パケットの気持ちになって辿る Amazon VPC のルーティング

岡本 京 (okamoh@amazon.co.jp)
ソリューションアーキテクト
アマゾン ウェブ サービス ジャパン株式会社

aws SUMMIT ONLINE

自己紹介

岡本　京（おかもと　ひろし）

- 所属と職種
  - •部長／シニアソリューションアーキテクト
  - •製造、自動車業界のお客様を担当
- 経歴
  - •前職はネットワークメーカーの
プリセールスエンジニア
- 好きな AWS サービス
  - •AWS Hyperplane

本セッションの狙い

- Amazon Virtual Private Cloud (VPC) はクラウド内に論理的に分離された仮想ネットワークを自由に定義することができるサービスです

- きめ細かいルーティングやフィルタリングを実現可能ですが、オンプレミスのルーター/スイッチに慣れ親しんだ方には直観的に理解しづらい部分もあるかと思います

- 本セッションではパケットの動きを追うことで VPC における設定ポイントやルーティングの動作を説明し、今後の設計において下地となる「クラウドネットワーキングの感覚」を掴んで頂きます

目次

- Amazon Virtual Private Cloud (VPC)
- パケットの気持ちで辿り、設計を考えてみましょう
- まとめ

Amazon Virtual Private Cloud (VPC)

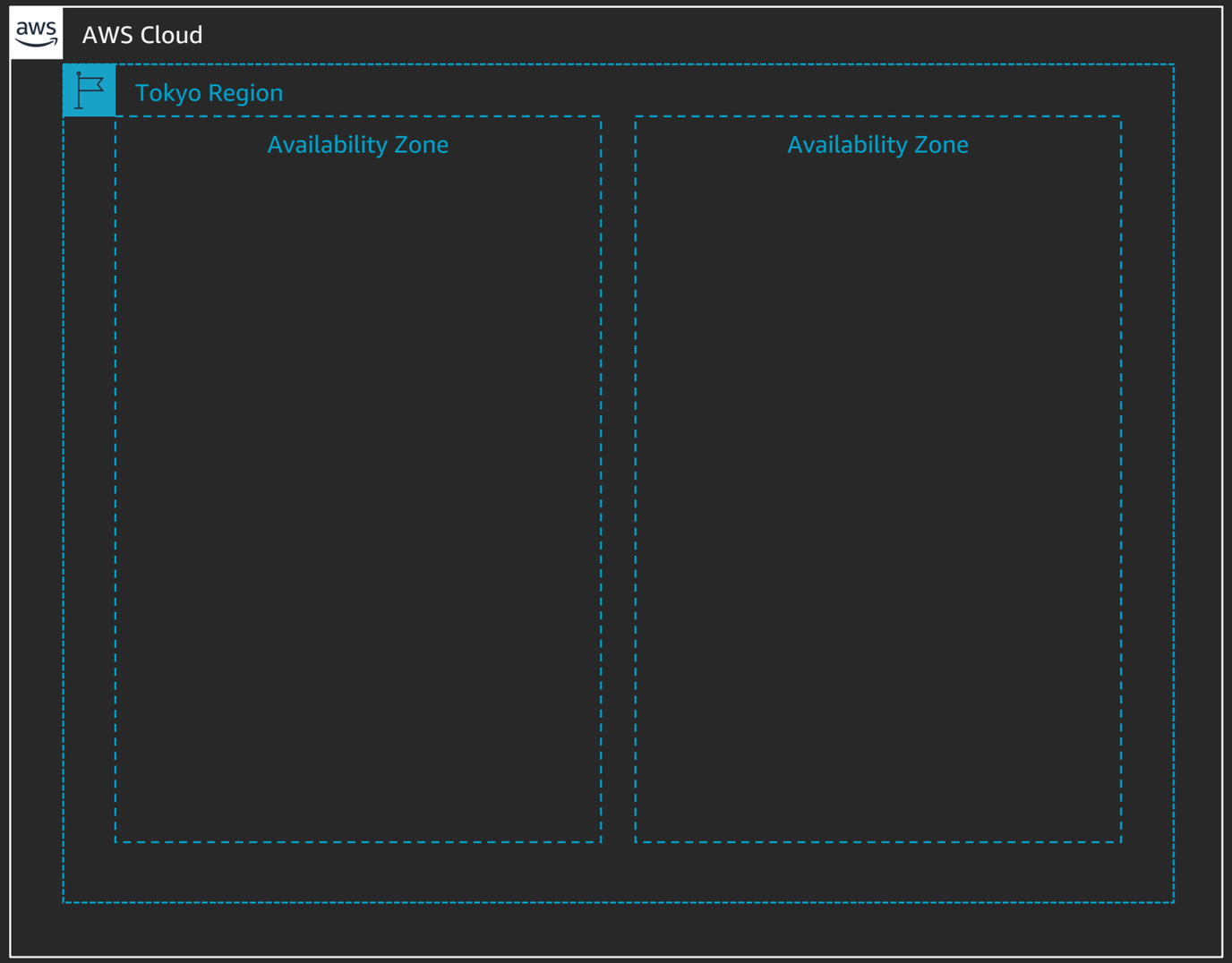
VPC の主要な概念

リージョン

- AWS サービスを提供するデータセンタークラスター

アベイラビリティゾーン (AZ)

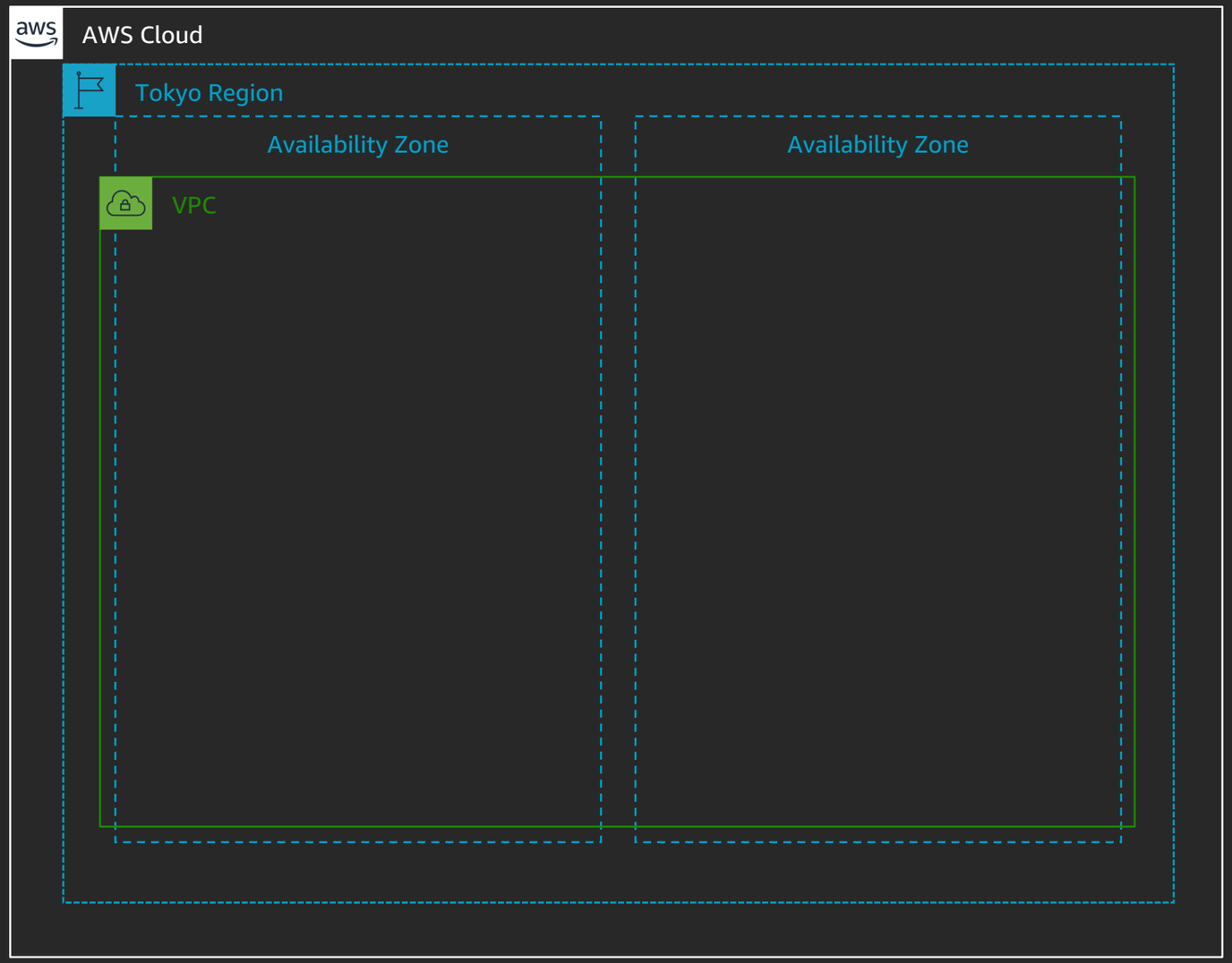
- リージョンのサブセット
- 各 AZ は電源、ネットワークなどが独立
- 高可用性のために Multi-AZ 構成を推奨



VPC の主要な概念

VPC

- 独立したプライベートネットワーク領域
- 組織/アクセス要件/サービス等に応じて作成

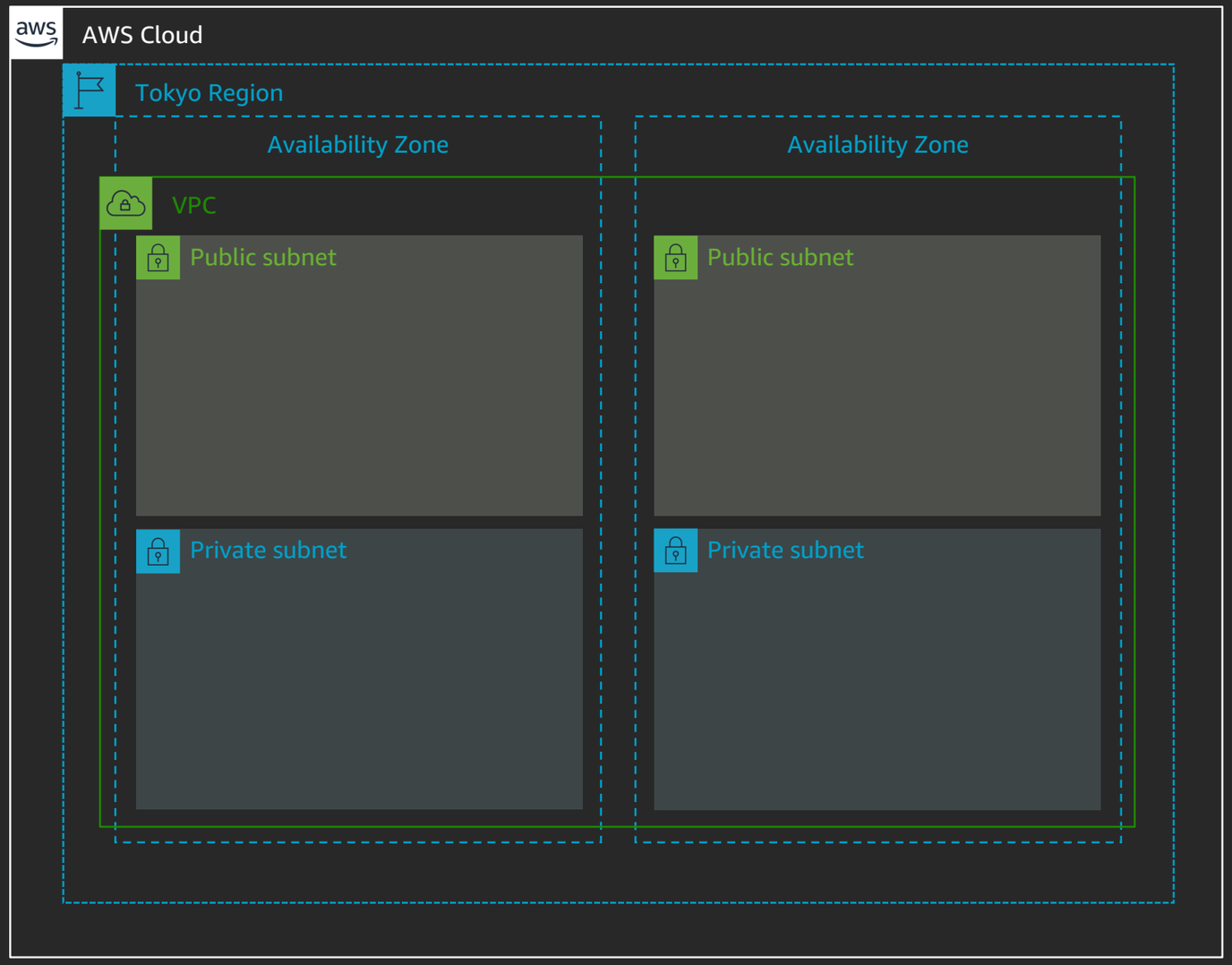


VPC の主要な概念

VPC

サブネット

- VPC 内でルーティングポリシーに従って切り出す
- 各 AZ に作成し可用性を担保



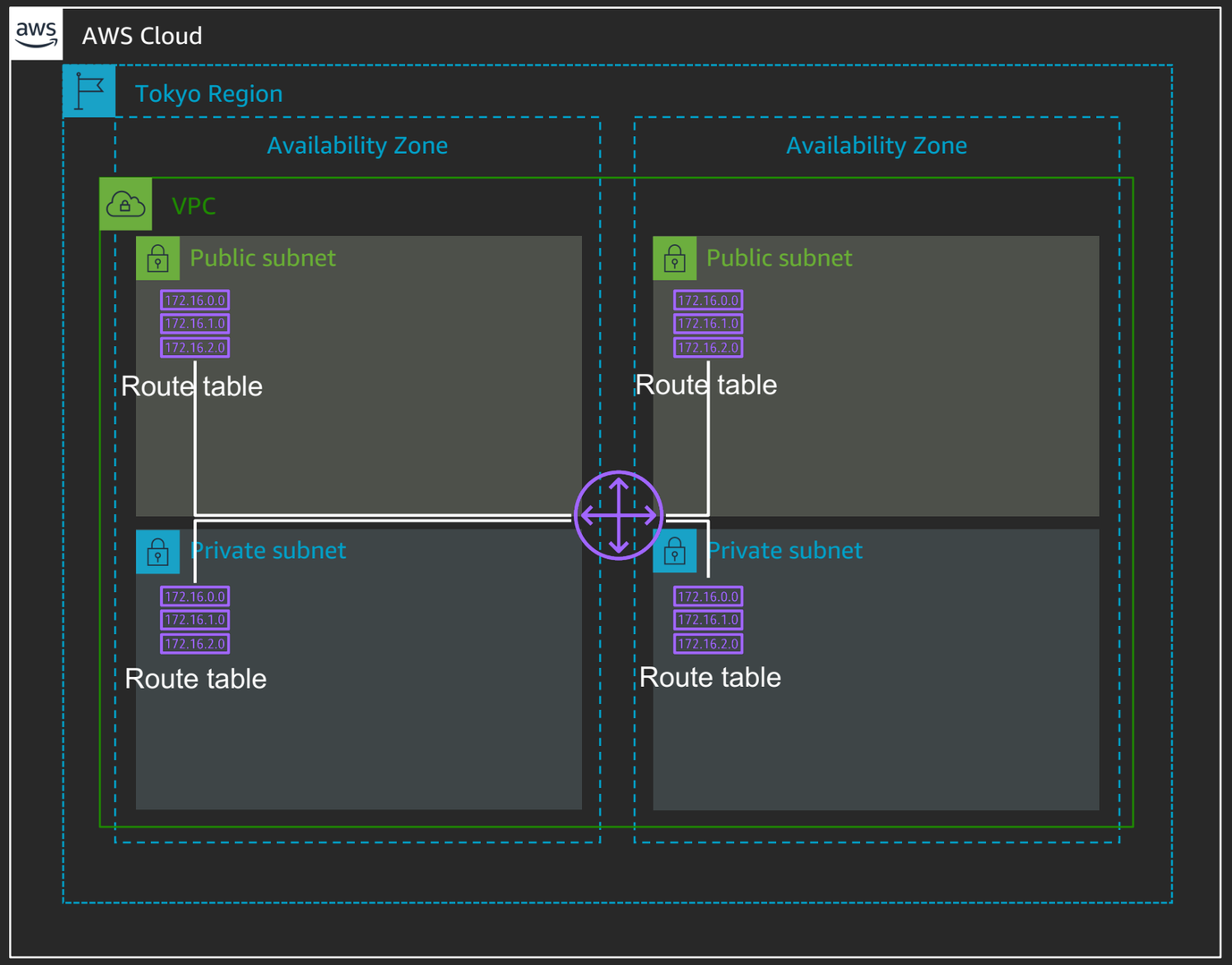
VPC の主要な概念

VPC

サブネット

ルートテーブル

- VPC に暗黙的に存在するルーター の経路設定
- EC2 インスタンスから見るとデフォルトゲートウェイにあたる



VPC の主要な概念

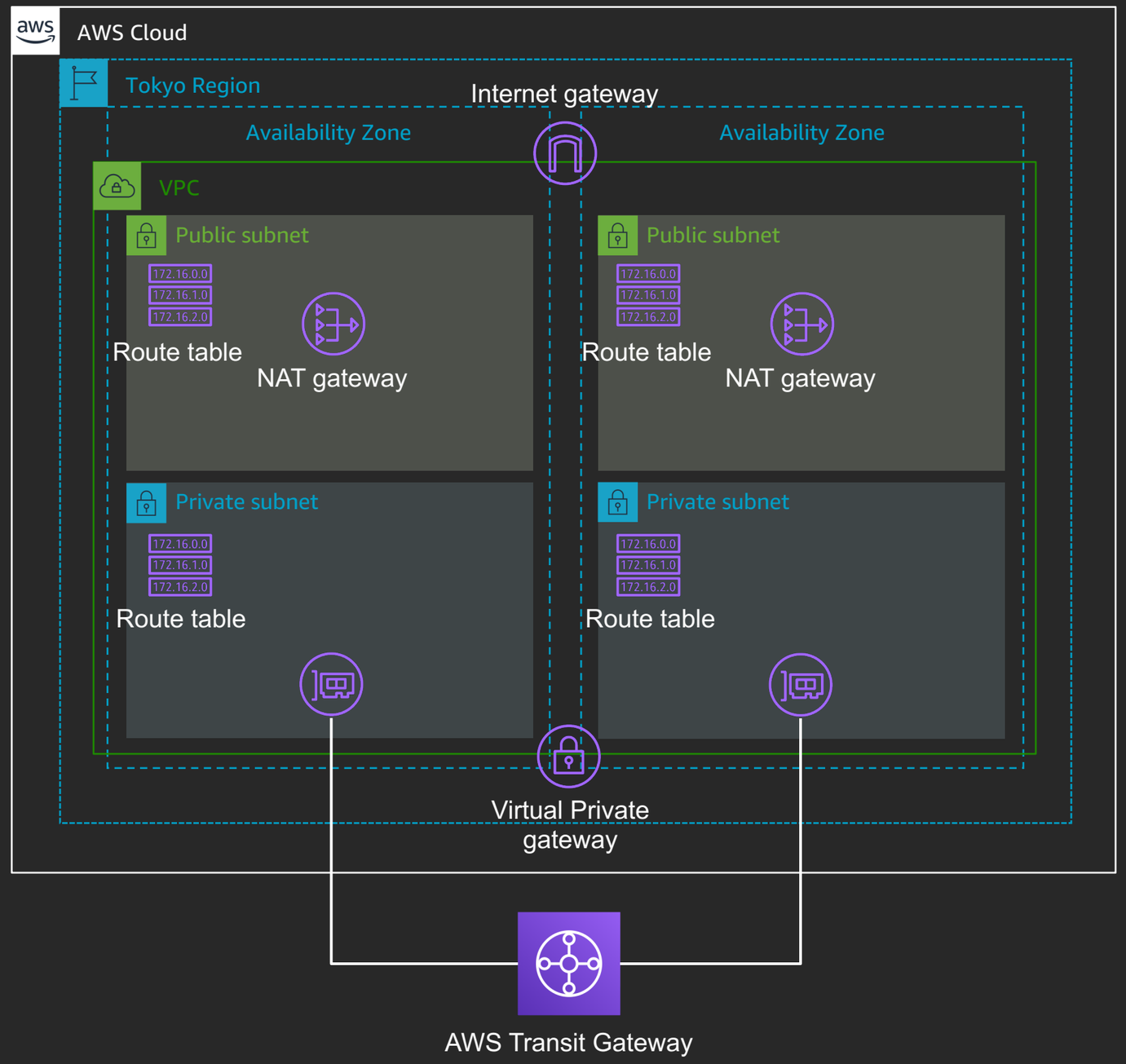
VPC

サブネット

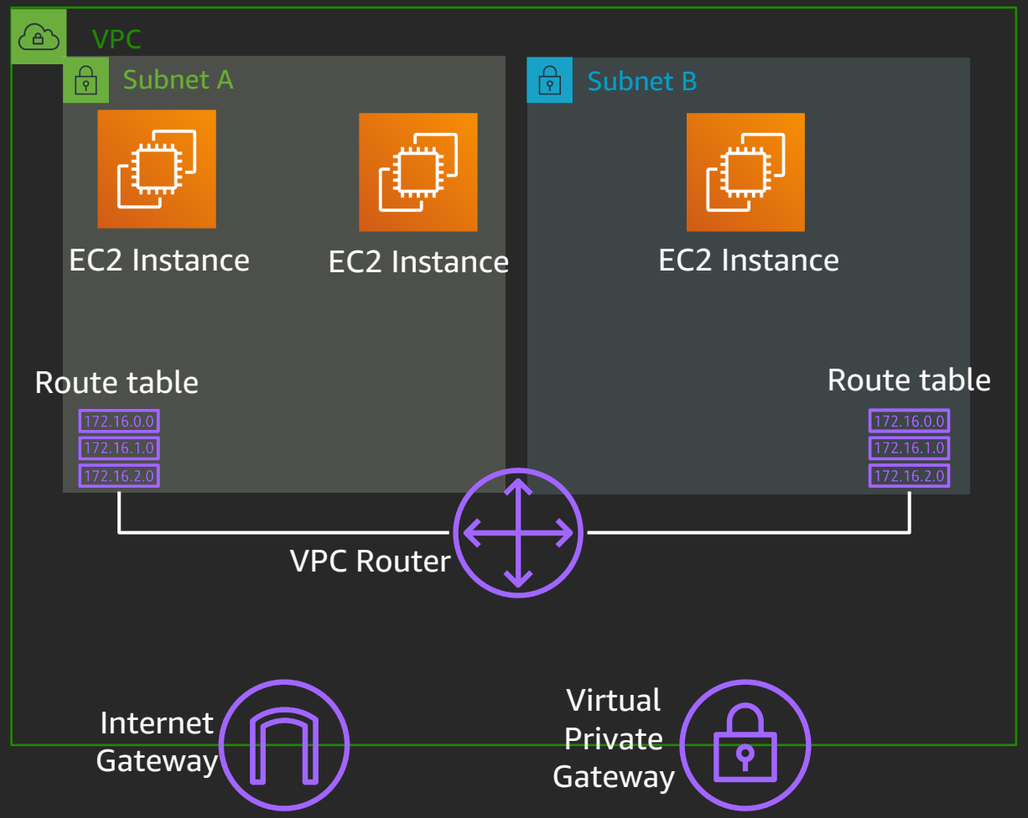
ルートテーブル

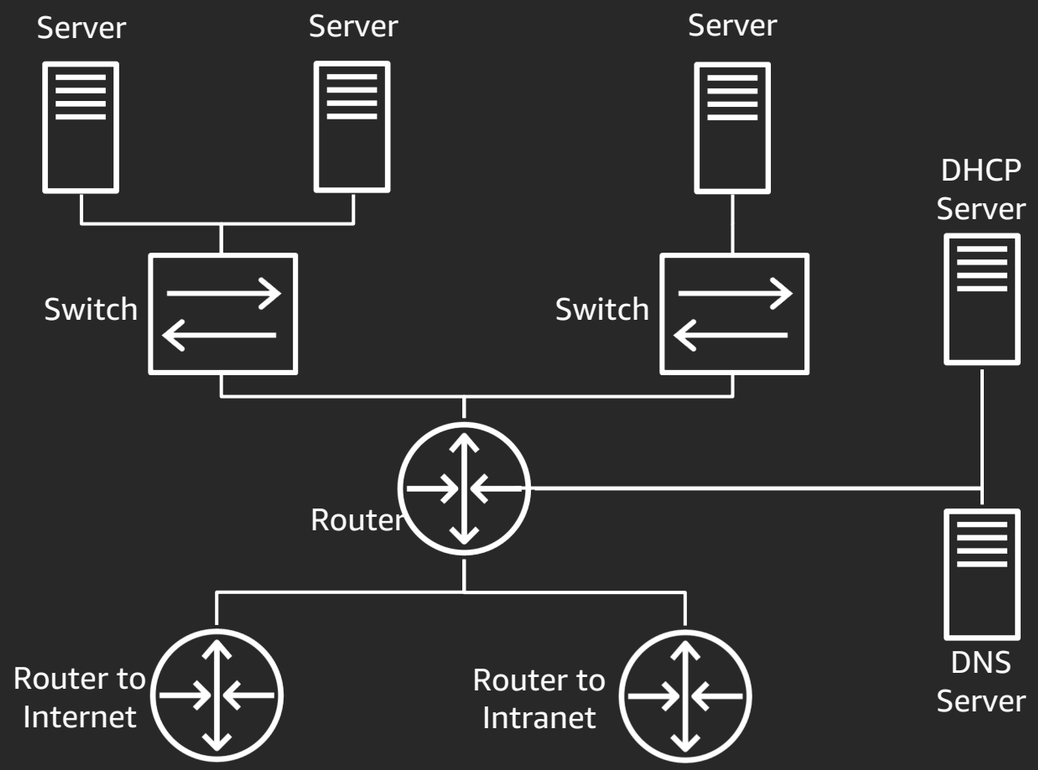
ゲートウェイ各種

- VPC にアタッチする論理コンポーネント
- 内部で冗長化済
- VPC 内で単一オブジェクトとして見えるタイプ or サブネット毎のエンドポイントを意識するタイプ



オンプレミスネットワークとの比較で理解する VPC





VPC を作る ≒ ルーターと DNS を設定する

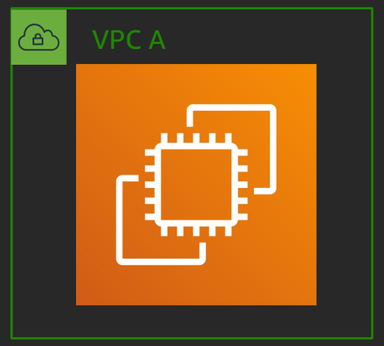
サブネットを作る ≒ スイッチ (VLAN) と DHCP を設定する

EC2 を作る ≒ サーバーに OS インストール及び初期設定を施し、スイッチに接続する

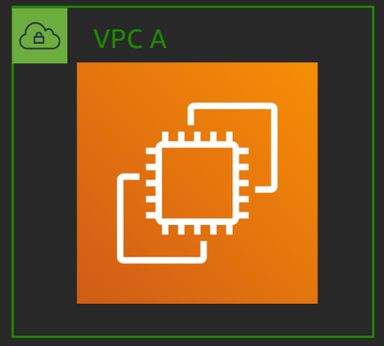
VPC 通信の実際

Dst Layer 2 ヘッダー
Dst Layer 3 ヘッダー

EC2 A



EC2 B



Host A

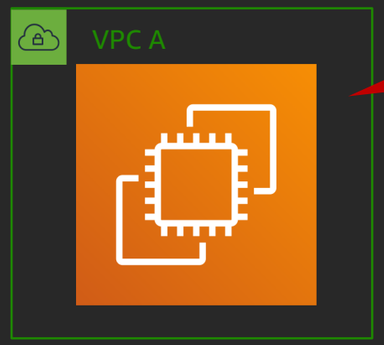
Host B

マッピングサービス

VPC 通信の実際

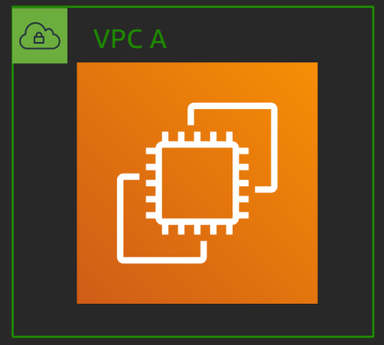
Dst Layer 2 ヘッダー
Dst Layer 3 ヘッダー

EC2 A



EC2 B に HTTP で
アクセスしたい！

EC2 B

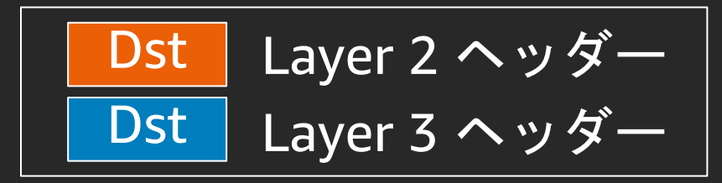


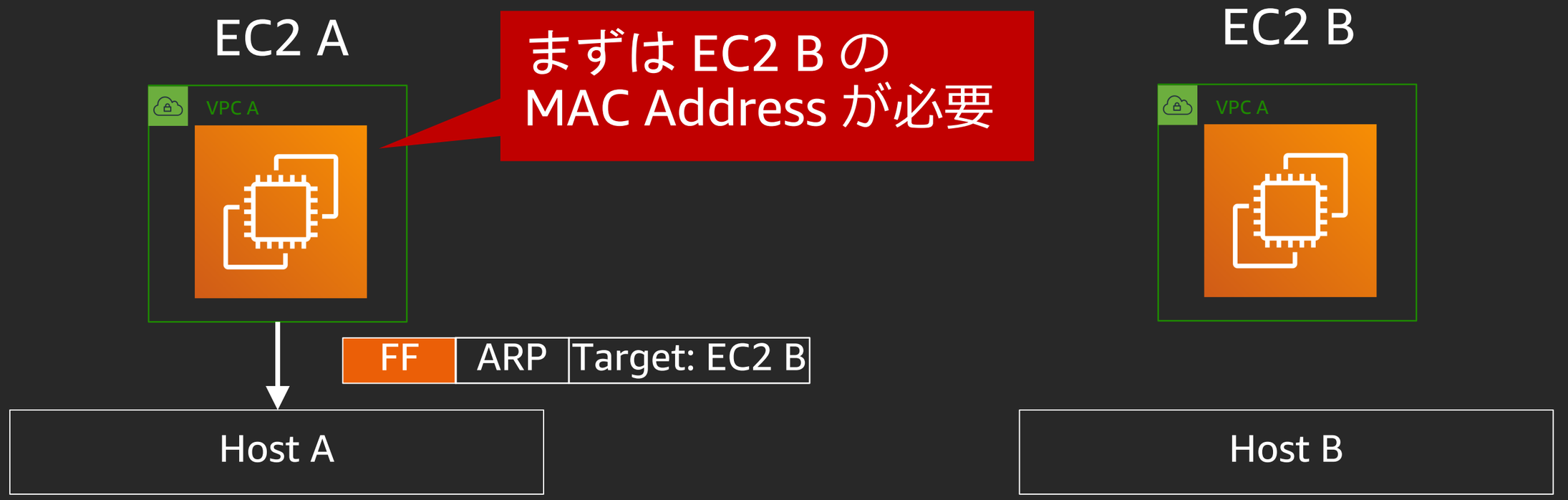
Host A

Host B

マッピングサービス

VPC 通信の実際





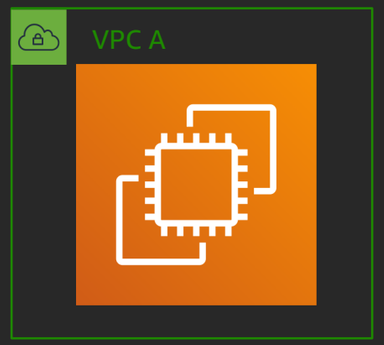
マッピングサービス

* EC2 B が異なるサブネットにある場合は EC2 A は Subnet A のゲートウェイ MAC Address を問い合わせる

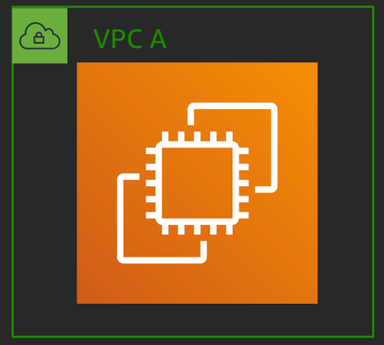
VPC 通信の実際

Dst Layer 2 ヘッダー
Dst Layer 3 ヘッダー

EC2 A



EC2 B



ARP を受けて
マッピングサービス
に問い合わせ

Host A

Host B

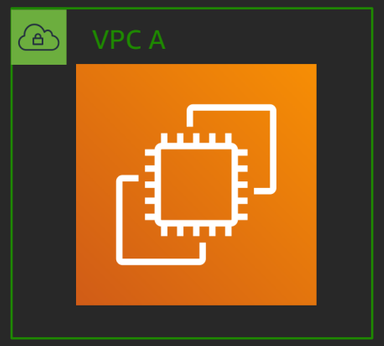
Query: EC2 B の MAC Address と
Host の IP Address

マッピングサービス

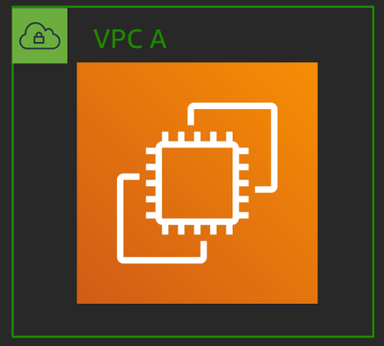
VPC 通信の実際

Dst Layer 2 ヘッダー
Dst Layer 3 ヘッダー

EC2 A



EC2 B



Host A

Host B

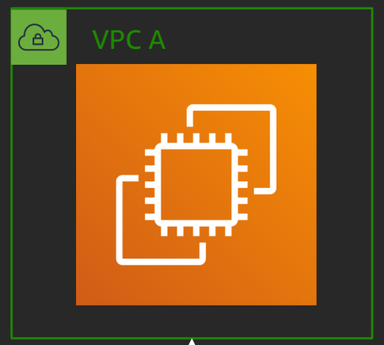
Reply: EC2 B の MAC Address と
Host の IP Address

マッピングサービス

VPC 通信の実際

Dst Layer 2 ヘッダー
Dst Layer 3 ヘッダー

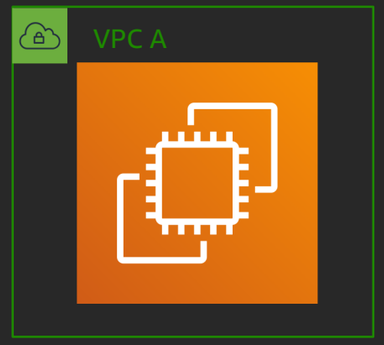
EC2 A



EC2 A | ARP | EC2 B MAC

Host A

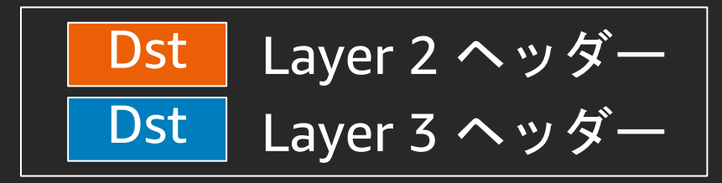
EC2 B



Host B

マッピングサービス

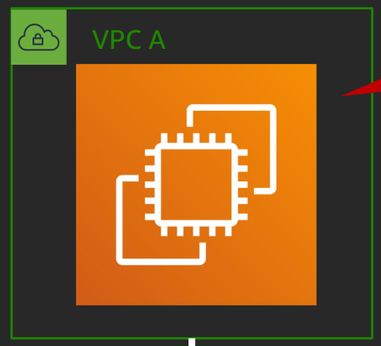
VPC 通信の実際

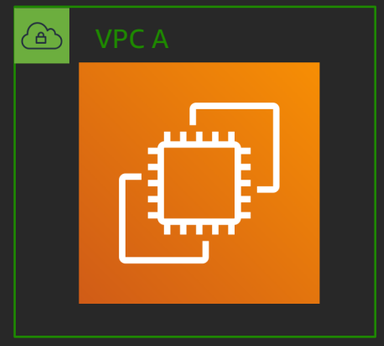


EC2 A

EC2 B に HTTP リクエストを送信！

EC2 B





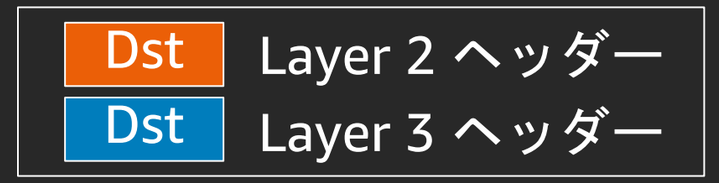
Host A

Host B

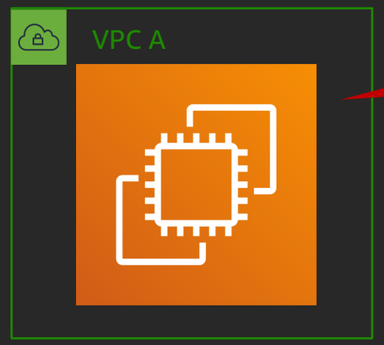
マッピングサービス

* EC2 B が異なるサブネットにある場合は Destination MAC は Subnet A のゲートウェイ MAC Address にセットされる

VPC 通信の実際

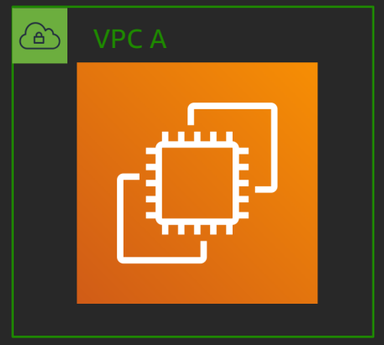


EC2 A



レスポンスを待機

EC2 B



Host A

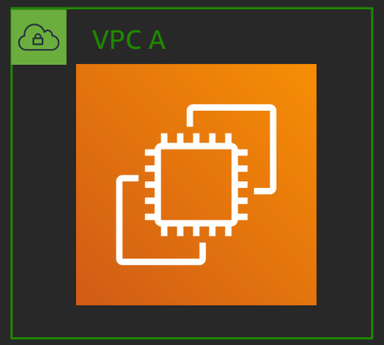
Host B

マッピングサービス

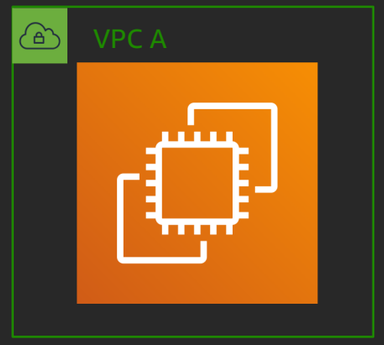
VPC 通信の実際

Dst Layer 2 ヘッダー
Dst Layer 3 ヘッダー

EC2 A



EC2 B



正規の通信かどうか
Validation

Host A

Host B

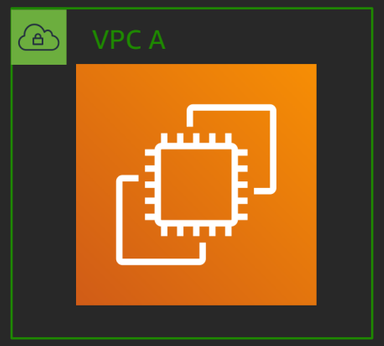
Query: EC2 A は Host A 上の
正規のインスタンスか?

マッピングサービス

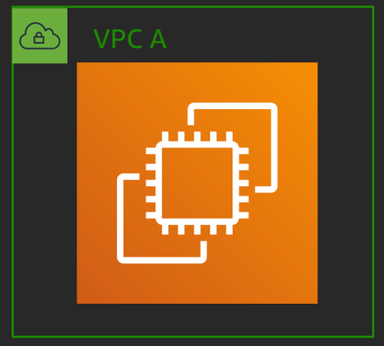
VPC 通信の実際

Dst Layer 2 ヘッダー
Dst Layer 3 ヘッダー

EC2 A



EC2 B



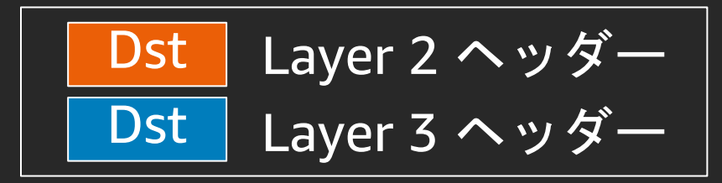
Host A

Host B

Reply: EC2 A は Host A 上の
正規のインスタンス

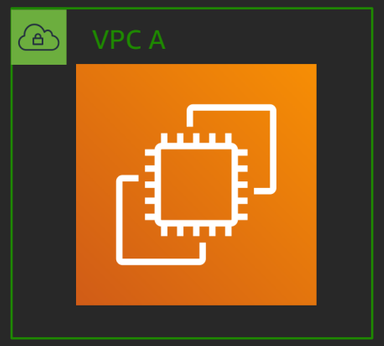
マッピングサービス

VPC 通信の実際

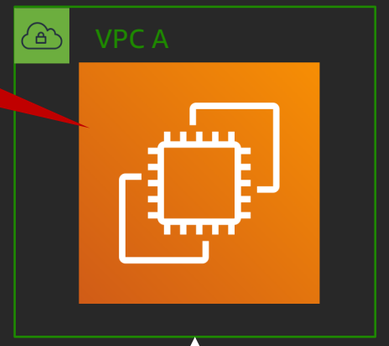


EC2 A

EC2 B



リクエストを受信！



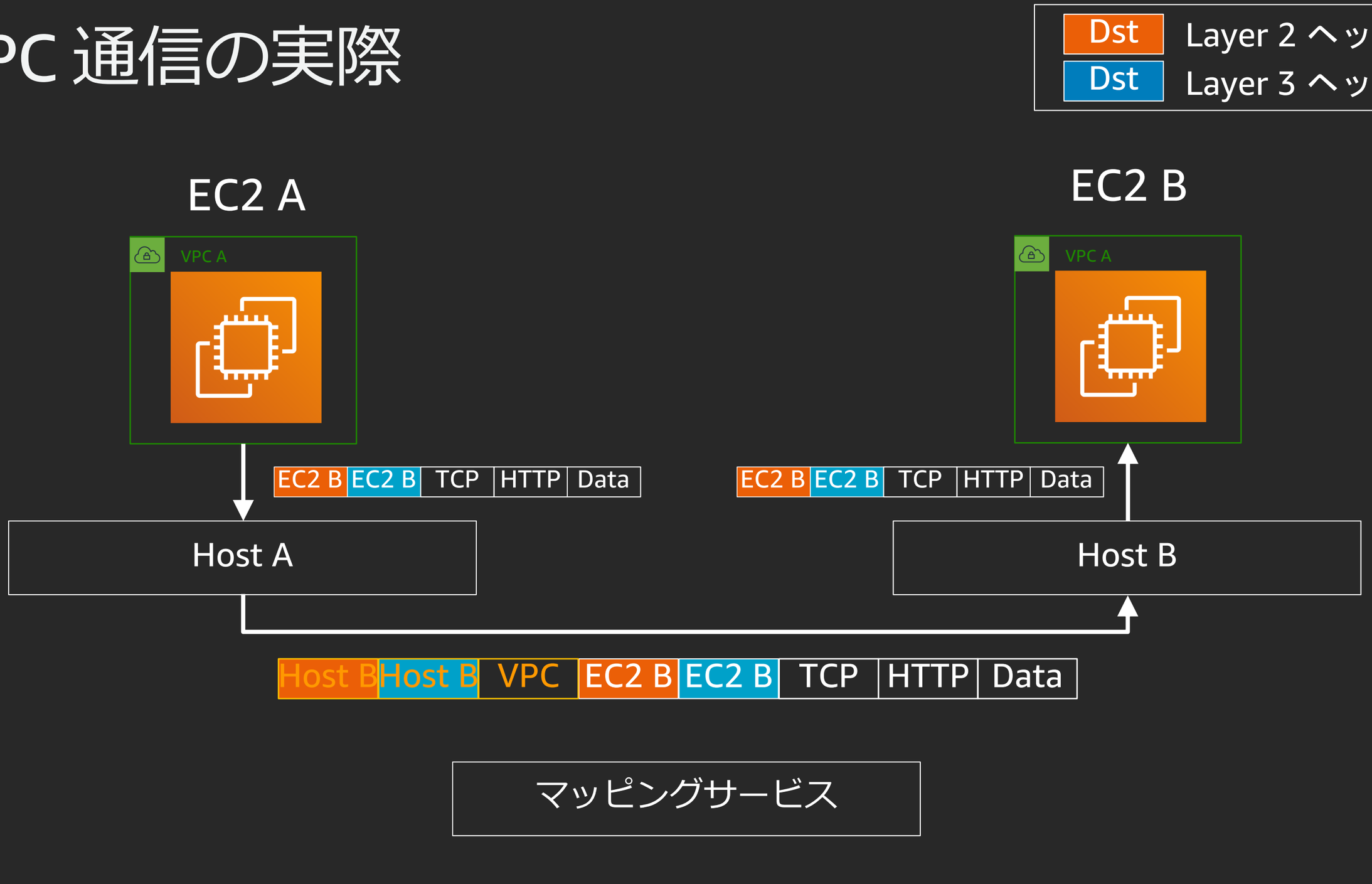
Host A

Host B

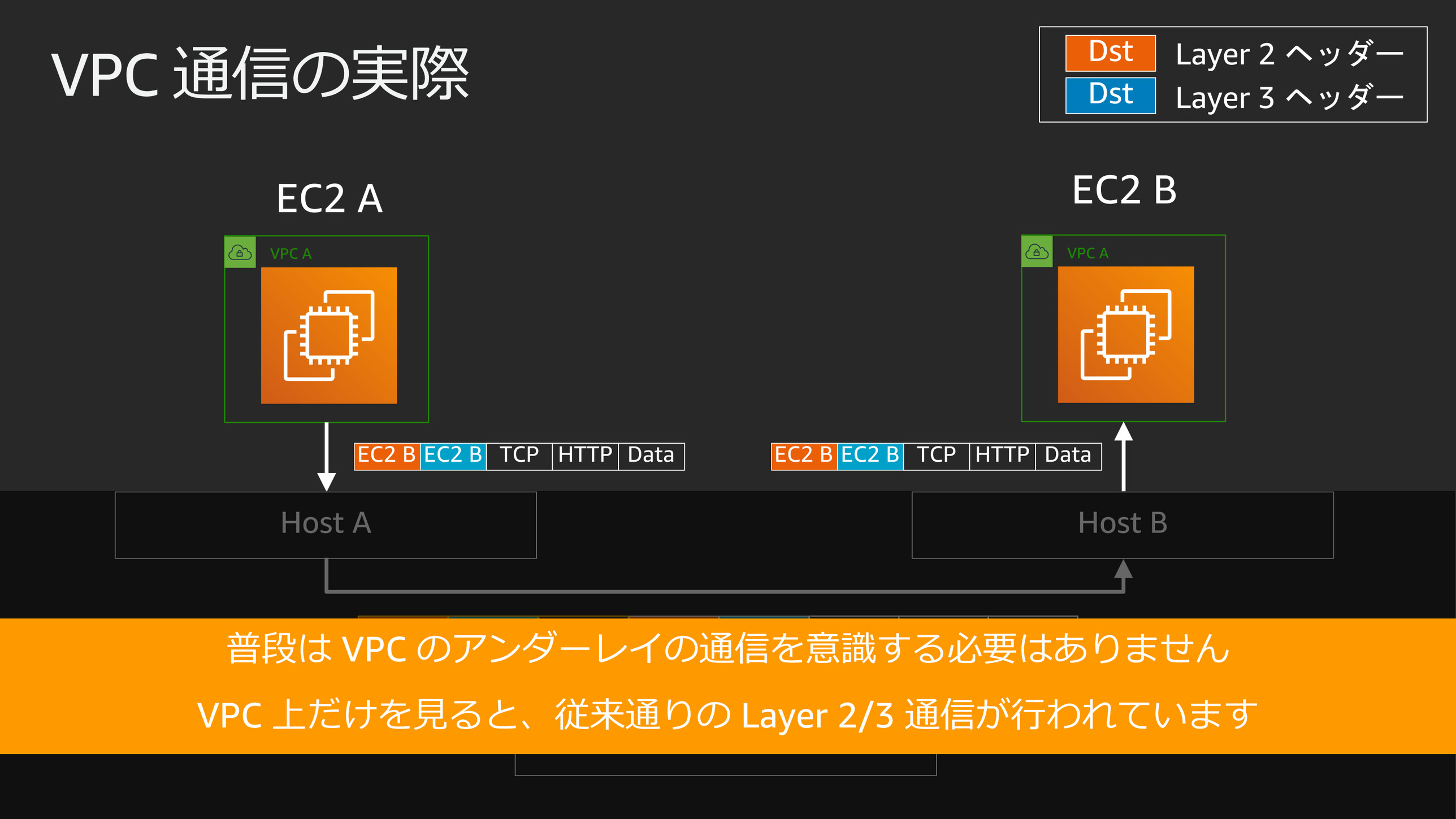
マッピングサービス

* EC2 B が異なるサブネットにある場合は Source MAC を Subnet B のゲートウェイ MAC Address に、Destination MAC を EC2 B MAC Address に変換

VPC 通信の実際


Dst
Layer 2 ヘッダー
Dst
Layer 3 ヘッダー
EC2 A
EC2 B
VPC A
VPC A
EC2 B
EC2 B
TCP
HTTP
Data
EC2 B
EC2 B
TCP
HTTP
Data
Host A
Host B
Host B
Host B
VPC
EC2 B
EC2 B
TCP
HTTP
Data
マッピングサービス

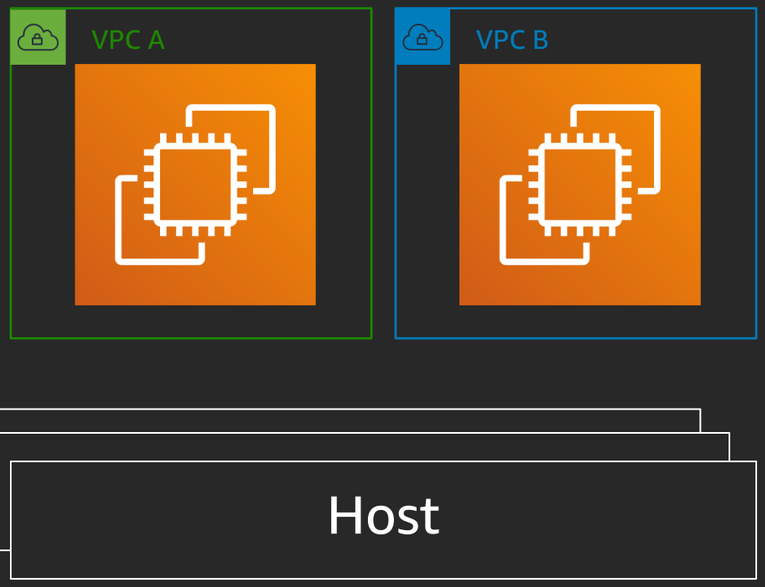
VPC 通信の実際

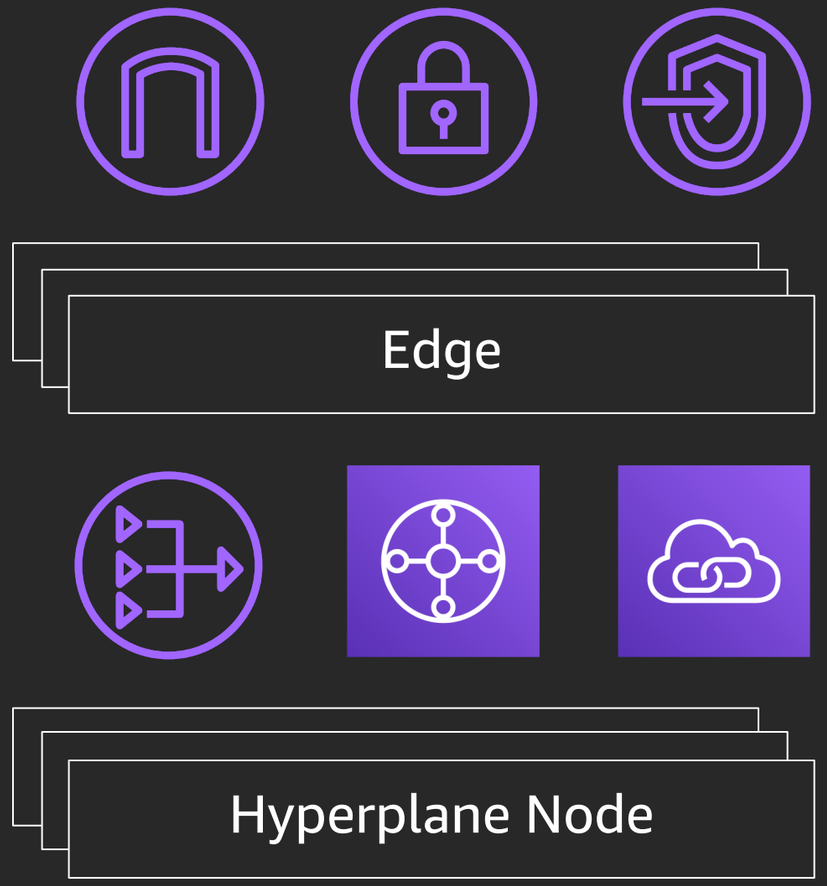


普段は VPC のアンダーレイの通信を意識する必要はありません

VPC 上だけを見ると、従来通りの Layer 2/3 通信が行われています

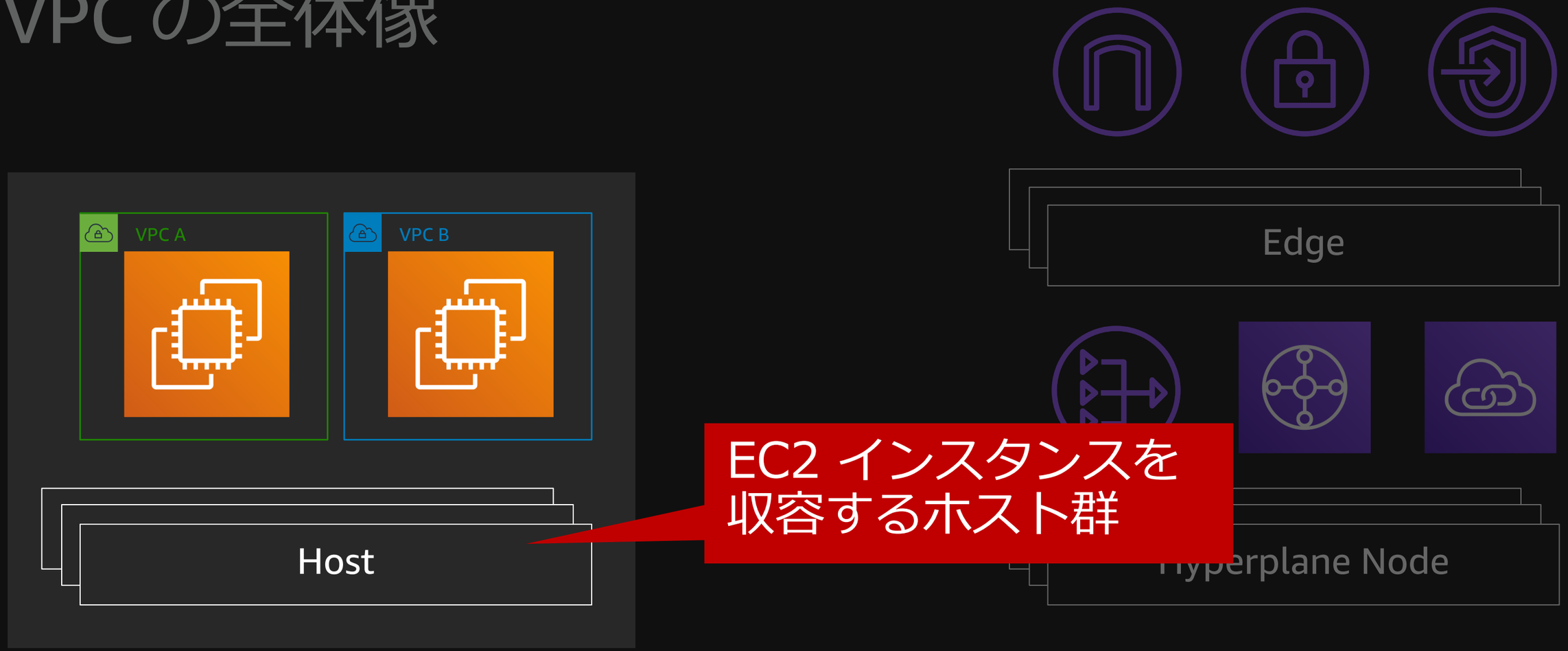
VPC の全体像





マッピングサービス

VPC の全体像

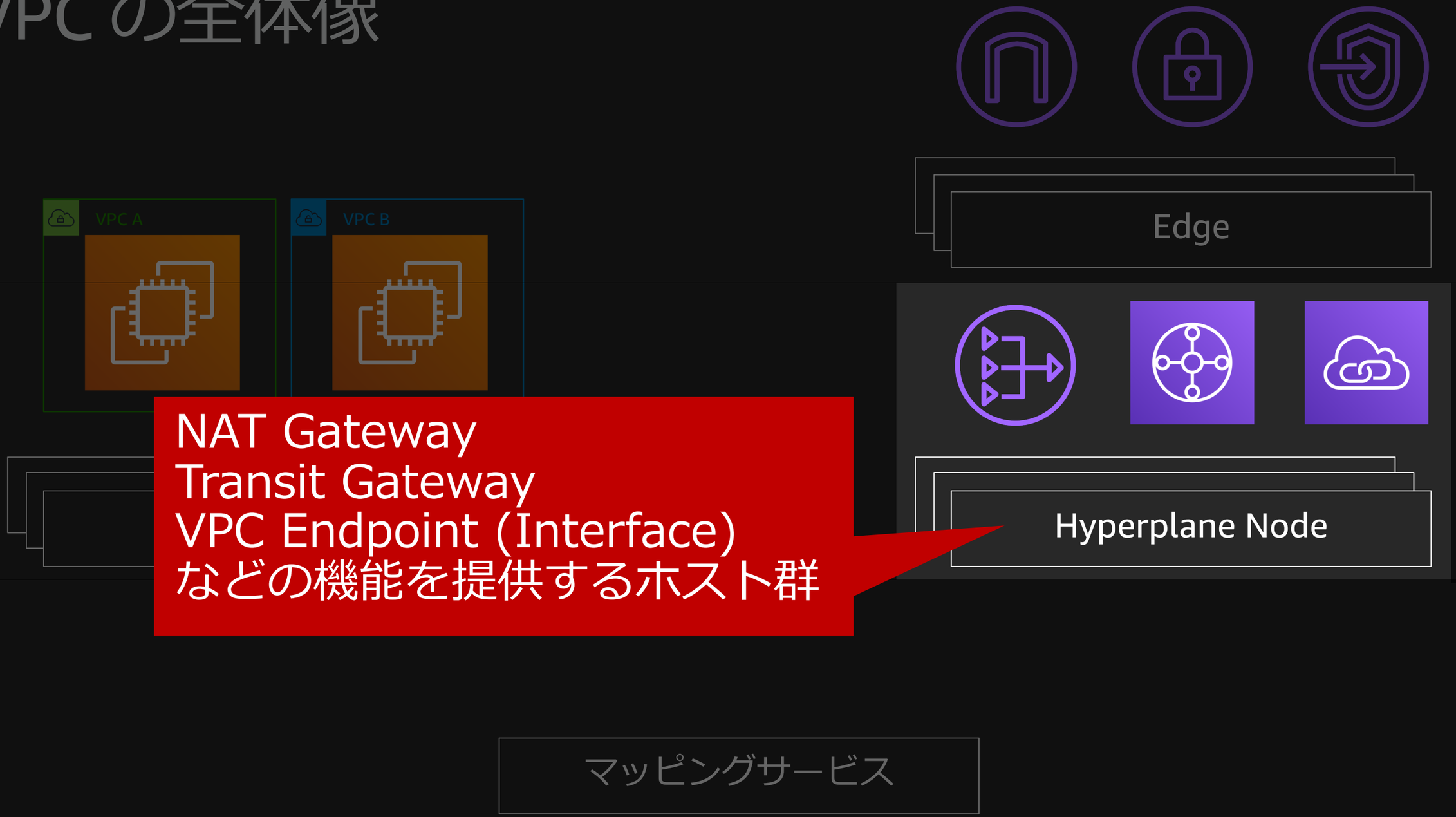


マッピングサービス

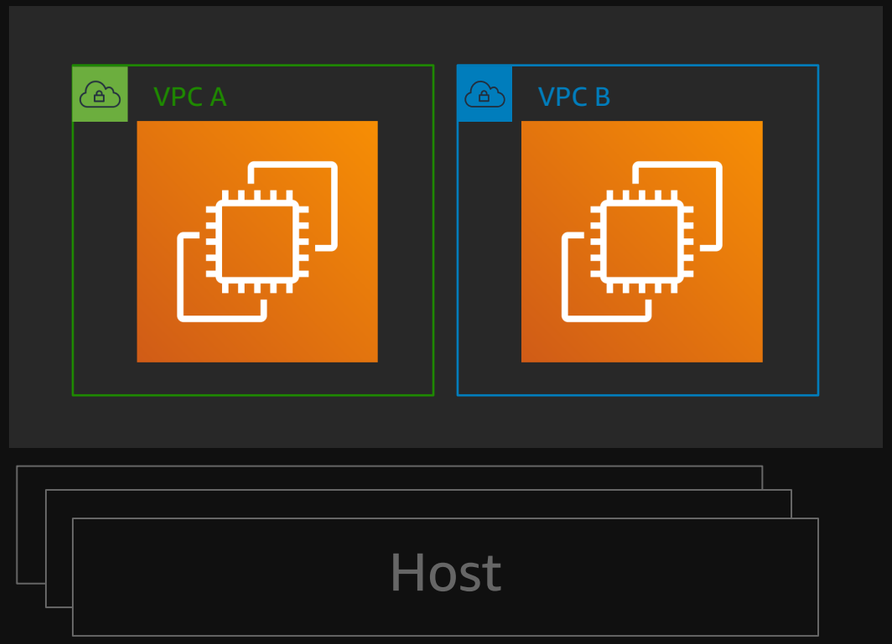
VPC の全体像

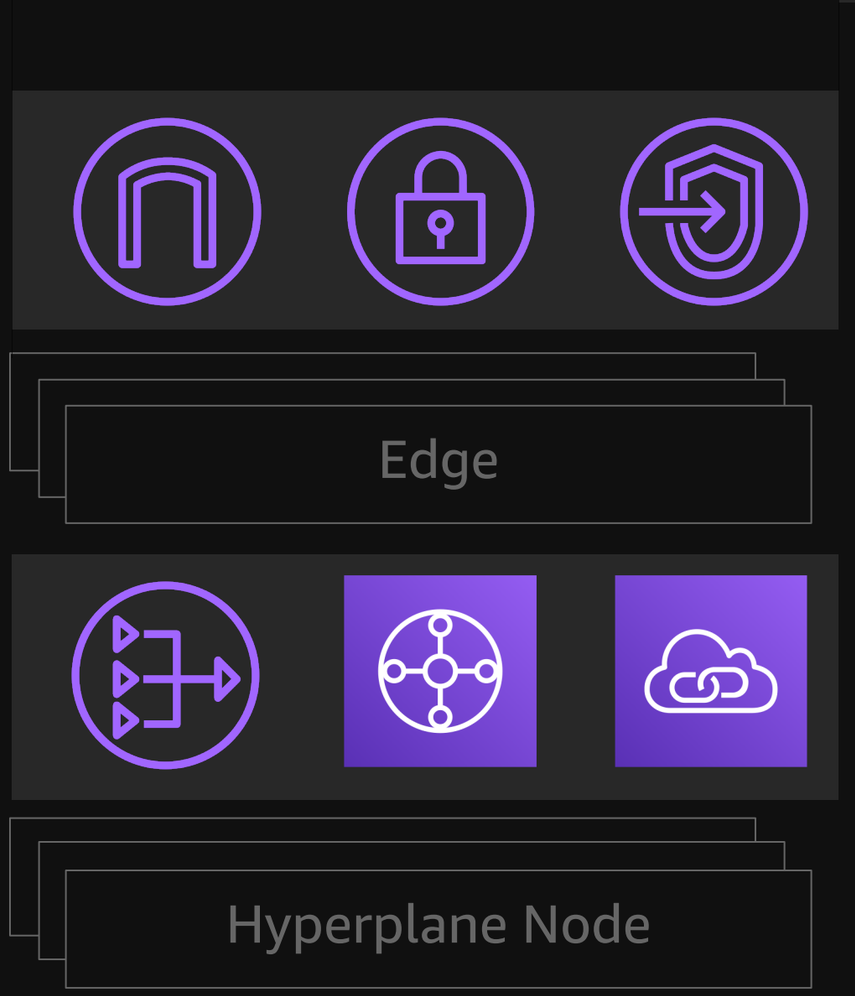

Internet Gateway
Virtual Private Gateway
VPC Endpoint (Gateway)
などの機能を提供するホスト群
VPC A
Edge
Host
Hyperplane Node
マッピングサービス

VPC の全体像


VPC A
VPC B
Edge
NAT Gateway
Transit Gateway
VPC Endpoint (Interface)
などの機能を提供するホスト群
Hyperplane Node
マッピングサービス

VPC の全体像



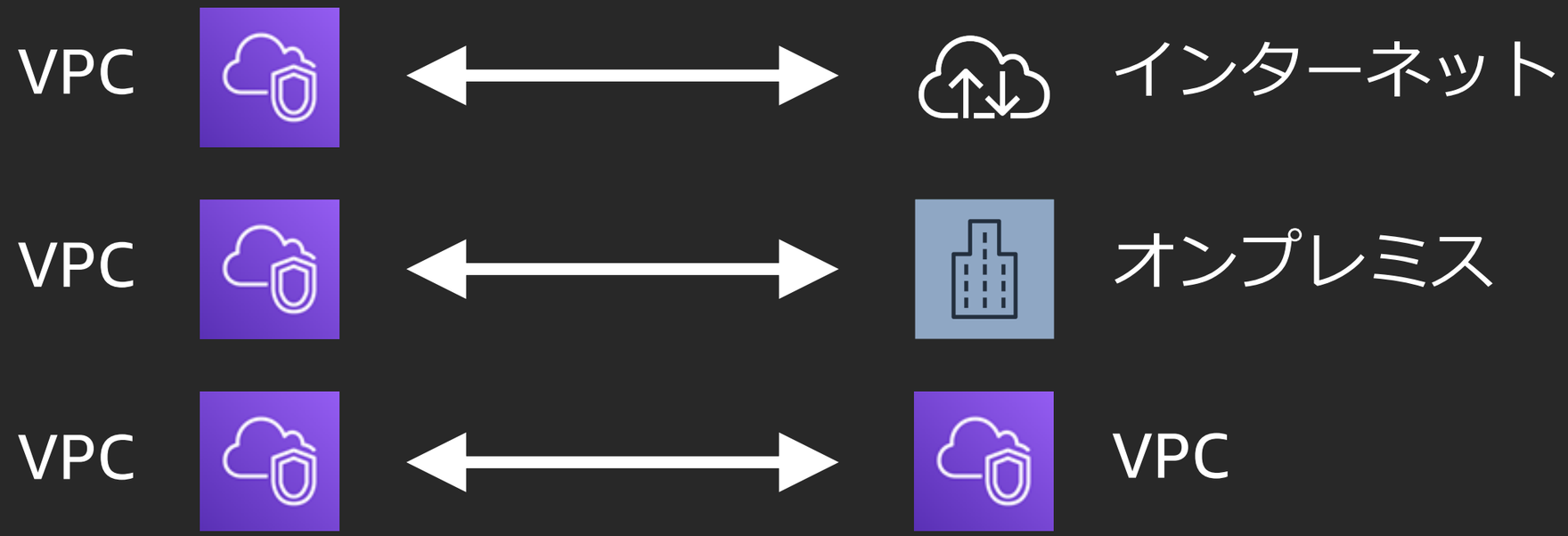


普段は VPC のアンダーレイの通信を意識する必要はありません

VPC 上だけを見ると、従来通りの Layer 2/3 通信が行われています

クラウドネットワークを設計するときは

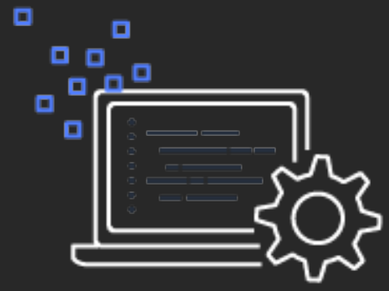
主な設計パターン



「ルーターがどこにあるか」「どういう経路情報を持っているか」が鍵です

パケットの気持ちになって経路を辿り、どんな設計、設定が必要かを考えてみましょう

パケットの気持ちで辿り
設計を考えてみましょう



設定項目

ヒント

※ 冗長化構成は省略します

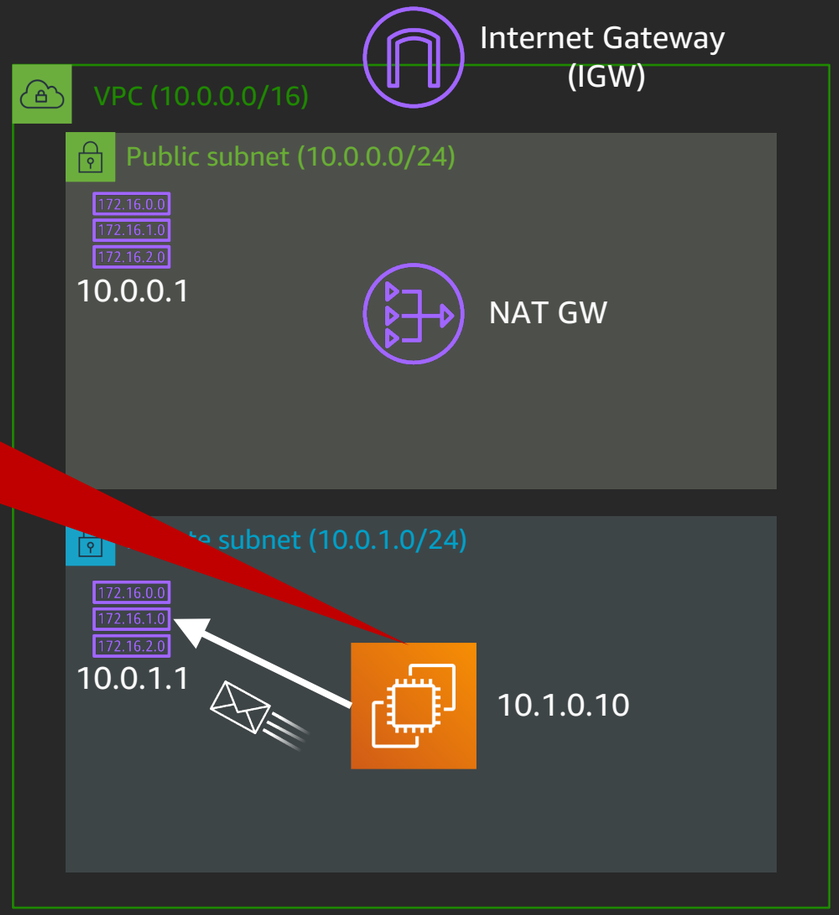
※ 戻りの通信は省略します

1. VPC 内から NAT GW を介して Internet にアクセス

1. www.amazon.co.jp に通信したい
2. DNS により解決した 104.78.77.230 に通信したい
3. 自サブネット外なのでデフォルトゲートウェイに送信

DNS は VPC 内の暗黙の DNS Resolver (Route 53 Resolver) に問い合わせることで実現

DNS の具体的なシーケンスはここでは割愛します



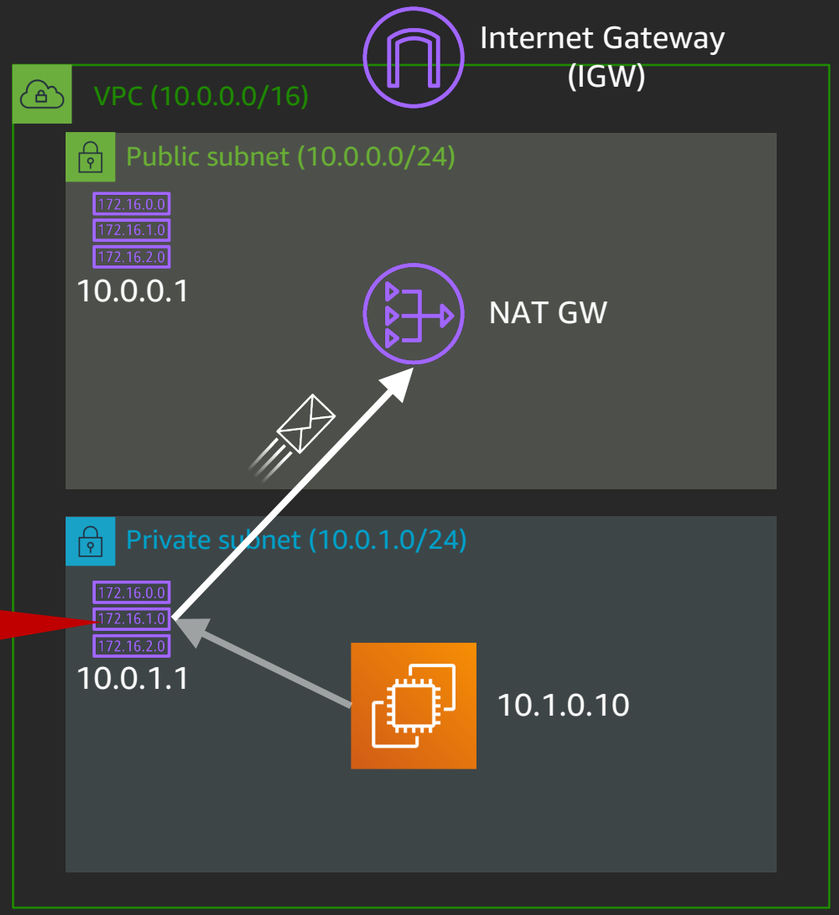
1. VPC 内から NAT GW を介して Internet にアクセス

Private Subnet のルートテーブルに VPC 外宛通信用のエントリーを作成

4. ルートテーブルを参照

送信先	ターゲット
10.0.0.0/16	local
0.0.0.0/0	**NAT GW**

5. 一致したエントリーに従い NAT GW に送信

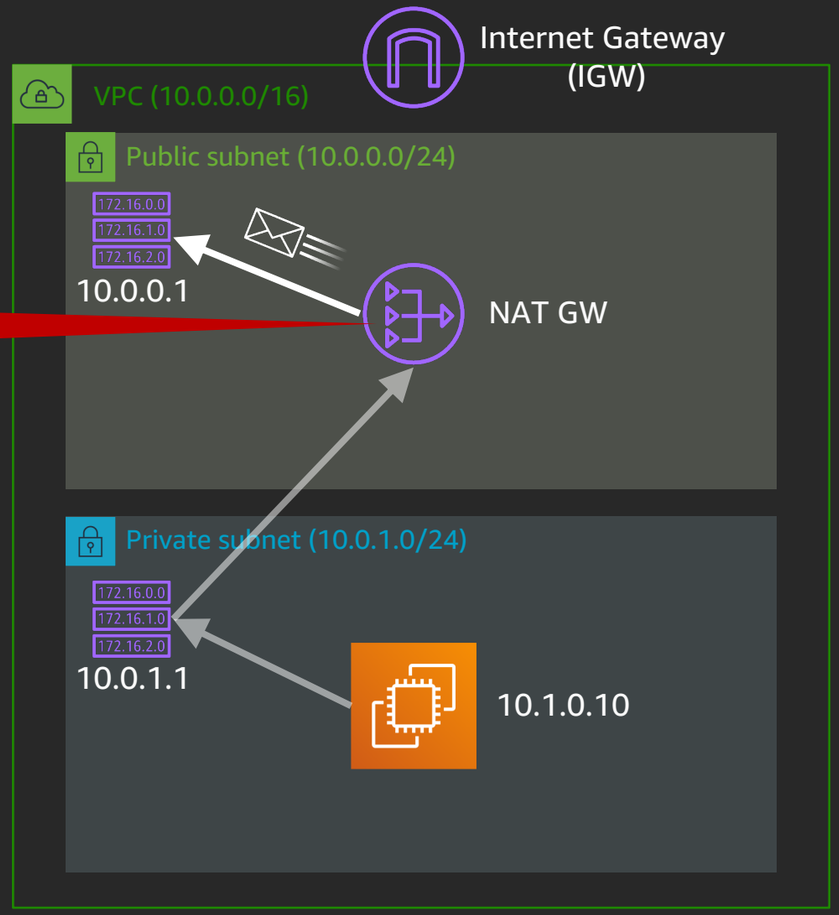


1. VPC 内から NAT GW を介して Internet にアクセス

6. 自サブネット外なのでデフォルトゲートウェイに送信

NAT GW で Src IP Address が NAT GW の Public IP Address に変換される

NAT GW はルーティングは行わない



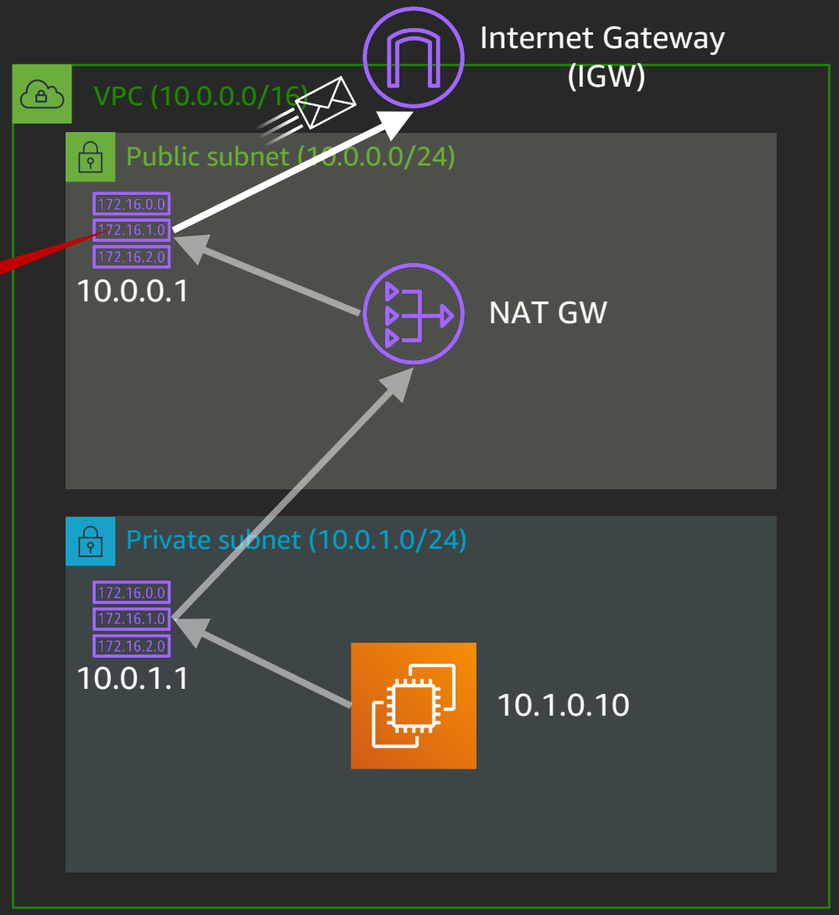
1. VPC 内から NAT GW を介して Internet にアクセス

Public Subnet のルートテーブルに VPC 外宛通信用のエントリーを作成

7. ルートテーブルを参照

送信先	ターゲット
10.0.0.0/16	local
0.0.0.0/0	**IGW**

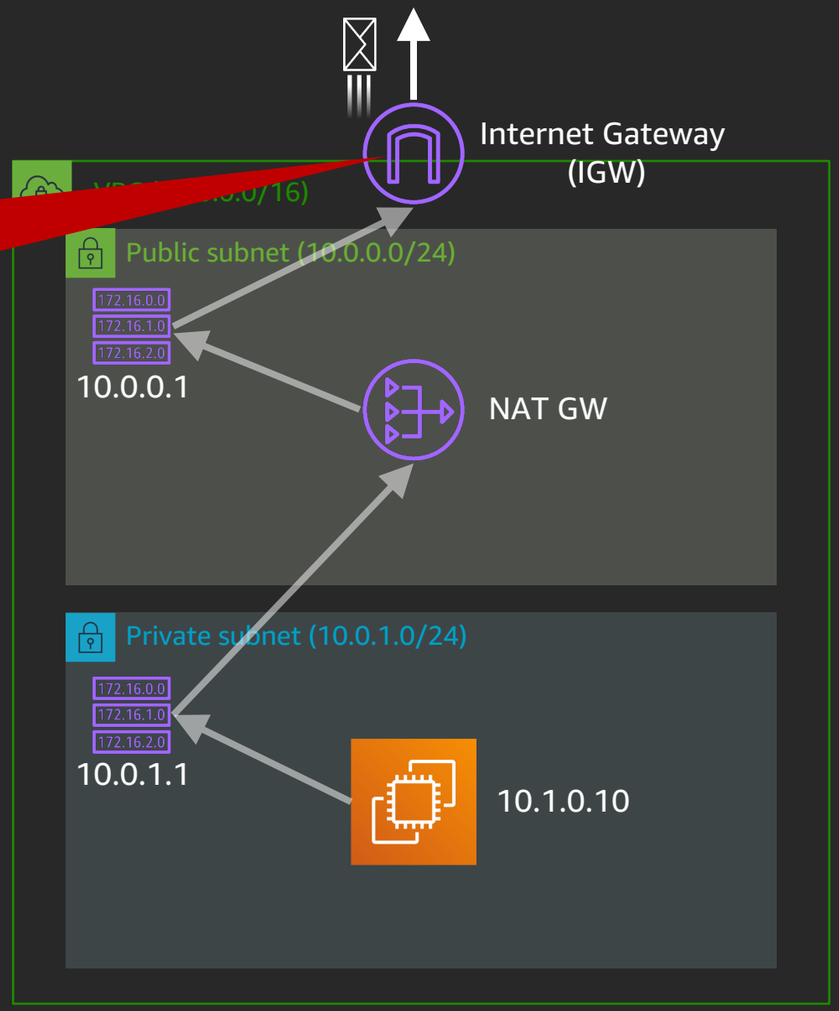
8. 一致したエントリーに従い IGW に送信



1. VPC 内から NAT GW を介して Internet にアクセス

9. 自身の経路情報を参照
10. 一致したエントリーに従い
    Next hop に送信
11. それを繰り返し
    104.78.77.230 に到達

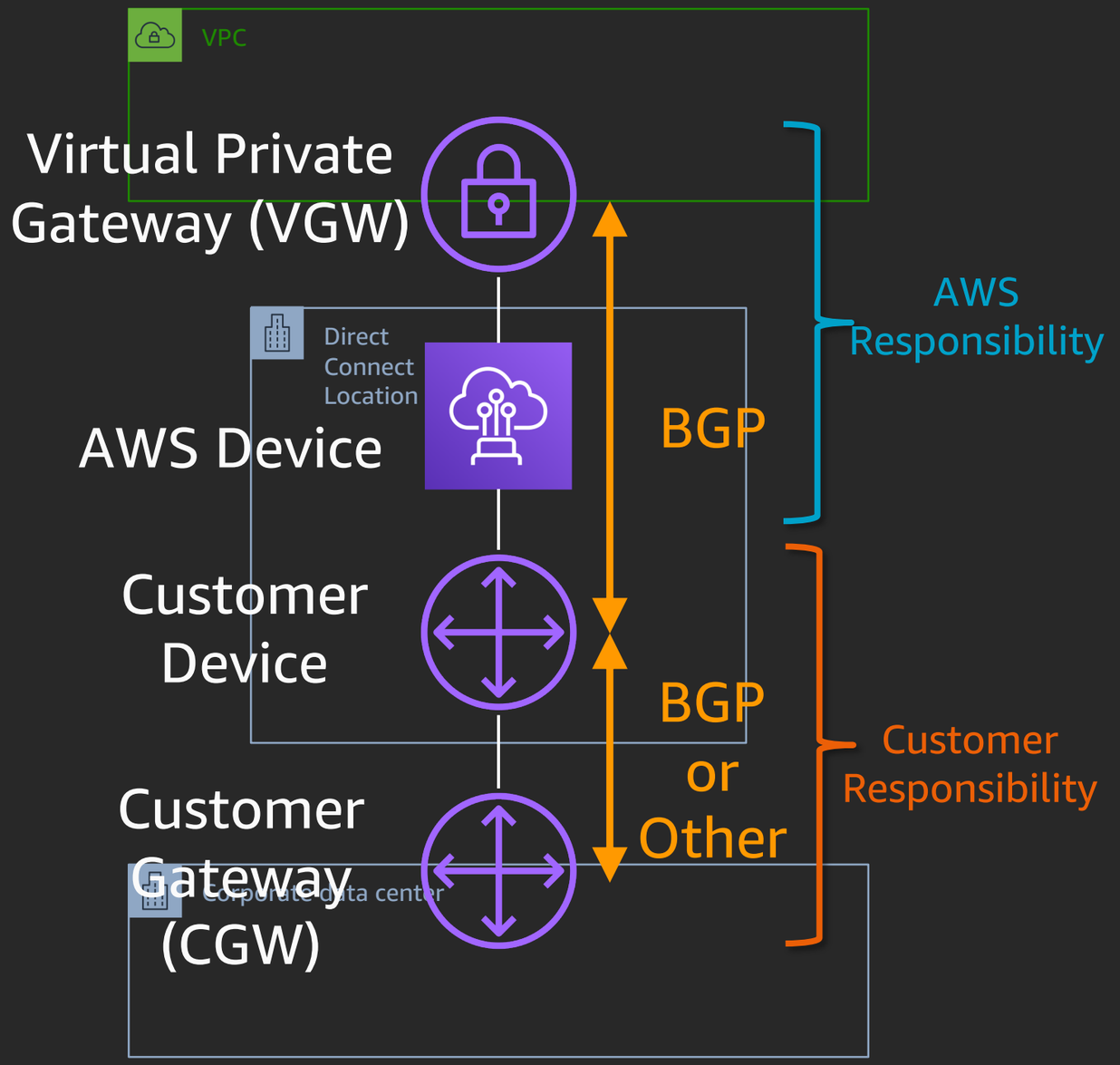
デフォルトでは IGW が持つ経路情報はユーザーには参照できない



2. VPC 内から VGW を介してオンプレミスにアクセス

AWS Direct Connect の基本

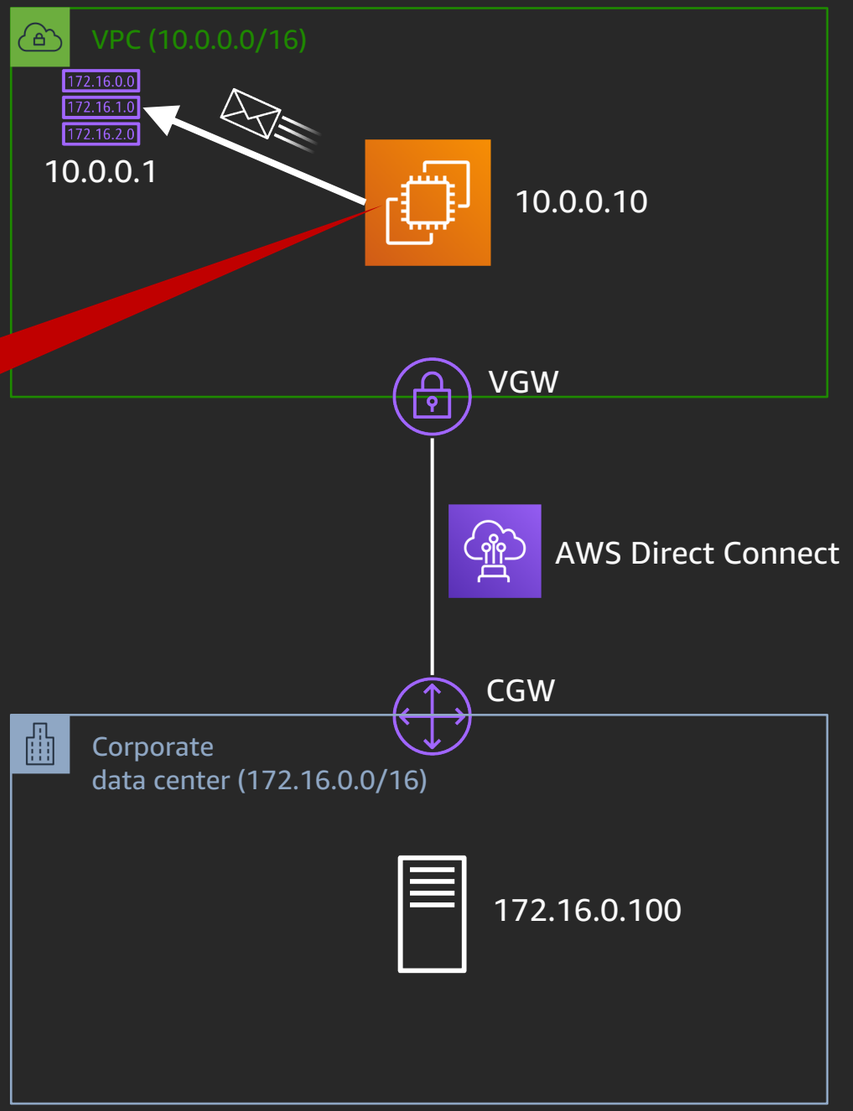
- 専用線や閉域網で自社拠点と AWS を接続するためのサービス
- Direct Connect Location を介して物理接続
  - Location 冗長化を推奨
- その上に仮想インターフェースを定義し、 BGP で経路交換
- AWS 側の終端は VGW または Direct Connect Gateway の選択肢がある(本セッションでは前者を扱います)



2. VPC 内から VGW を介してオンプレミスにアクセス

Direct Connect の疎通と BGP ネイバーの確立

1. 172.16.0.100 に通信したい
2. 自サブネット外なのでデフォルトゲートウェイに送信



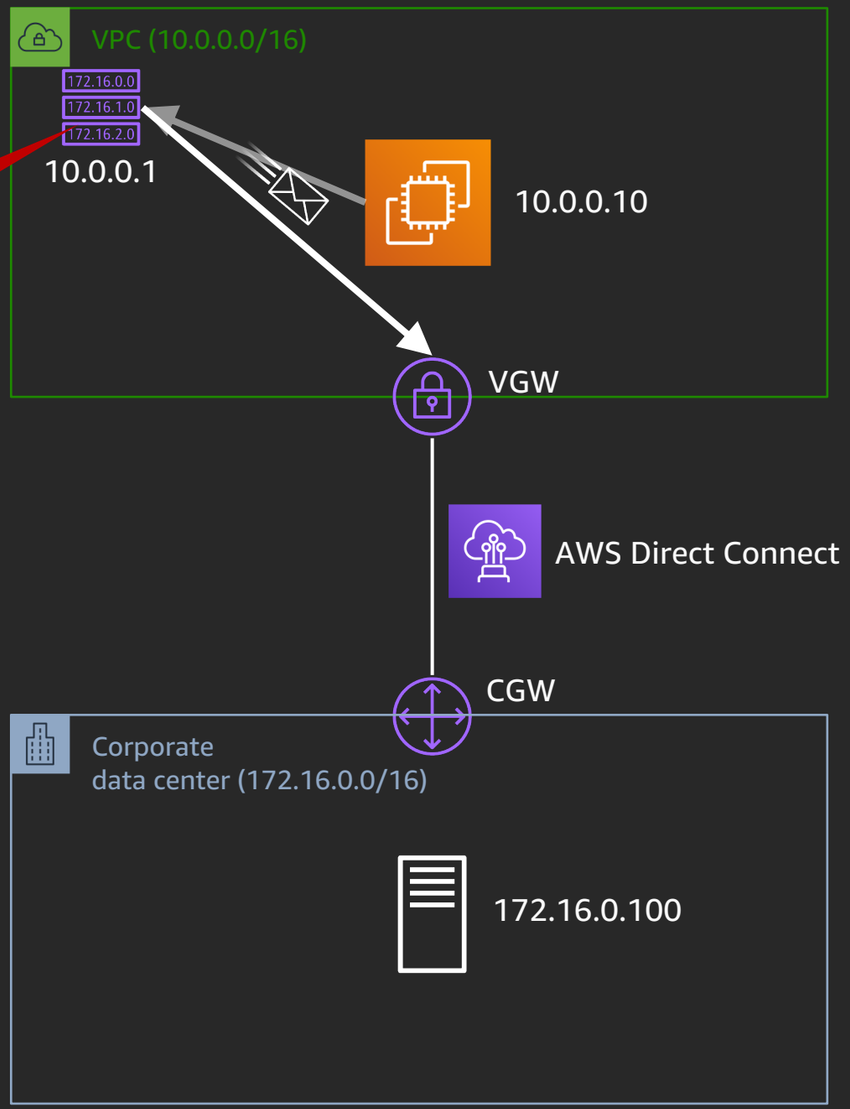
2. VPC 内から VGW を介してオンプレミスにアクセス

VPC のルートテーブルにオンプレミス宛通信用のエントリーを作成

3. ルートテーブルを参照

送信先	ターゲット
10.0.0.0/16	local
172.16.0.0/16	**VGW**

4. 一致したエントリーに従い VGW に送信



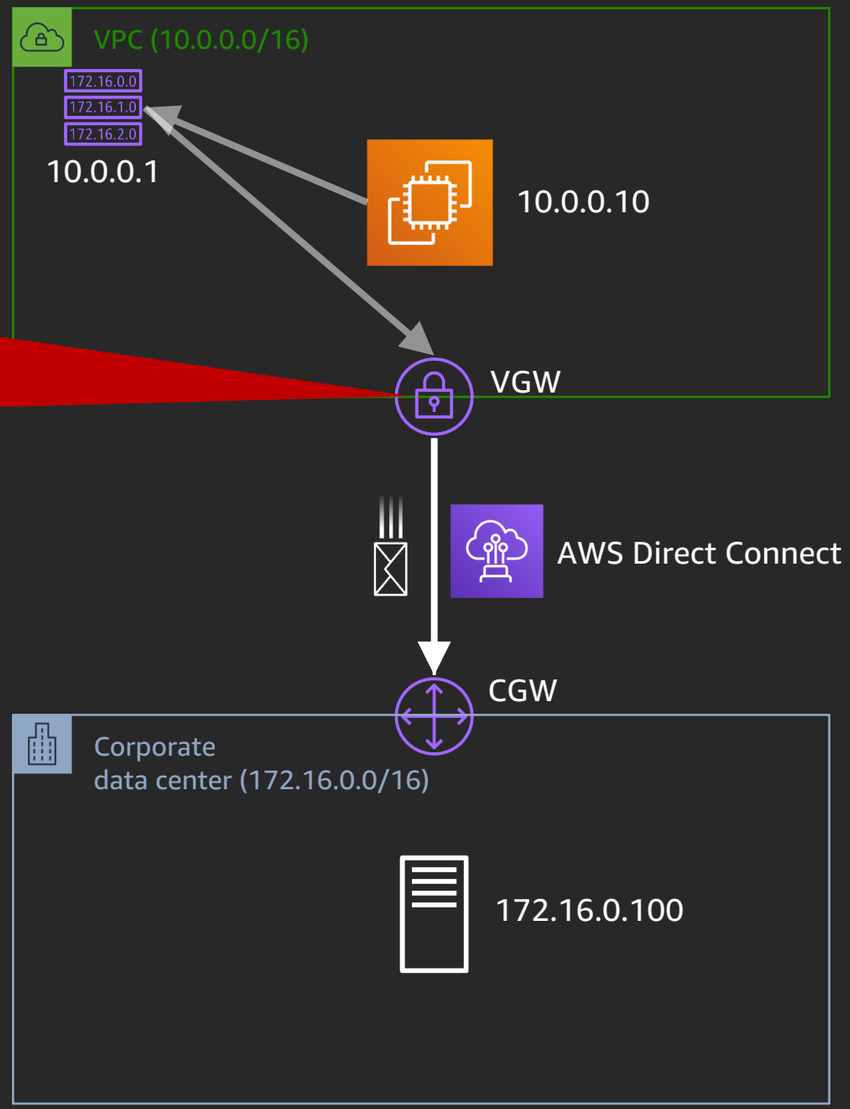
2. VPC 内から VGW を介してオンプレミスにアクセス

5. 自身の経路情報を参照

送信先	ターゲット
10.0.0.0/16	VPC Router
172.16.0.0/16	CGW (via BGP)

6. 一致したエントリーに従い CGW に送信

デフォルトでは VGW が持つ経路情報はユーザーには参照できない



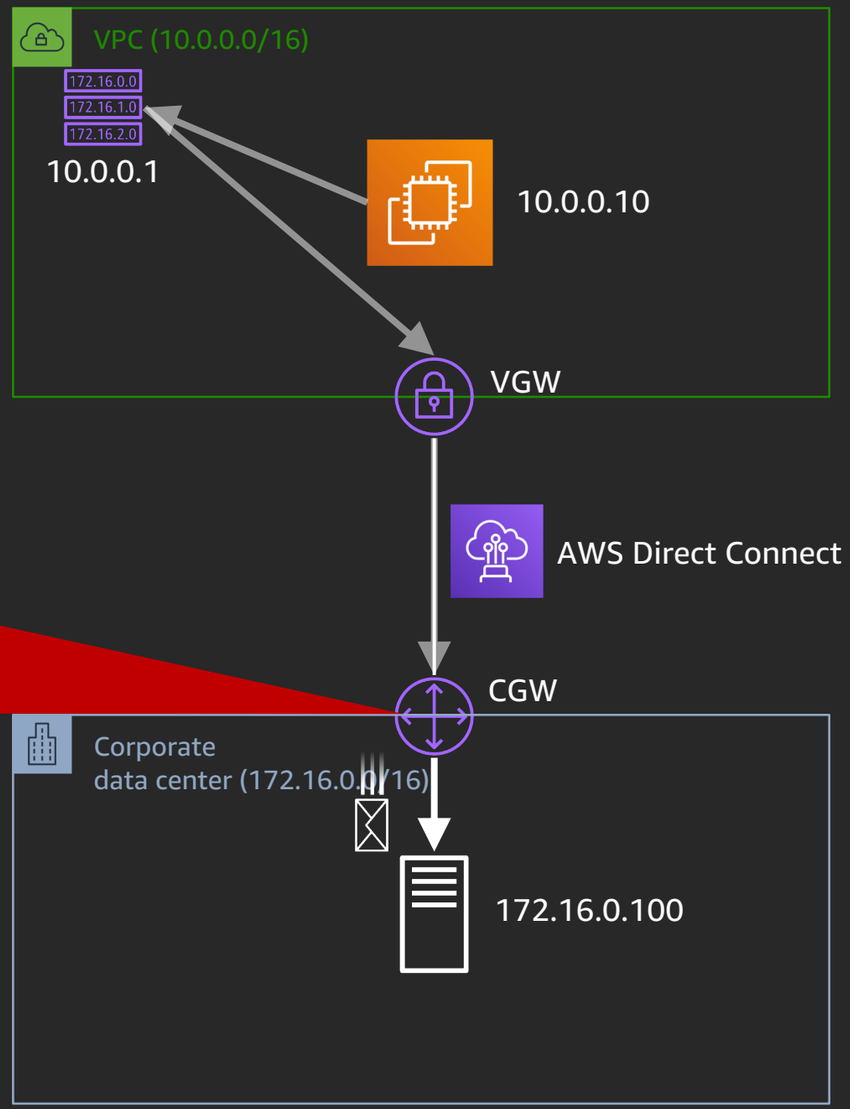
2. VPC 内から VGW を介してオンプレミスにアクセス

ここまで来たらオンプレミスの
世界です

7. 自身の経路情報を参照

送信先	ターゲット
172.16.0.0/16	Core Router
10.0.0.0/16	VGW (via BGP)

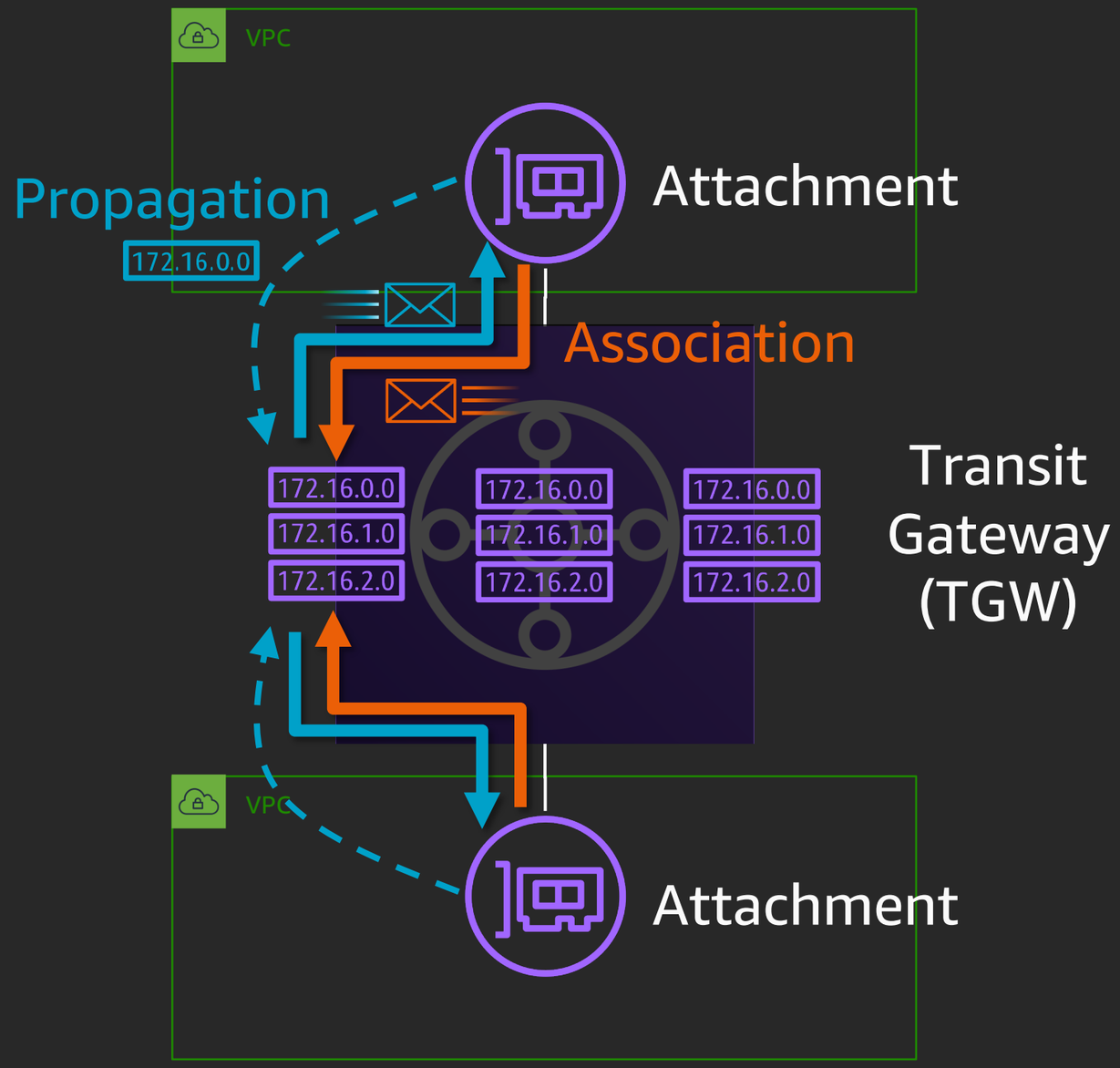
8. 一致したエントリーに従い
   Next hop に送信
9. それを繰り返し
   172.16.0.100 に到達



3. VPC 内から TGW を介して他の VPC にアクセス

Transit Gateway (TGW) の基本

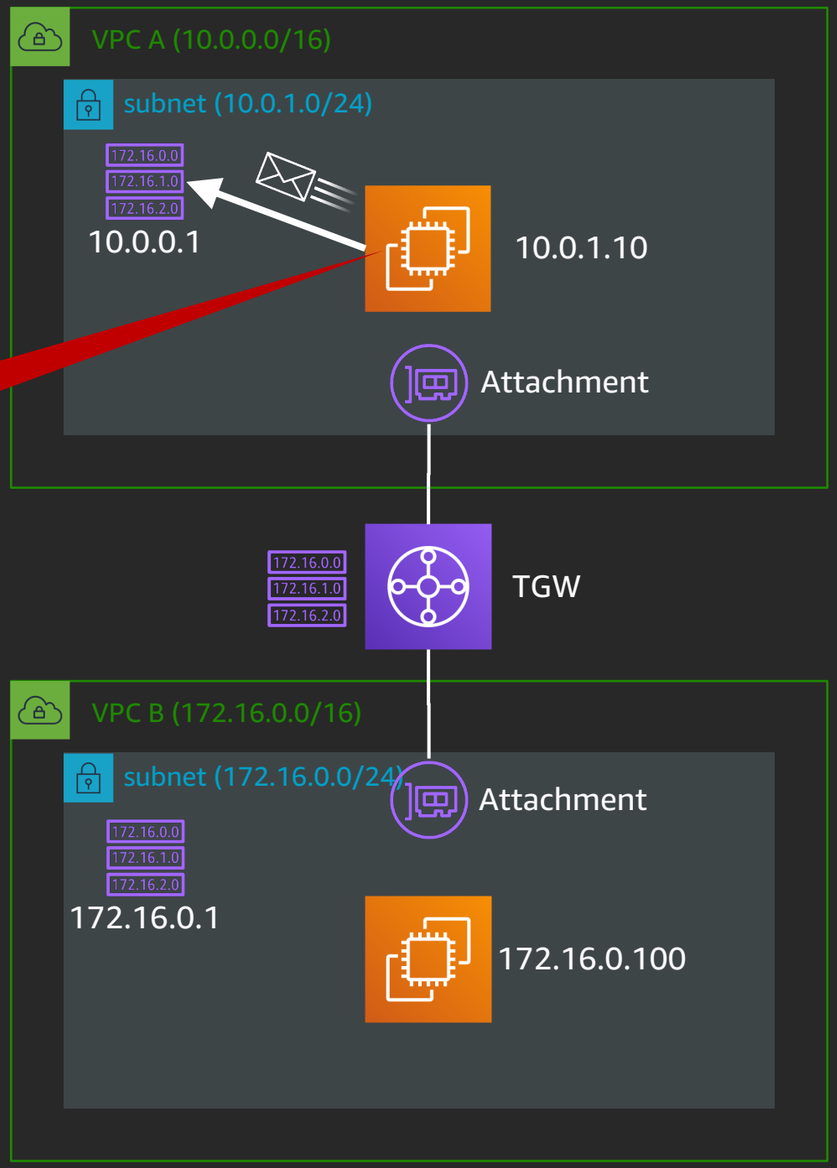
- ルートテーブルを必要数作成
  - デフォルトで 1 つ存在
- TGW と繋ぎたい VPC はまずアタッチメントを作成
  - AZ 毎に 1 つずつの専用サブネットとアタッチメントの作成を推奨
- アソシエーションで VPC からの Outbound 通信を制御
  - そのアタッチメントが参照する TGW ルートテーブルを指定する
- プロパゲーションで VPC への Inbound 通信を制御
  - そのアタッチメントが経路を広告する TGW ルートテーブルを指定する（複数可能）



3. VPC 内から TGW を介して他の VPC にアクセス

TGW の作成と VPC A/B のアタッチメント、アソシエーション、プロパゲーションを設定

1. 172.16.0.100 に通信したい
2. 自サブネット外なのでデフォルトゲートウェイに送信



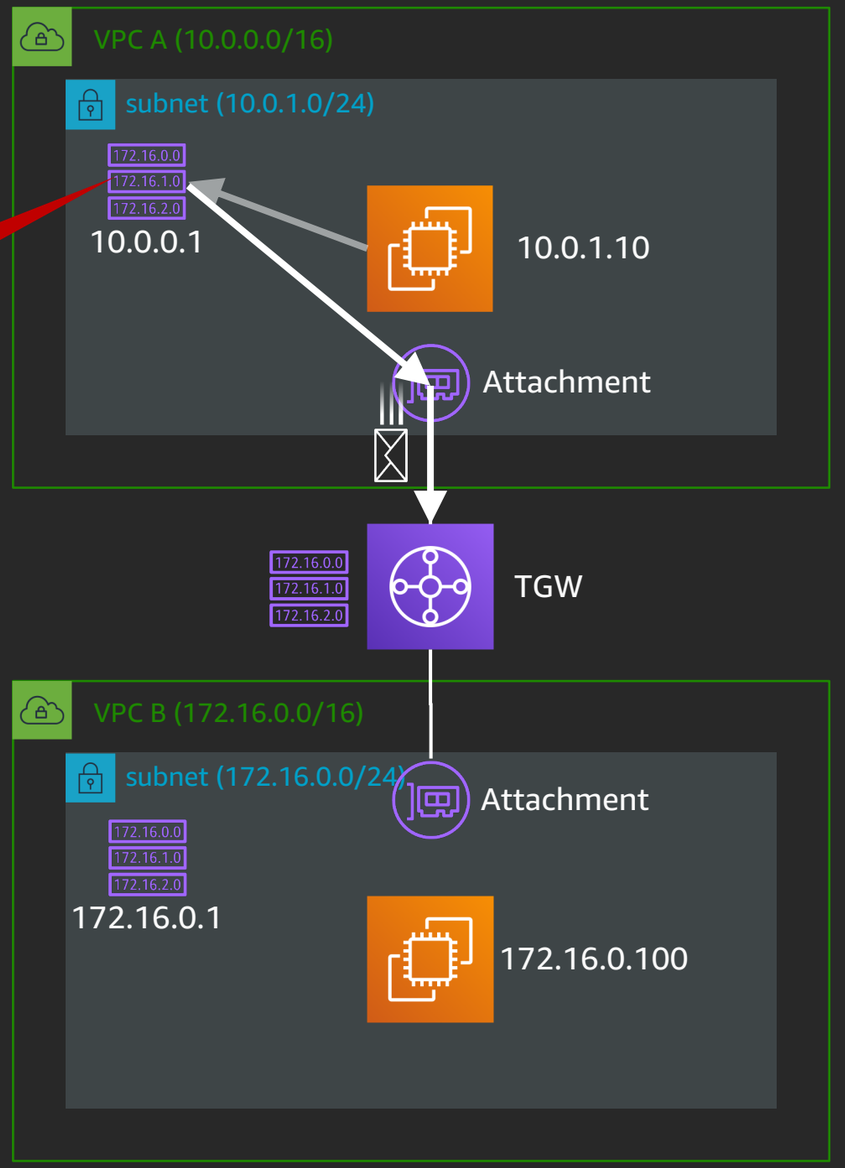
3. VPC 内から TGW を介して他の VPC にアクセス

VPC のルートテーブルに VPC B 宛通信用のエントリーを作成

3. ルートテーブルを参照

送信先	ターゲット
10.0.0.0/16	local
172.16.0.0/16	**TGW**

4. 一致したエントリーに従い TGW に送信

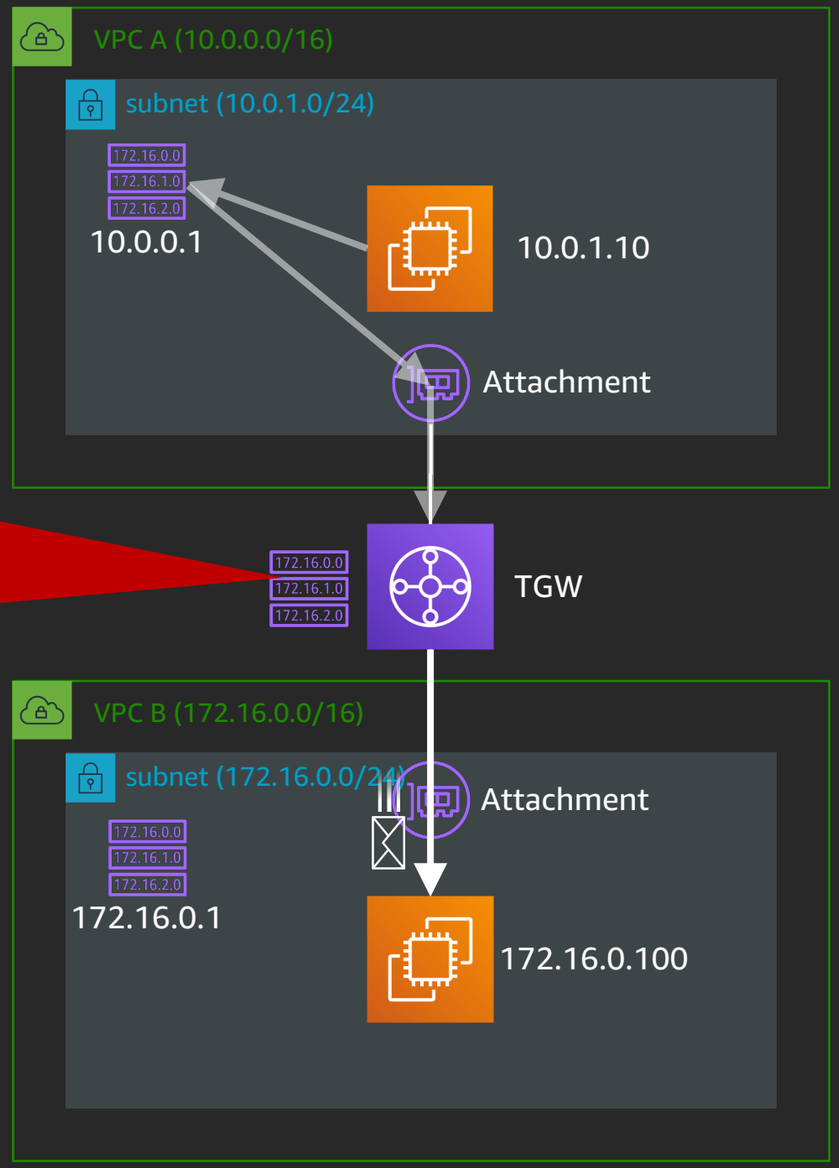


3. VPC 内から TGW を介して他の VPC にアクセス

5. TGW はアソシエーションに従ってルートテーブルを参照

送信先	ターゲット
10.0.0.0/16	VPC A (propagated)
172.16.0.0/16	VPC B (propagated)

6. 一致したエントリーに従い VPC B のアタッチメントから 172.16.0.100 に到達



次のステップ

要件が複雑化しても同じ手順で考えることができます

- 外部接続 VPC を作り全ての Outbound 通信を集約する (TGW)
- VPC への全ての Inbound 通信をアプライアンスに通す (Ingress Routing *)
- オンプレミスから海外リージョンに接続 (Direct Connect Gateway *)
- オンプレミスと AWS の多対多接続 (Direct Connect Gateway * + TGW)

もしご希望があれば Webinar 等で続編としてご案内します

* 本日ご紹介できなかった機能

まとめ

- AWS 上のルーティング設計に必要な知識はオンプレミスと同じ
  - ルーターの場所とパケットの動きを理解しましょう
- 可用性やスケーラビリティの設計はクラウドに任せられる
  - 経路設計に集中できる
- クラウドの利点を活用してクールなネットワークを設計しましょう
  - 必要な時にすぐに作成し、不要になれば削除できる
  - API による自動化、Infrastructure as Code の運用も可能
  - VPC 自体は無料。データ転送は国際回線も含めてオンデマンドの従量課金

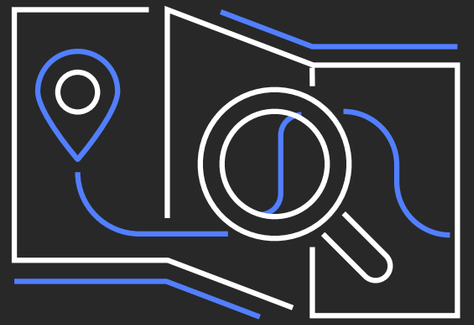
参考資料

- AWS Black Belt Online Seminar
  - Amazon VPC
    - Basic: https://d1.awsstatic.com/webinars/jp/pdf/services/20190313_AWS-BlackBelt-VPC.pdf
    - Advanced: https://d1.awsstatic.com/webinars/jp/pdf/services/20190417_AWS-BlackBelt-VPC-Advanced.pdf
  - AWS Direct Connect
    - Basic: https://d1.awsstatic.com/webinars/jp/pdf/services/20181114-AWS-Blackbelt-DirectConnect.pdf
    - Redundancy: https://d1.awsstatic.com/webinars/jp/pdf/services/20200219_BlackBelt_Onpremises_Redundancy.pdf
  - AWS Transit Gateway
    - https://d1.awsstatic.com/webinars/jp/pdf/services/20191113_AWS-BlackBelt_Transit_Gateway.pdf
  - Amazon Route 53
    - Hosted Zone: https://d1.awsstatic.com/webinars/jp/pdf/services/20191105_AWS_Blackbelt_Route53_Hosted_Zone_A.pdf
    - Resolver: https://d1.awsstatic.com/webinars/jp/pdf/services/20191016_AWS_Blackbelt_Route53_Resolver.pdf

Thank you!

aws SUMMIT ONLINE

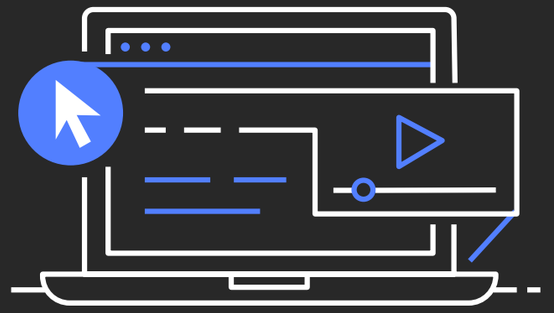
AWS トレーニングと認定



クラウド人材の育成

AWS トレーニングを活用し、ビジネスを牽引する人材の育成と組織作りを促進する

AWS トレーニング活用事例 »



自習コンテンツの活用

ウェビナーやのデジタルトレーニングを受講して、個人のスキルアップを目指す

AWS デジタルトレーニング »

AWS 認定取得を目指す

認定取得を目指して知識を底上げし、AWS の経験とスキルを証明する

AWS 認定の詳細 »

学習パスをお探しの方に

日本語版ランプアップガイドを公開しました。AWS ウェブページ、無料のデジタルトレーニング、クラスルームコース、動画、ホワイトペーパー、認定等を含んだ、9 種の役割別学習ガイドをご覧いただけます。 詳細を見る »

aws.amazon.com/training