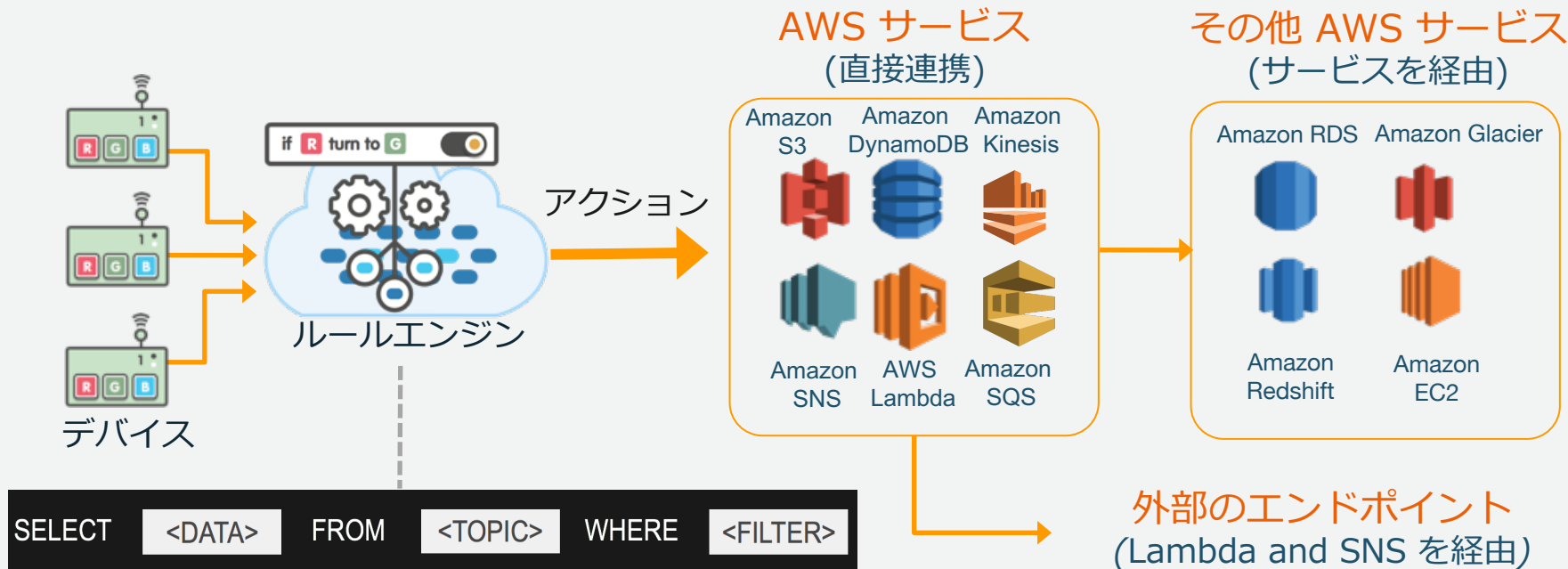
MQTT Pub-Sub メッセージ交換モデル

Topic を介してメッセージを双方向に送受信できる



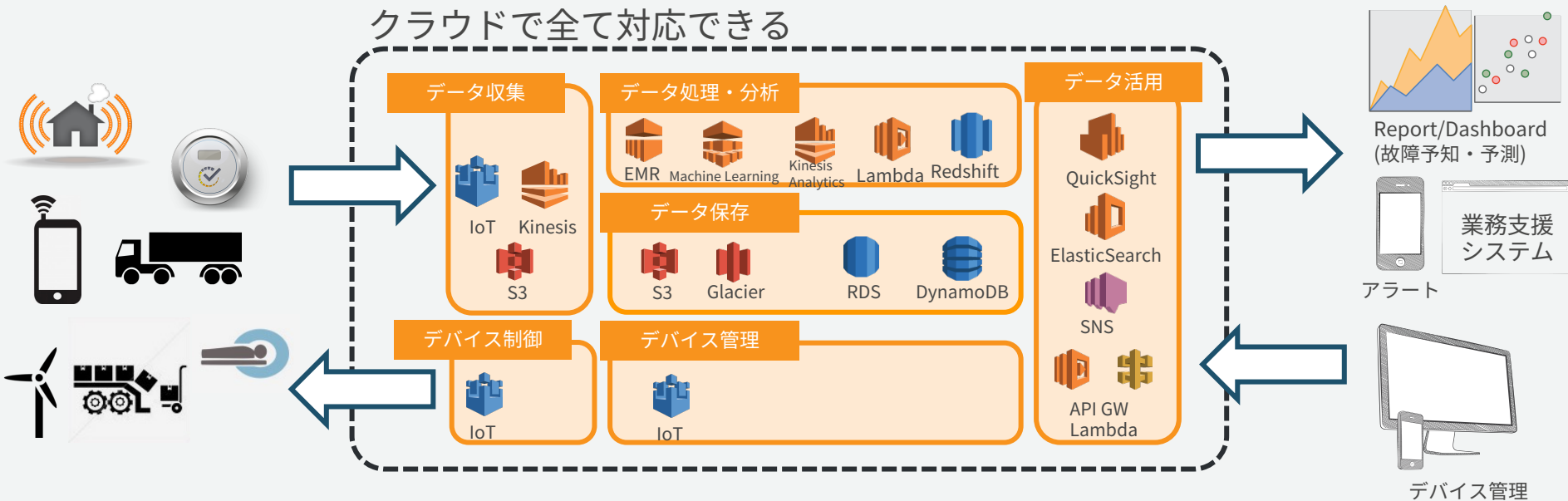
ルールエンジンとアクション

AWS IoT から AWS サービスや外部のエンドポイントの呼び出しを行なう



クラウドで構築するIoTプラットフォーム

ビルディングブロックのように組み合わせることで素早く構築が可能



参考資料

mcframe SIGNAL CHAIN

<https://www.mcframe.com/>

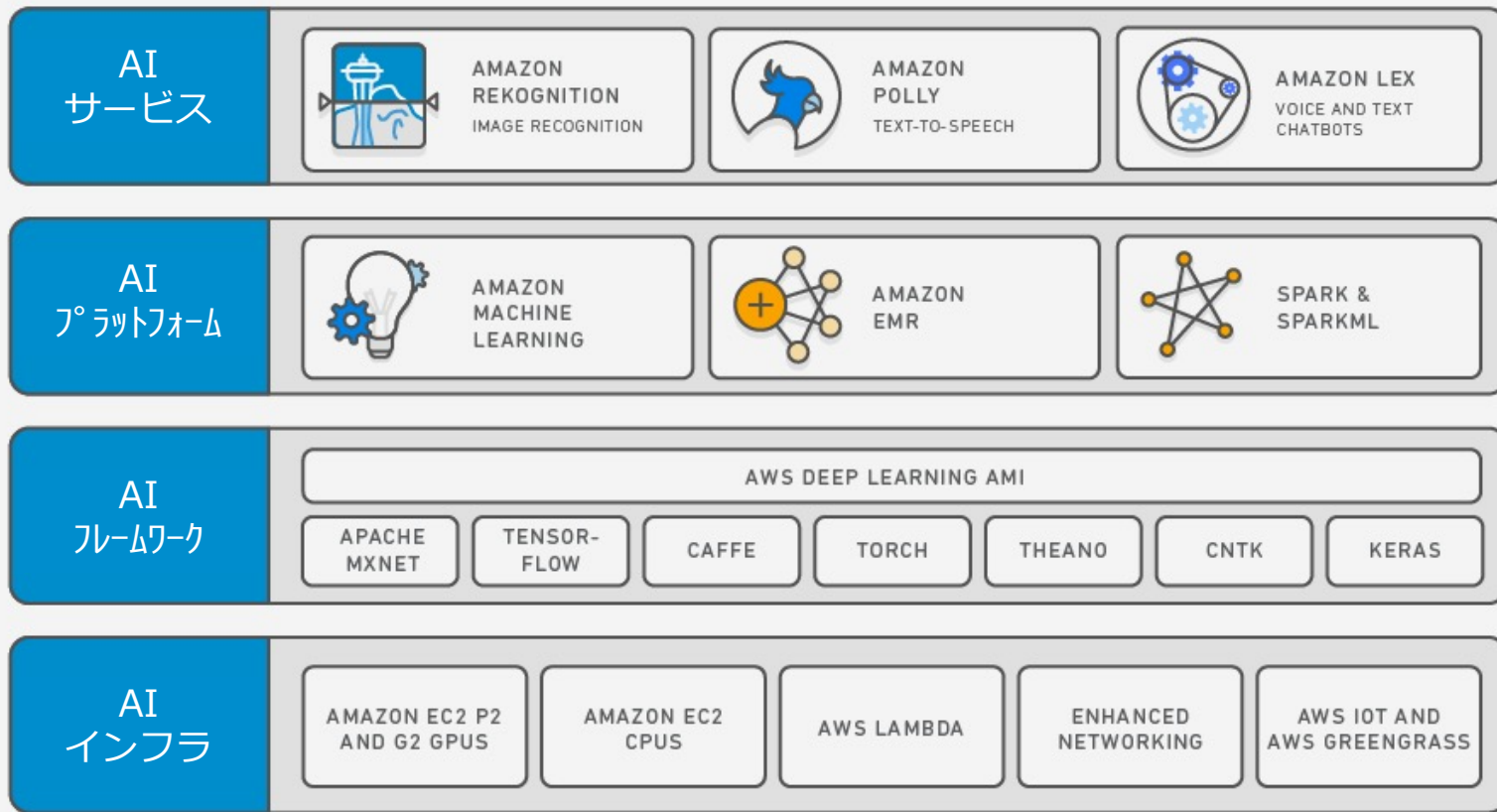
ブレインズテクノロジー株式会社 様 - Impulse

<https://www.brains-tech.co.jp/impulse/>

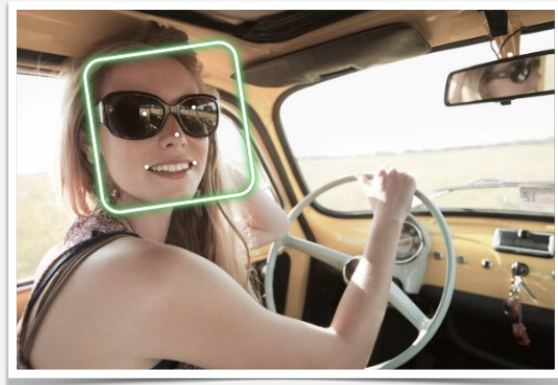
<https://www.brains-tech.co.jp/news/20170630-2.html>

AI・機械学習のトレンド

4層から構成される機械学習サービス

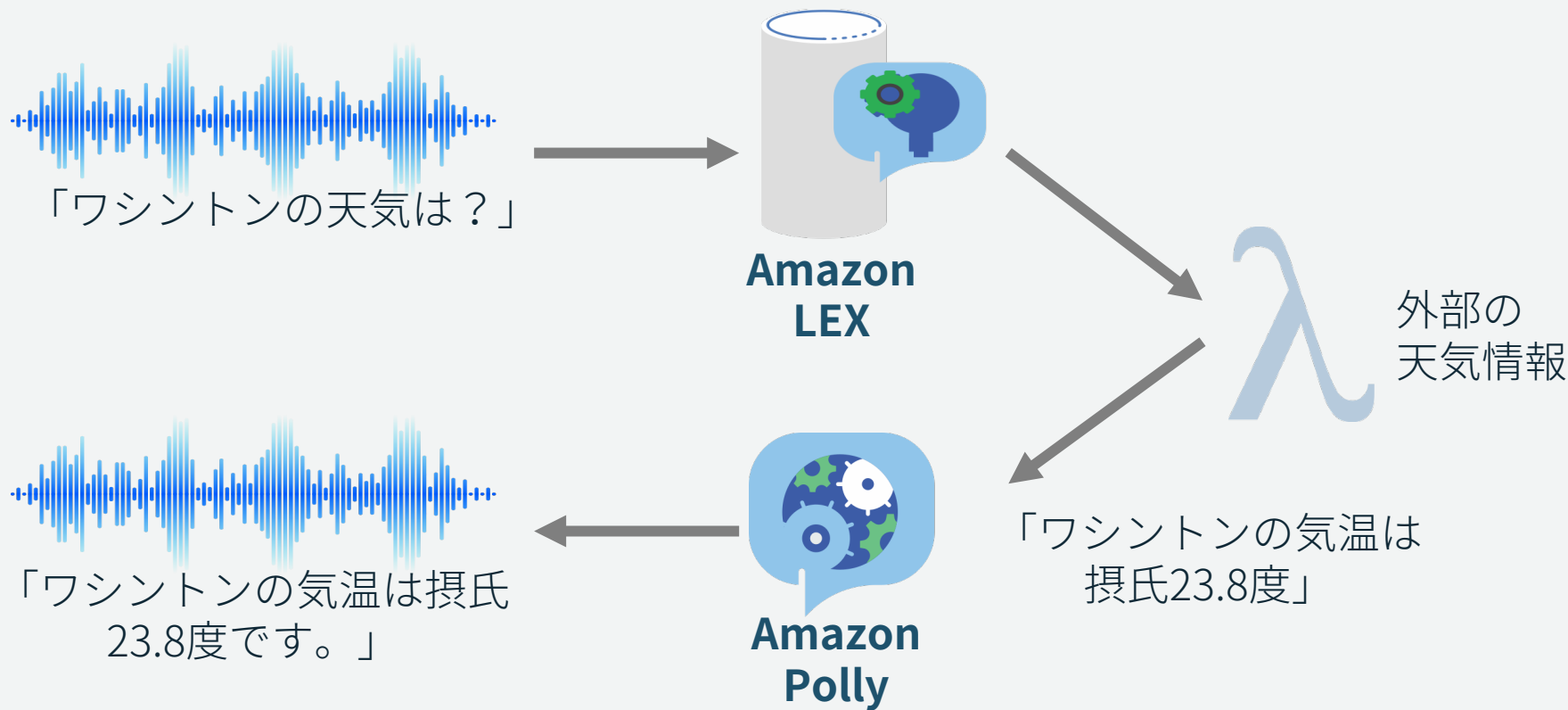


画像認識



物体と風景	顔
車	女性
屋外	笑顔
日中	サングラス
運転中	

音声言語処理



Apache Spark



高速な分散処理フレームワークとして

- メモリキャッシュ、なるべくディスクを使わない
- 繰り返し処理(機械学習等)で効果が高い

フレームワーク上で、更にアプリケーションが展開

- SQL,ストリーム,機械学習(MLlib),グラフ

MLlibで利用できるアルゴリズムの例

- 分類と回帰
 - 線形モデル (線形回帰、ロジスティック回帰、SVM)
 - 決定木
- 協調フィルタリング
- クラスタリング
 - k-means
- 次元削減
 - 特異値分解
 - 主成分分析

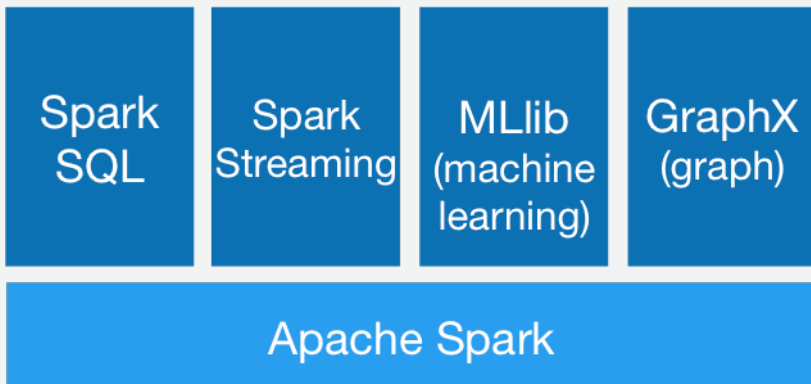
詳細はこちら

<http://mogle.web.fc2.com/spark/spark160/mllib-guide.html>

```
text_file = spark.textFile("hdfs://...")

text_file.flatMap(lambda line: line.split())
            .map(lambda word: (word, 1))
            .reduceByKey(lambda a, b: a+b)
```

Word count in Spark's Python API



<http://spark.apache.org/>

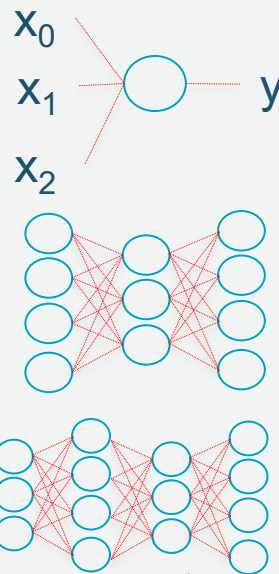
ニューラルネットと Deep Learning

ニューラルネットの栄枯盛衰

- 人間の神経組織(ニューロン)をモデル化した学習器
- 十分なノード数の中間層があれば、3層ニューラルネットであらゆる関数を近似できる
- 誤差逆伝播法の確立
- 様々な要素によって一時的に衰退
 - 次元の呪いによる計算量爆発
 - 誤差消失問題と局所収束問題
 - SVM等局所収束の無い学習器の登場

Deep Learningの隆盛

- CNN、DBM、DBN等4層以上でも学習が可能なアルゴリズムの登場
- 画像認識において顕著な実績
- 自然言語処理への応用
- 計算機能力の大幅な向上+GPUの活用
- クラウドとビッグデータ
- 豊富なライブラリ



パーセプトロン
多層NN

誤差逆伝播法

SVM等他のアルゴリズムが隆盛

ボルツマンマシンを用いた
深層自己符号化器

YouTubeの動画から猫の画像の認識

ILSVRCにおいてCNNで大差での優勝

1960

1980

2006

2012



Deep Learning AMI で各種パッケージを手軽に利用

OS はAmazon Linux と Ubuntu の2種類、各種フレームワークがプリインストールされる

MXNet v0.9.3 tag

Tensorflow v1.0.1 tag

Theano rel-0.8.2 tag


Caffe rc5 tag

Caffe2 v0.6.0 tag

CNTK v2.0 beta 12.0 tag

Torch master branch

Keras 2.0.2 tag



Deep Learning AMI Amazon Linux Version

Sold by: Amazon Web Services

The Deep Learning AMI is an Amazon Linux image supported and maintained by Amazon Web Services for use on Amazon Elastic Compute Cloud (Amazon EC2). It is designed to provide a stable, secure, and high performance execution environment for deep learning applications running on Amazon EC2. It includes popular deep learning frameworks, including MXNet, Caffe, Tensorflow, Theano, CNTK and Torch as well as packages that enable easy integration with AWS, including launch configuration tools and many popular AWS libraries and tools. It also includes the Anaconda Data Science Platform for Python2... [Read more](#)

Customer Rating	★★★★★ (11 Customer Reviews)
Latest Version	2.0 (Other available versions)
Operating System	Linux/Unix, Amazon Linux 2015.05.04
Delivery Method	64-bit Amazon Machine Image (AMI) (Read more)
Support	See details below
AWS Services Required	Amazon EC2, Amazon EBS
Highlights	<ul style="list-style-type: none">New in Version 2.0 - CPU Instance Type Support, MXNet built with MKL support.6 Deep Learning Frameworks - contains the most popular Deep Learning Frameworks (MXNet, Caffe, Tensorflow, Theano, Torch and CNTK) all prebuilt and pre-installed.Pre-installed components to speed productivity include Nvidia drivers, CUDA, cuDNN, Anaconda, Python2 and Python3.

Product Description

The Deep Learning AMI is an Amazon Linux image supported and maintained by Amazon

[Continue](#)

You will have an opportunity to review your order before launching or being charged.

Pricing Information

Use the Region dropdown selector to see software and infrastructure pricing information for the chosen AWS region.

For Region

US East (N. Virginia)

Free Tier Eligible EC2 charges for Micro instances are free for up to **750 hours** a month if you qualify for the **AWS Free Tier**.

Additional Taxes May Apply

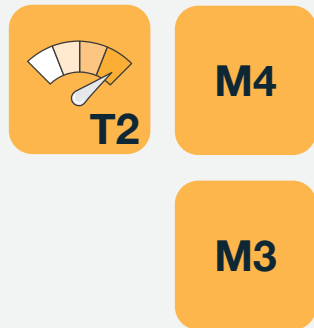
Pricing Details

Software pricing is based on your chosen options, such as subscription term and AWS region. Infrastructure prices are estimates only. Final prices will be calculated according to actual usage and reflected on your monthly report.

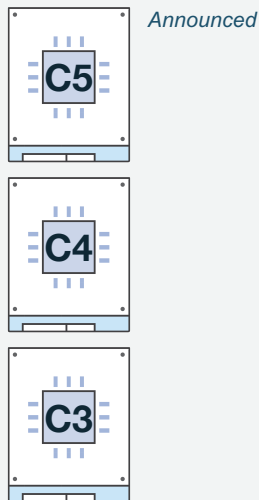
https://aws.amazon.com/marketplace/pp/B01M0AXXQB?qid=1475211685369&sr=0-1&ref=srh_res_product_title

EC2 コンピュートインスタンスタイプ

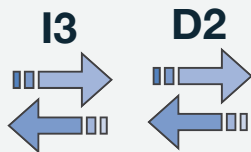
汎用



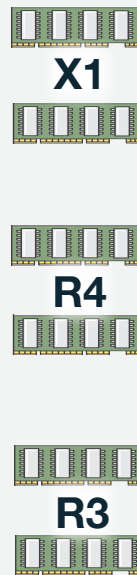
コンピュート
最適



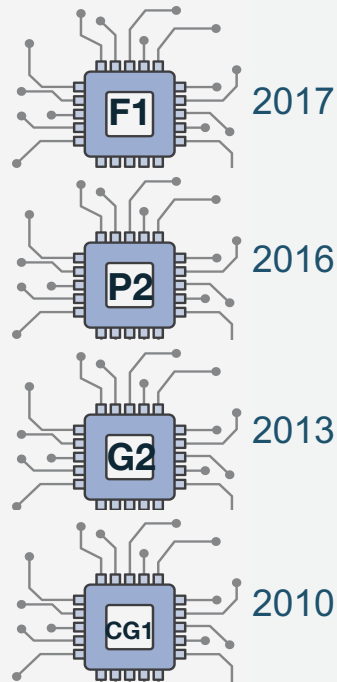
ストレージ
I/O 最適



メモリ最適



GPU / FPGA
アクセラレーテッド



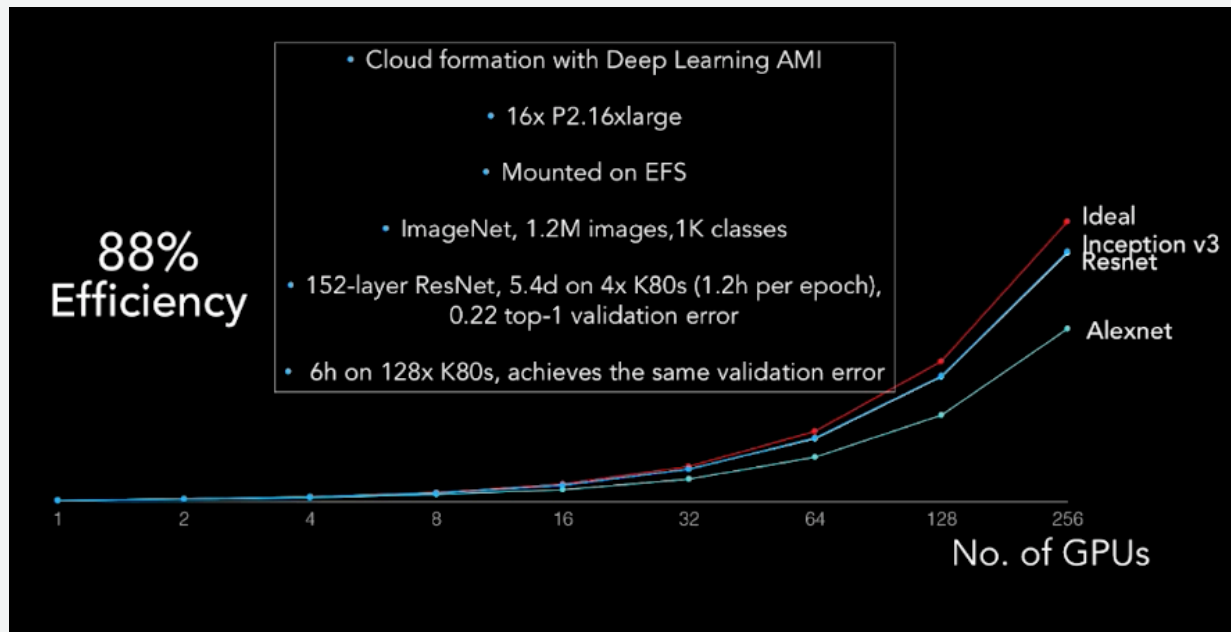
MXNet Deep Learning on GPU

P2 GPU インスタンスでDL学習と推論において高速な性能を実現

P2インスタンスでのMXNet学習処理：

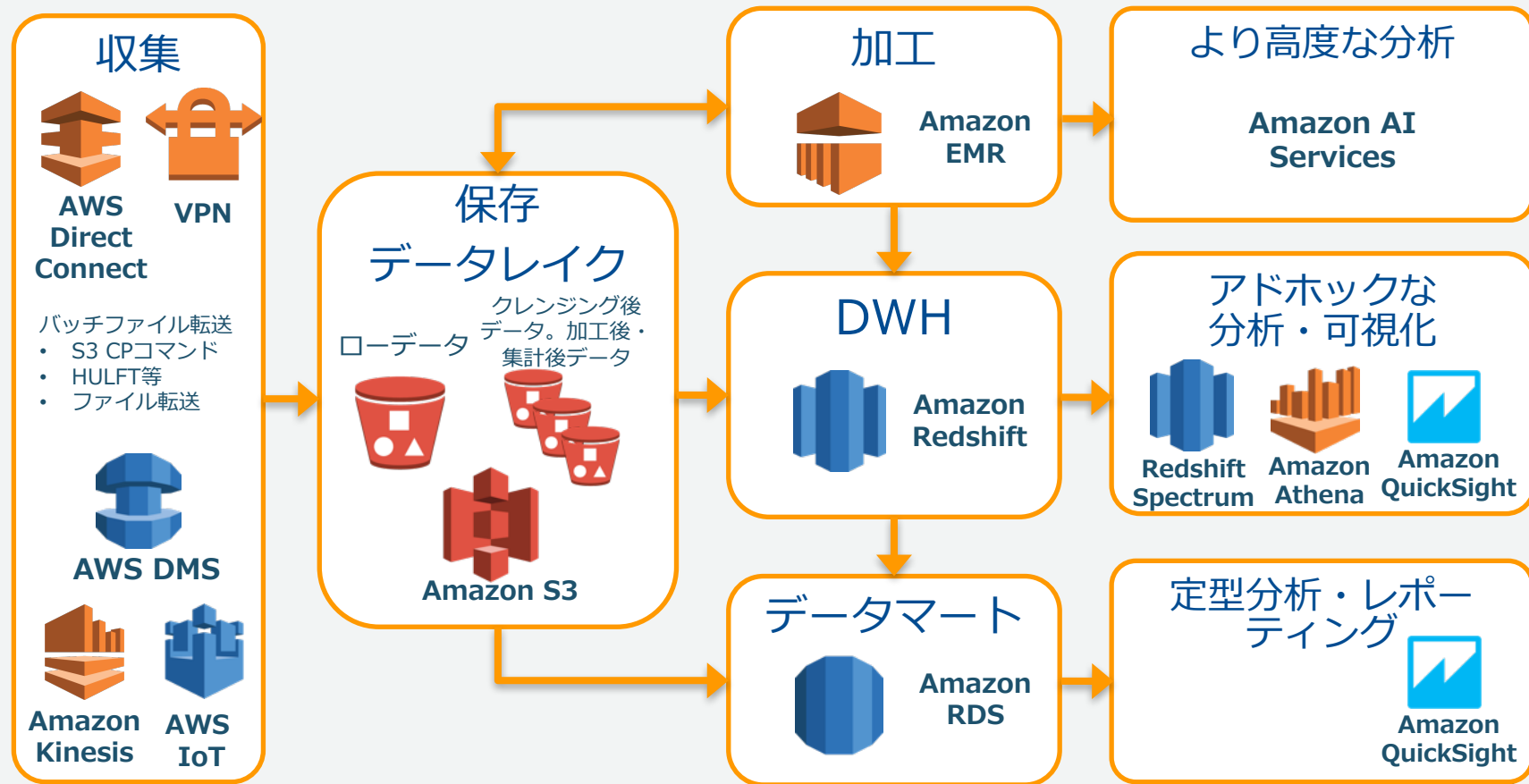
画像解析アルゴリズムInception v3を
MXNetで実装しP2で実行した結果

P2.16xlargeを16台で構成したクラス
ター
計256GPUまで並列に実行した場合、
85%の実行効率で性能向上



まとめ

各種サービスのマッピング例



まとめ

- クラウドの特徴
 - 従量課金、高い可用性と拡張性、コスト最適化
 - 高いセキュリティレベル、豊富なサービス群
- データ処理
 - 収集、保存、分析、活用（4つのステップ）
 - それぞれにマッチしたクラウドサービスがある
 - ビルディングブロックのように組み合わせてスピード開発
 - 豊富な事例
- トレンド
 - AI、機械学習

Thank you