aws SUMMIT

2-H1-2-10

AWS を支える ネットワークインフラと要素技術

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　岡本　京

自己紹介

岡本　京（おかもと　ひろし）

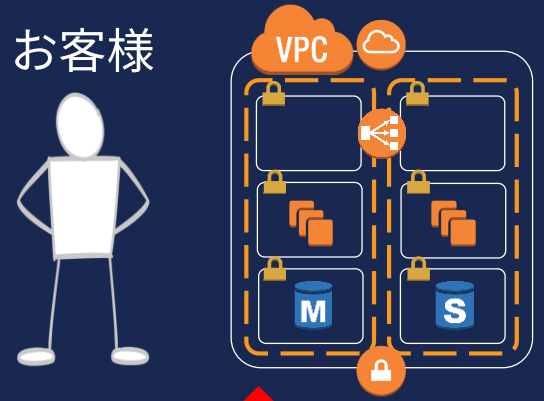
- 所属と職種
  - ストラテジックアカウント本部
    ソリューションアーキテクト

- 経歴
  - 前職はプリセールスエンジニア
    ネットワーク関連

- 好きな AWS サービス
  - Amazon VPC, AWS Direct Connect
  - 最近好き: AWS Cloud9

はじめに: AWS におけるネットワークの特徴

- オンデマンドで使用開始/撤退可能
  - リードタイム無し * 物理作業を伴う AWS Direct Connect の敷設を除く
  - 構築作業無し
- 従量課金
  - 初期投資無し
  - 無料でできることも多い
- **AWS により日々進化**
  - **規模の継続的な拡大**
  - **技術的なイノベーション**



本セッションの視点

本セッションのテーマ: AWS ネットワークの再発見

AWS を支えるネットワークインフラと
要素技術の進化の歴史

それらにより実現するサービス/機能の
使いどころ

目次

- リージョン内のネットワーク
  - Amazon VPC: スイッチング/ルーティング
  - AWS Hyperplane: ロードバランシング

- リージョン間のネットワーク
  - Amazon Global Network

aws SUMMIT

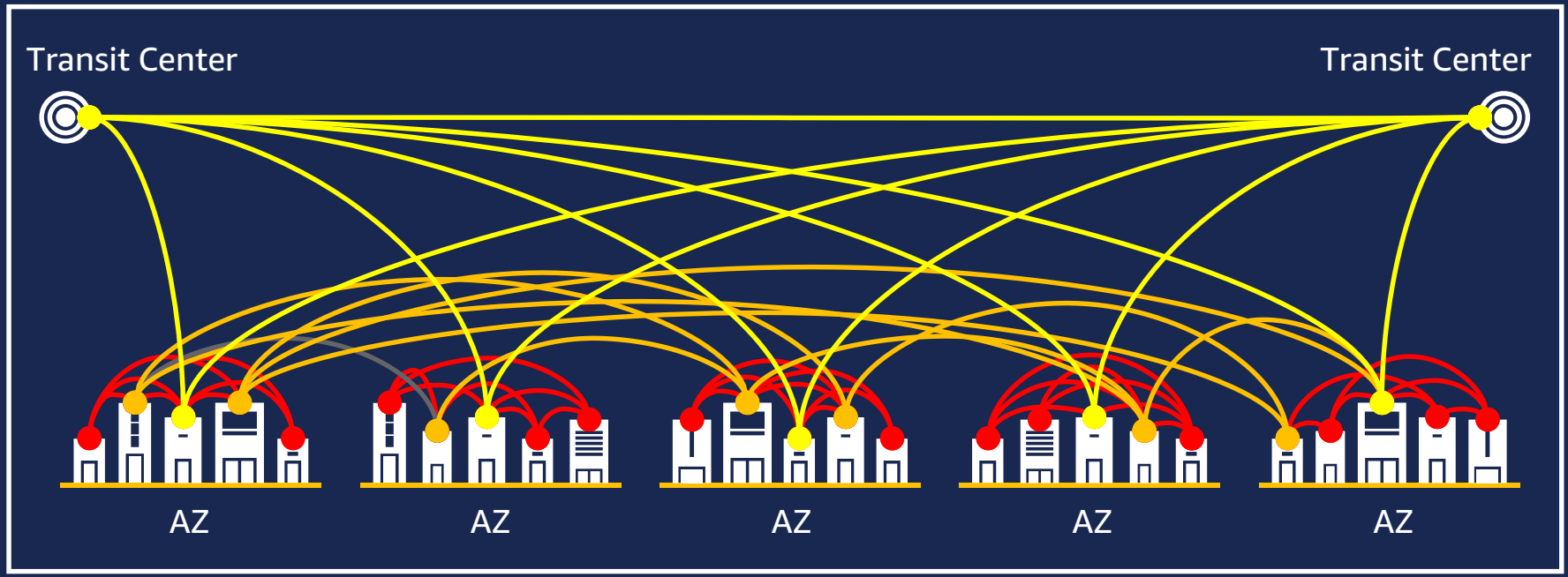
リージョン内のネットワーク

アベイラビリティゾーン（AZ）

- １つ以上のデータセンターで構成される、互いに独立したインフラ
  - 自然災害、電源、ネットワーク等を考慮
- 数十万台規模のサーバーを収容
- データセンター間は冗長化された独立したメトロファイバーで接続

リージョン

- 複数の AZ + 複数のトランジットセンターで構成されるインフラ
- AZ 間のレイテンシーは通常 1 ms、最大 2 ms



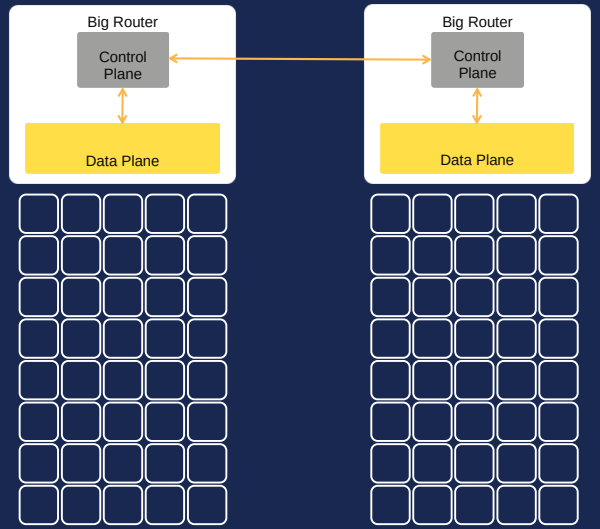
aws SUMMIT

Amazon VPC

イノベーション前夜

- 従来の手法 = 大容量ルーターによる収容

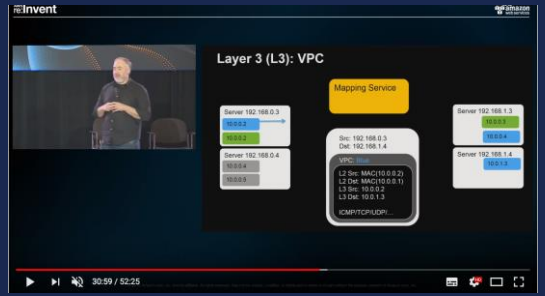
- 課題
  - VLAN, VRF の最大数
    - 数千が限界
  - 設定工数の増大
  - ベンダーの固有機能やリリースサイクルへの依存
    - マルチベンダー機器の管理は複雑
    - 不具合解決に長期間を要する



Amazon VPC

- クラウドスケールの分散型仮想ネットワーク
  - マッピングサービス
    - 「VPC ID/Instance ID – 物理ホスト」の管理とクエリ応答
    - インスタンス所在の Validation も実施
  - 物理ホストでの処理
    - インスタンスからの通信を捕捉しヘッダーの書き換えやマッピングサービスへのクエリを実施
    - EC2 が送信したパケットを自身のネットワーク情報で再カプセル化

参考 1
AWS re:Invent 2016:
Another Day, Another Billion Packets (NET401)



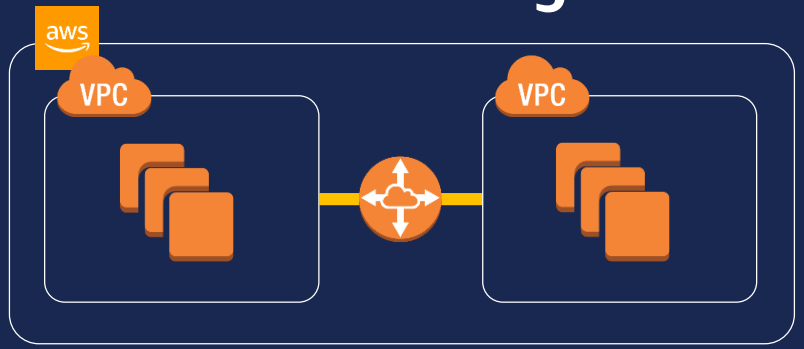
https://www.youtube.com/watch?v=St3SE4LWhKo

参考 2
AWSの深いところ見せちゃいます！
by AWSクラウドサポートエンジニア

https://codezine.jp/article/detail/9790

仮想ネットワークの特性を生かした機能の例

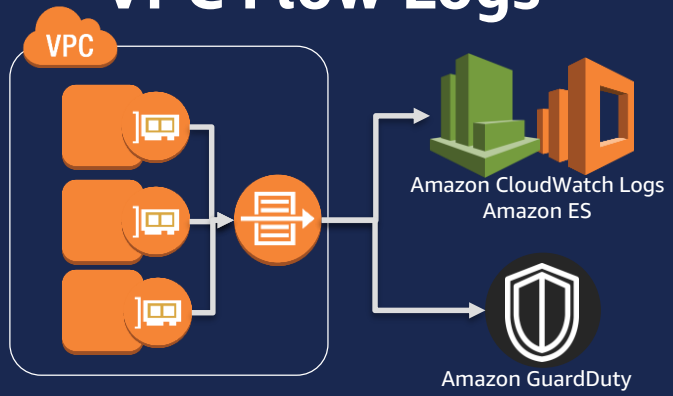
VPC Peering



- 2 つの VPC 間でルーティング
- 単一障害点、帯域幅のボトルネック無し

システム間をオンデマンドで接続
ホワイトリスト形式で管理

VPC Flow Logs



- VPC 内の全ての送受信トラフィックのログを取得

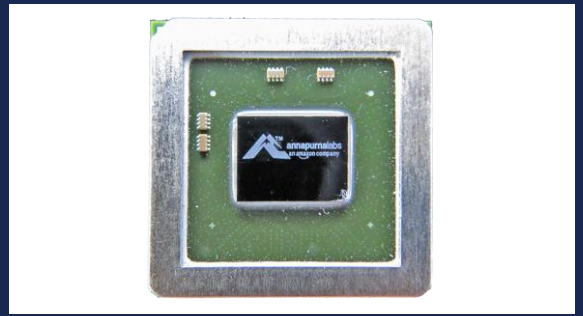
トラブルシューティング
意図しない通信や攻撃の検知

VPC を支えるカスタムハードウェア (1/2)

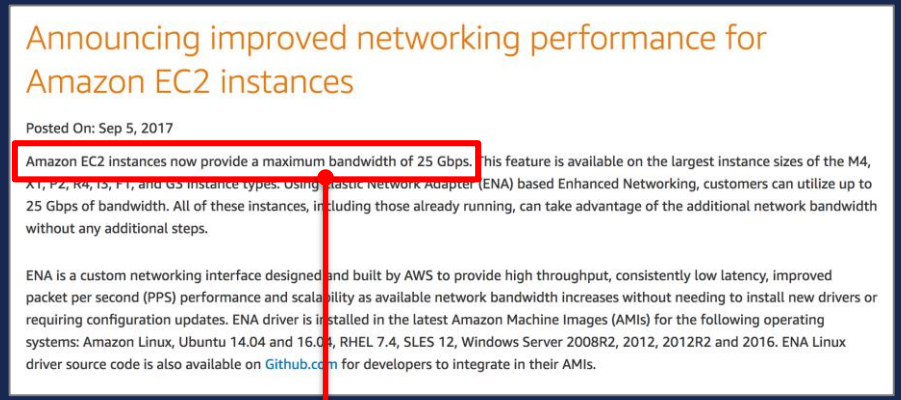
- ルーター
  - 当時主流の 10/40 GbE ではなく 25 GbE に Commit
    - 100 GbE = 25 GbE x 4 を見据えた判断
  - カスタム Broadcom Tomahawk ASIC 搭載
    - 128 ports of 25GbE
    - 1RU, 22lbs, <310W

VPC を支えるカスタムハードウェア (2/2)

- ネットワークアダプター
  - 2012 年世代
    - ハードウェアオフロードを開始
    - カスタム 10 GbE NIC とプロセッサ
    - SR-IOV, Enhanced Networking
  - 2016 年世代
    - 2 x 25 GbE
    - Elastic Network Adapter
      - インスタンスの最大帯域が 20 GbE に
      - 2nd Gen Enhanced Networking
    - Amazon Annapurna ASIC
      - AWS が HW/SW をコントロール



継続的なキャパシティ強化

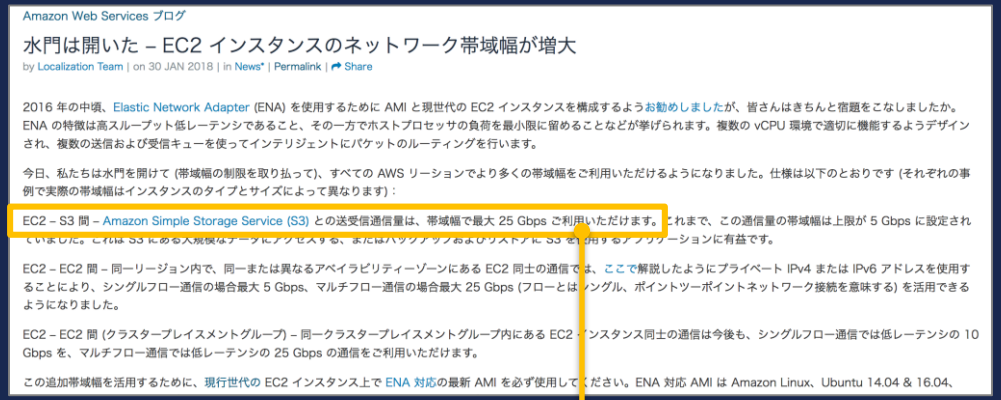


Announcing improved networking performance for Amazon EC2 instances

Posted On: Sep 5, 2017

Amazon EC2 instances now provide a maximum bandwidth of 25 Gbps. This feature is available on the largest instance sizes of the M4, X1, P2, R4, I3, F1, and G3 instance types. Using Elastic Network Adapter (ENA) based Enhanced Networking, customers can utilize up to 25 Gbps of bandwidth. All of these instances, including those already running, can take advantage of the additional network bandwidth without any additional steps.

ENA is a custom networking interface designed and built by AWS to provide high throughput, consistently low latency, improved packet per second (PPS) performance and scalability as available network bandwidth increases without needing to install new drivers or requiring configuration updates. ENA driver is installed in the latest Amazon Machine Images (AMIs) for the following operating systems: Amazon Linux, Ubuntu 14.04 and 16.04, RHEL 7.4, SLES 12, Windows Server 2008R2, 2012, 2012R2 and 2016. ENA Linux driver source code is also available on Github.com for developers to integrate in their AMIs.



Amazon Web Services ブログ

水門は開いた – EC2 インスタンスのネットワーク帯域幅が増大

by Localization Team | on 30 JAN 2018 | in News* | Permalink | Share

2016 年の中頃、Elastic Network Adapter (ENA) を使用するために AMI と現世代の EC2 インスタンスを構成するようお勧めしましたが、皆さんはきちんと宿題をこなしましたか。ENA の特徴は高スループット低レーテンシであること、その一方でホストプロセッサの負荷を最小限に留めることなどが挙げられます。複数の vCPU 環境で適切に機能するようデザインされ、複数の送信および受信キューを使ってインテリジェントにパケットのルーティングを行います。

今日、私たちは水門を開けて (帯域幅の制限を取り払って)、すべての AWS リーションでより多くの帯域幅をご利用いただけるようになりました。仕様は以下のとおりです (それぞれの事例で実際の帯域幅はインスタンスのタイプとサイズによって異なります) :

EC2 – S3 間 – Amazon Simple Storage Service (S3) との送受信通信量は、帯域幅で最大 25 Gbps ご利用いただけます。これまで、この通信量の帯域幅は上限が 5 Gbps に設定されていました。これは S3 にある大規模なデータにアクセスする、またはバックアップおよびリストアに S3 を利用するアプリケーションに有益です。

EC2 – EC2 間 – 同一リージョン内で、同一または異なるアベイラビリティーゾーンにある EC2 同士の通信では、ここで解説したようにプライベート IPv4 または IPv6 アドレスを使用することにより、シングルフロー通信の場合最大 5 Gbps、マルチフロー通信の場合最大 25 Gbps (フローとはシングル、ポイントツーポイントネットワーク接続を意味する) を活用できるようになりました。

EC2 – EC2 間 (クラスタープレイスメントグループ) – 同一クラスタープレイスメントグループ内にある EC2 インスタンス同士の通信は今後も、シングルフロー通信では低レーテンシの 10 Gbps を、マルチフロー通信では低レーテンシの 25 Gbps の通信をご利用いただけます。

この追加帯域幅を活用するために、現行世代の EC2 インスタンス上で ENA 対応の最新 AMI を必ず使用してください。ENA 対応 AMI は Amazon Linux、Ubuntu 14.04 & 16.04、

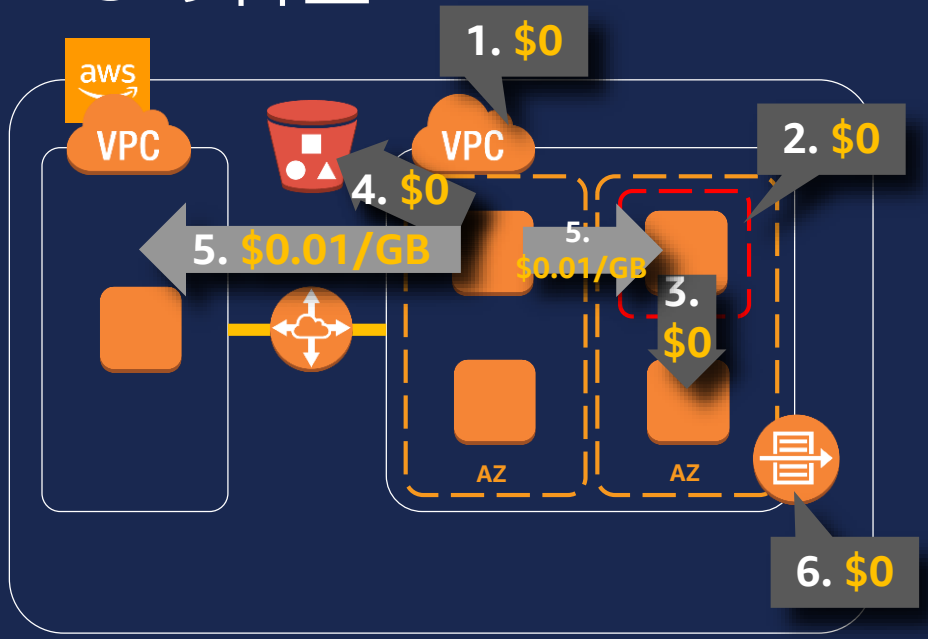
2017 年 9 月

Amazon EC2 instances now provide a maximum bandwidth of 25 Gbps.

2018 年 1 月

EC2 – S3 間 – Amazon Simple Storage Service (S3) との送受信通信量は、帯域幅で最大 25 Gbps ご利用いただけます。

VPC の料金



1. VPC
2. セキュリティグループ
3. AZ 内の通信
4. リージョン内の S3 等との通信
5. AZ 間/VPC Peering 間の通信
   *イン側にも同一の料金が発生
6. VPC Flow Logs の取得
   *保存には S3 等の料金が発生

大規模かつ可用性の高い日々進化するネットワーク環境を
低コストで使用可能

価格は 2018 年 5 月時点。特に注記がない場合東京リージョンの情報

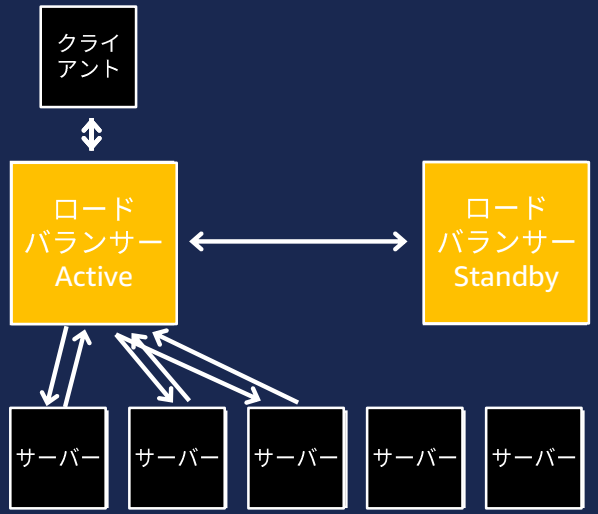
aws SUMMIT

AWS Hyperplane

イノベーション前夜

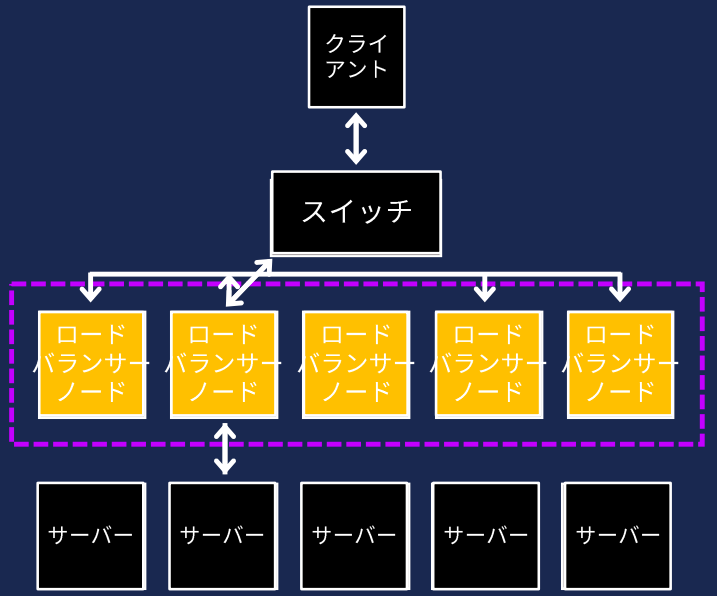
- 従来のアプローチ = ハードウェアロードバランサー
  - 構築が容易
  - 進化が活発（当時）

- 課題
  - トラブル解決が困難
    - 実装がブラックボックス
  - キャパシティ管理が困難
    - 1 台に複数サービスのエンドポイントを収容
    - 残リソースに応じた配置
  - コストが増大
    - Active-Standby で低い使用率



S3 Load Balancer

- 高度に分散したアーキテクチャーの採用
  - セッションステートは最低 3 ノードに格納
  - どのノードも全てのトラフィックの振り分け先を解決可能
    - ローカルで解決できないトラフィックは分散システムアルゴリズムを用いて解決
  - ノード障害時の影響が最小化
  - 高いリソース使用効率



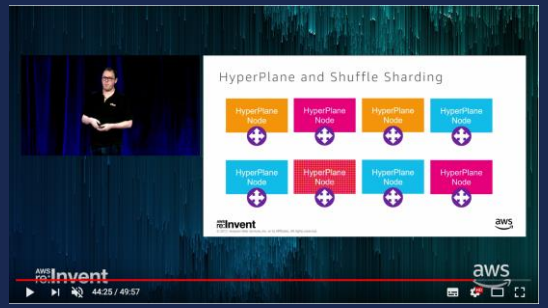
- コモディティハードウェアのみで構成可能

S3 に対する 37 Tb/s ものトラフィックボリュームに
対応できる安価なロードバランサーを開発

AWS Hyperplane

- Internal Network Load Balancing Service
  - S3 Load Balancer の経験を生かして開発
  - 現在は 4 つのサービスのコア技術として活用

参考
AWS re:Invent 2017:
Another Day, Another Billion Flows (NET405)



https://www.youtube.com/watch?v=8gc2DgBqo9U

Amazon EFS フルマネージド NFS

AWS Managed NAT VPC で使用する NAT ゲートウェイ

Network Load Balancer TCP ロードバランサー

AWS PrivateLink 任意の VPC に自身のサービスのエンドポイントを提供

Network Load Balancer

- Hyperplane ベースの TCP ロードバランサー
  - 数百万リクエスト /s に対応
  - Pre-warming なしに急激なスパイクにも対応
- NLB の料金
  - 起動時間あたりの料金 + LCU 料金
  - 例) 東京リージョン
    - 起動時間: $0.0243/時間 → **$17.50/月**
    - LCU: $0.006 USD/LCU 時間
      - 新規セッション数/維持セッション数/転送データ量の中の最大値で計算

スケーラビリティ、可用性などの考慮を AWS にオフロード
スモールスタートが可能な従量課金

aws SUMMIT

リージョン間のネットワーク

AWS リージョンの拡大

- 2006 – 2010: **4**

AWS リージョンの拡大

- 2006 – 2010: **4**
- 2011 – 2015: **+7**
- 2016 – 2018: **+11**

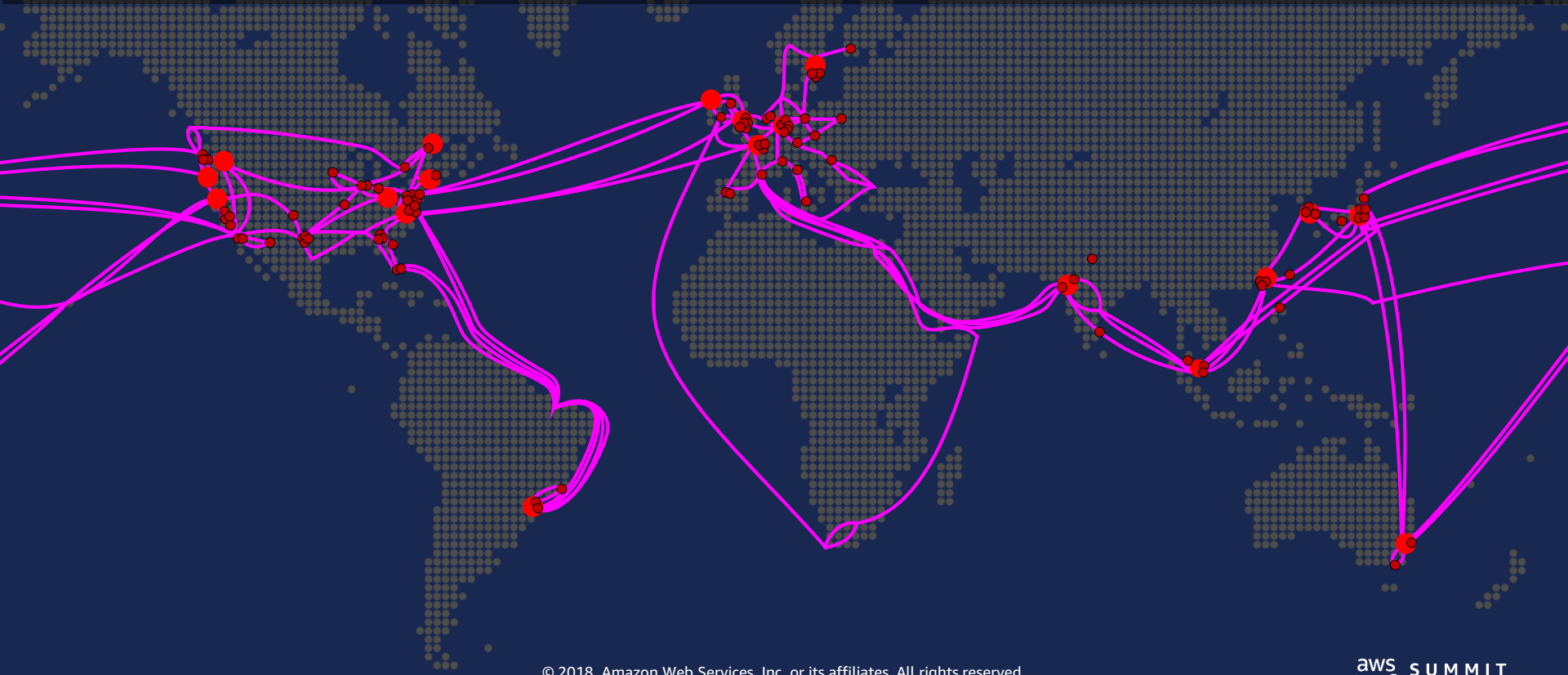
* 2018 年分は GA 前のリージョンを含む

AWS エッジロケーション

- 主に下記の機能を提供するインフラ
  - コンテンツキャッシュ (CloudFront)
  - DNS サービス (Route 53)
  - セキュリティ (AWS WAF, AWS Shield)
- グローバルで 100 箇所以上

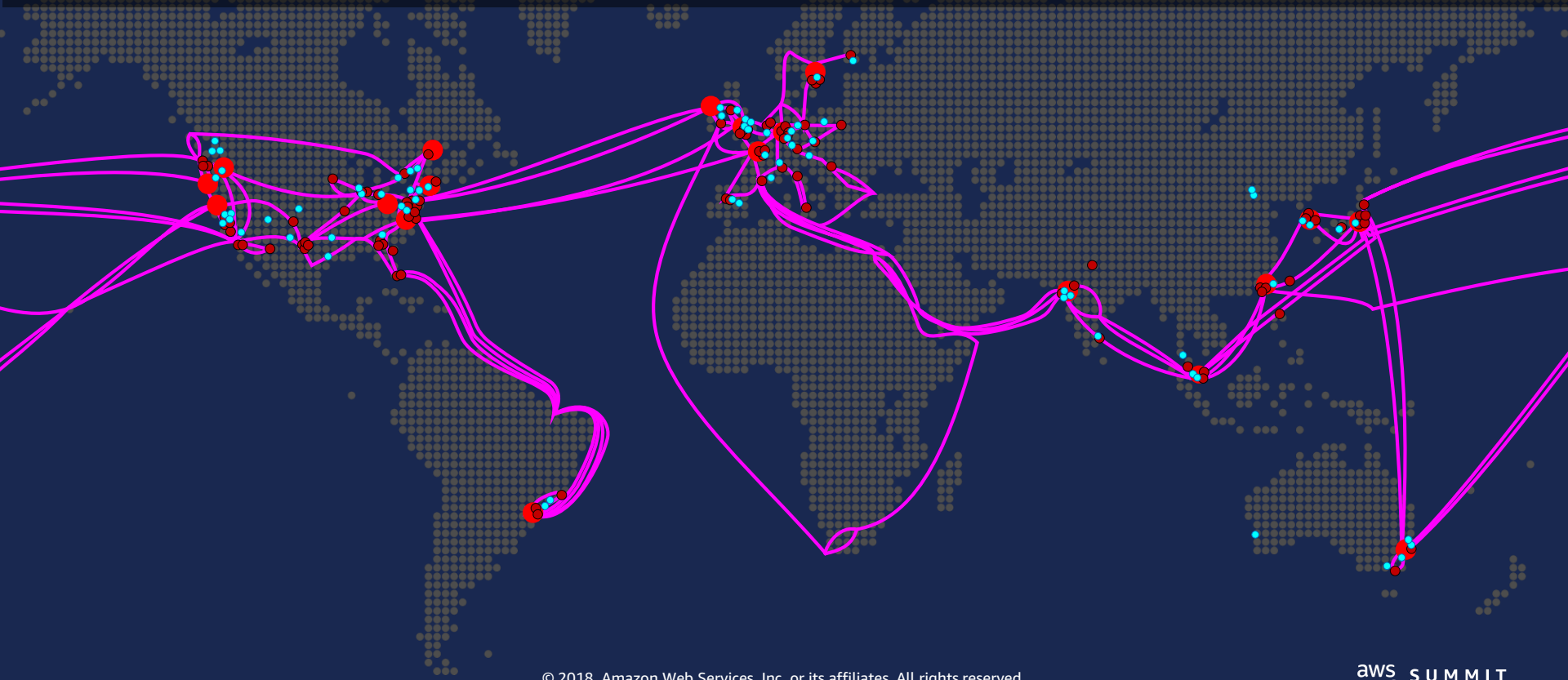
Amazon Global Network

- 冗長化された 100 GbE ネットワーク
- 中国を除く全てのリージョン間にプライベートキャパシティを提供

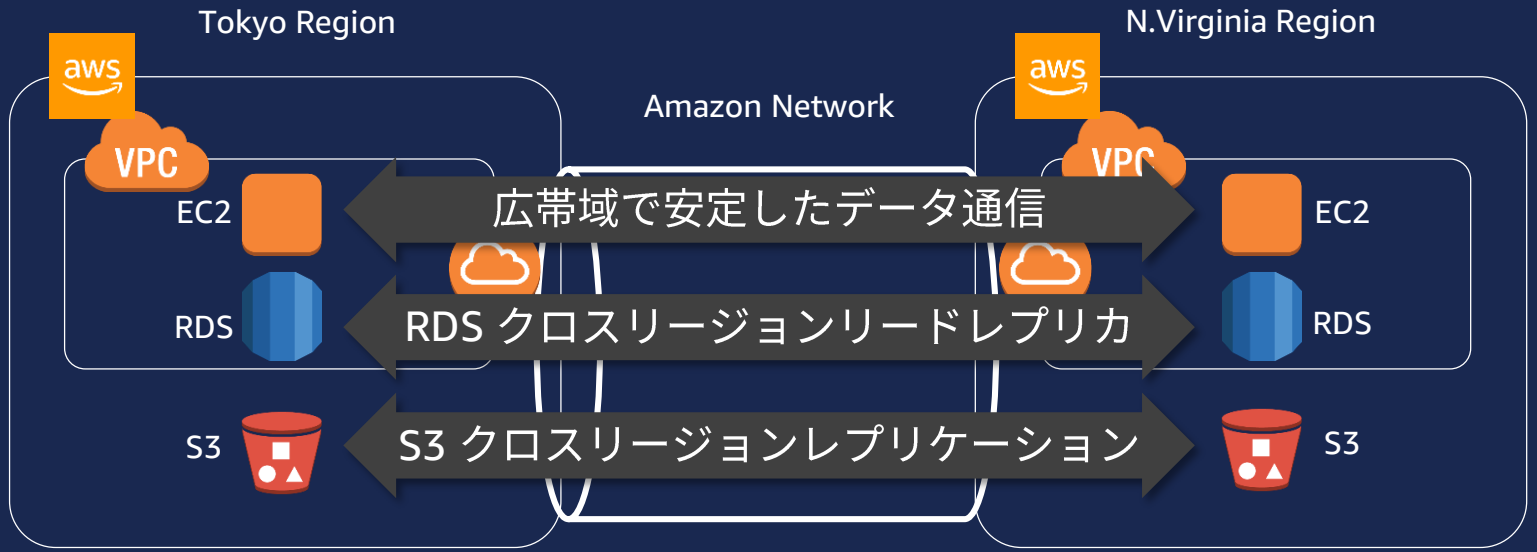


Direct Connect ロケーション

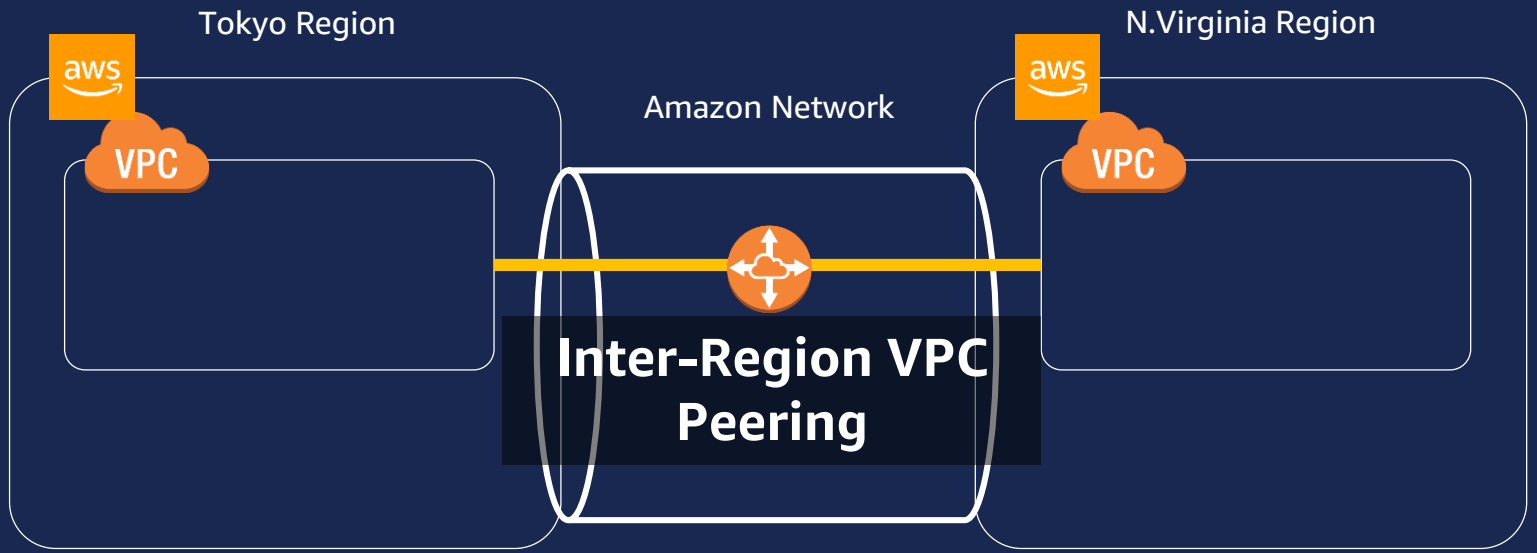
- お客様拠点と AWS リージョンを相互接続するポイント
- 日本には 3 箇所（Equinix TY2, OS1, アット東京 CC1）



Amazon Global Network の活躍どころ



活用を促進するアップデート 1/3



活用例	グローバルシステム連携 Disaster Recovery

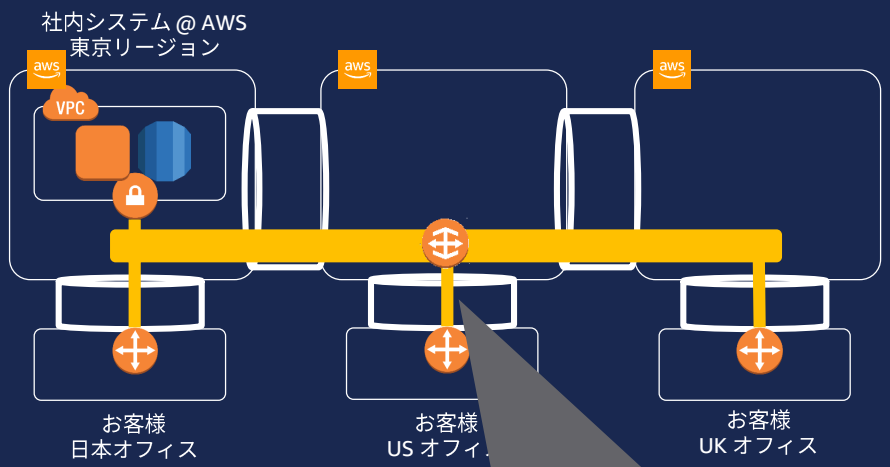
活用を促進するアップデート 2/3

活用例

グローバル拠点へのサービス提供
ハイブリッド構成のシンプル化

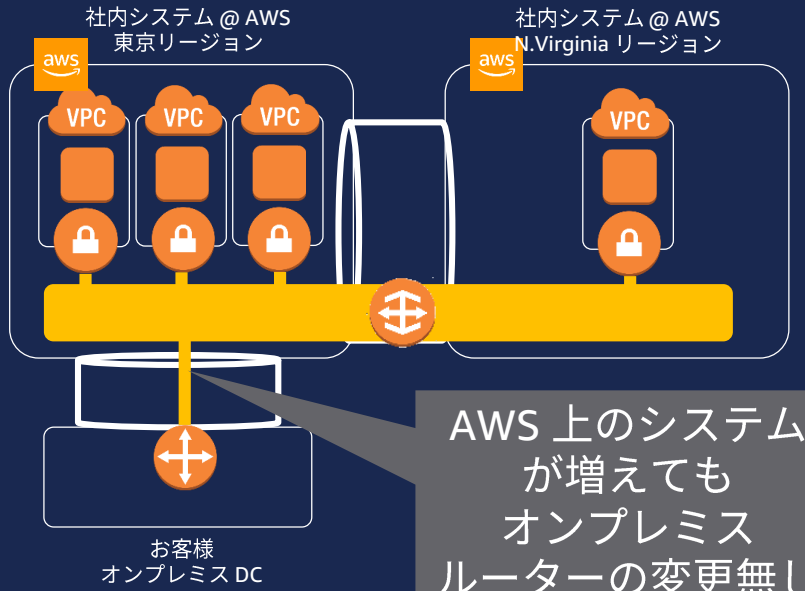
Direct Connect Gateway 活用例

グローバル拠点へのサービス提供



最寄りの AWS リージョンに
接続するだけで国際回線接続に

ハイブリッド構成のシンプル化



AWS 上のシステム
が増えても
オンプレミス
ルーターの変更無し

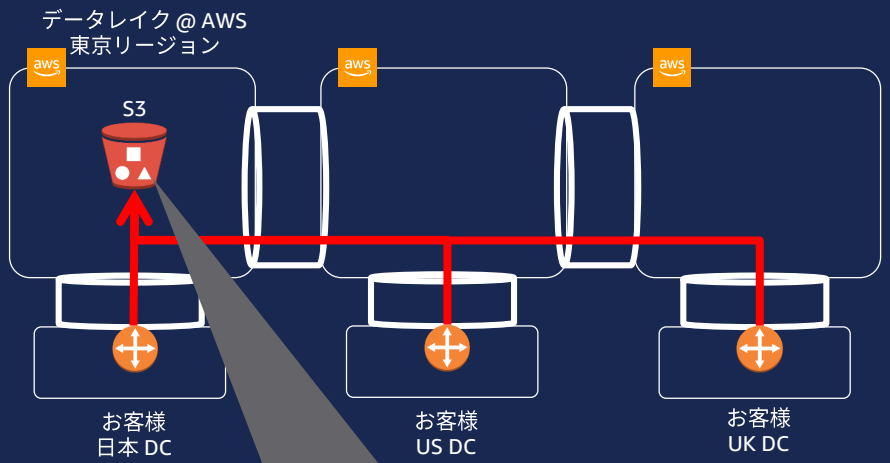
活用を促進するアップデート 3/3

活用例

グローバルデータレイク
グローバル Factory IoT

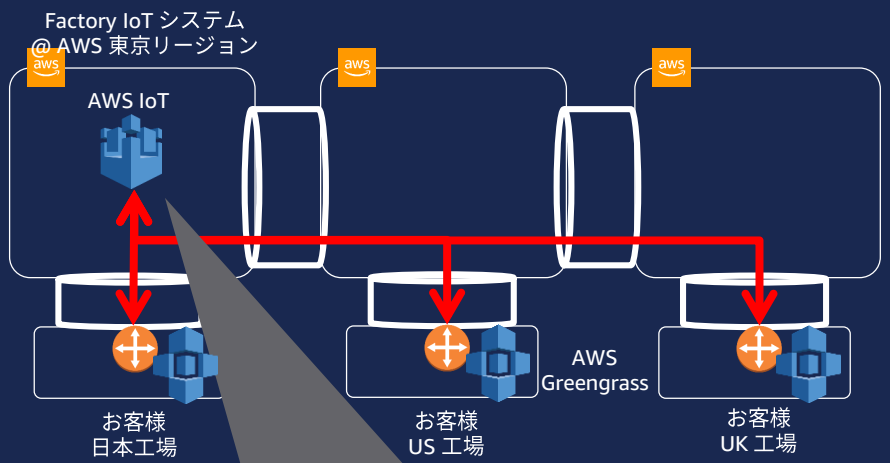
Public Connectivity for All Regions 活用例

グローバルデータレイク



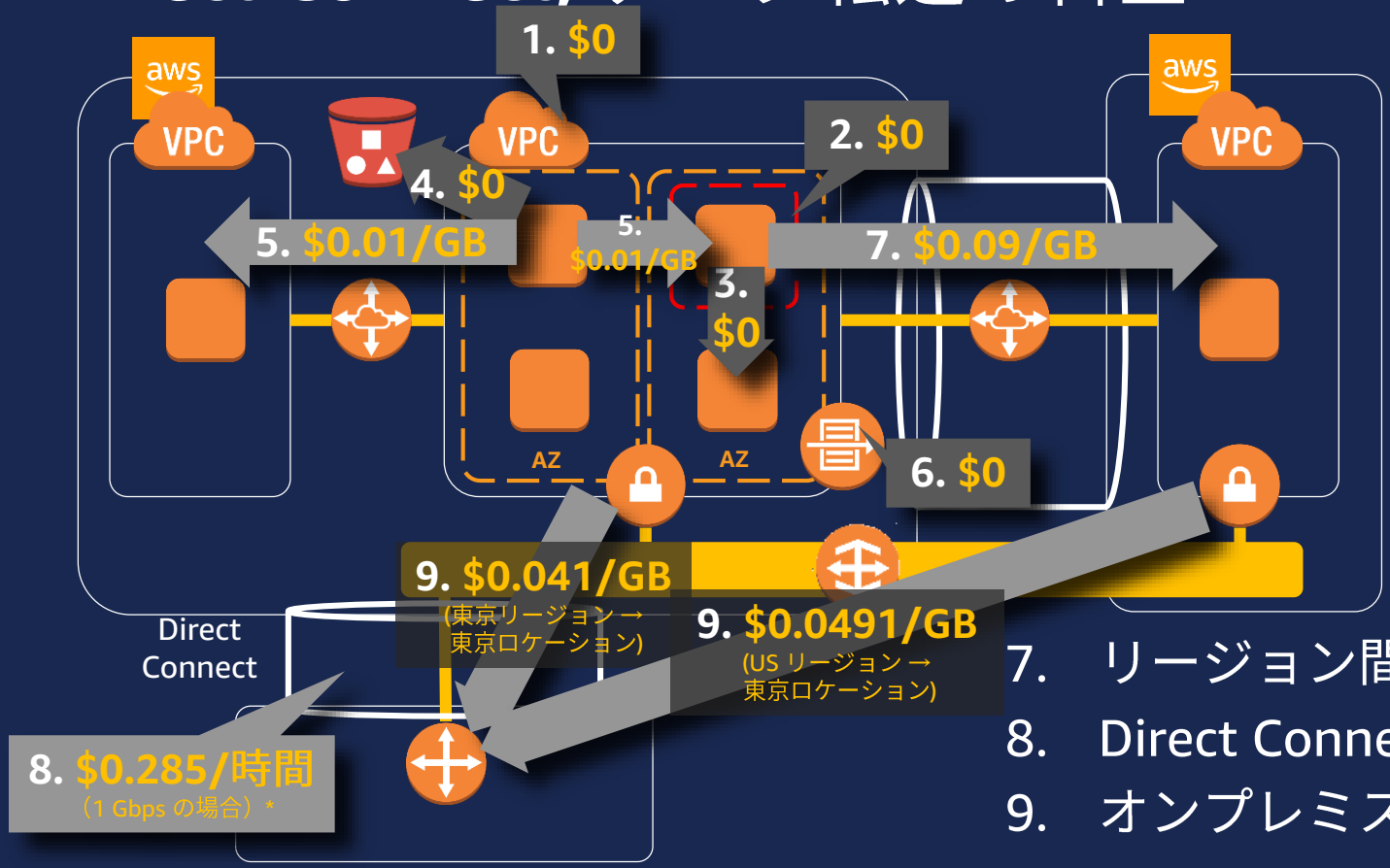
全社の情報を直接集約し
迅速に分析や活用

グローバル Factory IoT



リアルタイムに稼働状況を取得
クラウドで機械学習したモデルを
配布して工場で異常検知

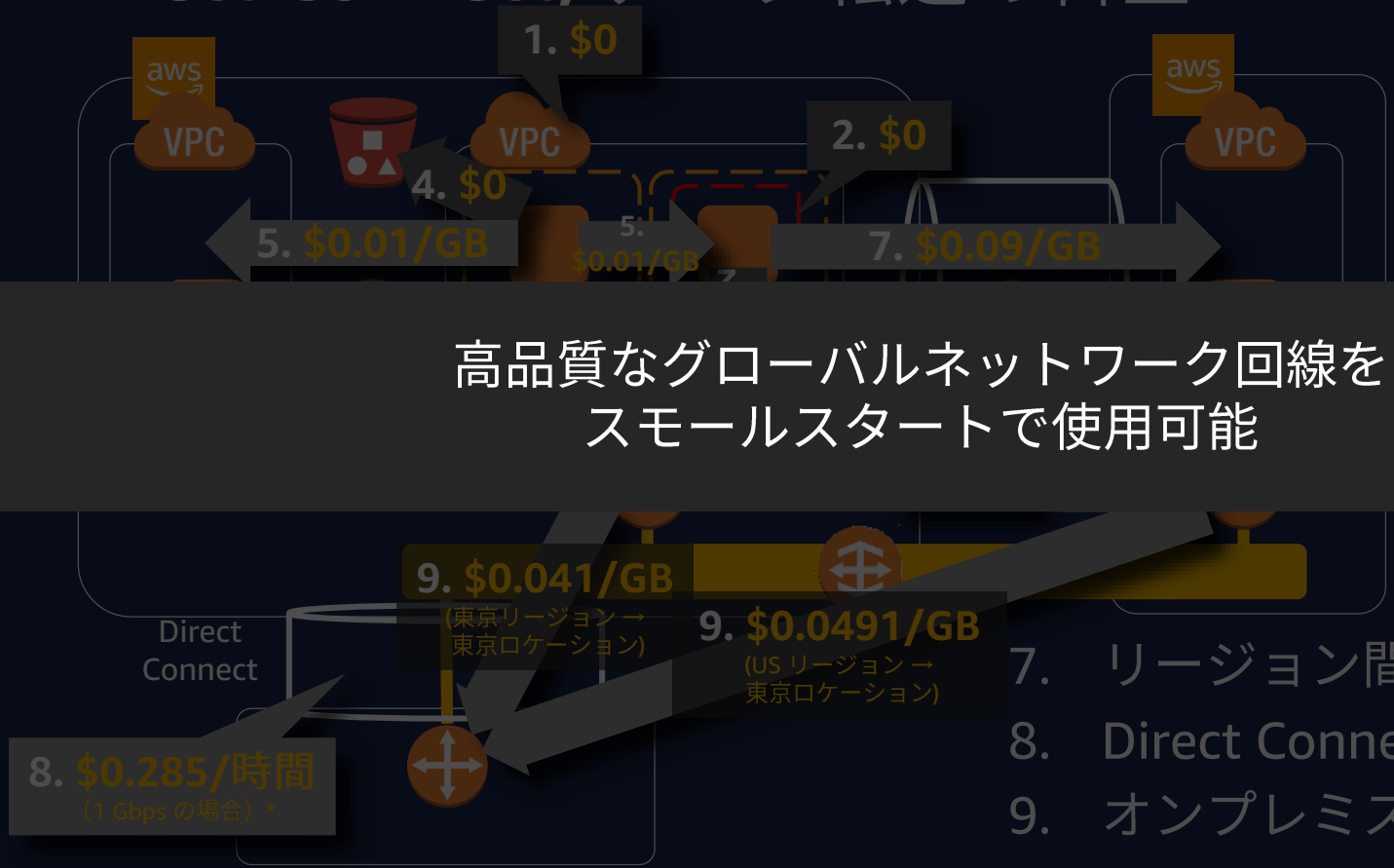
Direct Connect, データ転送の料金



7. リージョン間の通信
8. Direct Connect ポート料金
9. オンプレミスへの通信

価格は 2018 年 5 月時点。特に注記がない場合東京リージョンの情報
Direct Connect ロケーションから拠点までのネットワーク接続費用は含まれておりません

Direct Connect, データ転送の料金



高品質なグローバルネットワーク回線を
スモールスタートで使用可能

7. リージョン間の通信
8. Direct Connect ポート料金
9. オンプレミスへの通信

価格は 2018 年 5 月時点。特に注記がない場合東京リージョンの情報
Direct Connect ロケーションから拠点までのネットワーク接続費用は含まれておりません

まとめ（テーマの再掲）

- AWS ネットワークの特徴
  - オンデマンドで使用開始/撤退可能
  - 従量課金
  - **AWS により日々進化**
    - Amazon VPC
    - AWS Hyperplane
    - Amazon Global Network

これからも進化していく AWS ネットワーク
是非ご活用ください！

aws SUMMIT

Thank you!