

# AMD EPYC™ インスタンスで Google Cloudを 骨の髄まで使い倒す

Google  
Cloud  
Next

Tokyo

Proprietary





# 小林 宏行

日本AMD コマーシャル営業

## AMDの動向

(第5世代“Turin”の概要とパフォーマンス)

を踏まえ Google Cloud上でサービス提供が開始された

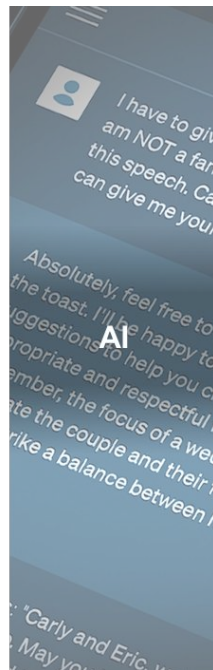
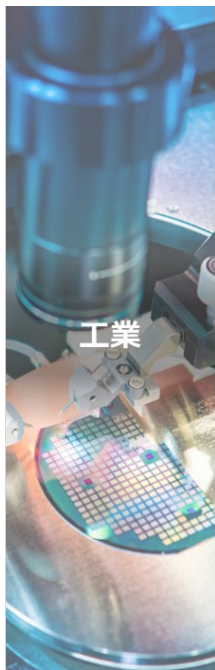
最新コンピュータエンジン C4DとH4Dのメリットと

機密データを扱う際に

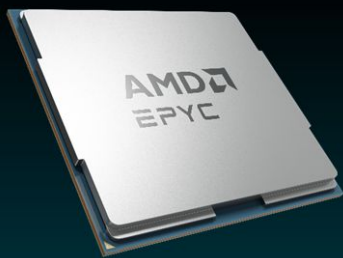
必須となるコンフィデンシャルコンピューティングに

ついて解説いたします。

# 数十億人の日常生活を支える AMD



# AMD データセンター ポートフォリオ



サーバー向け CPU

AMD  
EPYC



GPU アクセラレータ

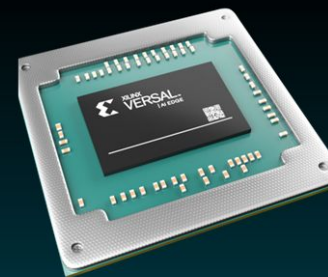
AMD  
INSTINCT



SmartNIC & DPU

AMD  
ALVEO

AMD  
PENSANDO

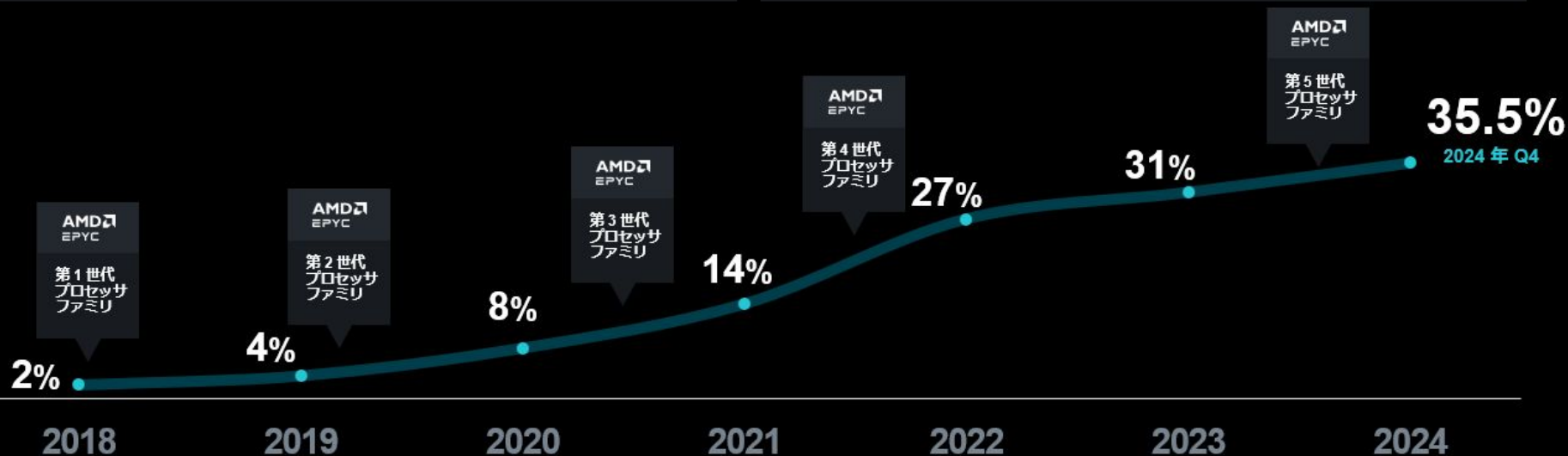


FPGA & アダプティブ SoC

AMD  
VERSAL

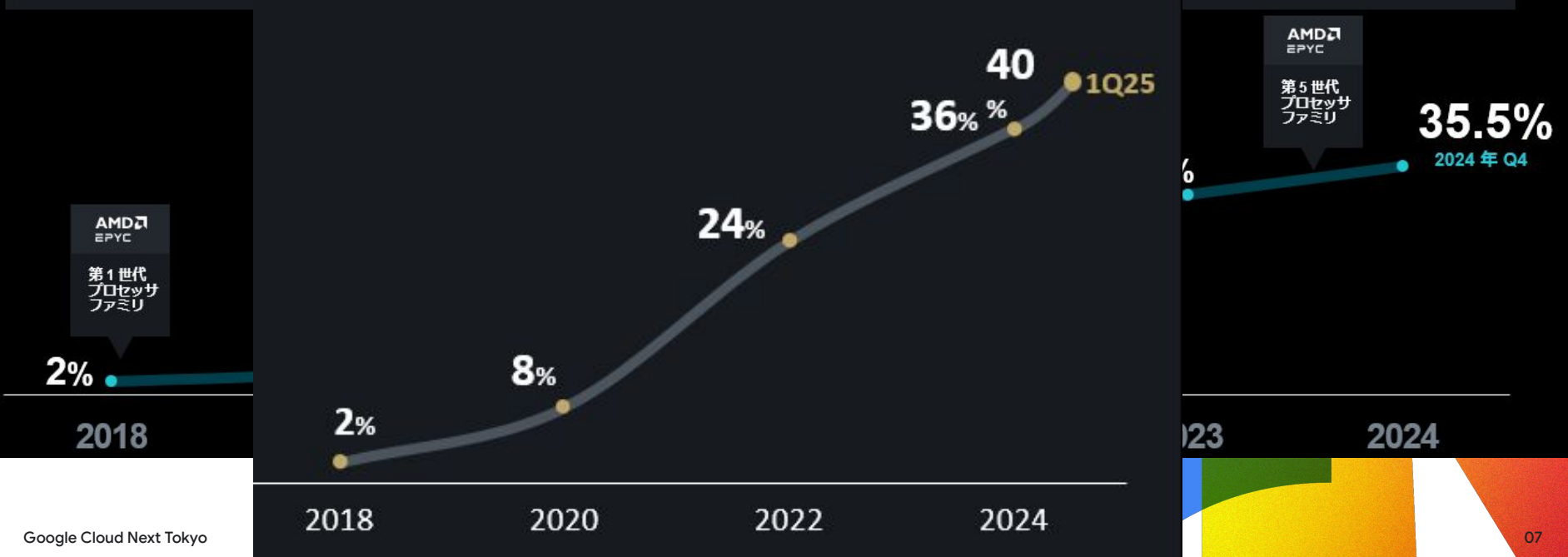
AMD  
ALVEO

# AMD EPYC™ はその信頼性により 世界のサーバーの 3 分の 1 以上に搭載

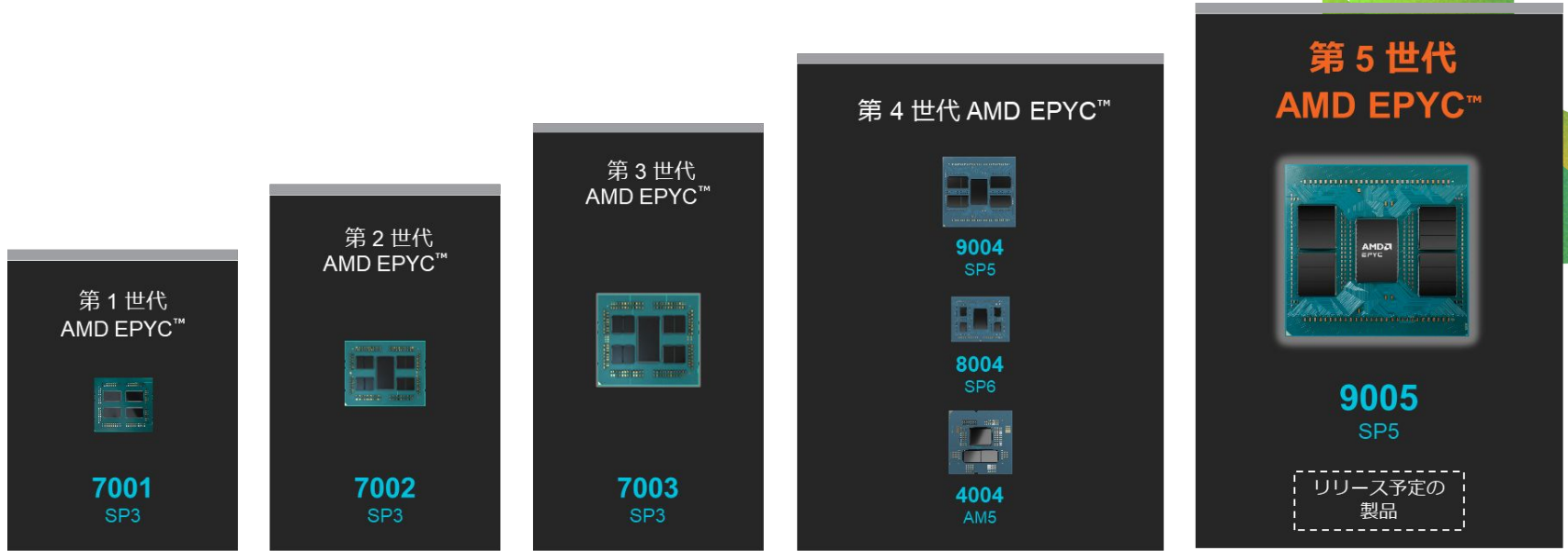


# ついに40%を超えました

>18x Server CPU Market Share Growth

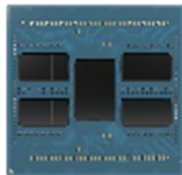


# AMD EPYC™ プロセッサと Google Cloud



# Google Cloud の驚異的な新製品を支える 第 5 世代 EPYC™ の性能

第 4 世代 AMD EPYC



最大 128 コア

最大 3.9 GHz

最大 4800 MT/秒

AVX512

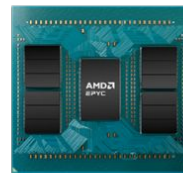
VM サイズの拡大で AI や HPC などの高負荷用途にも対応

世代間的大幅な性能向上と一貫したハイ パフォーマンス

メモリ帯域幅の向上で C シリーズ VM が要求する安定したパフォーマンスを提供

世代間の AI 推論性能の向上と暗号化/復号化パフォーマンスの強化

第 5 世代 AMD EPYC



最大 192 コア

最大 5 GHz (17% の IPC 性能向上)

最大 6400 MT/秒

AVX512 (フル 512b データパス)

# Google Cloud で提供される AMD EPYC™ 搭載マシン ファミリ

一般的用途

一貫したハイパフォーマンス	<b>C4D</b> ベアメタルオプション コンフィデンシャル コンピューティングオプション	<b>第 5 世代 AMD EPYC "Turin"</b> 最大 4.1 GHz   2 ~ 384 vCPU/VM   SMT=off オプション vCPU あたり 2、4、8 GB のメモリ、最大 3 TB 最新世代ローカル SSD、Hyperdisk Extreme 最大 50 万 IOPS、10 GB/秒 最大 200 Gbps のネットワーク帯域幅	トラフィックの多いウェブ、アプリ、 広告サーバー データベース インメモリ キャッシュ ゲーム サーバー ゲーム サーバー メディアストリーミングおよび トランスコーディング データ分析 CPU ベースの推論 SAP 認定 (C3D)
	<b>C3D</b> コンフィデンシャル コンピューティングオプション C3 より 10% 低価格	<b>第 4 世代 AMD EPYC "Genoa"</b> 最大 3.7 GHz   4 ~ 360 vCPU/VM   SMT=off オプション vCPU あたり 2、4、8 GB のメモリ、最大 2.9 TB ローカル SSD 最大 12 TB (4:1、8:1 のメモリ構成で利用可能)、Persistent Disk または Hyperdisk 最大 200 Gbps のネットワーク帯域幅	
高いコスト効率のフレキシブルシェイプ	<b>N2D</b> コンフィデンシャル コンピューティングオプション N2 より 13% 低価格	<b>第 3 世代 AMD EPYC "Milan"</b> 最大 3.5 GHz   2 ~ 224 vCPU/VM   SMT=off オプション vCPU あたり 1、4、8 GB のメモリ、最大 896 GB ローカル SSD 最大 9 TB、Persistent Disk 最大 100 Gbps のネットワーク帯域幅	低から中程度のトラフィックのウェブ/ アプリケーションサーバー デスクトップ仮想化 CRM アプリケーション 開発/テスト環境 SAP 認定
スケールアウト型のコストパフォーマンス	<b>T2D</b> N2 より 13% 低価格	<b>第 3 世代 AMD EPYC "Milan"</b> 最大 3.5 GHz   1 ~ 60 vCPU/VM   SMT=off (シングルスレッド) vCPU あたり 4 GB のメモリ、最大 240 GB Persistent Disk 最大 32 Gbps のネットワーク帯域幅	マイクロサービスのコンテナ化 圧縮/解凍 メディア トランスコーディング 大規模 Java アプリケーション SAP 認定

演算最適化

超ハイパフォーマンス	<b>H4D</b>	<b>第 5 世代 AMD EPYC "Turin"</b> 最大 4.1 GHz   192 vCPU/VM   SMT=off (シングルスレッド) vCPU あたり 4、8 GB のメモリ、最大 1.5 TB 新世代 3.75 TB ローカル SSD (8:1 のメモリ構成で利用可能) 200 Gbps ネットワーク帯域幅と RDMA ハイパーコンピューター クラスタ対応	密結合の HPC ワークロード <ul style="list-style-type: none"> <li>FEA/CFD</li> <li>CAE/EDA</li> <li>WRF</li> <li>モデリングとシミュレーション</li> <li>分子動力学/ゲノム科学</li> </ul>
	<b>C2D</b> コンフィデンシャル コンピューティングオプション C2 より 13% 低価格	<b>第 3 世代 AMD EPYC "Milan"</b> 最大 3.5 GHz   2 ~ 112 vCPU/VM   SMT=off オプション vCPU あたり 2、4、8 GB のメモリ、最大 896 GB ローカル SSD 最大 3 TB、Persistent Disk 最大 100 Gbps のネットワーク帯域幅	メディア トランスコーディング ハイパフォーマンスウェブ/ゲームサーバー AI/ML SAP 認定 (C2D)

# C4D – 第 5 世代 AMD EPYC™ 搭載で 一貫したハイパフォーマンスを実現

## 現在パブリックプレビュー中

- 第 5 世代 AMD EPYC のコア密度とコア IPC の性能向上により世代間の大幅なパフォーマンス向上を実現
- 2-384 vCPU シェイプを利用可能
- vCPU あたり 2、4、8 GB のメモリ、**インスタンスあたり最大 3 TB**
- ティア 1 のメンテナンスと信頼性
- **最大 200 GB** のネットワーキング (ティア 1)
- 最新世代の **LSSD**
- **Hyperdisk Extreme** で最大 50 万 IOPS、10 GB/秒
- **コンフィデンシャルコンピューティング** オプション
- C4D 初の AMD 搭載 **ベアメタル** Google Cloud インスタンス

### C4D に最適なワークロード:

- ウェブ、アプリ、広告サーバー
- データベース、キャッシュ
- データ分析
- メディアストリーミング
- ゲームサーバー
- ネットワークアプライアンス
- AI/ML 推論

### 初期提供:

**日本、東京でお使い  
頂けます！**

(北部)  
(シドニー)

**asia-northeast1**

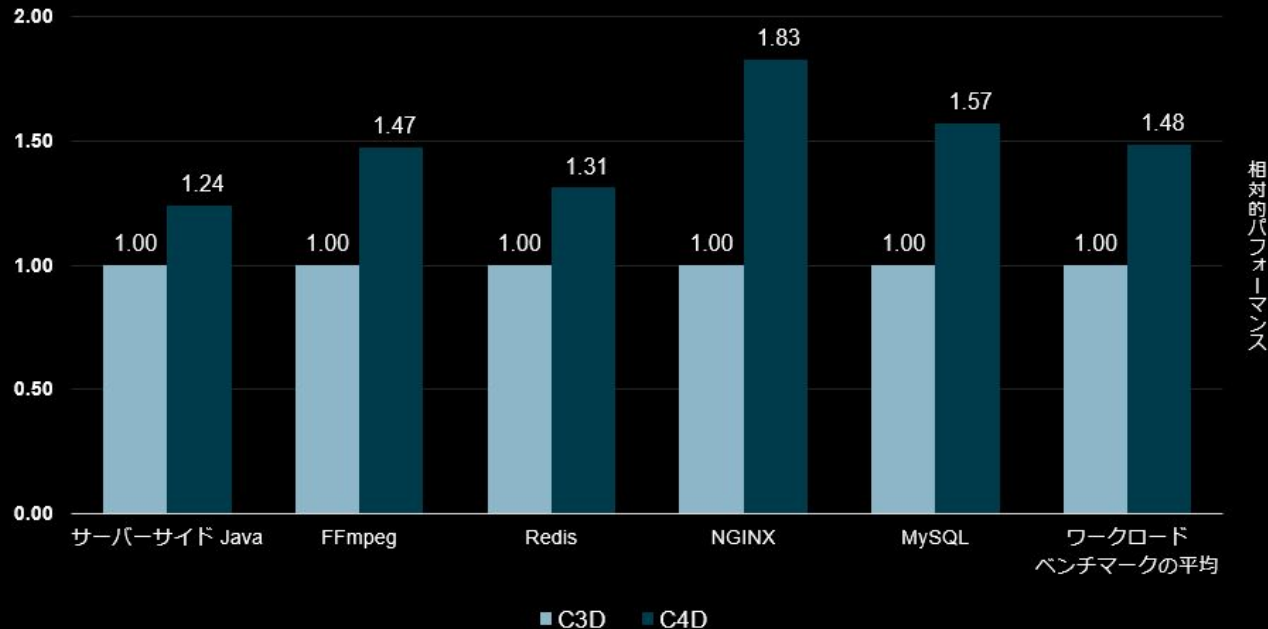
# C4D 一般用途パフォーマンスの向上

## 前世代 C3D との比較

### パフォーマンス向上の比較

業界標準のワークロードベンチマーク全体で、第4世代AMD EPYC搭載のC3D VMと比較して平均48%のパフォーマンス向上

VM パフォーマンス  
C4D 16 vCPU と C3D 16 vCPU の比較



# AMD EPYC™ CPU がデータセンターとクラウドにもたらすもの

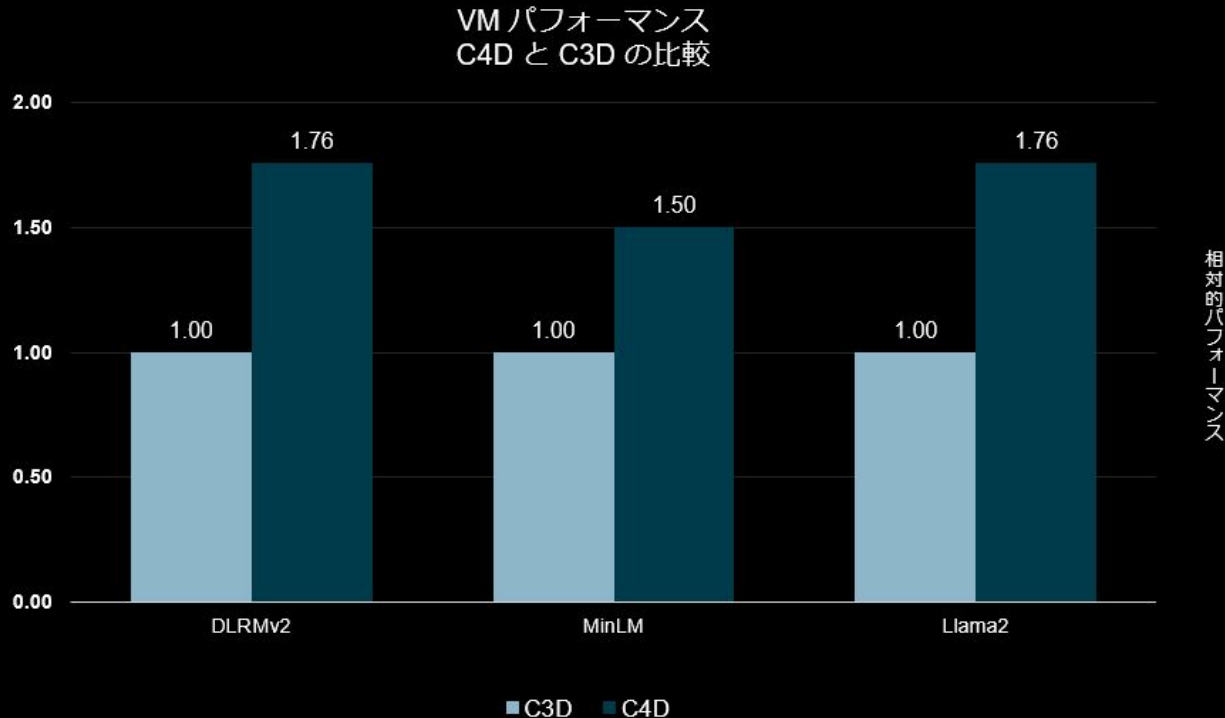
既存のクラウド資産で AI に対応	小/中規模モデル	大規模 AI モデルとトレーニング
<b>ワークロード統合</b>	<b>混合ワークロードとエンタープライズ AI</b>	<b>アクセラレーテッドインスタンス向けのホスト CPU</b>
<ul style="list-style-type: none"> <li>業界トップクラスのパフォーマンス</li> <li>卓越した効率性</li> <li>新しい AI ワークロードに必要なスペース、資金、エネルギーを確保</li> </ul>	<ul style="list-style-type: none"> <li>混在ワークロードの展開</li> <li>小～中規模モデルおよび従来の ML</li> <li>バッチ &amp; 小規模リアルタイム推論</li> </ul>	<ul style="list-style-type: none"> <li>ハイパフォーマンス</li> <li>広範なスケーラビリティ</li> <li>先進の GPU/アクセラレータで認定/認証済み</li> </ul>

# C4D AI 推論パフォーマンスの向上

## 前世代 C3D との比較

### パフォーマンス向上の比較

標準的な 3 つの AI 推論ベンチマーク全体で、第 4 世代 AMD EPYC 搭載の C3D VM と比較して 50% ~ 76% のパフォーマンス向上



# H4D – 第 5 世代 AMD EPYC™ 搭載の 超ハイ パフォーマンス

現在プライベートプレビュー中

- 第 5 世代 AMD EPYC のコア密度と性能向上で、密結合の HPC ワークロードと演算負荷の高いワークロードに対応
- **192 コア** (シングルスレッド) フルノードシェイプ
- vCPU あたり 4、8 GB のメモリーインスタンスあたり **最大 1.5 TB**
- **200 GB ネットワーキング**を備えた **RDMA** ベースのスケラビリティ
- メモリ比率が 8:1 のインスタンスで利用可能な **ローカル SSD**
- 展開、オーケストレーション、メンテナンスのためのツールを備えた、大規模で高密度なクラスター向けの **Google Cloud Cluster Director**
- **Dynamic Workload Scheduling** が、バースト負荷への対応を求められるお客様の HPC クラウド使用を最適化

## H4D に最適なワークロード:

- 密結合ハイパフォーマンスコンピューティング (HPC):
  - FEA/CFD
  - CAE/EDA
  - WRF
  - モデリングとシミュレーション
  - 分子動力学/ゲノム科学
- AI/ML

## 初期提供:

- us-central1 (アイオワ)
- europe-west4 (オランダ)



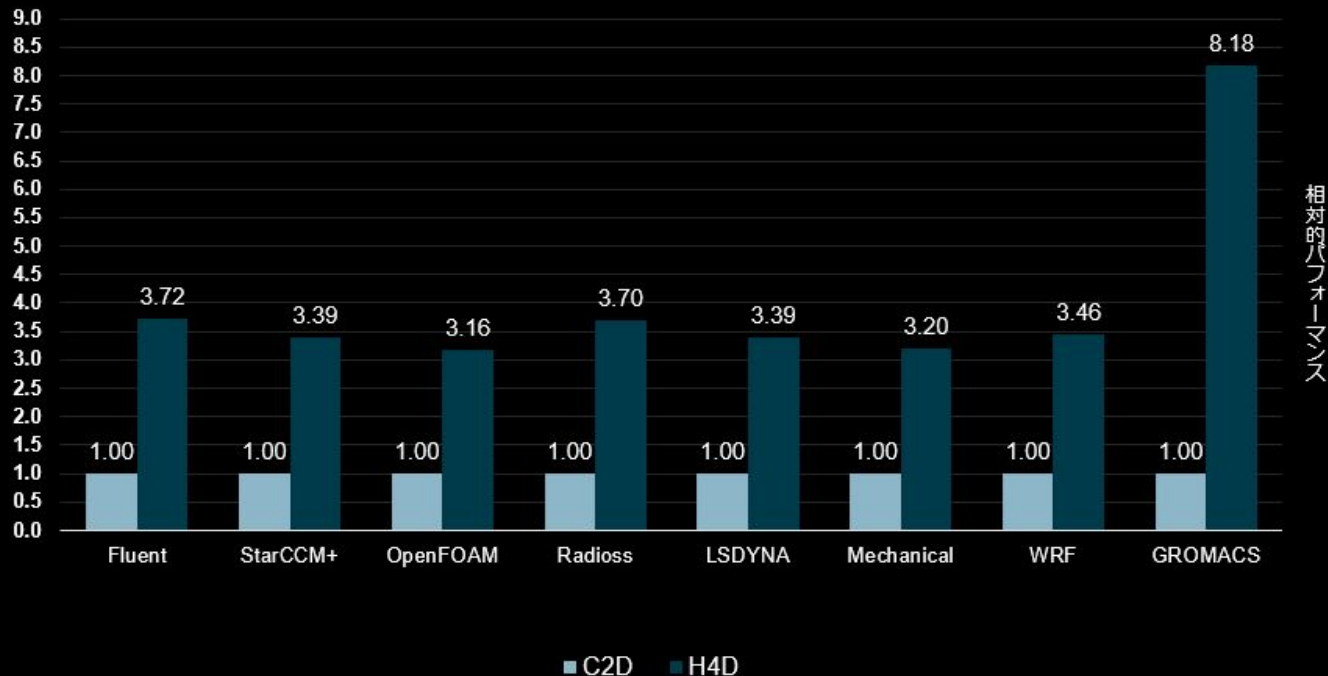
# H4D HPC パフォーマンスの向上

## 前世代の C2D との比較

### VM パフォーマンス H4D と C2D の比較

#### パフォーマンス向上の比較

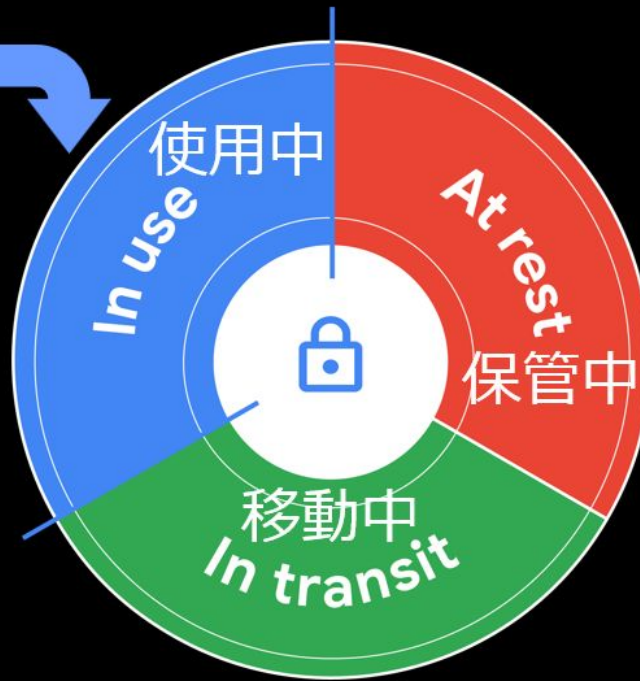
幅広い標準的 HPC ワークロード  
ベンチマーク全体で、第 3 世代 AMD EPYC  
搭載の C2D VM と比較して平均で  
約 4 倍のパフォーマンス向上



# Google CloudとAMDに依る コンフィデンシャルコンピューティング

機密性をデータ使用時にも拡張するブレイクスルー技術

AMD インフィニティ ガード  
ハードウェア暗号化がデータの  
使用中も有効に出来ます



# Google Cloud での AMD テクノロジを活用したコンフィデンシャルコンピューティング

AMD SEV (Secure Encrypted Virtualization) & AMD SEV-SNP (Secure Nested Paging)

Confidential  
VM

Confidential  
GKE

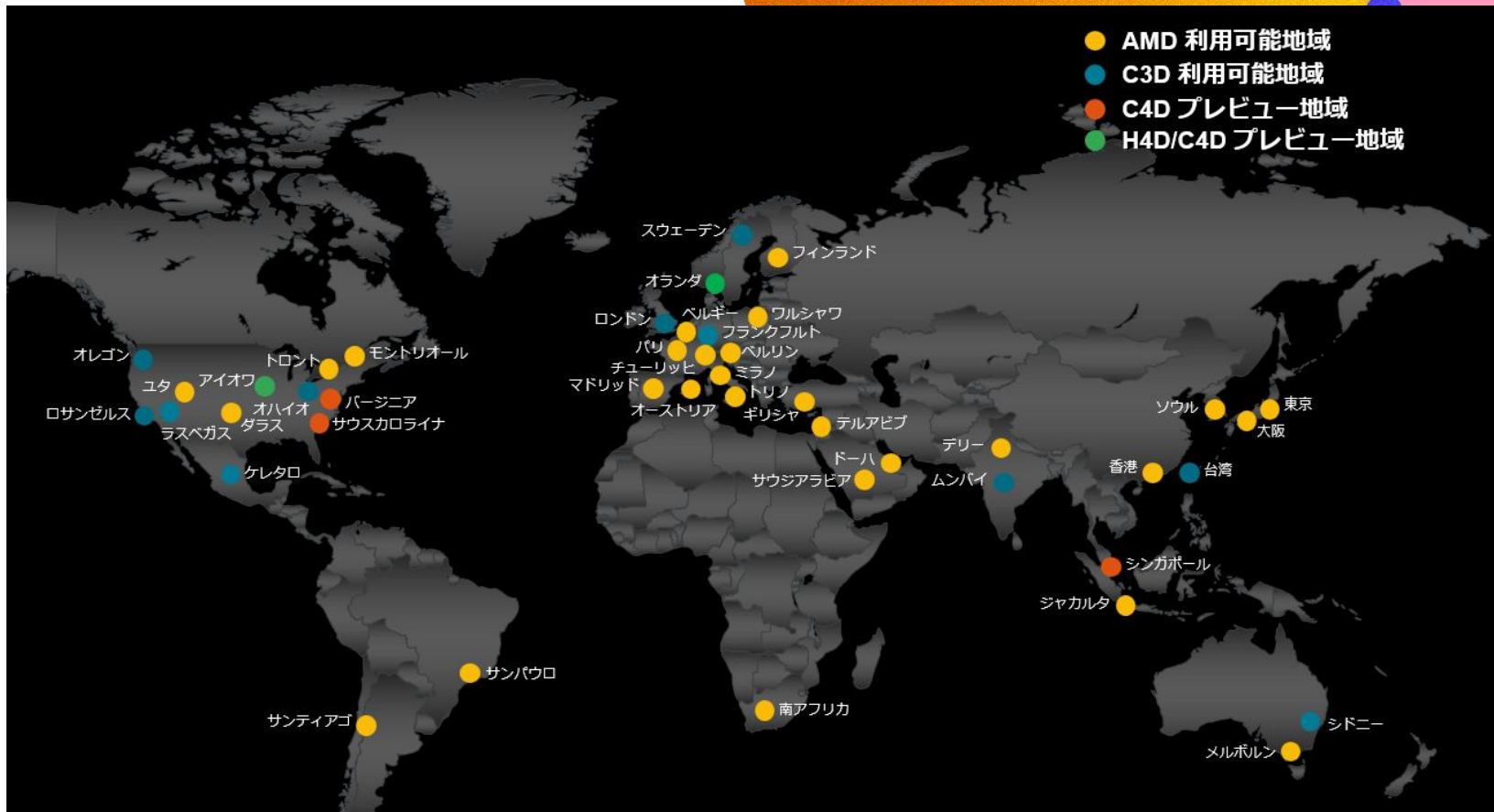
Confidential  
Dataproc

Confidential  
Dataflow

Confidential  
Space

Confidential VM	AMD SEV 有効化	AMD SEV-SNP 有効化
可用性	一般提供	一般提供
データのセキュリティ	機密性	機密性、整合性
証明	vTPM ベースの証明	ハードウェアに根ざした (コア ファームウェア) vTPM ベースの証明
マシン ファミリ	汎用の C4D、C3D、N2D 演算に最適化された C2D	汎用の N2D

# 利用可能な仮想マシンマップ



**AMD** 



ご清聴ありがとうございました。