

GTC2017

聴講に .

VOLTA アーキテクチャでスレッド毎に PC が付くっていうことの意味がよくわかってなかった , ということがよくわかった .

同じワーブの別スレッドが同じサイクルに別のインストラクションを実行できる , わけではなく ,

分岐したパス間での同期がちゃんと取れるようになった .

これで , Starvation Free アルゴリズムがかけるようになるとのこと .

GPU マシンで遊ぶ

組み立てた GPU マシンに TensorFlow などいれて遊んでみる .

知ってたけど速い !!

Tensorflow の手書き認識サンプルの LeNet では ,

CPU では , 95.8ms/iteration で , 合計 13 分 48 秒かかるのに対し ,

GPU を使うと , GPU:4.2ms/iteration で , 合計は 39 秒に .

RNN な seq2seq/translate.py でも ,

CPU の step-time が 0.52 ~ 0.54 秒なのに対して ,

GPU の step-time は 0.12 ~ 0.14 と 5 倍弱の高速化 .