



# NVIDIA L40

## データセンター向けのかつてない ビジュアルコンピューティング パフォーマンス



仮想ワークステーションアプリケーションから大規模なモデリングおよびシミュレーションに至るまで、最新のビジュアルコンピューティングおよび科学技術ワークフローは、複雑さと量の両方で大きく成長しています。企業は、これらのますます複雑化するワークロードの多様なコンピューティング要求を克服するために、究極のパフォーマンスと多用途の機能を備えた拡張性を提供できるデータセンターテクノロジーを必要としています。

NVIDIA® L40 GPU は、データセンターに前例のないビジュアルコンピューティングパフォーマンスを提供し、次世代のグラフィックス、コンピューティング、および AI 機能を提供します。革新的な NVIDIA Ada Lovelace アーキテクチャ上に構築された NVIDIA L40 は、最新世代の RT、Tensor、および CUDA コアのパワーを利用して、最も要求の厳しいデータセンターワークロードに対して画期的なビジュアライゼーションとコンピューティングパフォーマンスを提供します。

### NVIDIA Ada Lovelace アーキテクチャを採用

#### 第3世代 RT コア

強化されたスループットとコンカレントでのレイ トレーシングおよびシェーディング機能により、レイ トレーシングのパフォーマンスが向上し、製品設計と建築、エンジニアリング、および建設ワークフローのレンダリングが高速化されます。ハードウェアアクセラレーションによるモーション ブラーを使用してリアルなデザインの動作を確認し、驚異的なリアルタイム アニメーションを実現します。

#### 第4世代 Tensor コア

構造的スパーシティと最適化された TF32 形式のハードウェア サポートにより、すぐに使用できるパフォーマンスが向上し、AI とデータサイエンスモデルのトレーニングが高速化されます。DLSS を含む AI で強化されたグラフィックス機能を加速し、一部のアプリケーションでより優れたパフォーマンスで解像度を向上させます。

#### 大容量 GPU メモリー

48 GB の超高速 GDDR6 メモリーにより、データサイエンス、シミュレーション、3D モデリング、レンダリングなど、メモリを大量に消費するアプリケーションやワークロードに取り組みます。vGPU ソフトウェアを使用して複数のユーザーにメモリを割り当て、クリエイティブ、データサイエンス、およびデザインチーム間で大きなワークロードを分散します。

#### データセンター対応

NVIDIA L40 は、電力効率の高いハードウェアとコンポーネントを使用することで 24 時間 365 日のエンタープライズ データセンター運用向けに設計されており、大規模な展開に最適化されており、さまざまなデータセンターワークロードに対して最大のパフォーマンスを提供します。

L40 には、追加のセキュリティ層を提供するルートオブトラストテクノロジーを使用したセキュアブートが含まれており、NEBS レベル 3 に準拠して最新のデータセンター標準に適合しています。L40 は、デュアルスロット、パッシブ冷却、電力効率の高い設計でパッケージ化されており、大手 OEM ベンダーのさまざまな NVIDIA 認定システム (NVIDIA-Certified Systems™) で利用できます。

### 次世代のワークロードを 高速化

- > NVIDIA Omniverse™ Enterprise
- > レンダリングと3Dグラフィックス
- > NVIDIA RTX™ 仮想ワークステーション (RTX vWS) ソフトウェアを使用した高性能仮想ワークステーション
- > AI トレーニングとデータサイエンス
- > ストリーミングとビデオコンテンツ

## 技術仕様\*

GPU アーキテクチャ	NVIDIA Ada Lovelace アーキテクチャ
GPU メモリー	48GB GDDR6 ECC付き
メモリー帯域幅	864GB/s
インターコネク ト インタフェース	PCIe Gen4x16: 64GB/s 双方向
NVIDIA Ada Lovelace アーキテクチャのCUDA コア数	18,176
NVIDIA 第3世代 RT コア	142
NVIDIA 第4世代 Tensor コア	568
RT コア 性能 TFLOPS	209
FP32 TFLOPS	90.5
TF32 Tensor コア TFLOPS	90.5   181**
BFLOAT16 Tensor コア TFLOPS	181.05   362.1**
FP16 Tensor コア	181.05   362.1**
FP8 Tensor コア	362   724**
ピーク INT8 Tensor TOPS	362   724**
ピーク INT4 Tensor TOPS	724   1448**
フォームファクター	4.4" (H) x 10.5" (L) - デュアルスロット
ディスプレイポート	4 x DisplayPort 1.4a
最大消費電力	300W
電源コネク ター	16-pin
サーマル	パッシブ
仮想GPU (vGPU) ソフトウェア対応	Yes
vGPU プロファイル対応	<b>仮想GPUライセンスガイド<sup>1</sup></b> を参照
NVENC   NVDEC	3x   3x (AV1 Encode & Decode を含む)
ルート オブ トラスト対応セキュアブート	Yes
NEBS 準拠	Level 3
MIG サポート	No
NVLink サポート	No

\* 暫定的な仕様であり、変更される可能性があります

\*\* スパース性を含む

## 始める準備はできましたか？

NVIDIA L40 GPU のさらに詳しい情報は:

[www.nvidia.com/ja-jp/data-center/l40](http://www.nvidia.com/ja-jp/data-center/l40)

1. NVIDIA vGPU ソフトウェアの今後のリリースで提供予定

© 2023 NVIDIA Corporation and affiliates. All rights reserved. NVIDIA, the NVIDIA logo, NVIDIA-Certified Systems, NVIDIA Omniverse and NVIDIA RTX are trademarks and/or registered trademarks of NVIDIA Corporation and affiliates in the U.S. and other countries. Other company and product names may be trademarks of the respective owners with which they are associated. 2436245. FEB23

