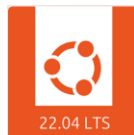


LLM

Large Language Models



LLM(大規模言語モデル)のトレーニングに
おススメ！
ハイスペックGPUを搭載したアプライド製
HPCのご紹介！



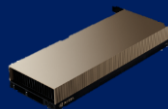
- CPU : Xeon Gold 6530 (32コア/64スレッド/2.1GHz/tb4.0GHz)
- メモリ : 2,048GB (64GB×32)
- ストレージ : 1.92TB SSD + 7,68TB SSD
- GPU : NVIDIA® H100 94GB
- OS : Ubuntu 22.04 LTS
- フレームワーク : TensorFlow / Pytorch / Chainer (DockerDesktop)
- 光学ドライブ : なし
- 電源 : [4基] 3,000W/200V- 冗長化仕様 (2+2)
- 80 Plus Titanium 認証
- 3年間センドバック方式ハードウェア保証

APPLIED
HPC Deep Type-
AS4UX2S8GP-AU019

16,800,000 (税込) 円

カスタマイズのご要望も承ります

NVIDIA H100 Tensor コア GPU 搭載



H100 では、推論が最大 30 倍高速化になる、レイテンシが最小限に抑えられるなど、機能が強化されます。それにより、市場をリードする NVIDIA の推論のリーダーシップをさらに拡大します。第 4 世代の Tensor コアは FP64、TF32、FP32、FP16、INT8 など、あらゆる精度をスピードアップします。Transformer Engine は FP8 と FP16 の両方を活用してメモリ消費を減らしてパフォーマンスを増やしつつ、大規模な言語モデルで精度を維持します。H100 は、倍精度 Tensor コアの毎秒浮動小数点演算 (FLOPS) を 2.6 倍にし、HPC で 50 teraFLOPS の FP64 コンピューティングを実現します。



<https://www.applied.ne.jp/rs/>
または



LLM(大規模言語モデル)運用におすすめ！ ハイスpek GPUを搭載したアプライド製 HPCのご紹介！



デュアル電源搭載キャスター付きフルタワー筐体

本モデルはXeonプロセッサを搭載したLLM運用向けワークステーションです。長時間の運用には熱対策が必須です。筐体内部の冷却には、前面に120mmの静音ファンを3基搭載しています。さらにプロセッサの冷却には、簡易水冷式プロセッサ・クーラーを搭載しており、他に類を見ない冷却性能を誇ります。また、ディープラーニングやグラフィック処理に必要なGPUを最大3基まで搭載することができます。GPUの安定した動作のためにシステム用の電源とは別に、GPU専用の電源を搭載したデュアル電源仕様になっています。筐体には4つのローラーを備えており、重くなりがちなワークステーションの移動を手助けします。

NVIDIA RTX6000 Ada 2枚と大容量メモリ256GBを搭載したワークステーション



- CPU : Xeon W9-3475X (36コア/2.2GHz/tb4.8GHz)
- メモリ : 256GB(32GB×8) DDR5-4800
- ストレージ : 2TB M.2 NVMe-SSD
- GPU: [2GPU] NVIDIA RTX6000 Ada 48GB-GDDR6
- OS : Ubuntu 22.04 LTS
- マザーボード : Intel W790 チップセット
- 電源 [2PSU] 1000W (80 Plus Platinumm 認証)
- 3年間センドバック方式ハードウェア保証

Applied
WST-XW93475XS3Q2TTNVM

5,998,000円^{税込}

NVIDIA RTX6000 Ada 1枚と大容量メモリ256GBを搭載したワークステーション



- CPU : Xeon W7-3465X (28コア/2.5GHz/tb4.8GHz)
- メモリ : 256GB(32GB×8) DDR5-4800
- ストレージ : 2TB M.2 NVMe-SSD
- GPU:NVIDIA RTX6000 Ada 48GB-GDDR6
- OS : Ubuntu 22.04 LTS
- マザーボード : Intel W790 チップセット
- 電源 [2PSU] 1000W (80 Plus Platinumm 認証)
- 3年間センドバック方式ハードウェア保証

Applied
WST-XW73465XS3Q2TTNVM

4,498,000円^{税込}