

新たな気象リスクへの 挑戦を可能にしたHPC on AWS

Kazuhisa Inuma
WNI Forecast Center
Section Leader
Weathernews

Agenda

Motivation

激甚化する気象災害
気象予測ができるまで

Rapid Update OWN Project

パフォーマンス
信頼性
コスト

激甚化する気象災害



2019年台風19号ウェザーリポート

新たな気象リスクへの対策

台風、ゲリラ雷雨、線状降水帯など
局地的・集中的な降雨リスクに対し
早期からのリスク対策が社会的に求
められる時代。よりよい対策のため
高い予測精度が求められる。

計画運休

BCP



weathernews

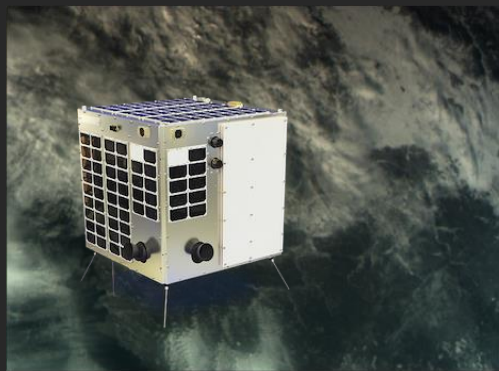
*Always **WITH** you!*

天気は眠らない。

ウェザーニューズは 24 時間 365 日あなたとともに。

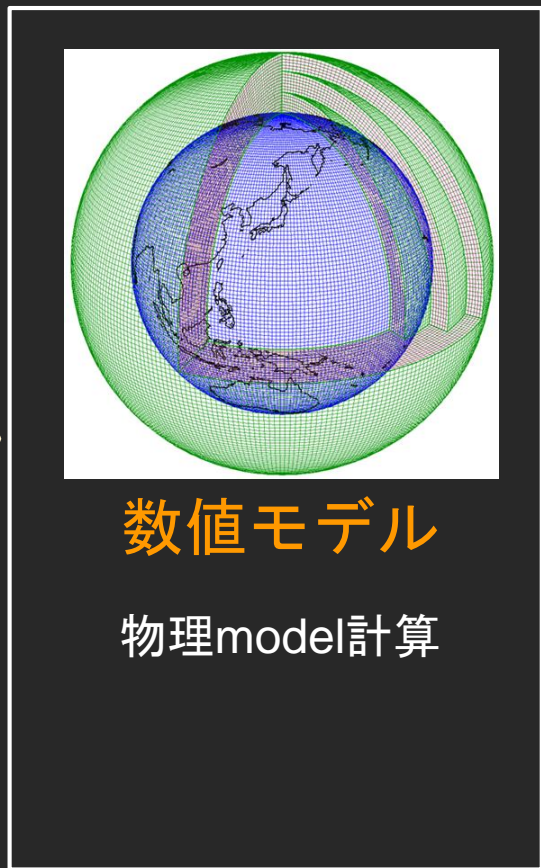
弊社では交通・インフラなど44市場のお客様に対して、24時間リスク対策情報を提供

気象予測が出来るまで



観測データ

- 公的機関の観測器
- 独自インフラ
- ウェザーリポート



数値モデル
物理model計算

High Performance
Computing



AI統計処理

補正処理



人による判断

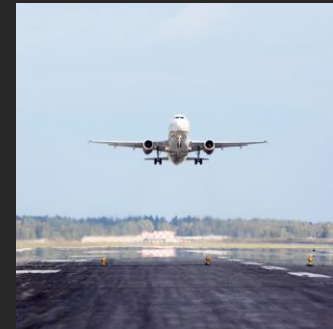
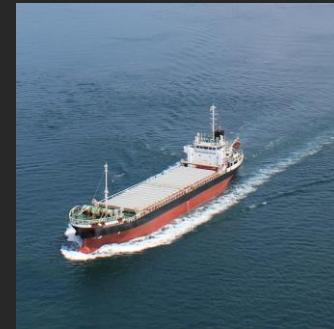
個別の補正
対応策への変換



台風19号による鉄道の影響予測

問題なし ◯ 影響あり △ 影響大 ×

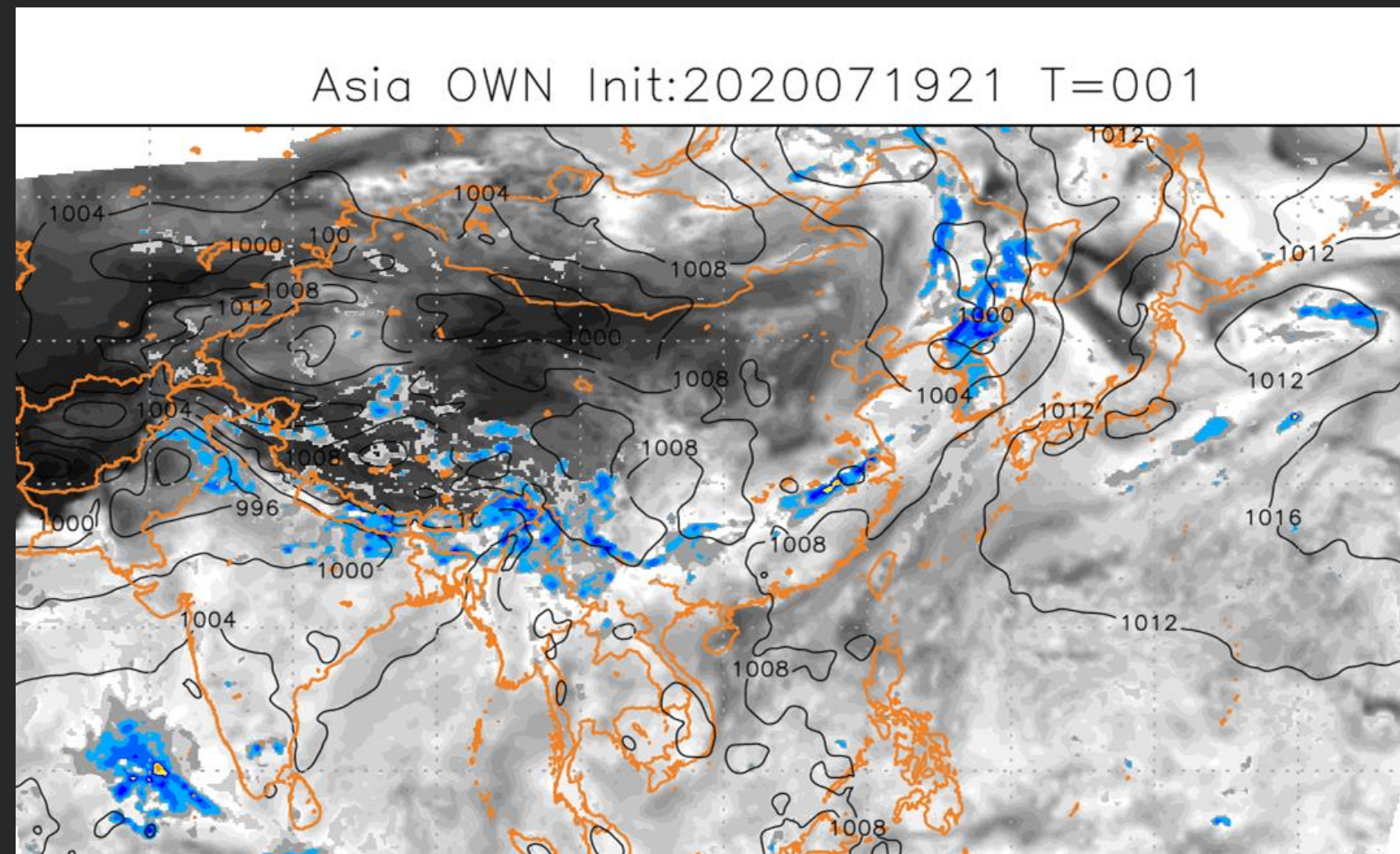
路線	12日(土)			13日(日)			14日(月)		
	午前	午後	夜	午前	午後	夜	午前	午後	夜
北海道	◯	◯	◯	◯	◯	◯	◯	◯	◯
東北・山形・秋田	◯	△	×	△	×	×	△	×	×
上越	◯	△	×	△	×	×	△	×	×
北陸	◯	△	×	△	×	×	△	×	×
東海	×	×	×	×	×	×	×	×	×
山陽	△	△	△	△	△	△	△	△	△
九州	△	△	△	△	△	△	△	△	△
北海道	◯	◯	◯	◯	◯	◯	◯	◯	◯
東北	×	×	×	×	×	×	×	×	×
関東	×	×	×	×	×	×	×	×	×
北陸	×	×	×	×	×	×	×	×	×
近畿	×	×	×	×	×	×	×	×	×
中国	△	△	△	△	△	△	△	△	△
四国・九州	△	△	△	△	△	△	△	△	△



コンテンツ・顧客判断支援

OWN on-premises cluster system

独自モデルでの数値シミュレーション

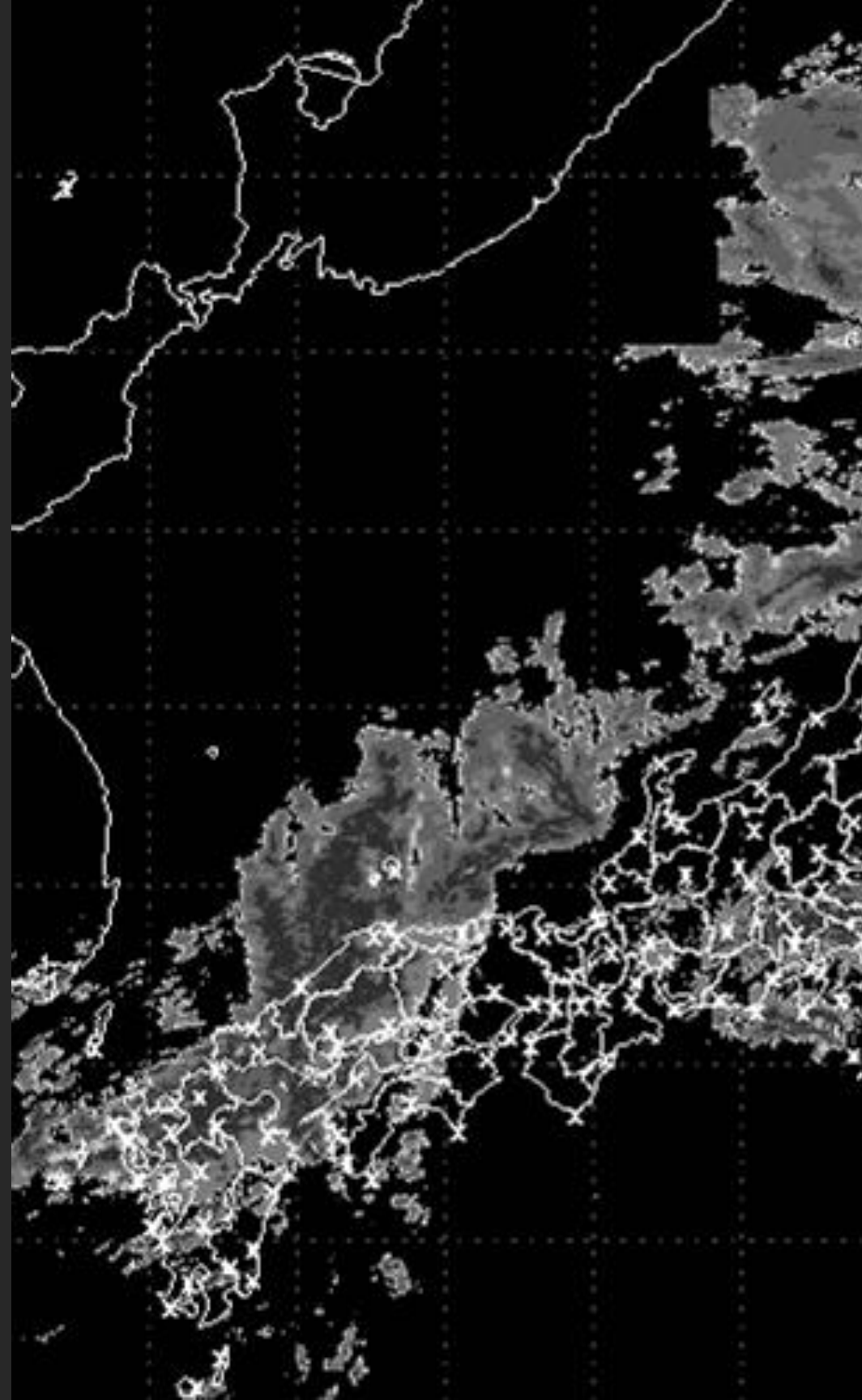


分散メモリ型
クラスターシステム
倍精度約15TFlops規模

新たな気象リスクに 対応するための OWNの高度化

現状のオンプレクラスシステムは、定時実行の安定性を優先しており、1日4回の予測更新が限度。

通常の降雨リスクに対してはこれで不足ないが、**新たな気象リスク**を考慮していくためには、必要な時に最高のパフォーマンスを発揮する瞬発力が必要になってきた。



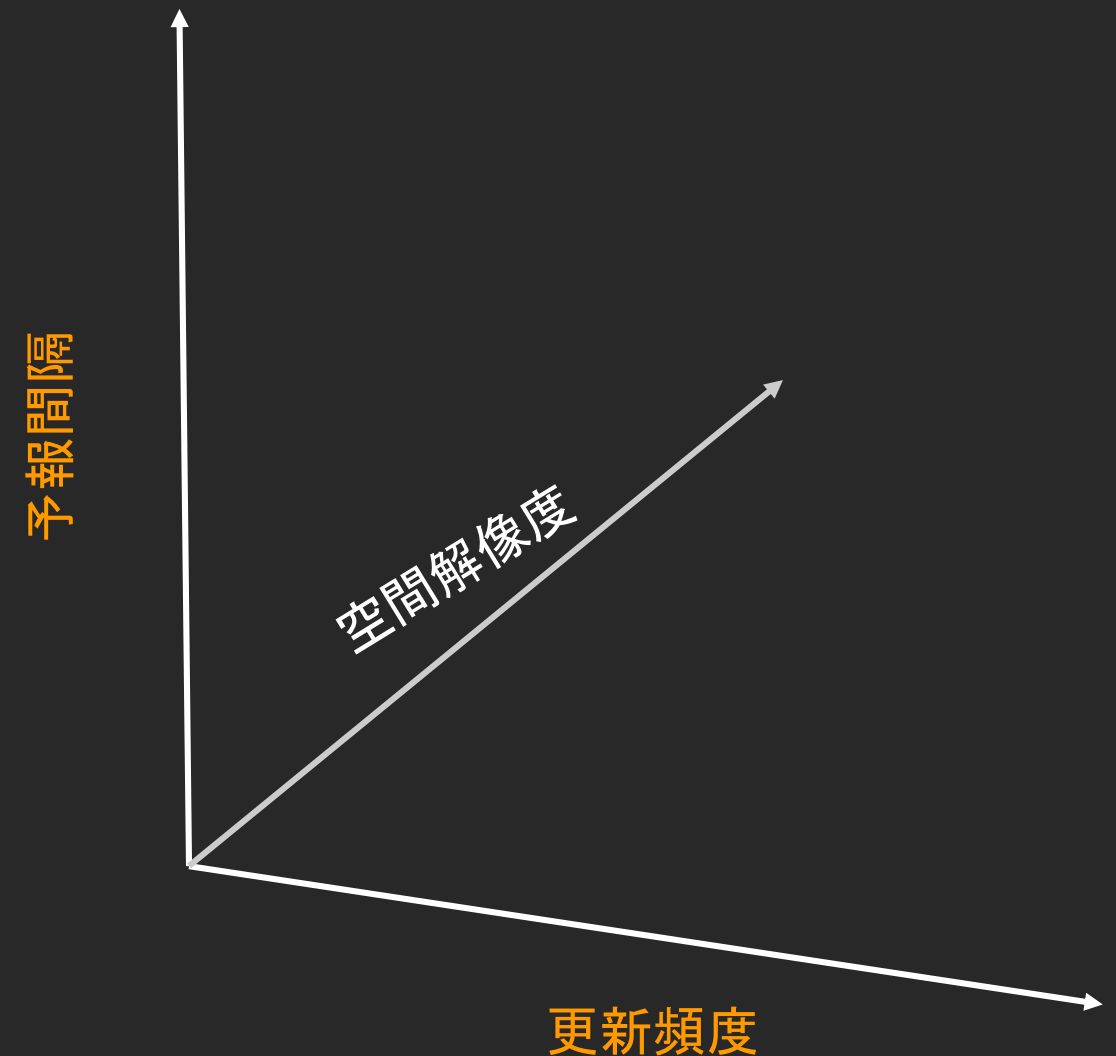
Rapid Update OWN Project 始動

OWN on AWS ParallelCluster

2018年からParallelCluster (CfnCluster) を用いてOWNのクラウド実装を検討

目的: 予報間隔と更新頻度の改善
運用化に向けた調査は2020年3月～

更新頻度	6時間→3時間
予報間隔	1時間→10分



OWN高度化 にあたって求められる要件や課題

パフォーマンス

- 60分以内に24時間分の予報の計算を完了させる
- 72vCPUのサーバー1台だと180分かかる問題サイズ

信頼性

- 現行のシステムと同等以上の可用性
- 許容範囲は年に数回処理が失敗する程度

コスト

- 適切なコストにおさえる
- オンデマンドやSavings Plansの適用以上にコストを削減したい

パフォーマンス

OWN高度化 にあたって求められる要件や課題

パフォーマンス

- 60分以内に24時間分の予報の計算を完了させる
- 72vCPUのサーバー1台だと180分かかる問題サイズ

信頼性

- 現行のシステムと同等以上の可用性
- 許容範囲は年に数回処理が失敗する程度

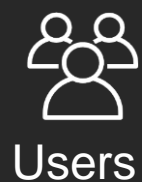
コスト

- 適切なコストにおさえる
- オンデマンドやSavings Plansの適用以上にコストを削減したい



AWS ParallelClusterとは

数コマンド操作でジョブ投入に応じて自動でスケールするクラスタを
AWS上に構築可能なAWS公式のオープンソースソフトウェア



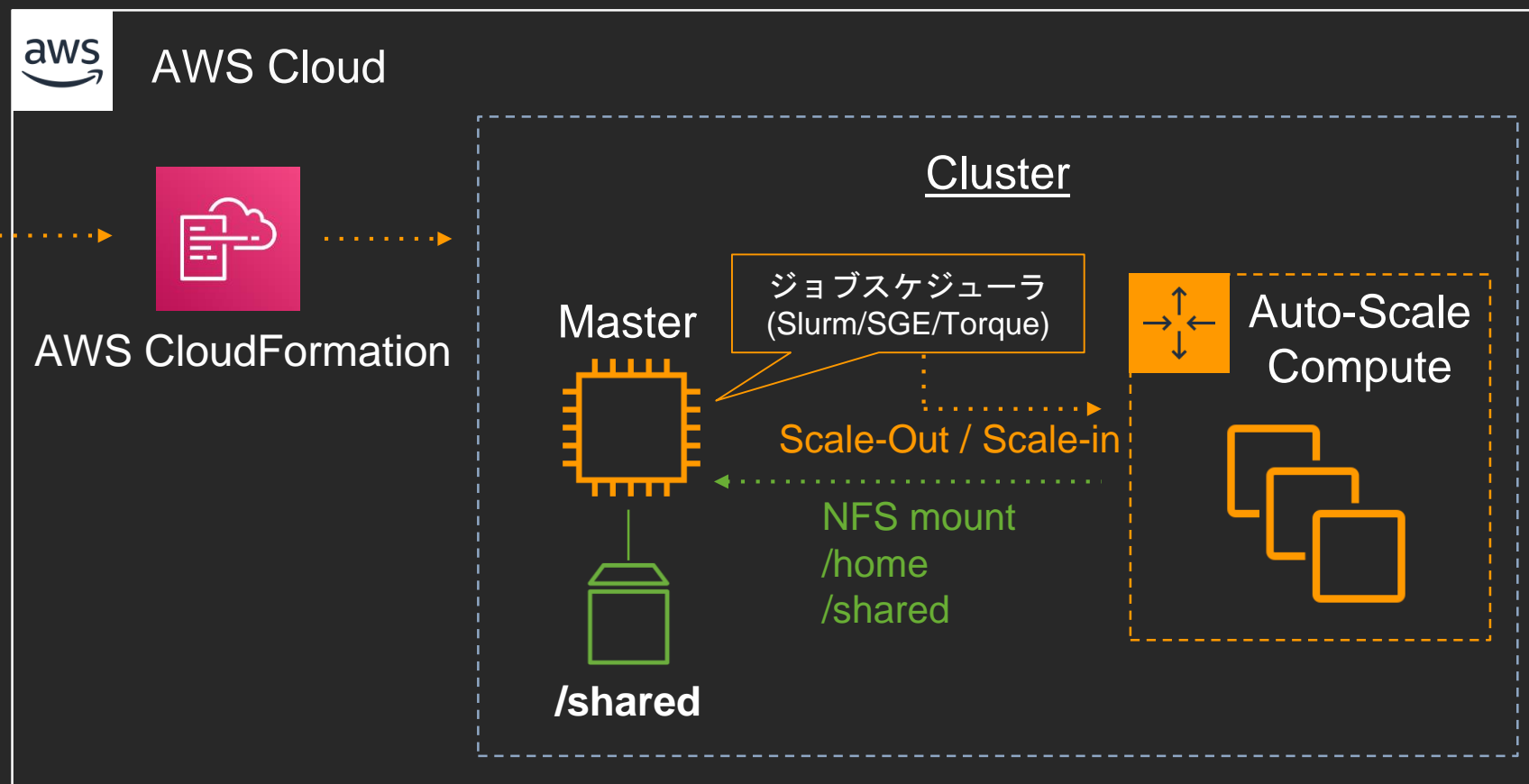
Users

```
$ pcluster create <NAME>
```

config ファイル

```
[aws]
aws_region_name = ap-northeast-1

[cluster slurm1]
master_instance_type = c5.large
compute_instance_type = c5.4xlarge
max_queue_size = 10
initial_queue_size = 0
scheduler = slurm
cluster_type = spot
```

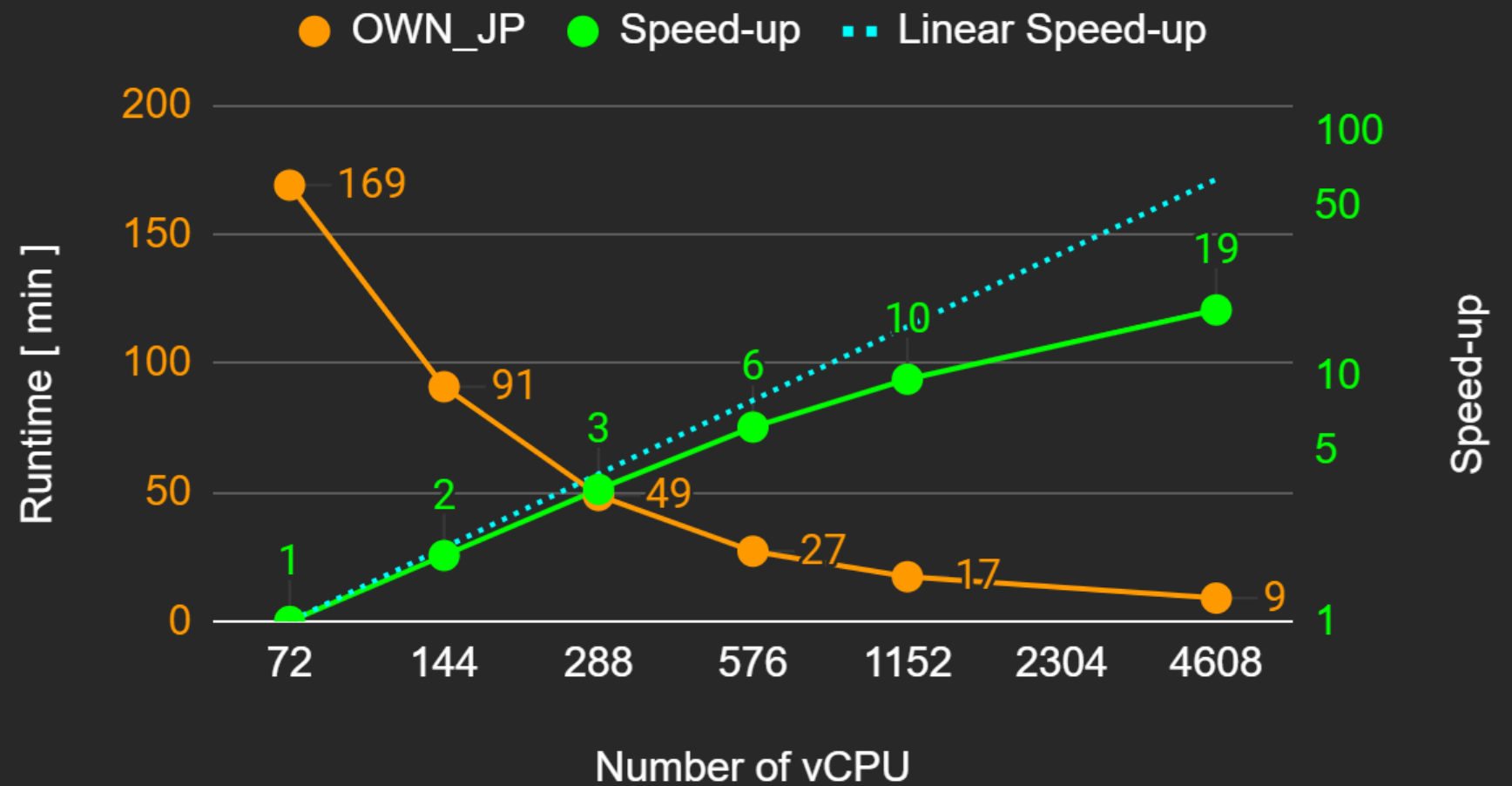


ParallelClusterでの並列計算

ParallelClusterでの
並列計算の結果

64Nodes 4608vCPUだと9分
で計算が完了する

OWN_JP : Strong Scaling (socket)



EC2の性能を最大限に引き出すために



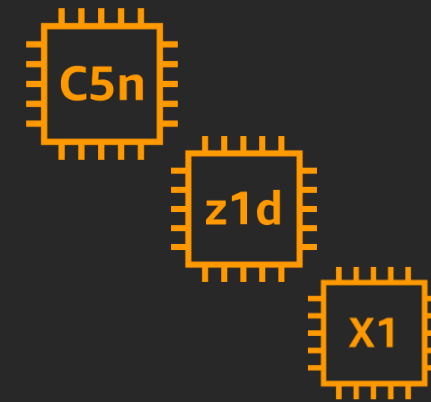
Elastic Fabric
Adapter (EFA)
の利用

ネットワーク
デバイスの変更



Amazon
FSx for Lustre
の利用

ファイルシステム
の変更



インスタンス
の選択

アプリケーションに
最適なインスタンス
タイプを調査

Parallel Clusterの設定変更レベルで対応

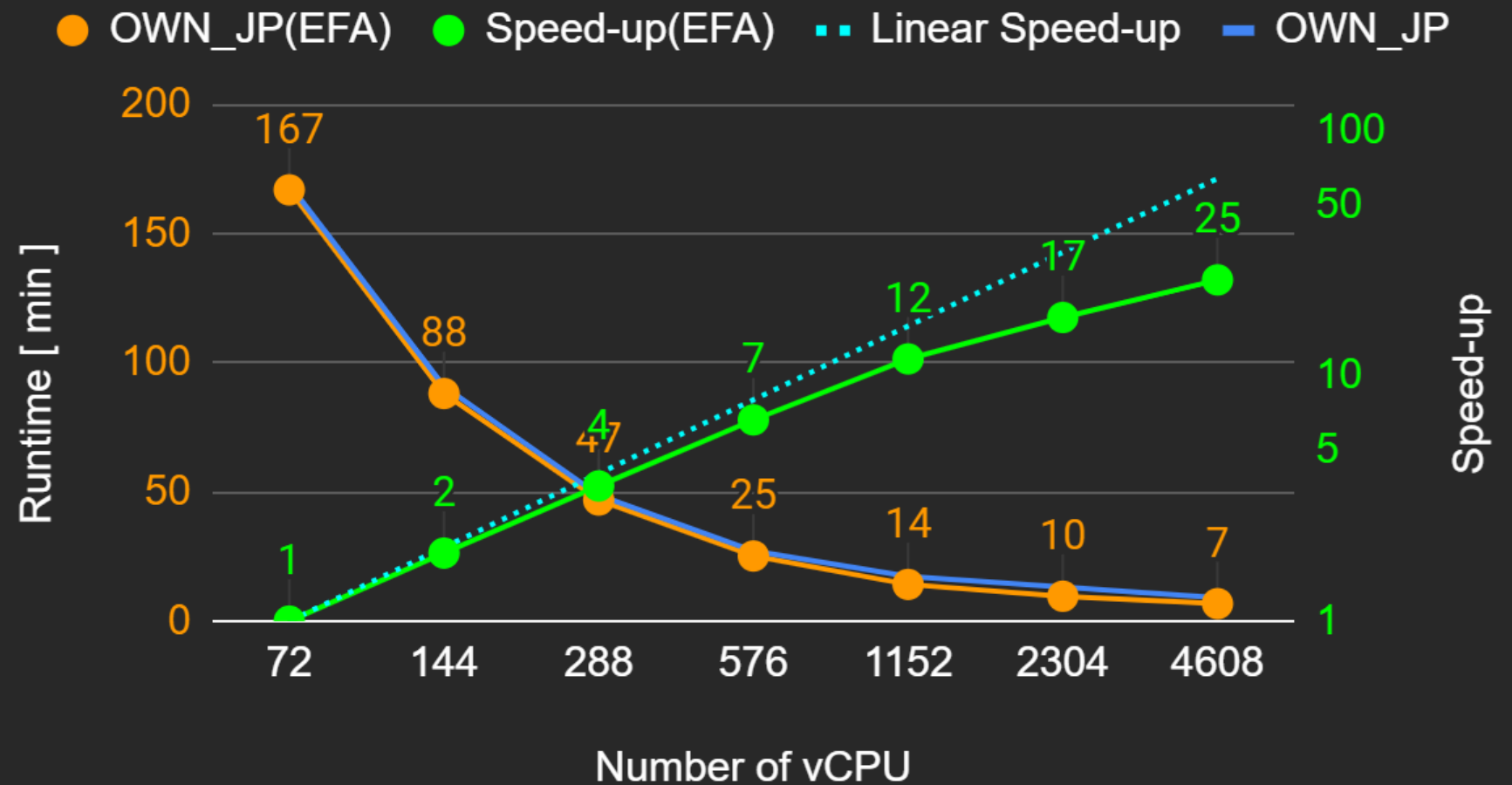
EFAの利用

Parallel Clusterの設定でEFAを用いた場合

特に並列数を増やした場合に**高速化**された

vCPU数が少ない場合は1~3%の改善だが4068まで並列化させた場合は**25%**の改善が見られた

OWN_JP : Strong Scaling (EFA)

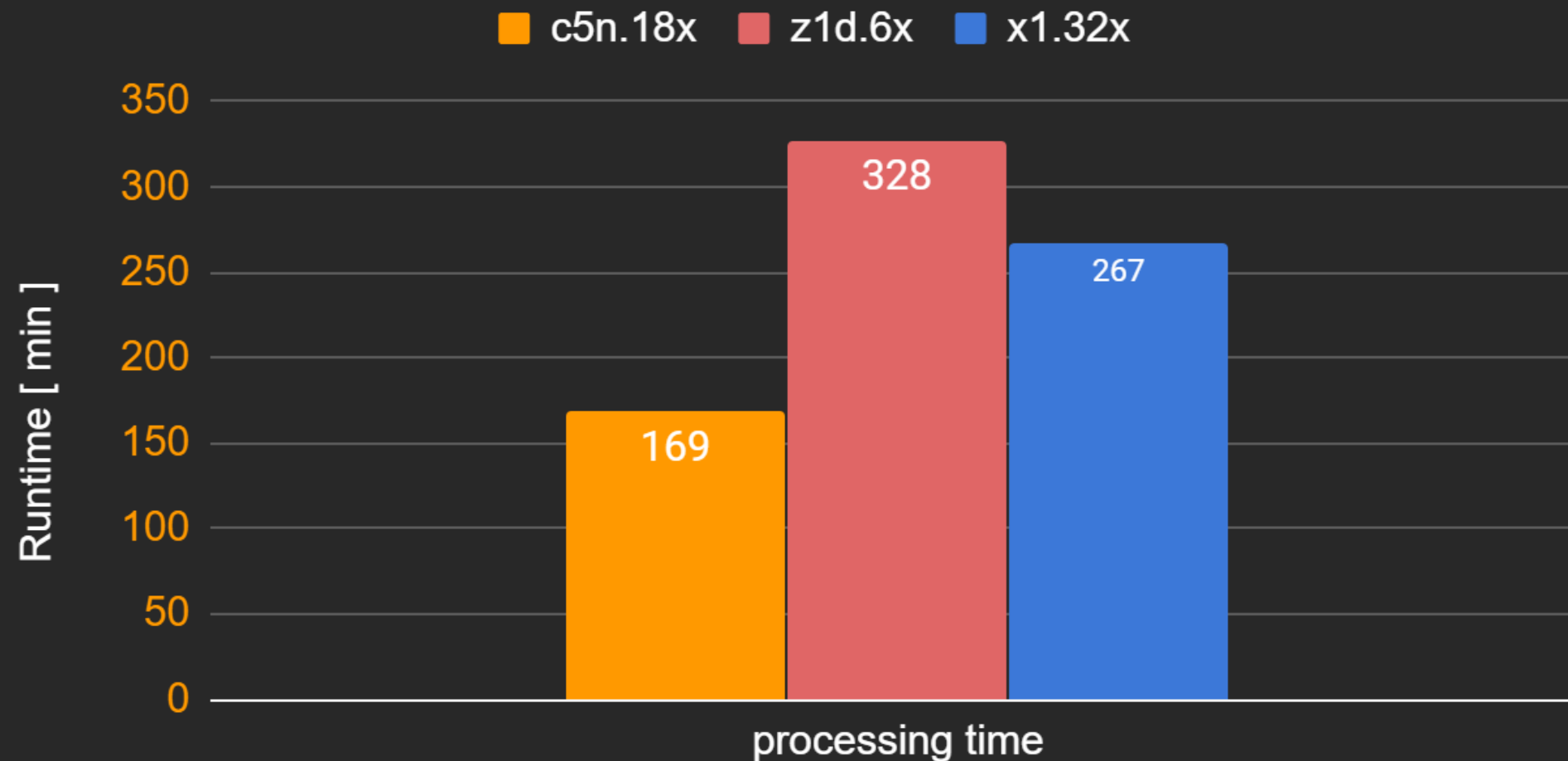


改善率(%)

1	3	4	7	18	25
---	---	---	---	----	----

インスタンスタイプの変更

JAPAN_05L : Single Node, Instance types



	vCPU	メモリ	周波数	network
c5n.18xlarge	72	192 GiB	3.4 GHz	100 Gbps
z1d.6xlarge	24	192 GiB	4.0 GHz	10 Gbps
x1.32xlarge	128	1,952 GiB	2.3 GHz	25 Gbps

信賴性

OWN高度化 にあたって求められる要件や課題

パフォーマンス

- 60分以内に24時間分の予報の計算を完了させる
- 72vCPUのサーバー1台だと180分かかる問題サイズ

信頼性

- 現行のシステムと同等以上の可用性
- 許容範囲は年に数回処理が失敗する程度

コスト

- 適切なコストにおさえる
- オンデマンドやSavings Plansの適用以上にコストを削減したい

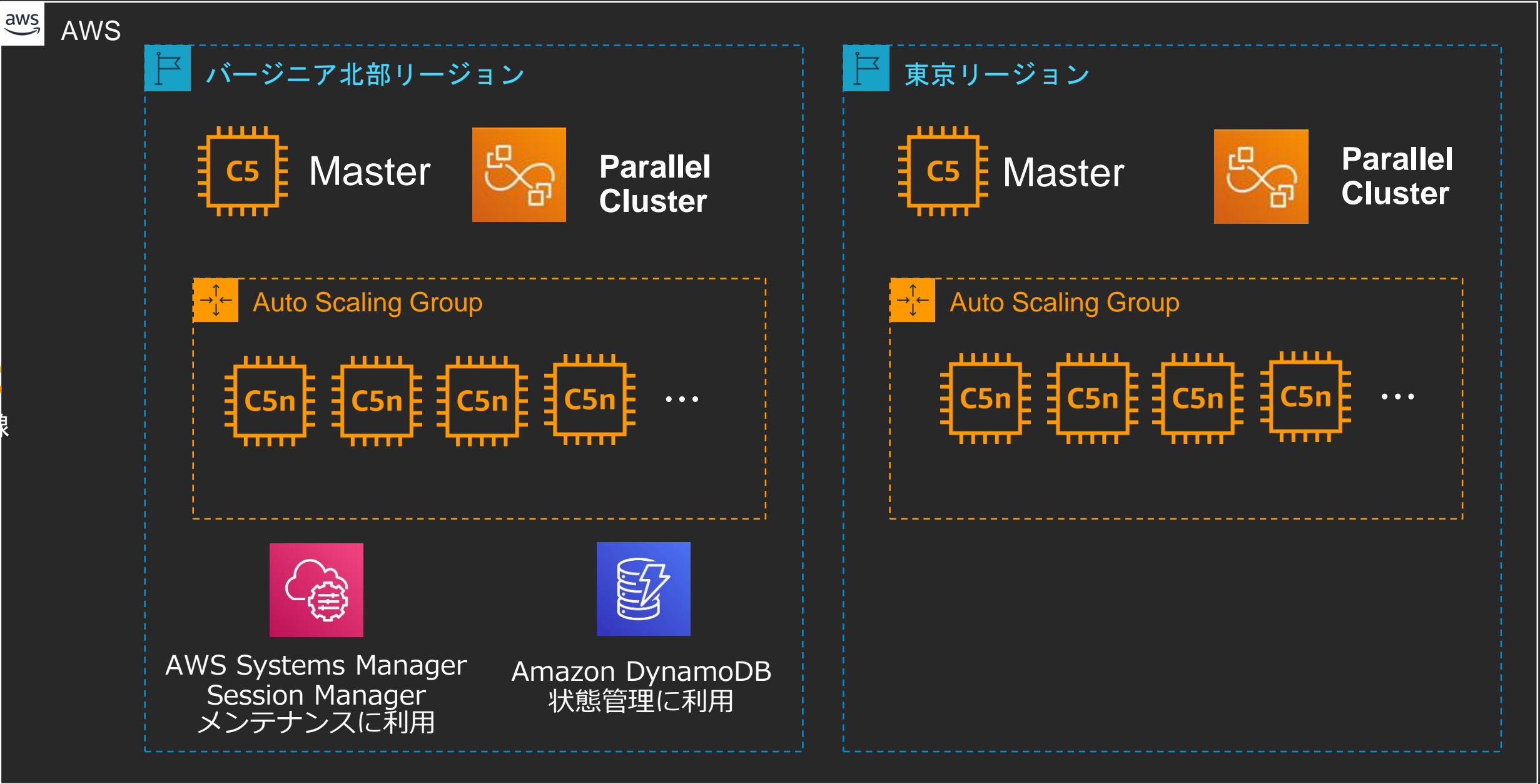
信頼性向上のための方針

- 基本は正系のインスタンスで処理を行い、失敗すれば副系のインスタンスで再度処理を試みる
- 複数のリージョンを利用したフェイルオーバーを実現する

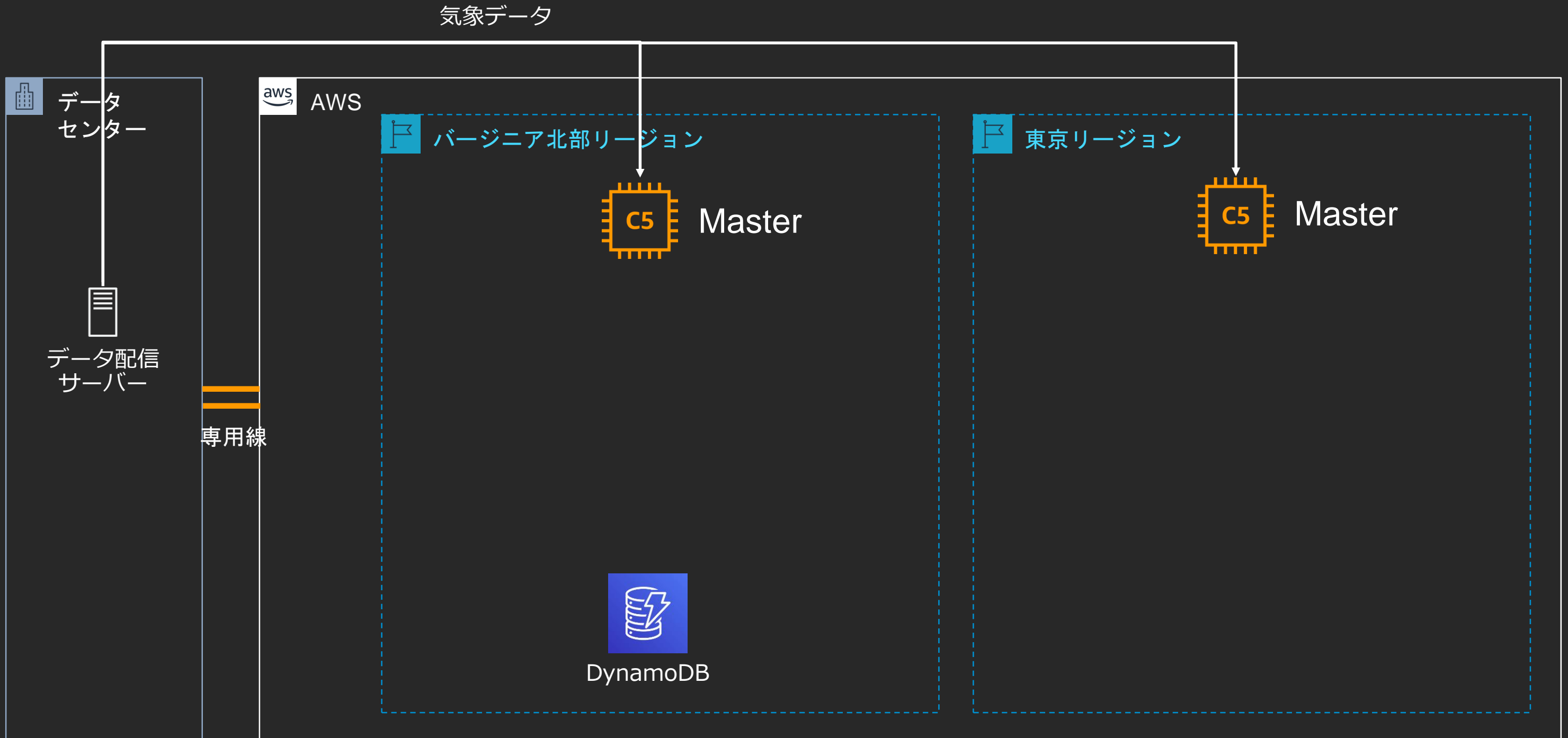
構成案

正系

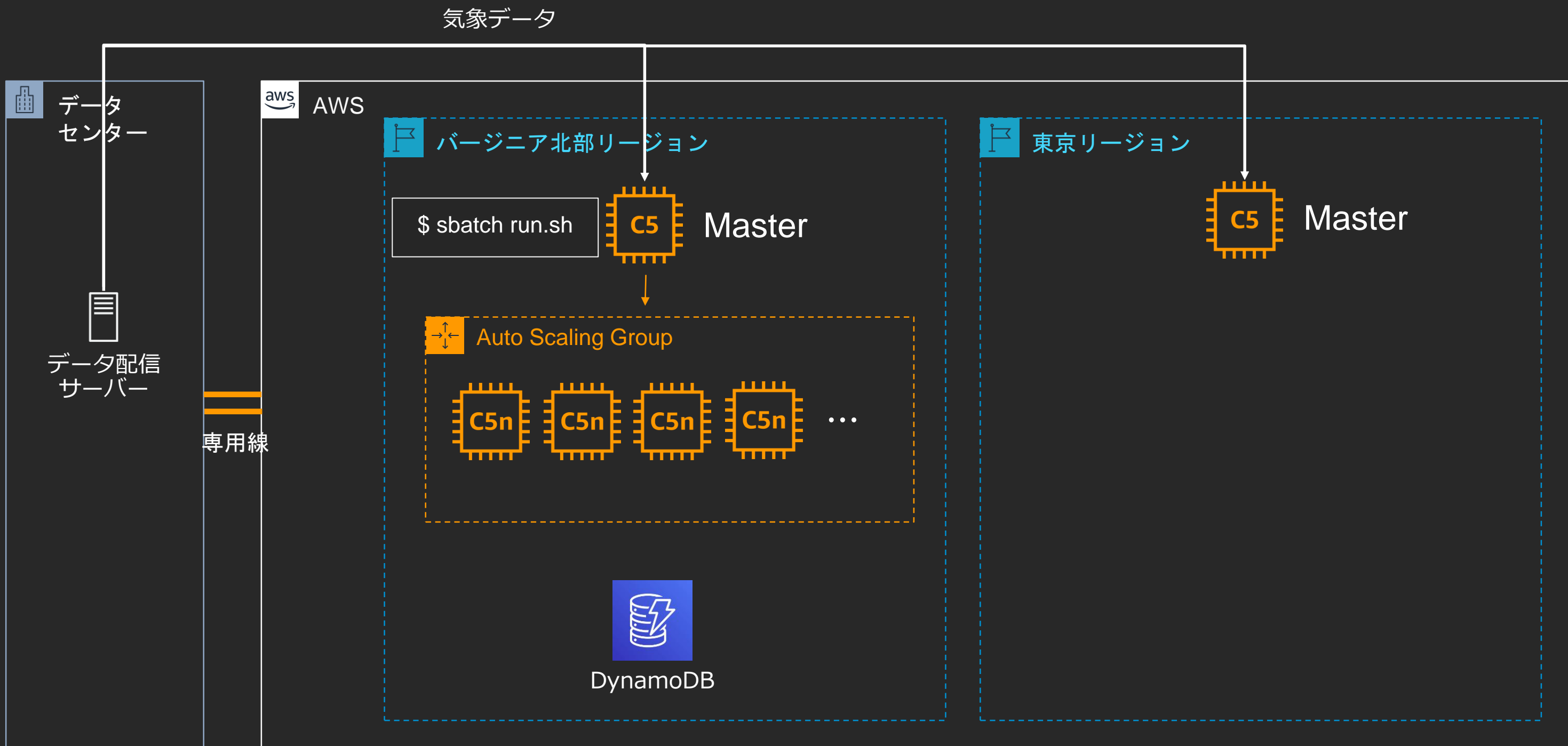
副系



待機時

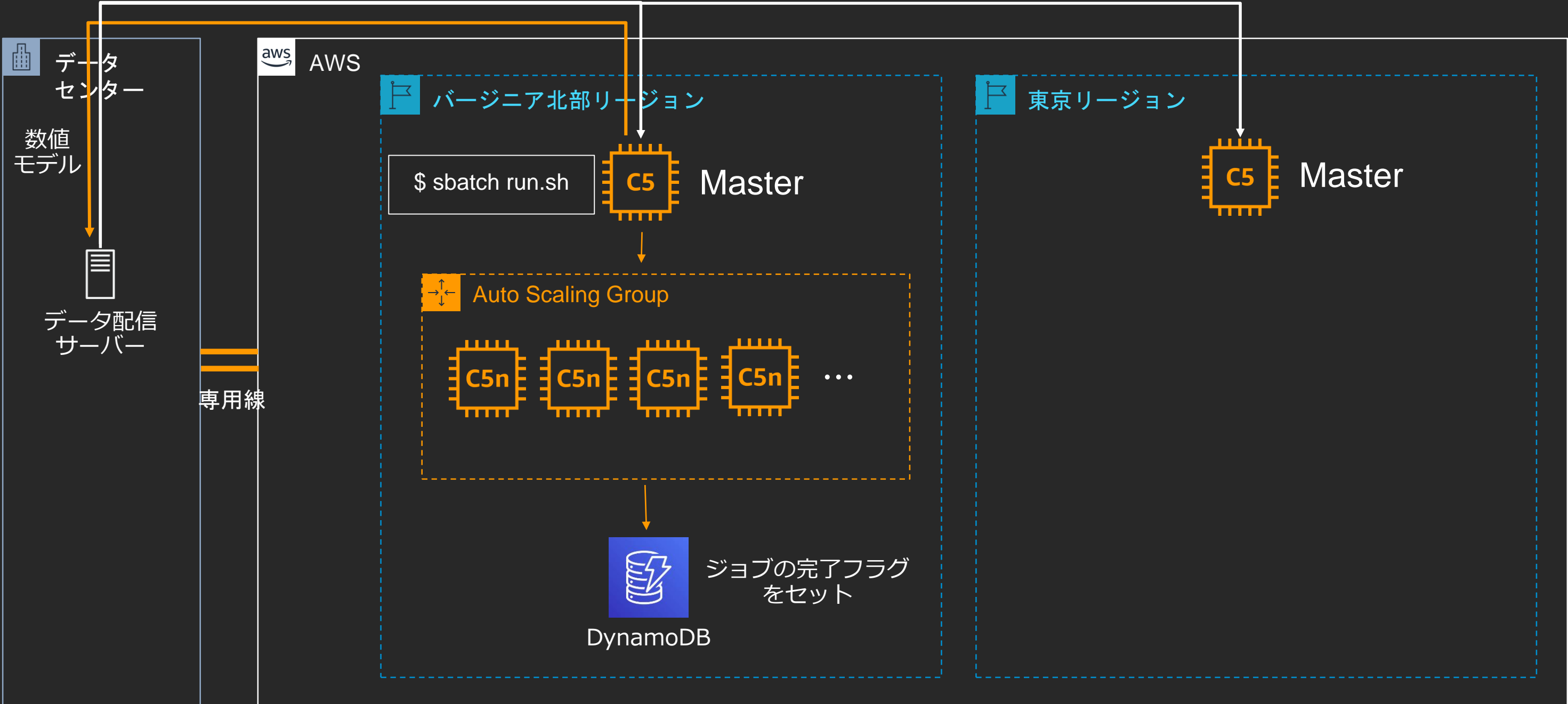


通常の計算処理の流れ

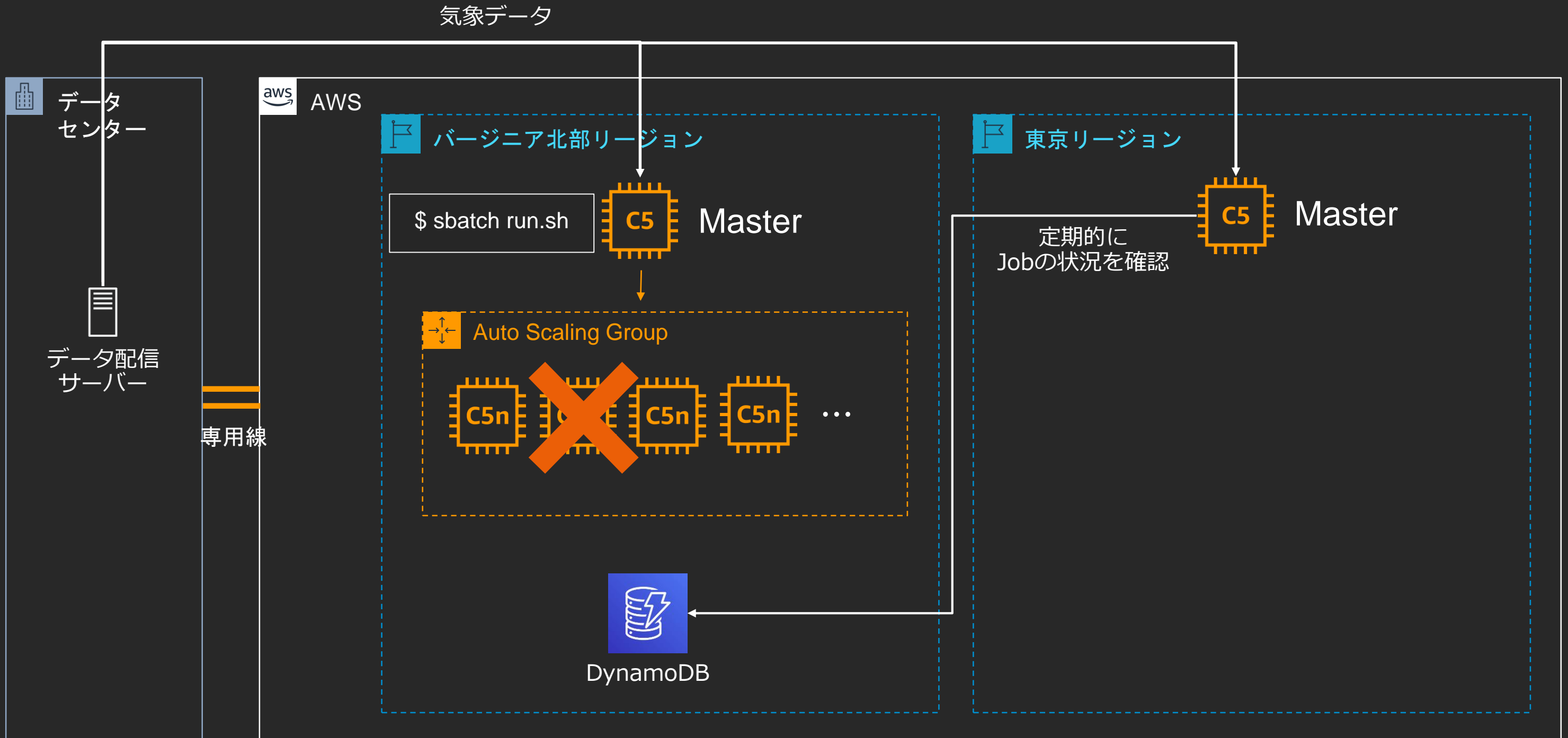


通常の計算処理の流れ

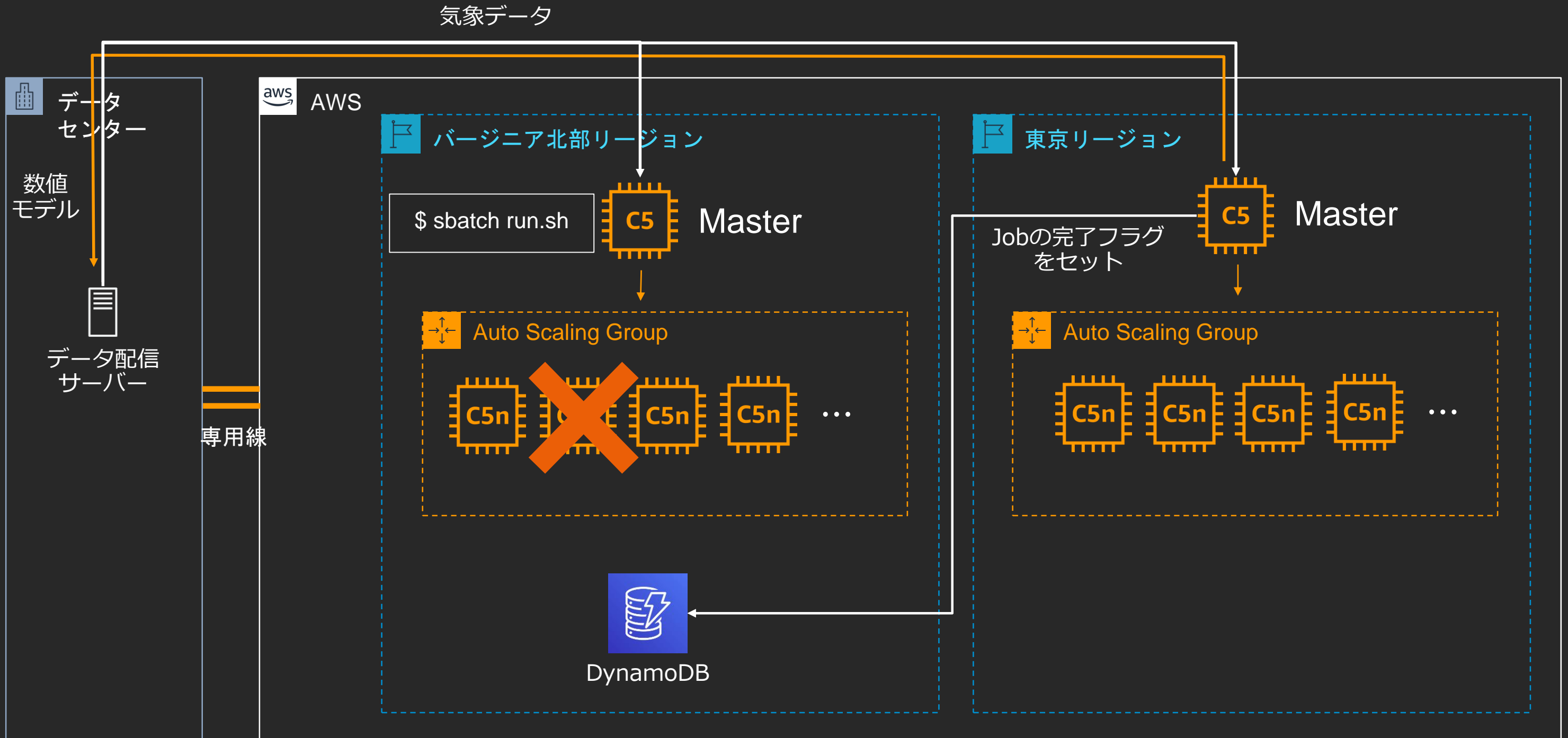
気象データ



副系での計算処理の流れ



副系での計算処理の流れ



コスト最適化

OWN高度化 にあたって求められる要件や課題

パフォーマンス

- 60分以内に24時間分の予報の計算を完了させる
- 72vCPUのサーバー1台だと180分かかる問題サイズ

信頼性

- 現行のシステムと同等以上の可用性
- 許容範囲は年に数回処理が失敗する程度

コスト

- 適切なコストにおさえる
- オンデマンドやSavings Plansの適用以上にコストを削減したい

様々な料金タイプ

オンデマンドインスタンス

任意のタイミングで起動・終了させるデフォルトの利用法
使った分だけの課金

Savings Plans

利用量をコミットすることで割引が受けられる
特に常時起動する場合に適している

スポットインスタンス

非常に安価にインスタンスが利用できる
インスタンスが処理途中で終了させられる可能性がある

正系・副系を組み合わせた料金タイプ

	コスト	正系の安定性
A案:正系・副系オンデマンド	○	○
B案:常時起動でSavings Plans + オンデマンド	△※	◎
C案:スポット + オンデマンド	◎	△

A案 正系・副系：オンデマンド

一番シンプルな構成

正副共に必要なタイミングで起動し、処理完了後に終了する

B案 正系：常時起動+Savings Plans 副系：オンデマンド

正系の安定性が強化される

※ 今回の処理のように一日の中の利用時間が短い場合は、オンデマンドに比べコスト高になる可能性があった 利用頻度が高い場合はコスト面でも優れている

C案 正系：スポット 副系：オンデマンド

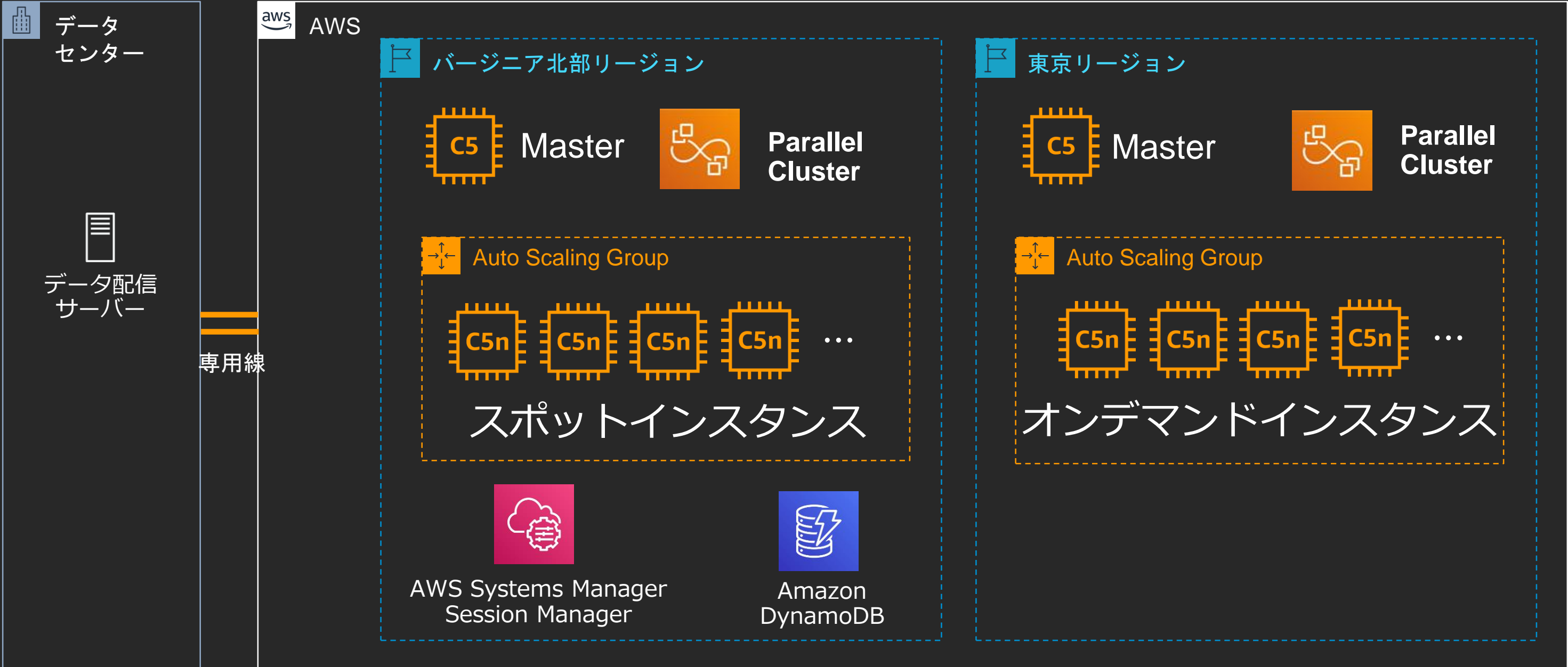
利用頻度の多い正系でスポットの低コストを生かせる

副系機で安定性をカバーする

最終的な構成

正系

副系



ここまでのまとめ

パフォーマンス

- 色々なレイヤーでの試行錯誤、並列化検証
- EFA、Lustre、インスタンスタイプの検討など

信頼性

- 基本スポットで処理を行い、失敗すればオンデマンドで再度処理を試みる
- 複数のリージョンを利用したフェールオーバーを実現する

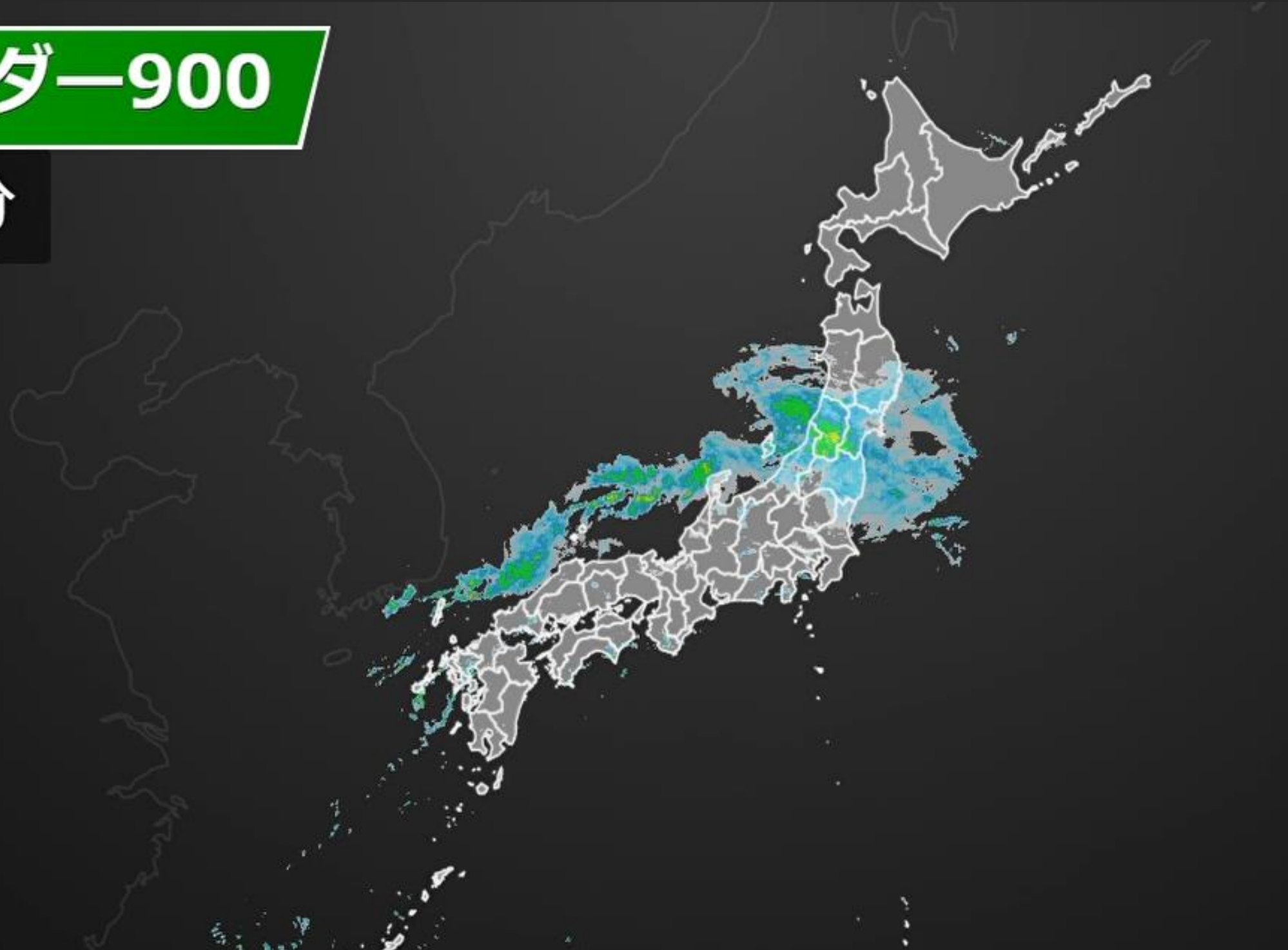
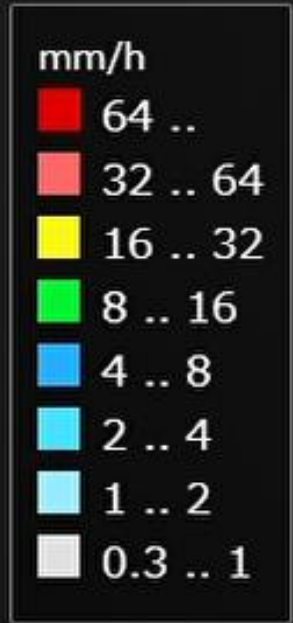
コスト

- スポットインスタンスを利用した大幅なコスト削減

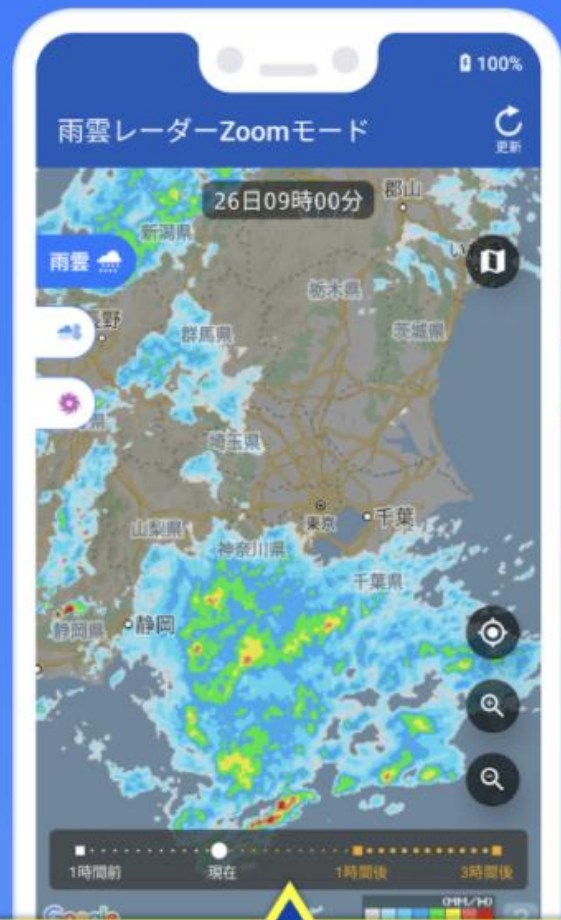
成果物

AIレーダー900

10時30分



雨雲・台風進路・雨雪の境目など
雨雲レーダー



雨が降る時間・止む時間が分かる!

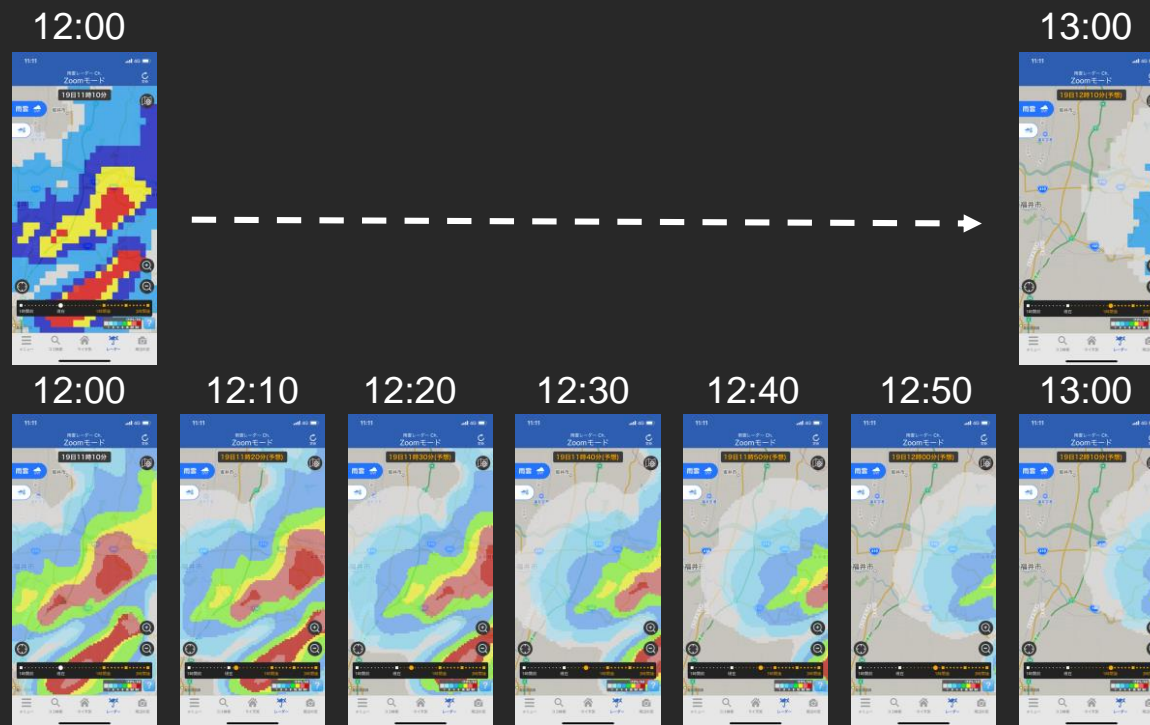
3時間先まで10分ごとの雨雲予想
 AI(人工知能)など最新技術を搭載

Rapid Update OWN Projectの成果

「ウェザーニュース」アプリで見られます!

他社
 1時間間隔

ウェザー
 ニュース
 10分間隔



※1時間より先の予測

	現在～30分後	35分後～1時間後	1時間後～
他社	250mメッシュ/5分間隔	1kmメッシュ/5分間隔	1kmメッシュ/1時間間隔
ウェザー ニュース	250mメッシュ/5分間隔	250mメッシュ/5分間隔	250mメッシュ/10分間隔

ダウンロード(無料)はこちら! (iOS版・Android版) →



Thank you!

Kazuhisa Inuma
WNI Forecast Center
Section Leader
Weathernews