



Amazon ECS Deep Dive

~ Capacity Providers ~

アマゾン ウェブ サービス ジャパン株式会社
落水 恭介

2021.09.07

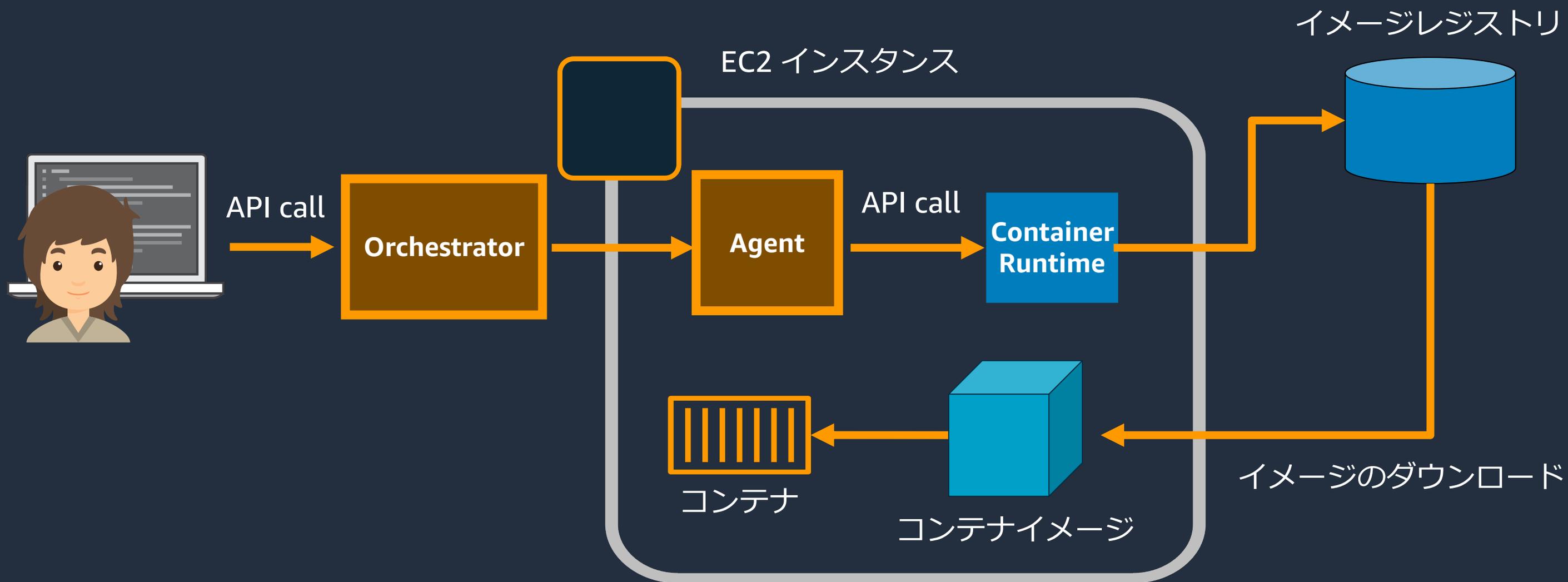


本日本話すること

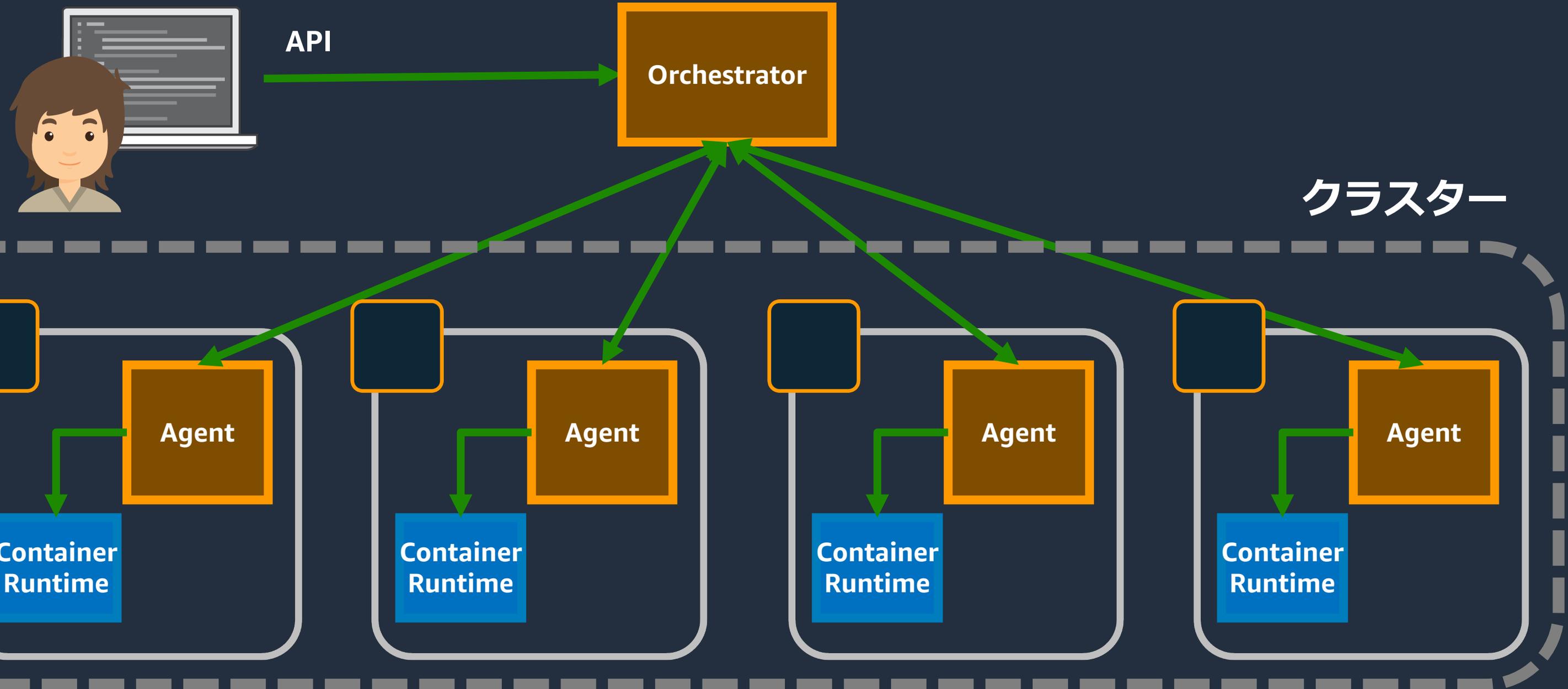
- Amazon ECS での Auto Scaling とは
 - コンテナインスタンス (EC2 インスタンスレイヤ)
 - ECS サービス (タスクレイヤ)
- ECS Cluster Auto Scaling
 - ECS Cluster Auto Scaling とは
 - Capacity Providers の概要

Amazon ECS での Auto Scaling とは

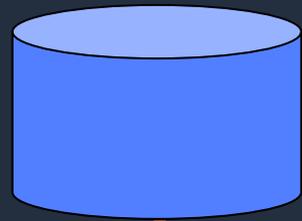
コンテナオーケストレーションの仕組み



コンテナオーケストレーションによるクラスター管理



イメージレジストリ



Amazon ECR



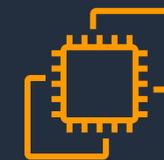
オーケストレータ



Amazon ECS



ホスティング

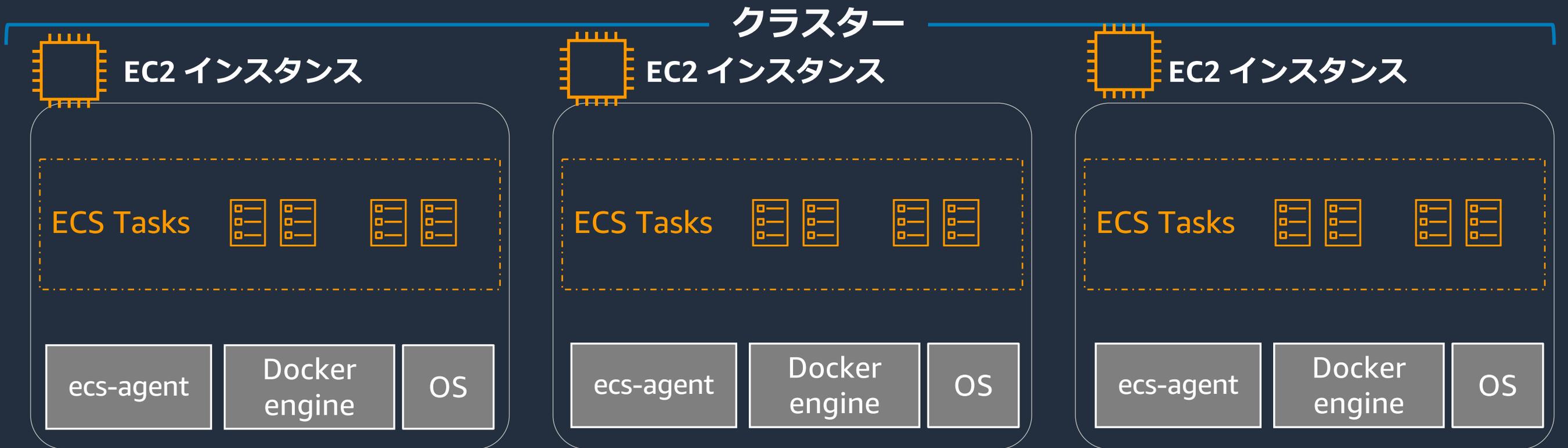


Amazon EC2



AWS Fargate

Amazon ECS の動作イメージ (on EC2)



Amazon ECS の動作イメージ (on Fargate)

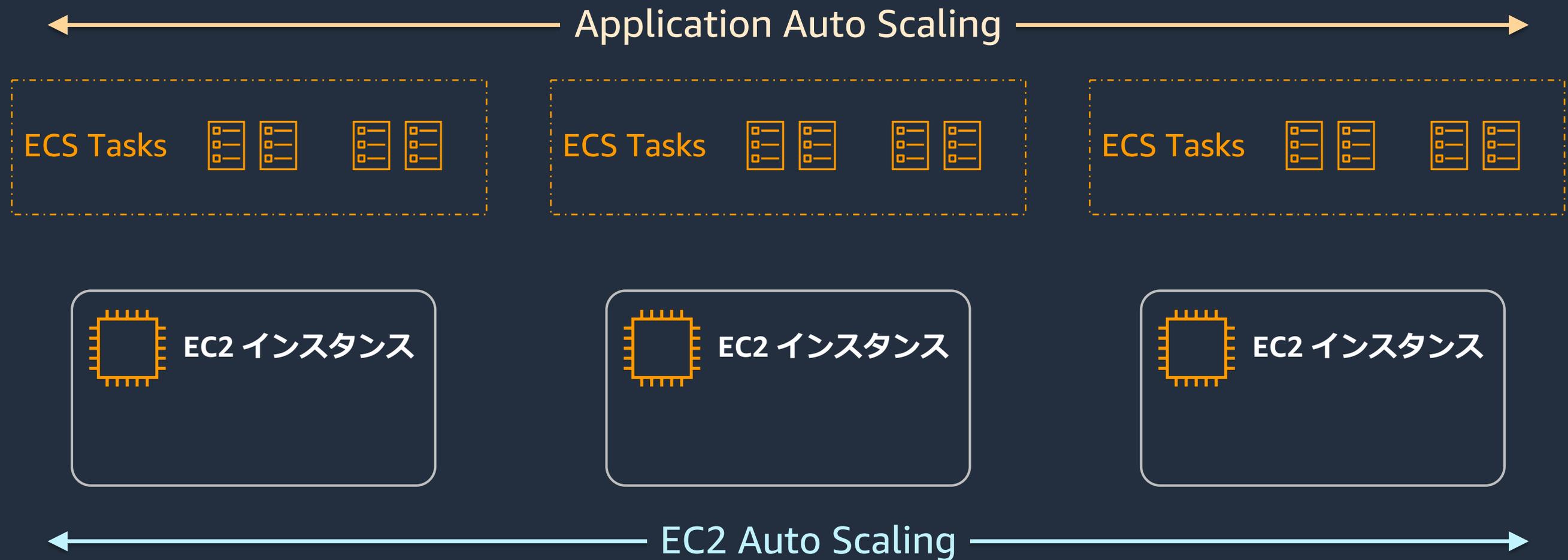


クラスター



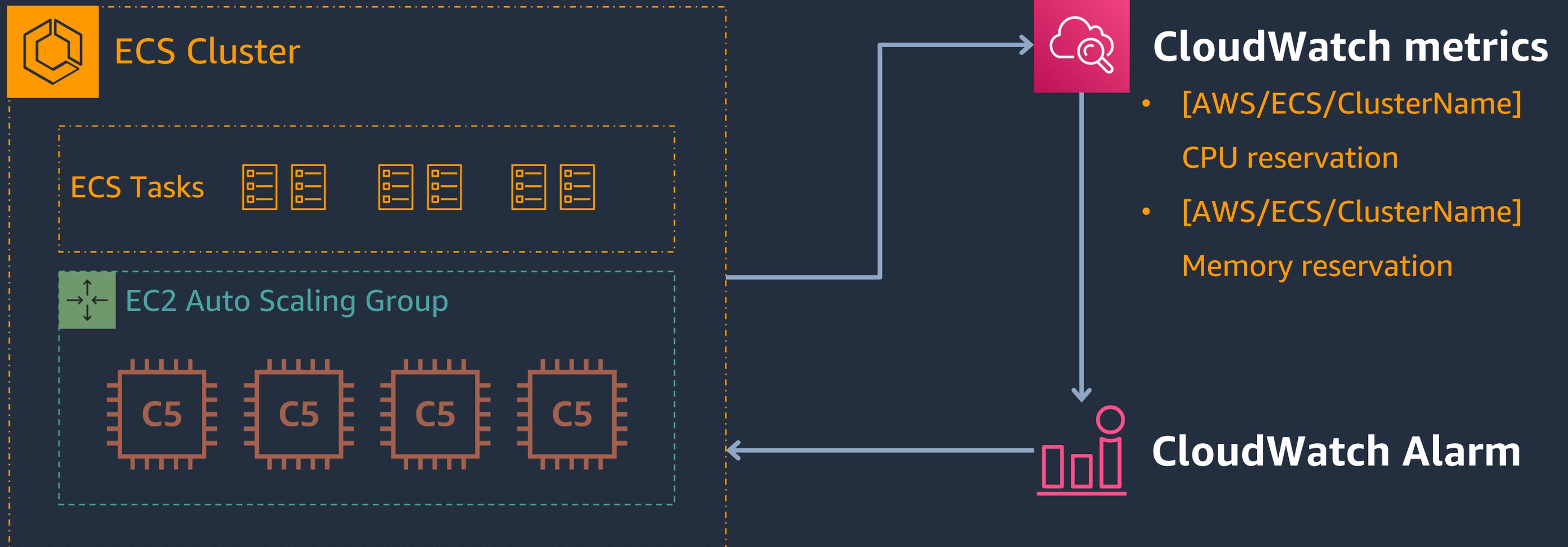
ECS on EC2 における Auto Scaling とは

EC2 インスタンス / ECS タスクのそれぞれのレイヤで Auto Scaling の設定が必要



EC2 インスタンスレイヤの Auto Scaling

ECS タスクの状況と連動するメトリクスをターゲットに設定することで、
EC2 インスタンスのキャパシティをコンテナワークロードの需要に追従させる



ECS タスクレイヤの Auto Scaling

Application Auto Scaling を活用して ECS サービスの必要タスク数を自動的に増減

ターゲット追跡スケールリングポリシー

- 指定したメトリクスがターゲットの値に近づくように自動的に調整

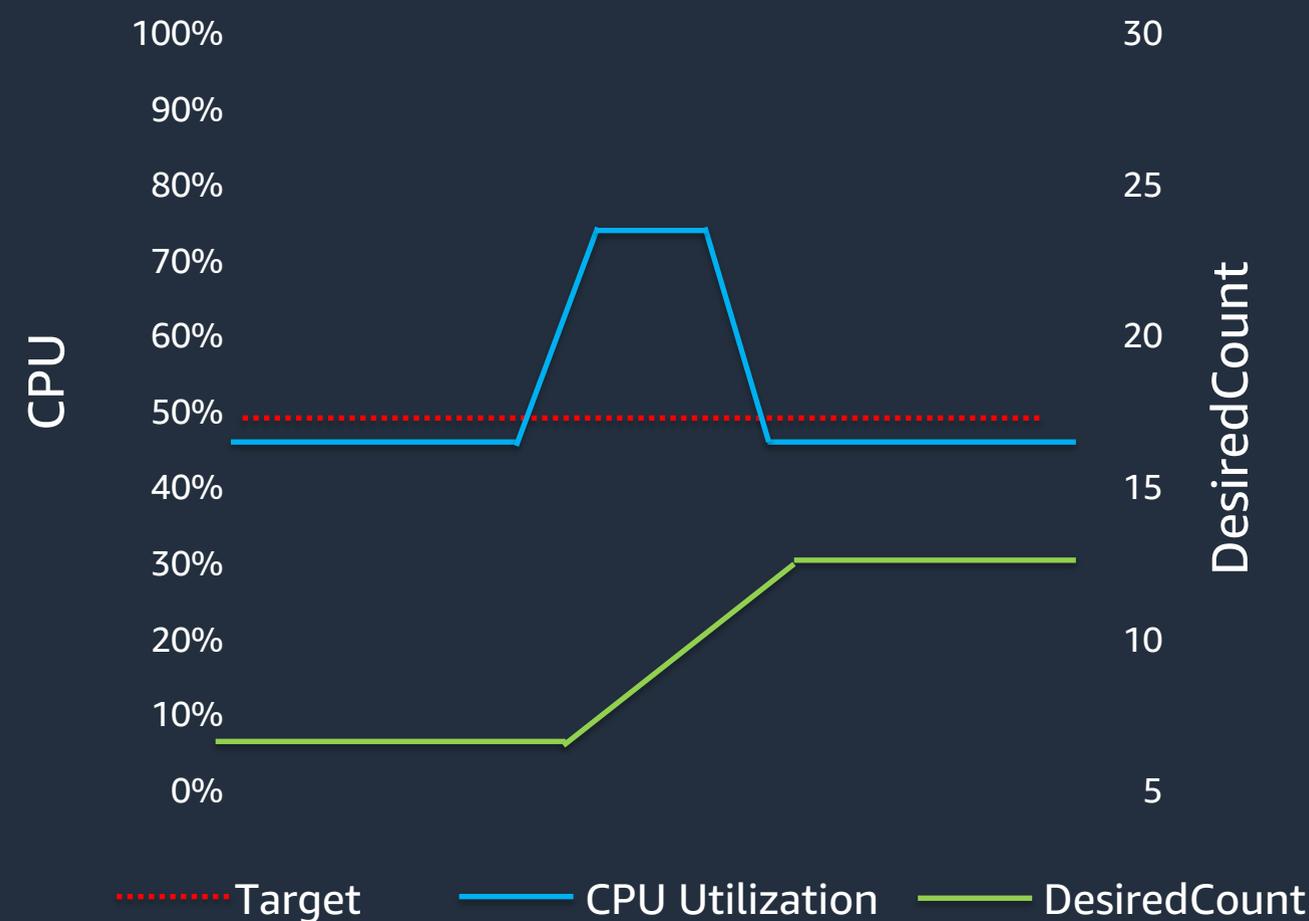
ステップスケールリングポリシー

- 設定されたスケールリングアクションに基づいて増減

スケジュールに基づくスケールリング

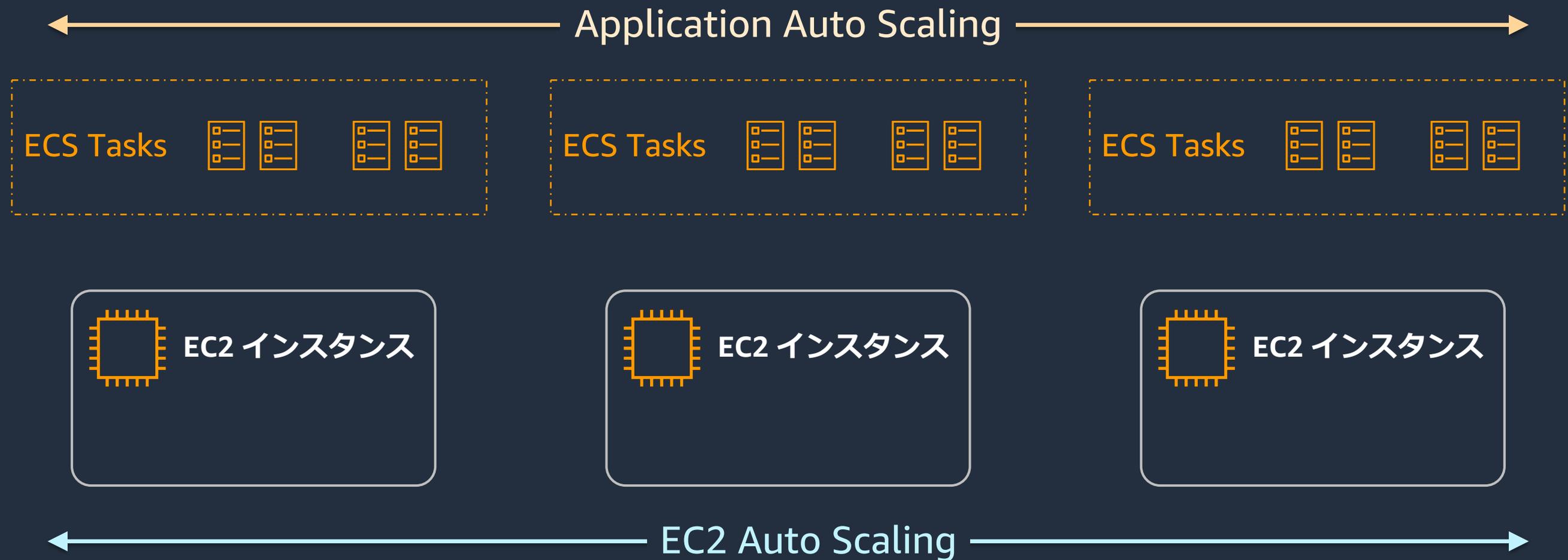
- 日付と時刻に基づいてタスク数を増減

例) ターゲット追跡スケールリングポリシー



ECS on EC2 における Auto Scaling の課題

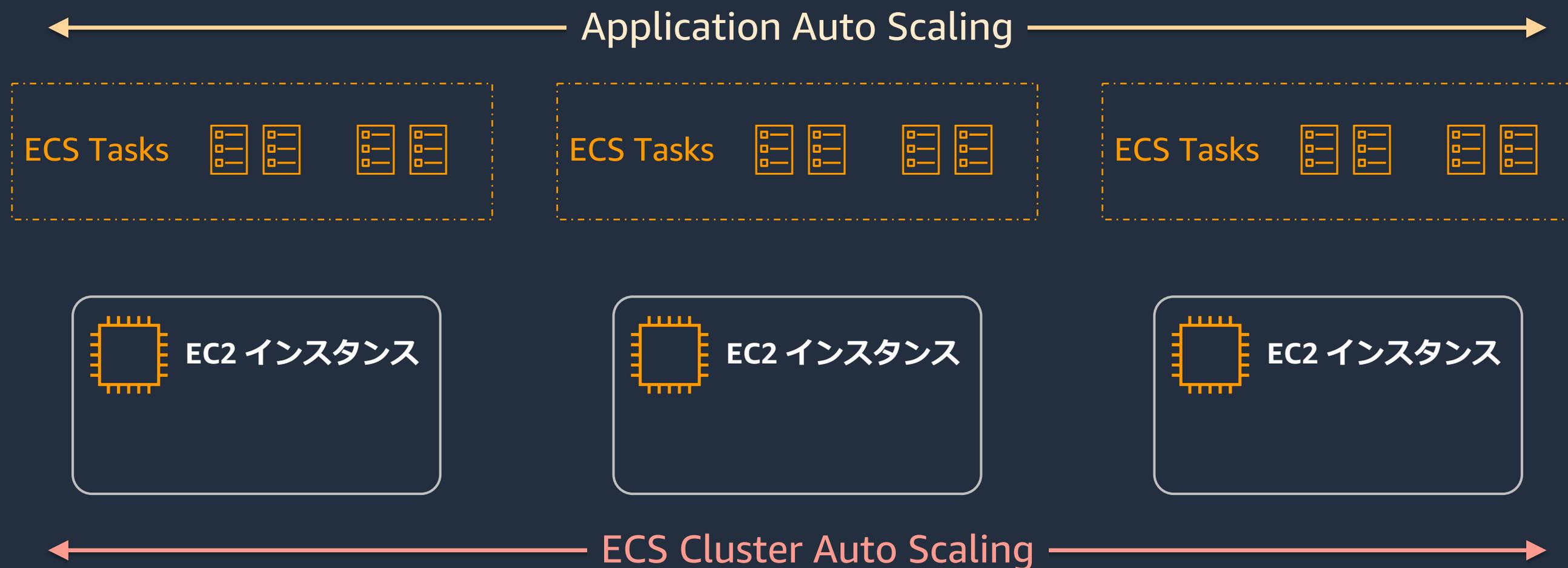
複数レイヤの Auto Scaling の管理は複雑になりやすく、運用コストの増加を招く



ECS on EC2 における Auto Scaling の課題

複数レイヤの Auto Scaling の管理は複雑になりやすく、運用コストの増加を招く

➤ **ECS Cluster Auto Scaling** により EC2 レイヤのマネージドスケールリングが可能に



Thank you!