INNOVATE

DATA EDITION

AUGUST 19, 2021

T 2 - 5

Log Analytics (Elasticsearch, Kinesis)
アプリログもインフラログも
IoT センサーデータも分析して可視化できる！
AWS のログ分析・可視化ソリューション

小林 航
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

内容についての注意点

- 本資料では2021年6月時点でのサービス内容および価格に基づいたスライドや説明になっています。最新の情報はAWS公式ウェブサイト(http://aws.amazon.com)にてご確認ください。
- 資料作成には十分注意しておりますが、資料内の価格とAWS公式ウェブサイト記載の価格に相違があった場合、AWS公式ウェブサイトの価格を優先とさせていただきます。

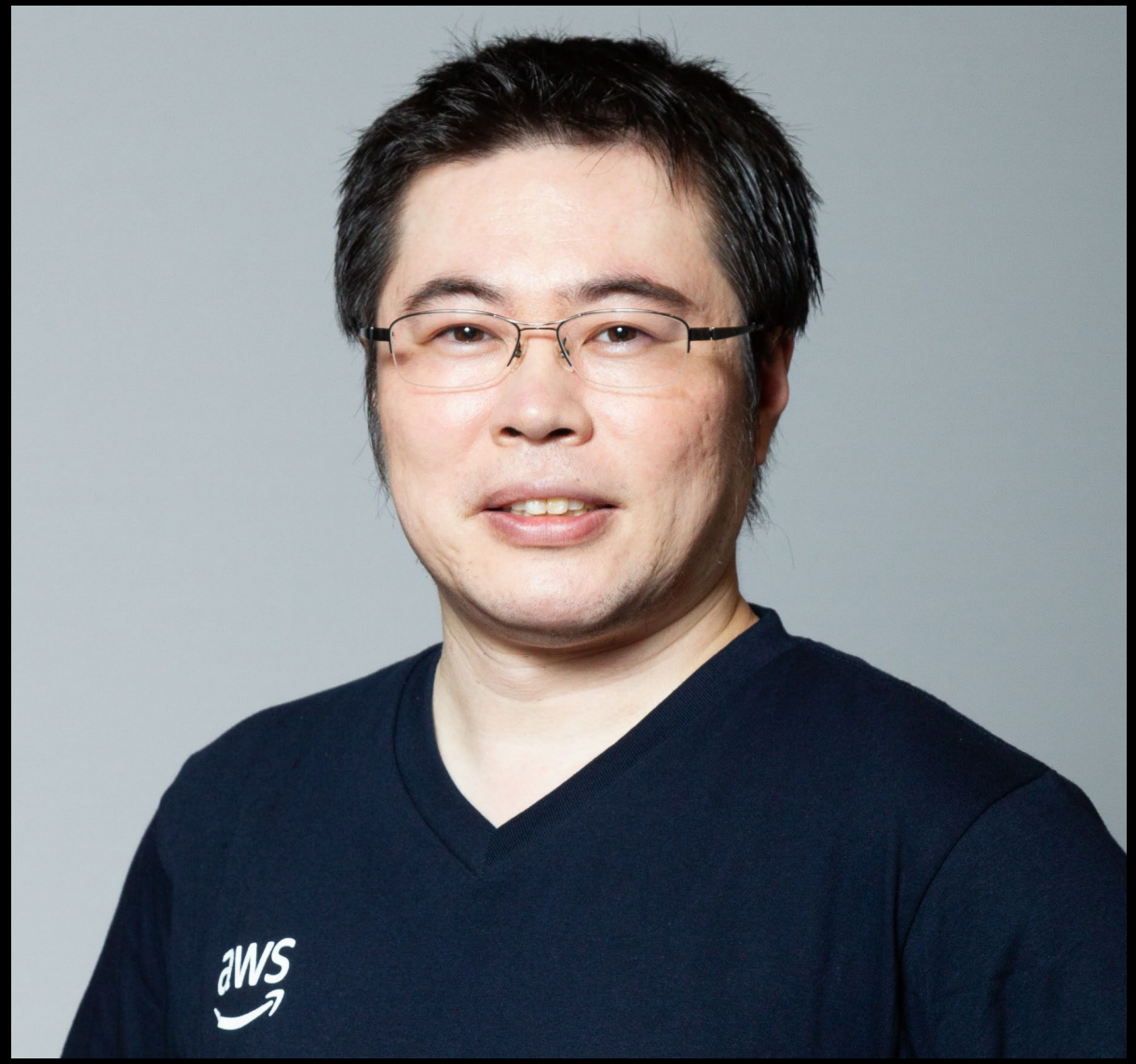
- AWS does not offer binding price quotes. AWS pricing is publicly available and is subject to change in accordance with the AWS Customer Agreement available at http://aws.amazon.com/agreement/. Any pricing information included in this document is provided only as an estimate of usage charges for AWS services based on certain information that you have provided. Monthly charges will be based on your actual use of AWS services, and may vary from the estimates provided.

自己紹介

小林 航

ソリューションアーキテクト

製造業のお客さまを担当
好きなサービス
- Amazon Elasticsearch Service
- Amazon Redshift
- Amazon RDS



アジェンダ

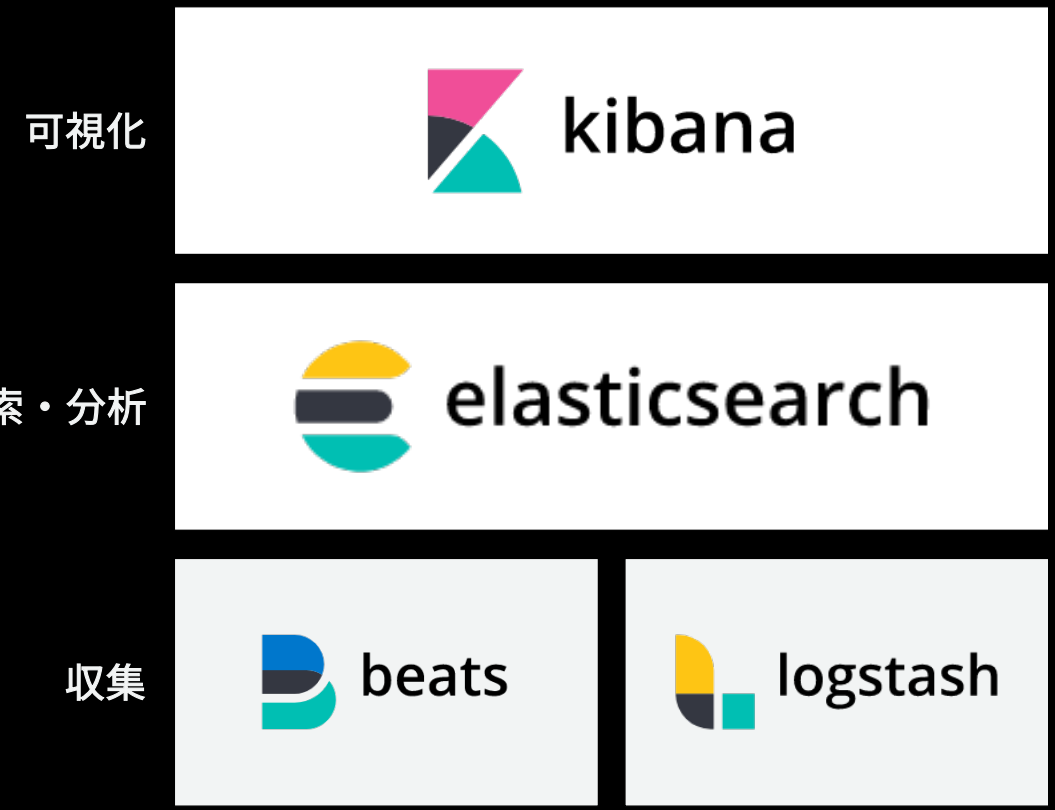
ログ分析の価値と Amazon Elasticsearch Service

Amazon Elasticsearch Service で実現するログ分析アーキテクチャ

SIEM on Amazon Elasticsearch Service のご紹介

Amazon Elasticsearch Service は Amazon OpenSearch Service にリブランド予定 (2021 夏)

- Elasticsearch は様々なログ分析のユースケースを解決する分散型 RESTful検索/分析エンジン
- Elasticsearch 準拠の Amazon Elasticsearch Service はv7.10までで、それ以降は Amazon OpenSearch Service にリブランドされ提供される
- OpenSearch は、Elasticsearch/Kibana 最後の Apache License 2.0リリースであるv7.10からのコミュニティ主導のオープンソースフォーク
- Elasticsearch バージョンから OpenSearch バージョンへのアップグレードはシームレスに実施される



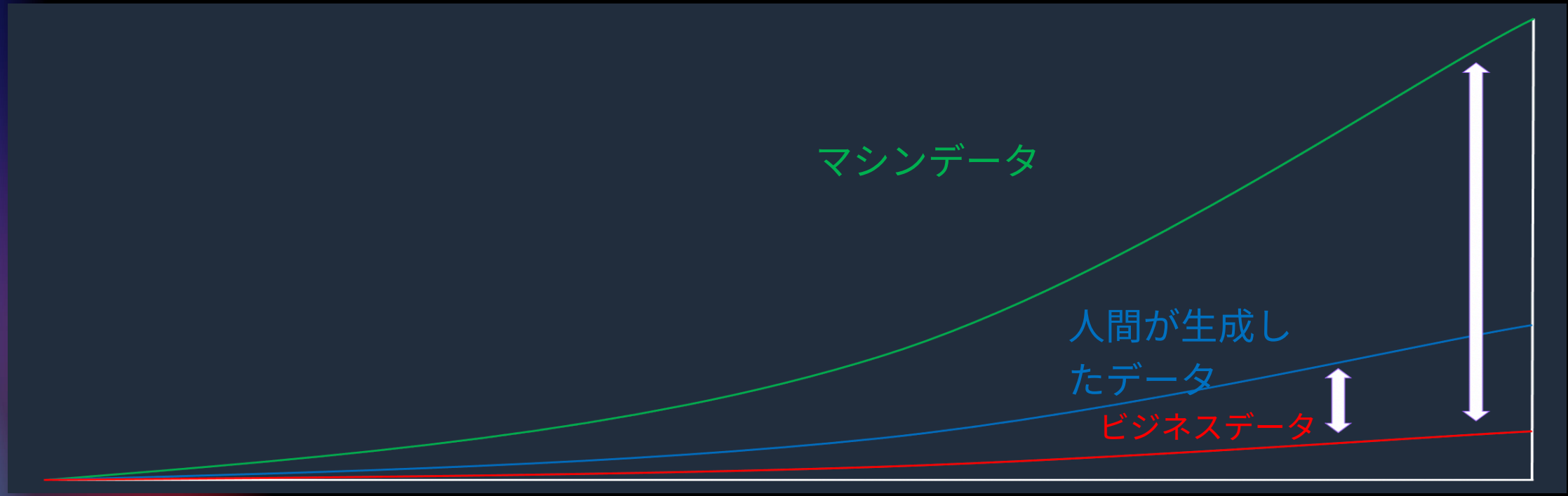
https://aws.amazon.com/jp/blogs/opensource/introducing-opensearch/

https://aws.amazon.com/jp/elasticsearch-service/the-elk-stack/what-is-opensearch/

ログ分析の価値と Amazon Elasticsearch Service

データの指数関数的成長

“人間とマシン生成されたデータは、従来のビジネスデータよりも全体で10倍速い成長率を経験しており、マシンデータは50倍の成長率でさらに急速に増加 “



Source: insideBigData—The Exponential Growth of Data, February 16, 2017

マシンデータ、ログ分析の価値

IoT、モバイル

- コネクティッドカー
- スマートホームデバイス
- センサーデバイス
- モバイルアプリケーション

IT、DevOps

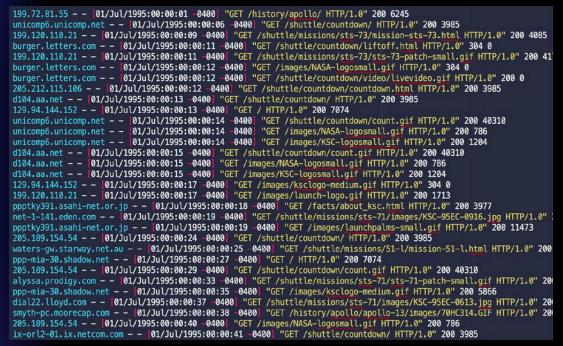
- データベース
- ロードバランサー
- ネットワーク
- サーバー

アプリケーションデータ

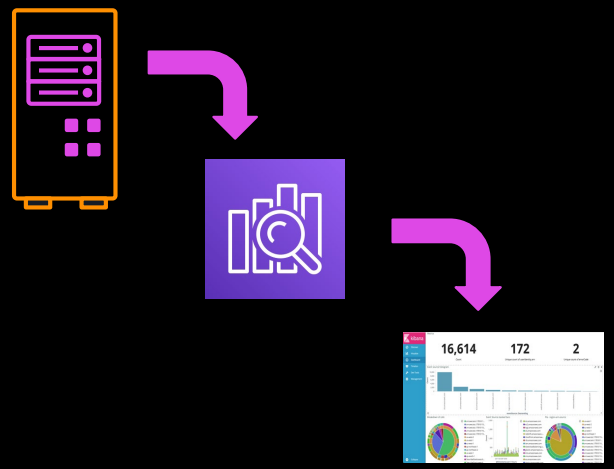
- サービス/マイクロサービス
- Web アプリケーション
- ビジネスアプリケーション
- API

Amazon Elasticsearch Service

- すべてのログを保存、検索、分析



課題：運用データやセキュリティデータ、デバイスデータは、複数のログがインフラストラクチャ全体に分散されて生成されます。これらを監視して対応するには、高レベルのビューが必要

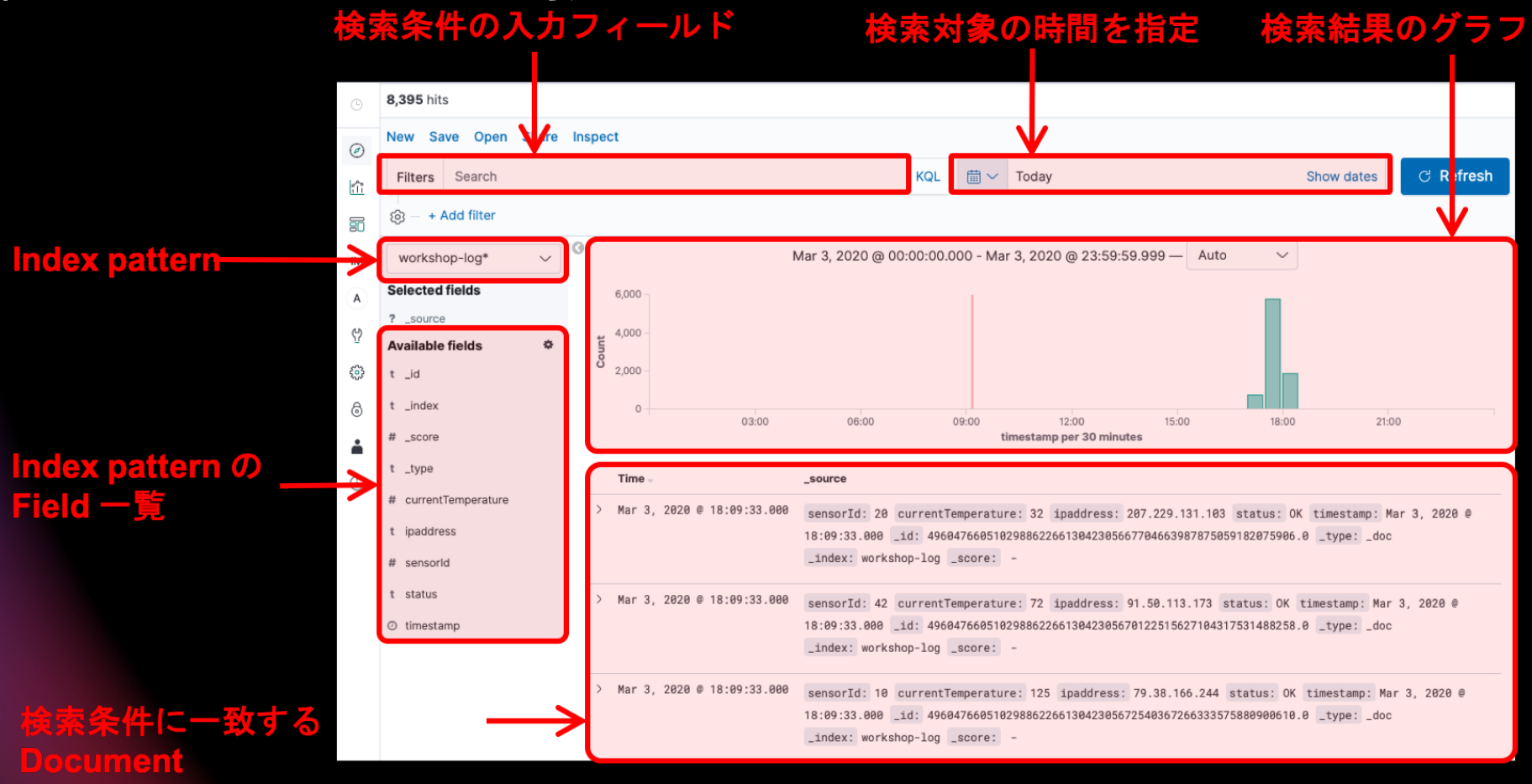


Amazon Elasticsearch Service とバンドルされているKibana を使用して、ログデータをほぼリアルタイムで分析し、インフラストラクチャ、アプリケーション、セキュリティ、IoT、AWS サービスを監視

Amazon Elasticsearch Service は、AWS クラウドでの Elasticsearch のデプロイ、スケーリング、モニタリング、保護を簡単にするフルマネージドサービス

Kibana によるドキュメント検索

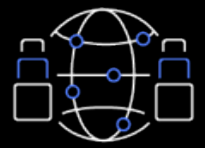
検索結果の CSV エクスポートに対応



Amazon Elasticsearch Service の利点

フルマネージド型

API や AWS コンソールを利用して、わずか数分でクラスターをデプロイできる



柔軟性

データの検索やログの分析が可能

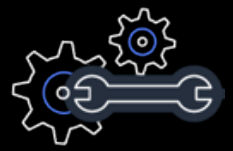
AWS とオープンソースの取り込みツールをサポートする

優れたコスト効率

従量課金制の支払い

運用コストの削減、適切なインスタンスタイプのサイジング

リザーブドインスタンスも選択可能



高可用性

セルフヒーリング、24 時間 365 日のモニタリング、1クリックでマルチ AZ活用、自動バックアップ、AWS サポートの利用、Amazon CloudWatch でのメトリクス取得

スケーラブル & 高性能

ワンクリックで、スケール、Elasticsearch バージョンのアップグレード、パッチ適用

セキュア

Amazon VPC 内へのデプロイ、Amazon Cognito によるログイン

HIPAA、FISMA、SOC、PCI、FedRamp に準拠

Snapshot

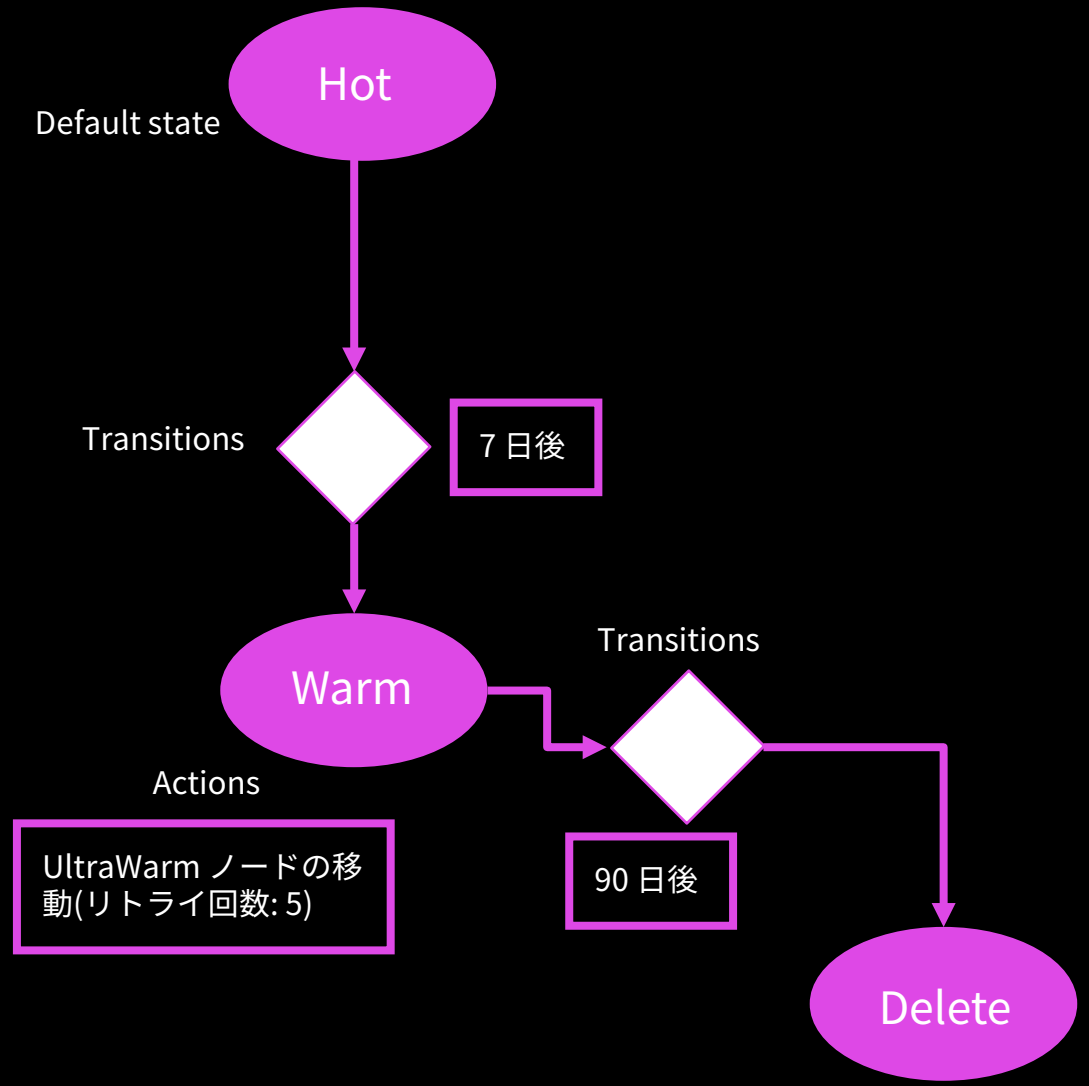
- スナップショットは、Index のバックアップ
- Amazon ES には、以下の 2 種類のスナップショットが存在する
- 基本的には自動スナップショットで足りるが、異なる Amazon ES にデータ移行したい場合は、手動での取得が必要

種類	用途	説明
自動スナップショット	バックアップ	- ES 5.3 以降の場合，1h ごとにスナップショットを取得し、14 日間保持（ES 5.1 以前は 1 日ごと） - 追加料金なし
手動スナップショット	- バックアップ - データ移行	- Elasticsearch の API を実行し、手動で Amazon S3 に対してスナップショットを作成 - Index State Management とも連携可能 - Amazon S3 利用料金が発生

https://docs.aws.amazon.com/ja_jp/elasticsearch-service/latest/developerguide/es-managedomains-snapshots.html

Index State Management (ISM) でインデックス管理の自動化

- ISM 機能により、日/週/月単位で新しく作られるインデックスの管理を自動化
- 指定したトリガに応じて Index に対して特定のアクションを実行する機能
- Kibana 上でインデックス管理ポリシーを設定・管理し、既存インデックスに適用
- レプリカの削減、 UltraWarm への移動、 RollOver、 RollUp、 削除などの処理を段階的に実行可能
- Index サイズベース、 経過時間ベース、 cron ベースのトリガをサポート



メトリクス

Amazon CloudWatch と連携したモニタリング用メトリクスを提供. 大きく分けて 以下 2 レベルのメトリクスが提供されている

Cluster メトリクス

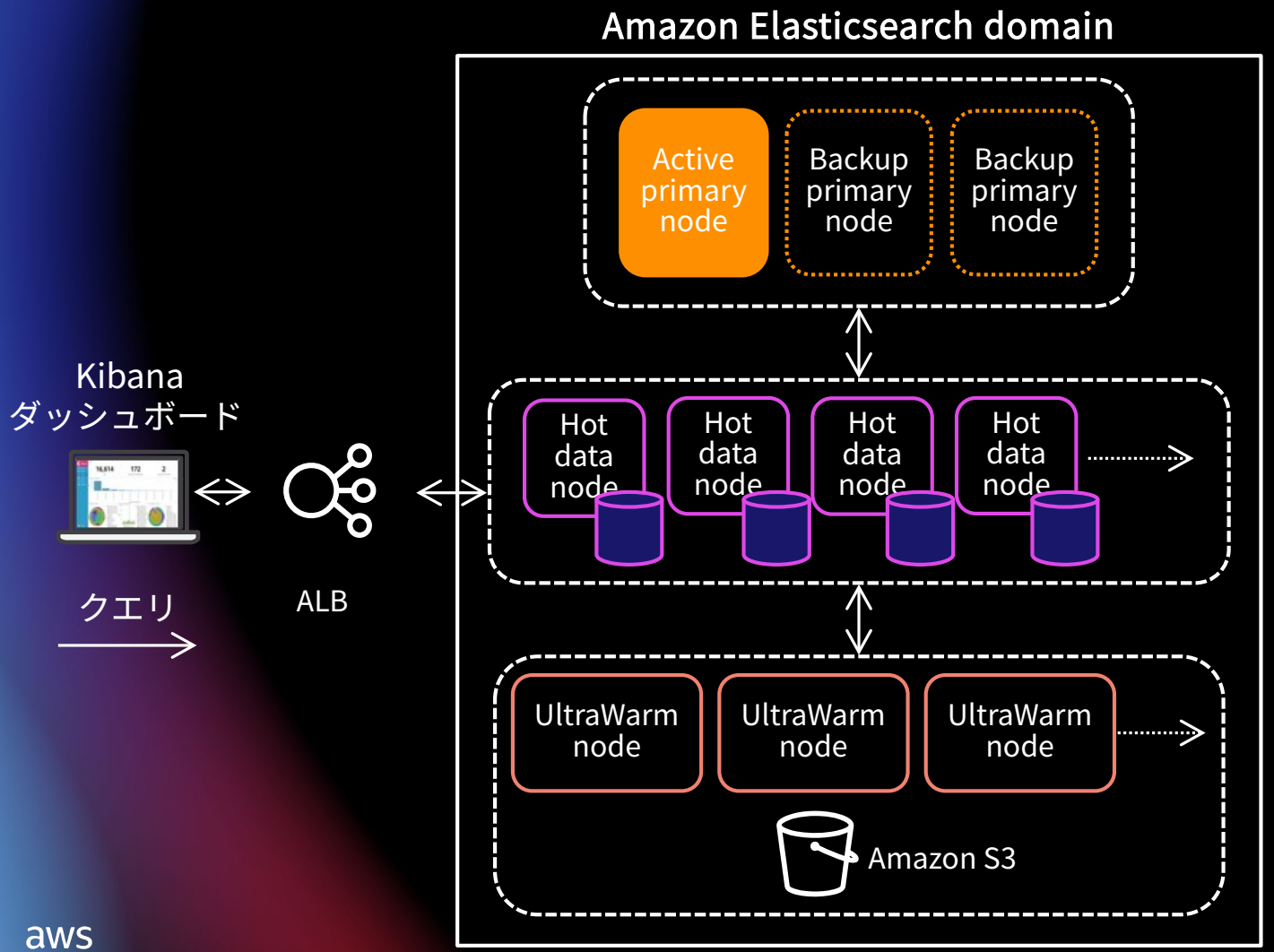
- Cluster 全体の状態、傾向を把握するためのメトリクス

Node メトリクス

- CPU 使用率など、 個々のノードのパフォーマンスを把握するためのメトリクス

UltraWarm for Amazon Elasticsearch Service

AMAZON ELASTICSEARCH SERVICEの新しいウォームストレージティア



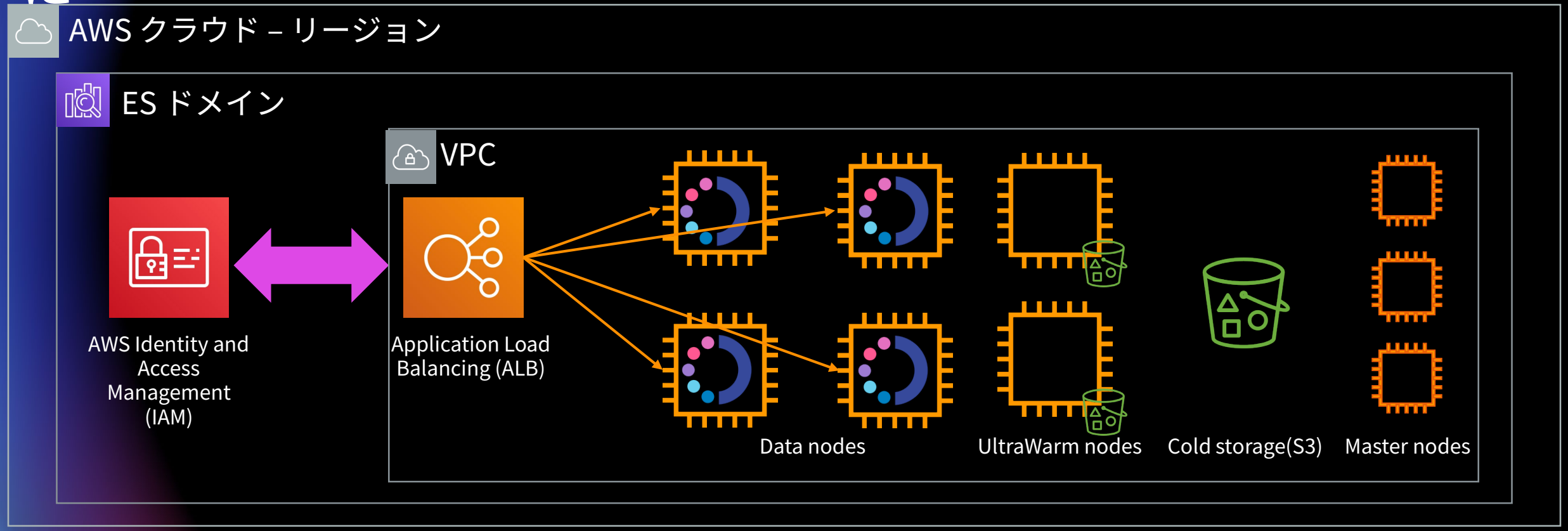
最大 90% のコスト削減

ドメインあたりで 3 PB までスケール

数年間の運用データを分析

インタラクティブにログの分析と可視化

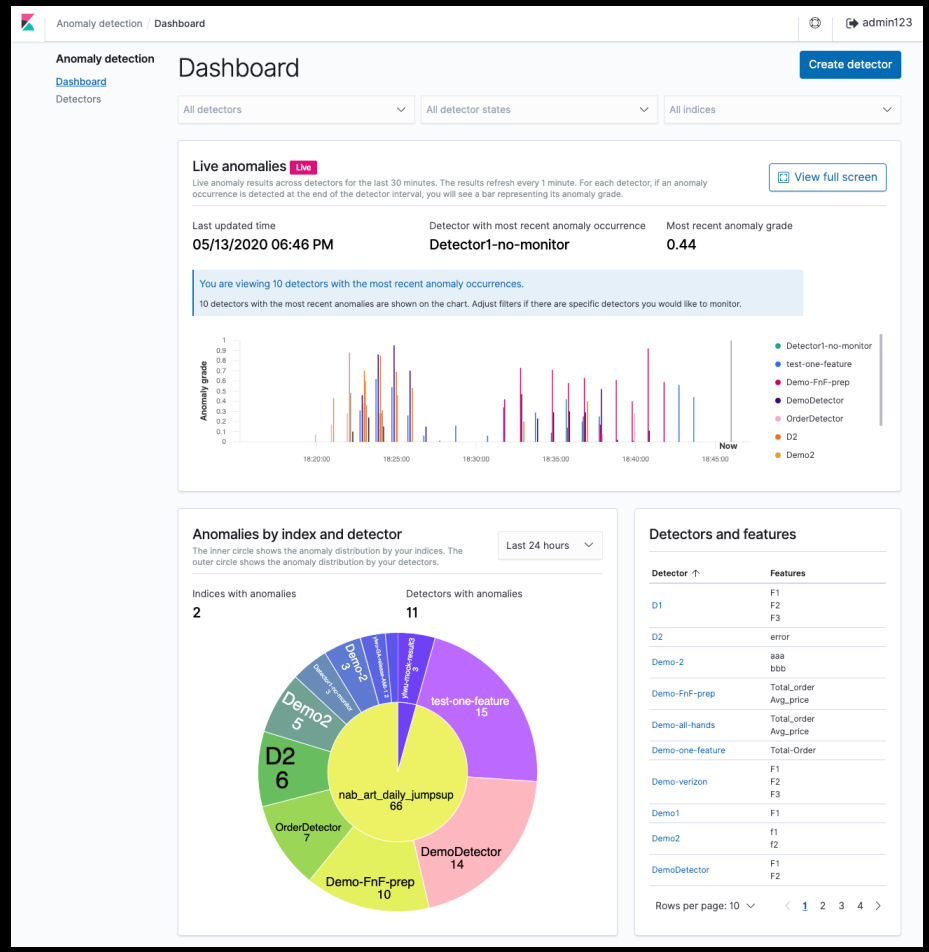
Amazon Elasticsearch Service でコールドストレージが利用可能に



コールドストレージは、アクセスの頻度が低いデータをオンデマンドで安全に保存および分析することを容易にするフルマネージドストレージ層であり、他のストレージ層よりも低コストである。コールドストレージは UltraWarm 上に構築されており、Amazon S3 でのデータの保存に特化したノードを提供。コールドストレージにより、使用していないときに UltraWarm からインデックスをデタッチし、コンピューティングを解放してコストを削減できるようになった
https://aws.amazon.com/jp/about-aws/whats-new/2021/05/amazon-elasticsearch-service-announces-a-new-lower-cost-storage-tier/

Anomaly Detection

- Random Cut Forest アルゴリズムを用いた、時系列の異常検知機能
- Kibana 上でインデックスやフィールド、メトリクスを指定して複数の Detector を作成
- フィールドの average() や max() 以外にも、クエリ構文で任意のメトリクスを作成することが可能
- アラート機能と連携して、異常な値が出たら Amazon SNS 等への通知が可能
- version7.4 以降で使用可能



https://docs.aws.amazon.com/elasticsearch-service/latest/developerguide/ad.html
https://opendistro.github.io/for-elasticsearch-docs/docs/ad/
https://www.semanticscholar.org/paper/Robust-Random-Cut-Forest-Based-Anomaly-Detection-on-Guha-Mishra/ecb365ef9b67cd5540cc4c53035a6a7bd88678f9

事例

アプリケーション、インフラストラクチャモニタリング

Autodesk ケーススタディ

- **アプリケーションとインフラストラクチャを監視する理由**
- アプリケーションとインフラストラクチャのログデータをプロアクティブに監視して、パフォーマンスの問題をより迅速に見つけ、運用状態を改善する必要があった
- **Amazon Elasticsearch Service がどのように役立ったか**
- リアルタイムの検索およびログ分析機能を提供して、パフォーマンスの問題を特定または予測し、チームがリアルタイムの根本原因とフォレンジック分析を実行できるようにした。これにより、平均検出時間（MTTD）と平均解決時間（MTTR）の問題が短縮された

AUTODESK

チャレンジ

- 高度に分散された組織 - メトリックスを収集および測定するための一貫した方法がない
- 他の AWS サービスと簡単に統合する必要
- 現在および将来の要件に対応するスケール
- 顧客に影響を与える問題を見つけて修正するための TBs のログデータ

ソリューション

- Amazon ESの採用。AWS 上に構築された統合ログデータ管理ソリューション。アプリケーション全体のログ分析のための単一のインターフェース
- Amazon Kinesis Firehose を介して Amazon S3 / Amazon Athenaおよび Amazon ESにアプリケーションログをストリーミング
- 10の i3.4xlarge Amazon ES データノード - 33TB。110 TBへ成長見込み
- ニアリアルタイムの分析とダッシュボードのための、Amazon ES に組み込まれているKibana

効果

- 全てマネージドサービス -「 manage less to gain more 」。顧客へのプロダクト開発にフォーカスできる
- 問題を診断して解決するための組織横断な共通言語。排除されたサイロ
- スケーラブルで費用効果が高い - TBあたりの大きな価値を提供するi3インスタンス
- 顧客の問題を見つけて修正する時間を短縮することにより、顧客体験が向上

“最終的には、ログデータをリアルタイムで可視化できるため、ソフトウェア製品を改善し、お客様により良いサービスを提供できている”

Tommy Li
Senior Software Architect,
Autodesk

aws

IoT での活用事例: KEMPPI 様 クラウドにパラメータを保存する IoT 溶接装置

データストアに Amazon ES を使用。「これは、溶接の資格や手順に関連した自由形式の問い合わせを最も高速に処理できる方法でした。」

背景：

Kemppiはフィンランドの会社で、自社を “溶接のパイオニア企業” と位置付け、溶接装置とアプリケーションソフトウェアを設計製造

課題：

- 需要増大と経験豊富な溶接工の不足によるスキルギャップ
- もっと頻繁に新しい機能を本番環境に移したい

ソリューション

- 遠隔地の作業者に作業指示書を配信、進捗および品質を管理
- クラウドに保存された溶接パラメータを自動的に適用
- スマートカードの認証と組み合わせ、資格ある作業員が作業した履歴を保存
- AWS IoT Core、Lambdaなどを用いたIoT対応溶接機

効果

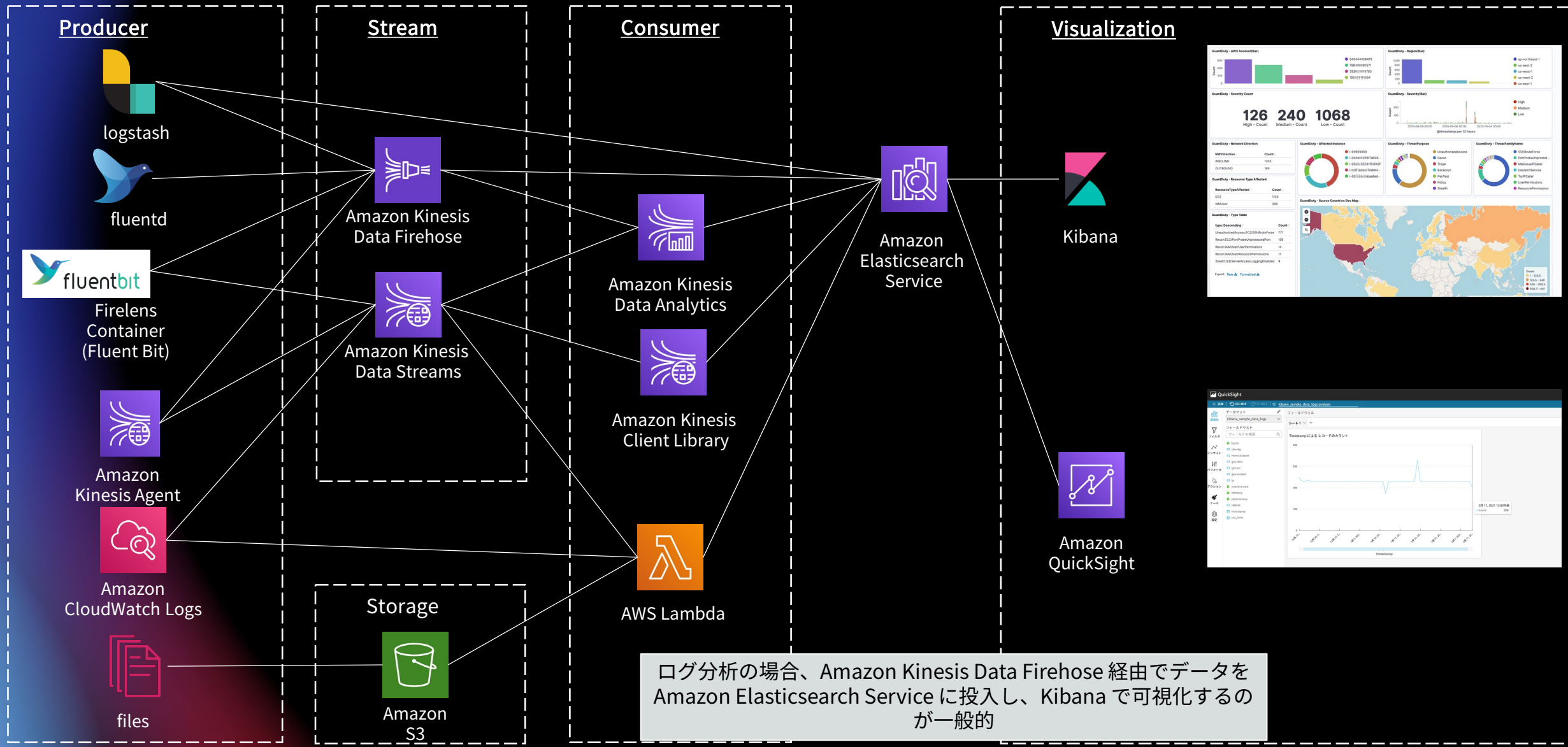
- ソフトウェア開発と配信のコストを約 50% 削減
- 市場投入サイクルを12ヶ月→6ヶ月に、機能更新を最大で週に10回実施

https://aws.amazon.com/jp/solutions/case-studies/kemppi/

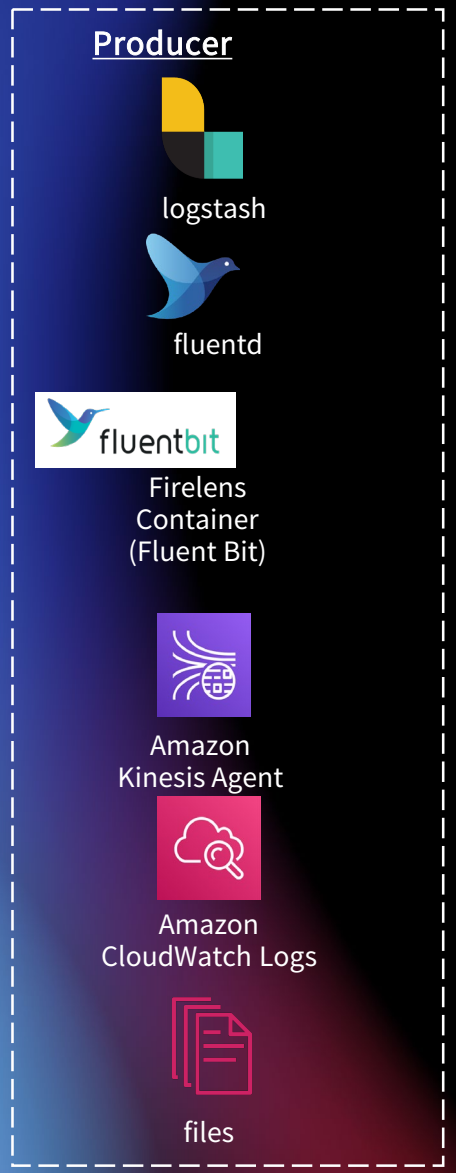
https://www.youtube.com/watch?v=Jb7_4OMtLJk

Amazon Elasticsearch Service で実現するログ分析アーキテクチャ

Amazon Elasticsearch Service で実現するログ分析



Amazon Elasticsearch Service で実現するログ分析

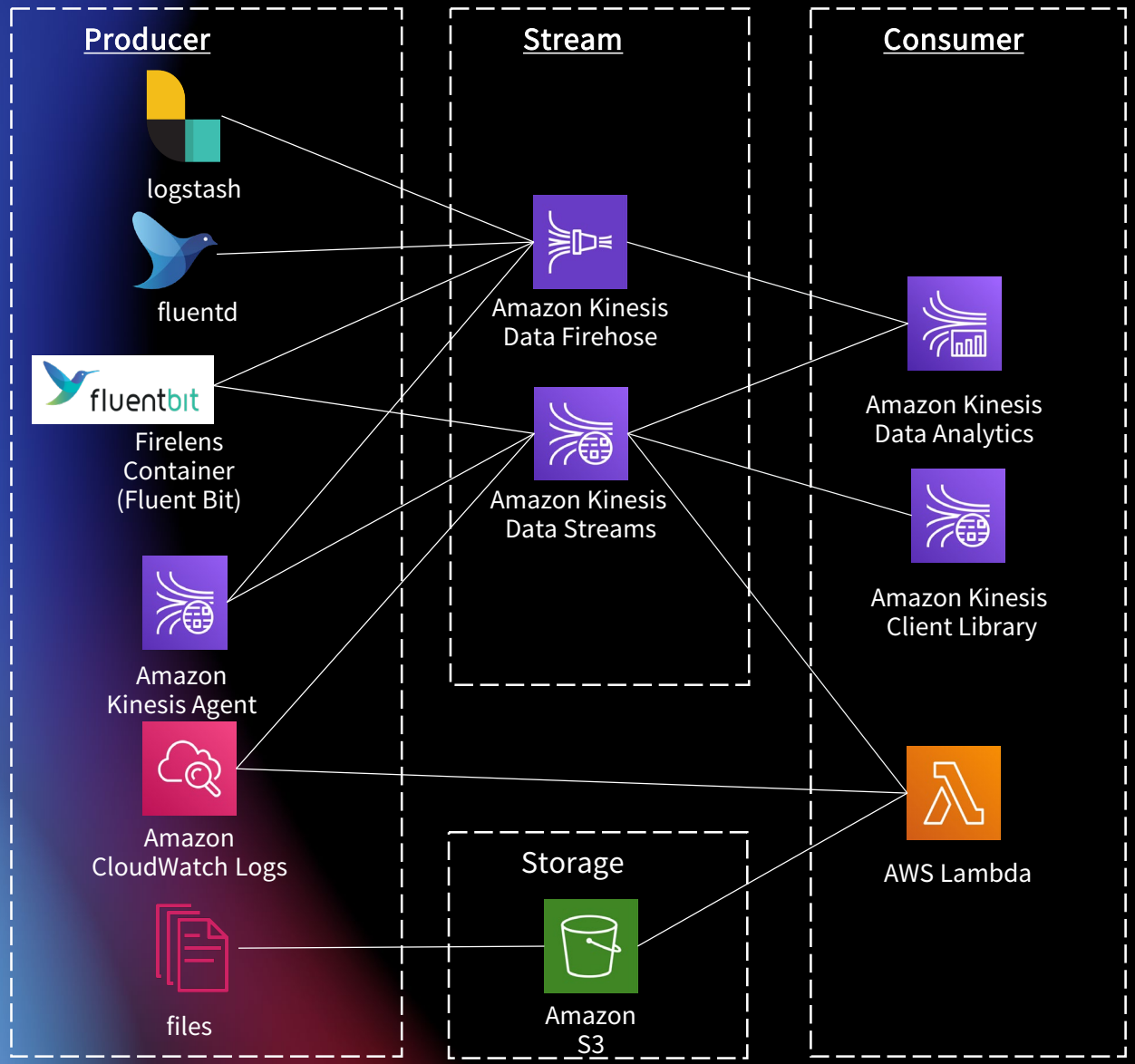


様々なデータが分析可能。例えば、アプリケーションインスタンスや IoT デバイス、あるいは AWS サービス

アプリケーションサーバからログ転送する場合、Kinesis Agent や CloudWatch Agent を配置

コンテナ周りのログ収集には、fluentd や fluentbit は、ログデータを収集するための一般的な選択肢

Amazon Elasticsearch Service で実現するログ分析



Producer から Amazon Elasticsearch Service にデータを直接送信も可能だが、同時実行による Elasticsearch の過負荷を回避するには、同時接続数を減らすレイヤーが必要

例えば、Amazon Elasticsearch Service に Stream 配信できる Kinesis Data Firehose が代表例

複雑な Stream 処理では Kinesis Data Analytics for Apache Flink 等の利用も検討

SIEM on Amazon Elasticsearch Service のご紹介

SIEM on Amazon Elasticsearch Service のご紹介

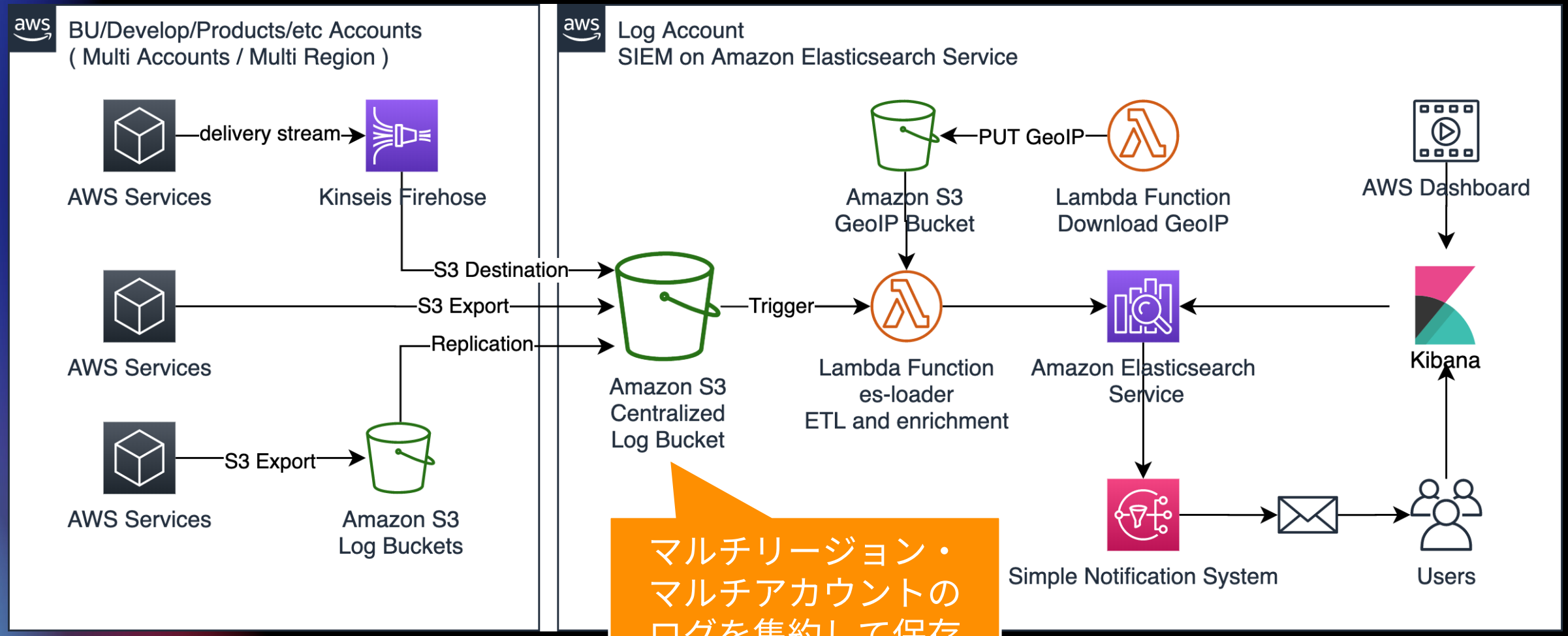
- AWS サービスのセキュリティ監視をするためのスクリプトやダッシュボードのサンプル。オープンソースソフトウェアとして公開
- 日本のセキュリティ ソリューションアーキテクトや Analytics ソリューションアーキテクトが中心となって開発

■特徴

- マネージドサービスとサーバーレスのみで構成
- マルチアカウント・マルチリージョン対応
- AWS サービス専用の正規化、ダッシュボード
- AWS CloudFormation / CDK によるデプロイ。約30分で完了
- クラウドサービスをご利用した分だけの従量制料金

SIEM on Amazon ES のアーキテクチャ

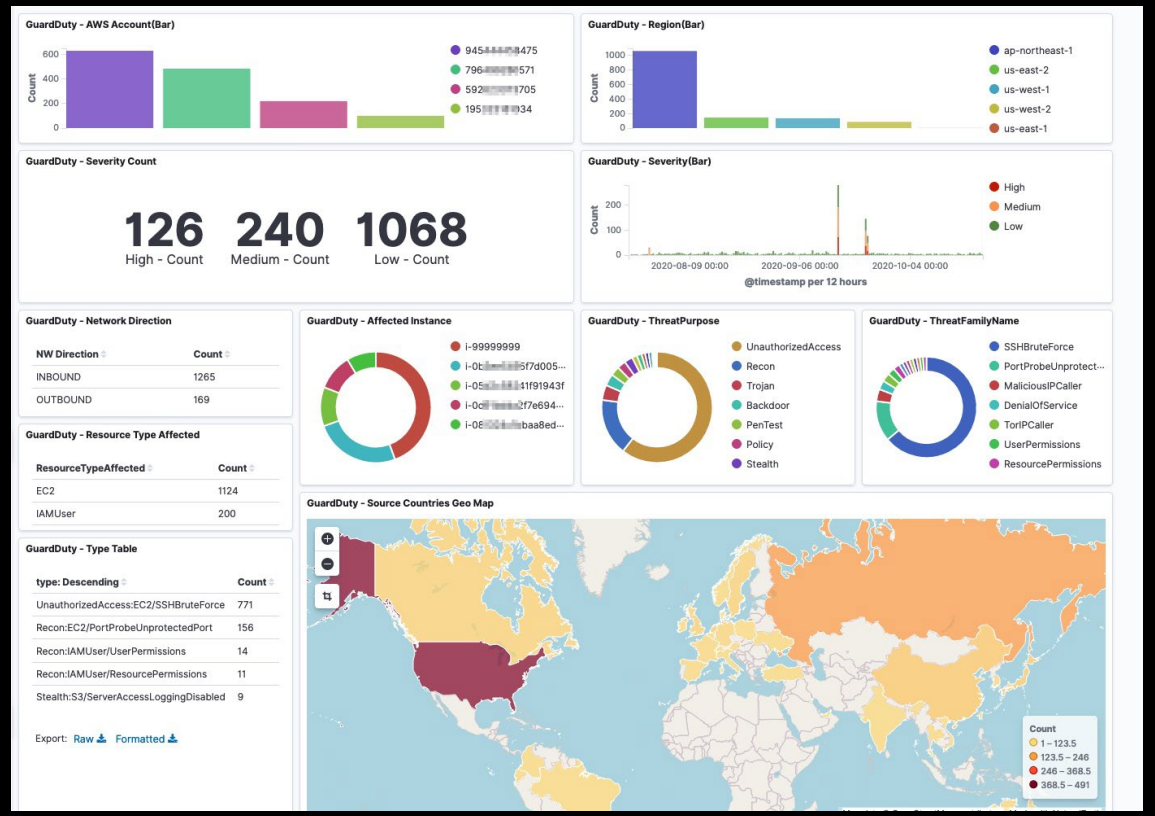
マネージドサービスとサーバーレスのみで構成



AWS ログ分析専用ダッシュボードのテンプレート

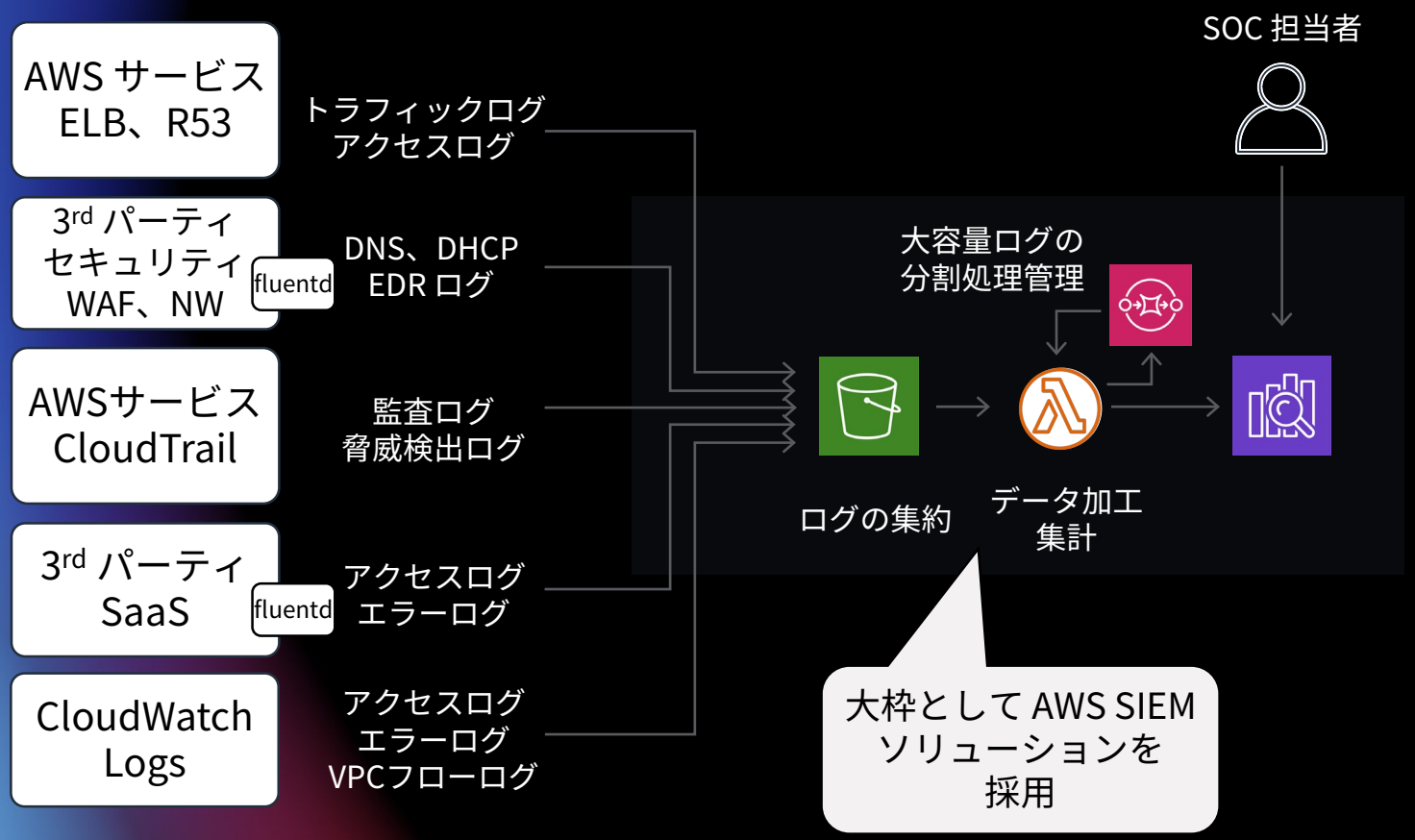
ログをセキュリティ分析の観点から
時系列、地理情報などで視覚化

AWS CloudTrail



Amazon GuardDuty

Sansan 様
統合セキュリティ監視（SIEM）



1TB/Day のログ処理と可視化

- 課題
- ✓ サービスプロバイダとしての
  セキュリティ担保の責任
- ✓ 対象ログデータは常に増加傾向
- ✓ 運用工数にも大きく影響

- 効果
- ✓ 短期実装: 構想から **2ヶ月**で最初の
  リリース。以降、対象ログを拡張。
- ✓ コスト最適化: 商用SIEMと比べて
  ライセンス - **80% 以上**削減
- ✓ オペレーション最適化
  - インフラ管理タスク – **ほぼ自動化**
    （サーバーレス化）

https://aws.amazon.com/jp/blogs/news/siem-on-amazon-elasticsearch-service/

SIEM on Amazon ES の利用方法

1. GitHub に移動 https://github.com/aws-samples/siem-on-amazon-elasticsearch/blob/main/README_ja.md
2. AWS CloudFormation のテンプレートを選択

3. CloudFormation でパラーメータを4つ入れて実行。約30分後にデプロイ完了
4. 可視化・分析したいログを S3 バケットに保存

まとめ

- 増大し続けるログデータやマシンデータを分析、可視化する基盤として Amazon Elasticsearch Service を提供している
- Amazon OpenSearch は、Elasticsearch / Kibana 最後の Apache License 2.0リリースであるversion7.10からのコミュニティ主導のオープンソースフォークであり、クラウドならではの様々な追加機能を提供している。既存の Amazon Elasticsearch Service からもアップグレード可能
- ログ分析の場合、Amazon Kinesis Data Firehose 経由でデータを Elasticsearch に投入し、Kibana で可視化するのが一般的
- AWS サービスのセキュリティ監視をするためのスクリプトやダッシュボードのサンプルを、SIEM on Amazon Elasticsearch として提供している。30分でデプロイ可能

ご清聴ありがとうございました

小林 航（わたる）
wataruko@amazon.co.jp

Please complete the session survey