Speed up your AI
FIXSTARS

Performance Engineering Platform

Fixstars AI Booster

活用事例集 1

無料でお試し！

https://www.fixstars.com/ja/ai/booster

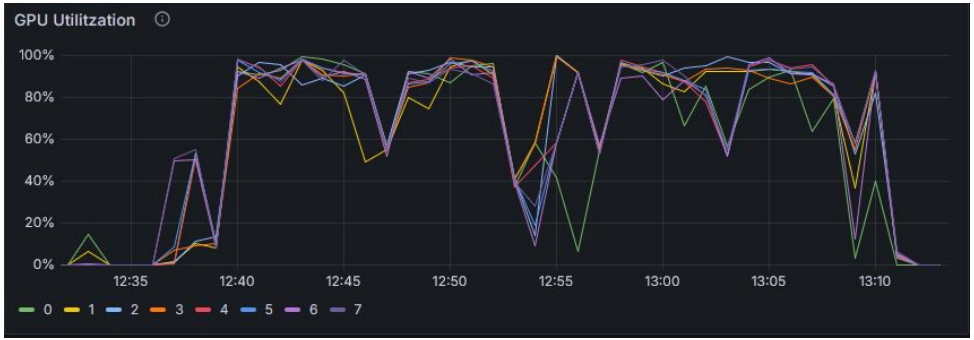
活用例1：AI 学習プロセスのワークロードを常時可視化して異常検知

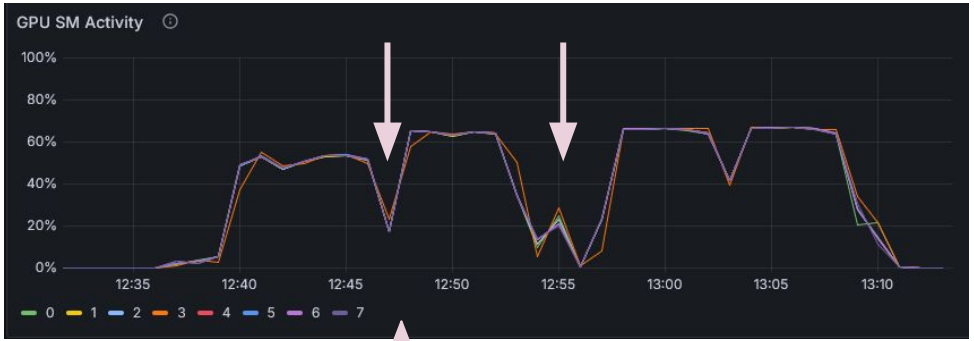
AI ジョブ実行中における、GPU利用状況の時系列変化を定常的に確認することで、パフォーマンスの低下に気づく

→計測なしに改善はできない

類似のGPU利用状況可視化ツールの問題点

ツール名	問題点
MLFlow	・イテレーション／エポック単位での取得(監視頻度が荒い)
Nsight Systems	・ユーザにsudo権限が必要 ・データ取得を手動で実行する必要あり
PyTorch Profiler	・オーバヘッドが大きい ・取得結果の解析が必要





・エポック間のSM Activity低下が無視できない
・valのSM Activity低下が著しい

Copyright© Fixstars Group

活用例2：Flame Graph を活用してソフトウェアボトルネックを特定

- ダッシュボードの”Flame Graph” 可視化機能で、プロセスをライブラリ・関数レベルでブレイクダウン
- ジョブ全体のボトルネックを特定し、その後、細かな単位でプロファイルを取得することで改善に繋げる

Flame Graphとは?
ソフトウェアの処理にかかった時間を視覚化したもので、どの関数や処理がどれだけの時間を消費しているかを直感的に表します。
縦軸は処理の階層、横軸は実行時間の割合を示し、長いバーほど時間を多く使っていることを意味します。
性能改善やボトルネックの特定に役立ちます。

Checkpoint 処理 (AI学習の途中経過を保存することで、学習が未完で終わった際にその地点から再処理するための処置)が**測定期間の12%**を占めている。
→状況にもよるが、削除したり、頻度を低くすることを検討する

Copyright© Fixstars Group

AI Booster 実践チュートリアル:

Llama4 Scoutの継続事前学習の処理速度を上げよう

Step 1. ダッシュボードを確認して現状の問題点を把握

GPU Utilizationは
一見してずっと高い

GPUコアの利用効率は
20%台と低い

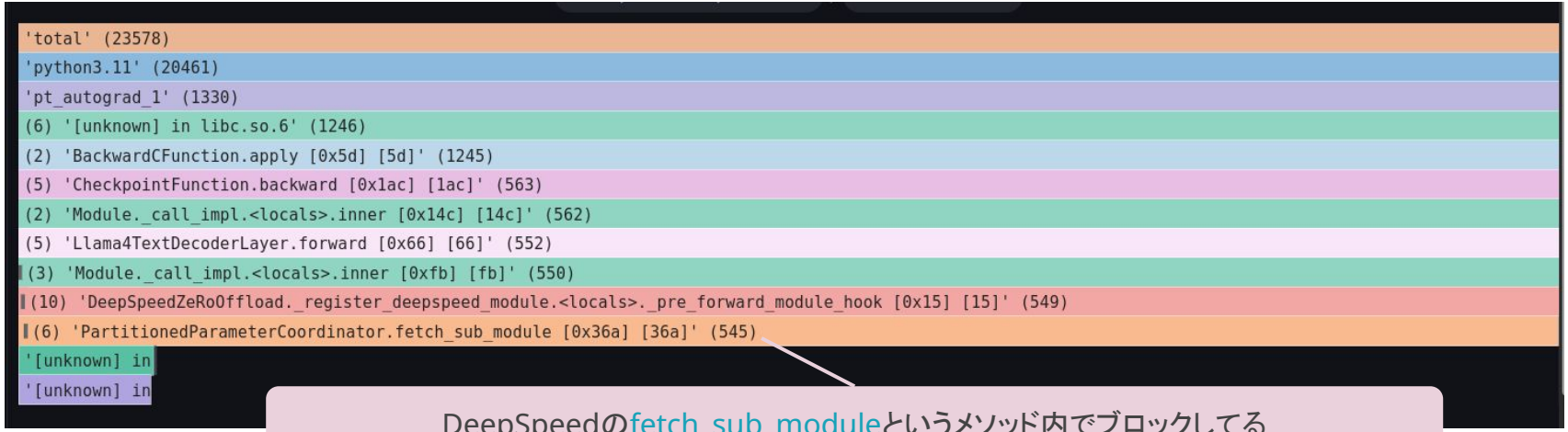
一定の間隔で
GPUがストール

インターコネクトの使用
帯域はSend/Recvとも
に平均6GB/s程度
(理論帯域は25GB/s)

AI Booster 実践チュートリアル:

Llama4 Scoutの継続事前学習の処理速度を上げよう

Step 2. Flame Graph を確認し、ボトルネックを特定



DeepSpeedのfetch_sub_moduleというメソッド内でブロックしてる

該当メソッドのコメント

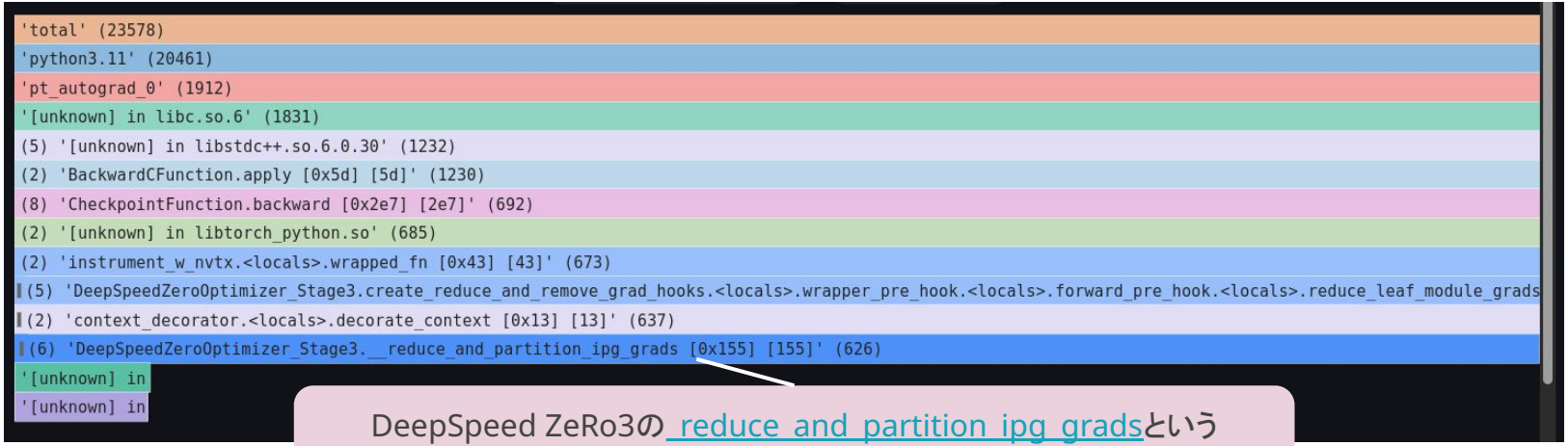
```
This method does the following (in order):
1. kick off fetch for parameters in immediately required sub module
2. kick off fetch for next few parameters we will need later (prefetch)
3. block on parameters in immediately required sub module
```

各ノードに分散配置されている
パラメータの到着を待っているのでは？

AI Booster 実践チュートリアル:

Llama4 Scoutの継続事前学習の処理速度を上げよう

Step 2. Flame Graph を確認し、ボトルネックを特定（他の箇所）



DeepSpeed ZeRo3の_reduce_and_partition_ipg_gradsというメソッド内でブロックしてる

該当メソッドから抜粋

```
while self.param_reduce_events and self.param_reduce_events[0].query():
    self.param_reduce_events.popleft()
if len(self.param_reduce_events) > self.max_param_reduce_events:
    self.param_reduce_events.popleft().synchronize()
```

勾配集約の完了イベントを待っているのでは？

AI Booster 実践チュートリアル:

Llama4 Scoutの継続事前学習の処理速度を上げよう

Step 3. 得られた情報の分析

- GPUのフルパワーを出し切れていない
- チェックポイント書き出しと思われるGPUストールがあるが、
  全体からすると影響はごくわずか
- フレームグラフとコードの情報を総合すると、
  各ノードに分散したデータの集約でブロックしている雰囲気がある
- 一方でインターコネクトの帯域には余裕あり

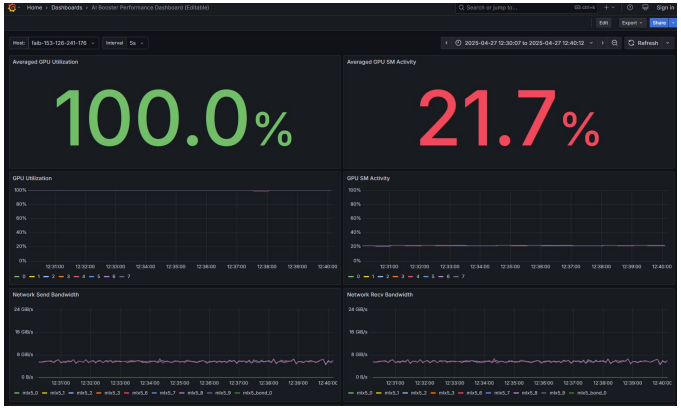
DeepSpeed ZeROに非効率性が存在するのでは？

AI Booster 実践チュートリアル:

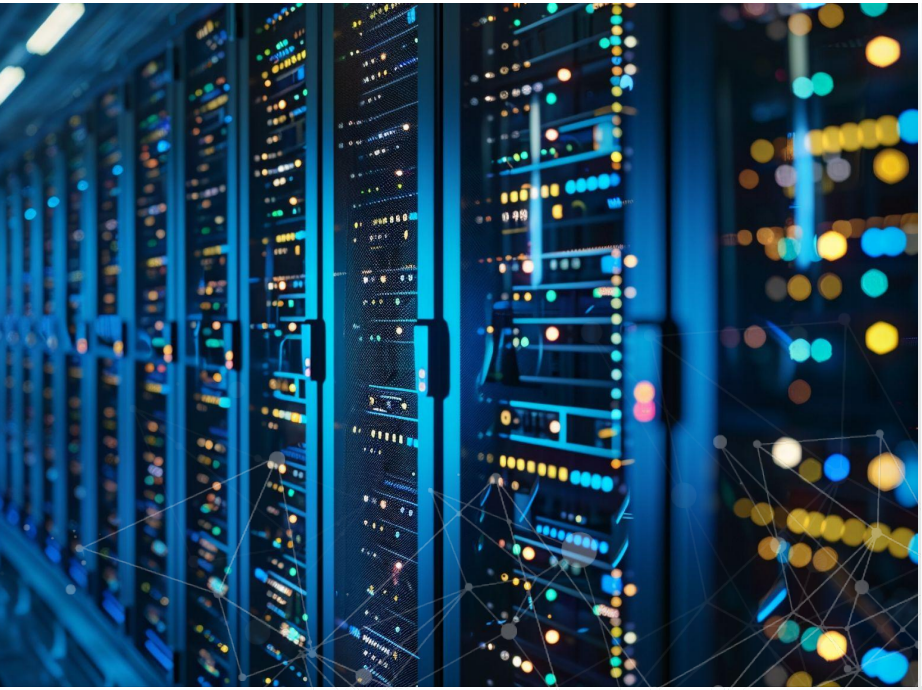
Llama4 Scoutの継続事前学習の処理速度を上げよう

Step 4. 処理の最適化

- 分析結果から、DeepSpeed ZeROの改善に取り組む
- AI Booster に同梱されているツールの一つ、DeepSpeed ハイパラチューナーの全40回の試行で得られたパラメータを使用



- 処理時間が**50[s/itr]から40[s/itr]**に、**20%**削減
- 通信帯域が**6.0→7.5[GiB/s]**に、GPUのプロセッサ効率も**22→28%**に改善



お問い合わせ

Fixstars AI Boosterの詳細については
下記までお問い合わせください。

株式会社フィックスターズ

Email support-ai@fixstars.com

Web https://www.fixstars.com/ja/ai/booster

Speed up your AI FIXSTARS

Copyright© Fixstars Group