

AIの性能をアクセラレート：「MLPerf Training v3.0」における Dell PowerEdge XE9680のパフォーマンス測定結果

著者: Liz Raymond, Naye Yoni, Frank Han, Delmar Hernandez

エグゼクティブサマリー

PowerEdge XE9680は、人工知能、機械学習、およびHPCワークロードにおいて、妥協なきパフォーマンスの実現のために設計および最適化された高性能サーバです。高度な機能を備えた革新的な 8 GPU プラットフォームが、Dell PowerEdge ラインアップに追加されました。

- NVIDIA H100 80GB 700W SXM GPU×8 または A100 80GB 500W SXM GPU×8
- 第4世代インテル® Xeon® スケーラブル・プロセッサ×2
- DDR5 DIMM at 4800MT/s ×32
- PCIe Gen 5 x16 FHスロット×10
- SAS/NVMe SSD (U.2 or E.3) × 8ベイおよび NVMe RAID対応のBOSS-N

PowerEdge XE9680の卓越した性能に関する洞察を提供するレポートを共有できることを嬉しく思います。本レポートは、MLCommonsのMLPerf 3.0ベンチマークを使用した厳密なテストと評価を通じ、AIモデルの学習におけるPowerEdge XE9680の卓越した性能を詳細に分析しています。

MLPerf 3.0 MLPerfは、機械学習（ML）ワークロードの性能を評価するベンチマーク群で、MLのライフサイクルで重要な2側面、学習と推論に焦点を当てます。本書では学習の側面について詳しく解説します。

Dell Performance Labs

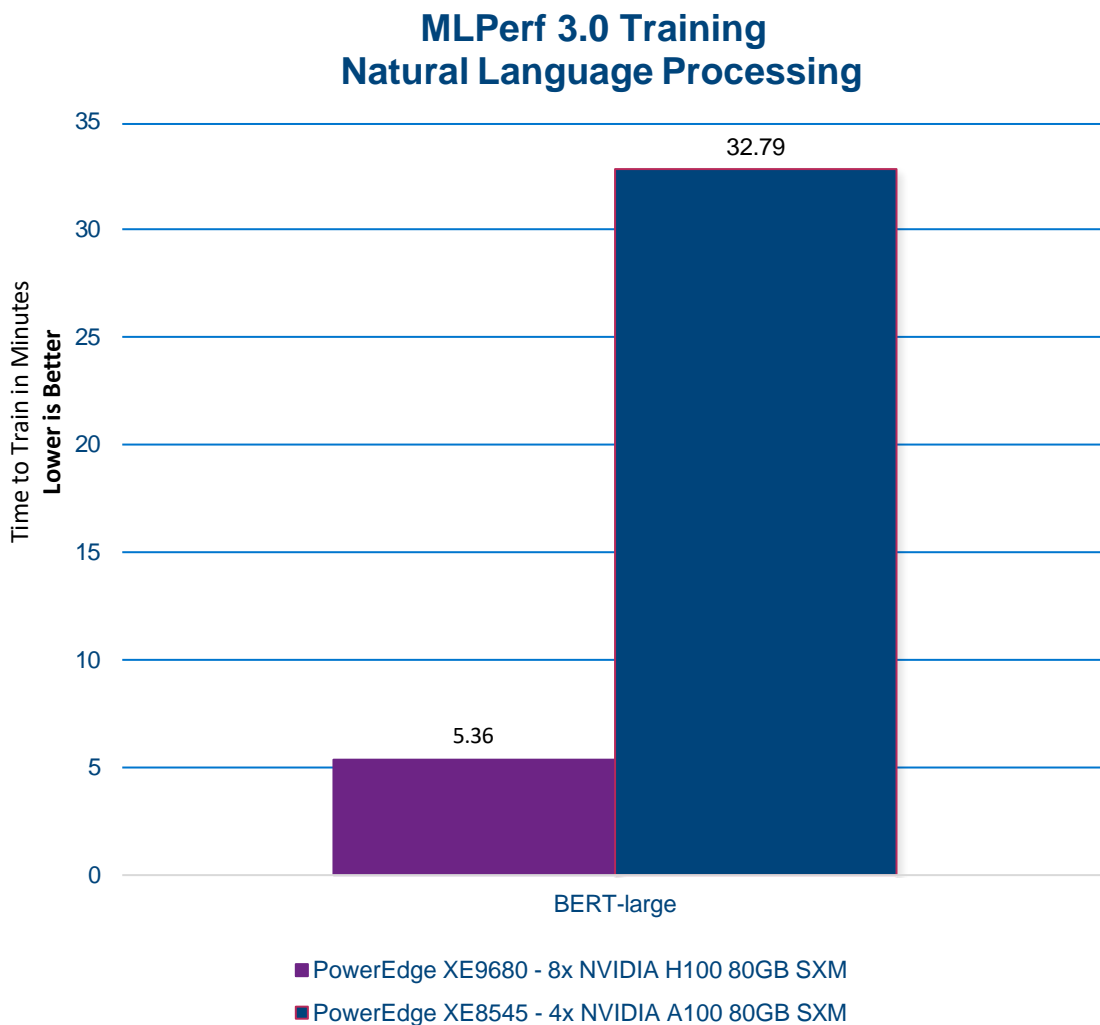


パフォーマンス

今回Dellのパフォーマンスラボでは、8基のNVIDIA H100 80GB SXM GPUを搭載した最新のPowerEdge XE9680を使用して、MLPerf 3.0 Trainingベンチマークを実施しました。比較のため、4基のNVIDIA A100 80GB SXM GPUを搭載した前世代のPowerEdge XE8545でも同様のテストを実行しました。

BBERT（Transformersの双方向エンコーダー表現）は、Googleが2018年に発表したトランスフォーマーベースのニューラルネットワークモデルです。このモデルは、各シーケンス内の単語の文脈と意味を捉えることで人間のようにテキストを理解し生成をするよう設計されています。我々は、H100 GPUを搭載したPowerEdge XE9680が、Wikipediaデータセットを使用したBERT-largeモデルを用いたMLPerf NLPベンチマークで、6倍ものtime-to-train性能の向上を実現したことを嬉しく思います。この結果は、モデルの学習を大幅に高速化することに貢献し、価値実現までの時間を短縮できることを示します。

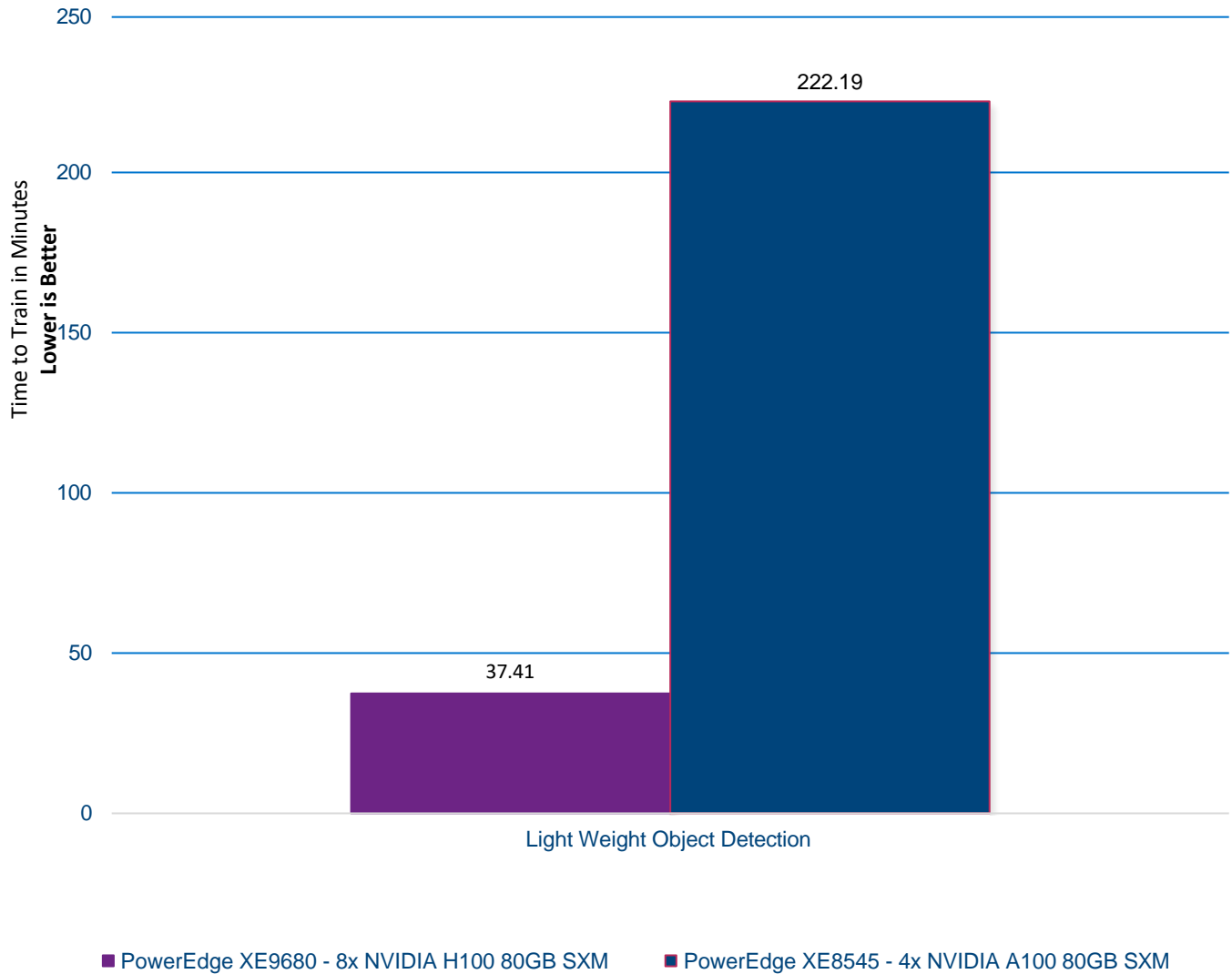
このレポートでは全体を通して、time-to-train、つまり学習完了までの時間の値が低いほど、PowerEdge XE9680がAIモデルの学習を高速化し、スピードと効率性が向上していることを示していることに留意してください。



MLPerf 3.0においてRetinaNetモデルは、数百万枚の多様な画像を含むOpen Imagesデータセットを活用します。このベンチマークで、RetinaNetモデルの学習時間が6倍近くまで向上することを確認しました。

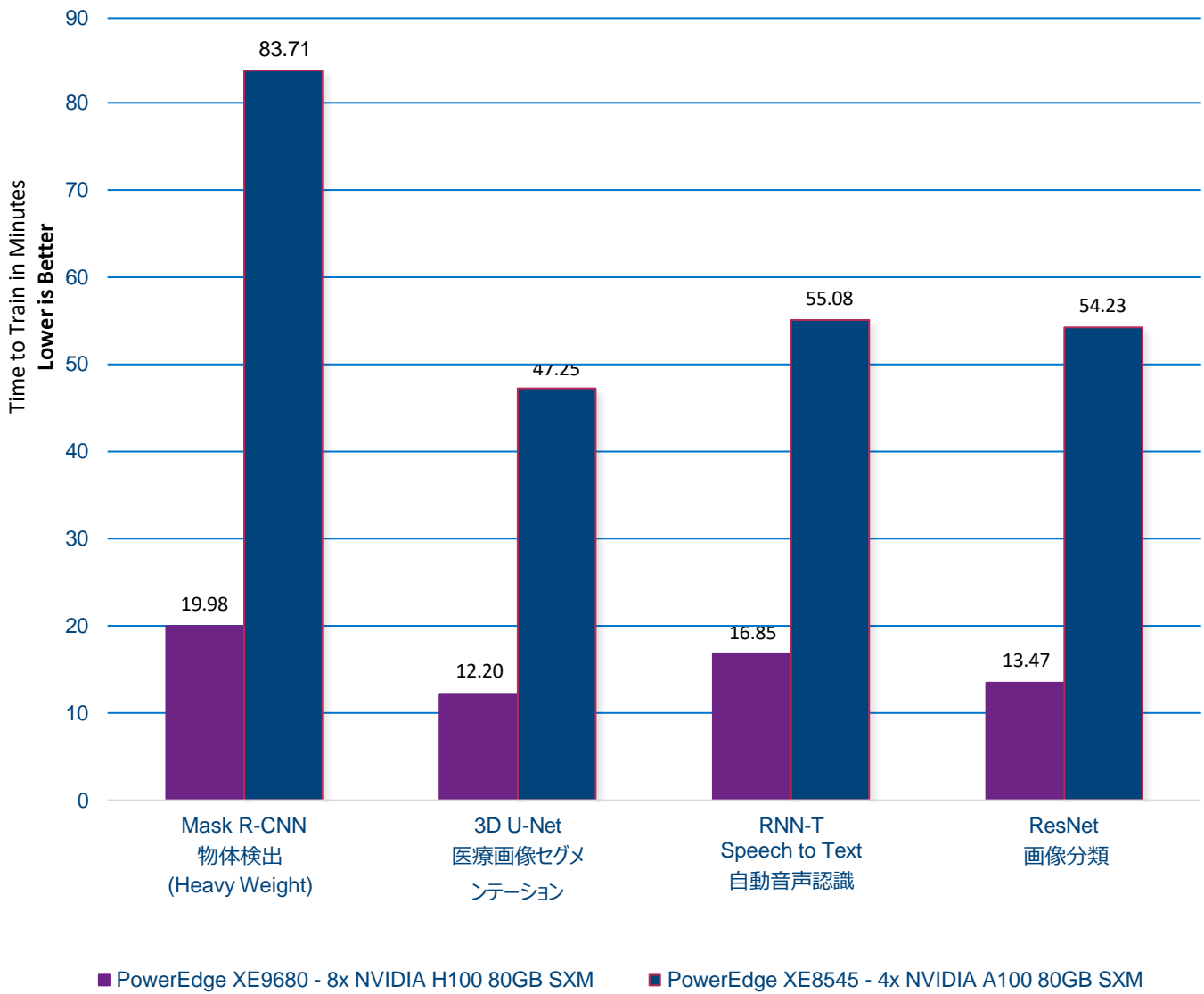
RetinaNetモデルとOpen Imagesデータセットを活用することで、MLPerfではシステムの能力の包括的な評価と比較が可能になります。データセットの規模と多様性は、様々なドメインや物体カテゴリでの物体検知性能の評価に堅牢さもたらします。

MLPerf 3.0 Training Object Detection



以下の図に示すように、PowerEdge XE9680はMLPerf 3.0 Training ベンチマークスイート全体で、一貫して優れた結果を出しています。この堅牢なパフォーマンスは、要求の厳しい幅広い機械学習タスクに対応するサーバーの卓越した能力と信頼性を裏付けています。

MLPerf 3.0 Training



結論

PowerEdge XE9680サーバは、前世代の製品を最大6倍のパフォーマンス向上により凌駕しています。この驚異的な進化により、AIモデルの学習が大幅に高速化され、チームは学習タスクをより迅速に完了できるようになります。このサーバーの詳細については、専任のアカウントエグゼクティブにお問い合わせいただくか、www.dell.com にアクセスすることをお勧めします。

表 1. サーバーの構成

	PowerEdge XE8545	PowerEdge XE9680
CPU	2x AMD EPYC 7763 64-Core Processor	2x Intel® Xeon® 8470 52-core Processor
GPU	4x NVIDIA A100-SXM-80GB (500W)	8x NVIDIA H100-SXM-80GB (700W)

- 1) 2023年6月にDellが実施したテストに基づきます。テストは8基のNVIDIA H100 SXM4-80GB 搭載の PowerEdge XE9680 および 4基の NVIDIA A100-SXM-80GB 搭載の PowerEdge XE8545 で実施しました。MLPerf v3.0 BERT Large、Mask R-CNN、ResNet、RetinaNet、RNN-T、3D U-Netモデルの学習結果です。MLPerf の名称およびロゴは、米国およびその他の国における MLCommons Association の商標です。無断複写・転載を禁じます。無断使用は固く禁じられています。詳しくは www.mlcommons.org をご覧ください。個々の結果は異なることがあります。

参考情報

MLPerf - <https://mlcommons.org/en/training-normal-30/>



For more info,
visit the [Servers
Info Hub](#)



[Contact us](#) for
feedback and
requests



Follow us for
PowerEdge
news