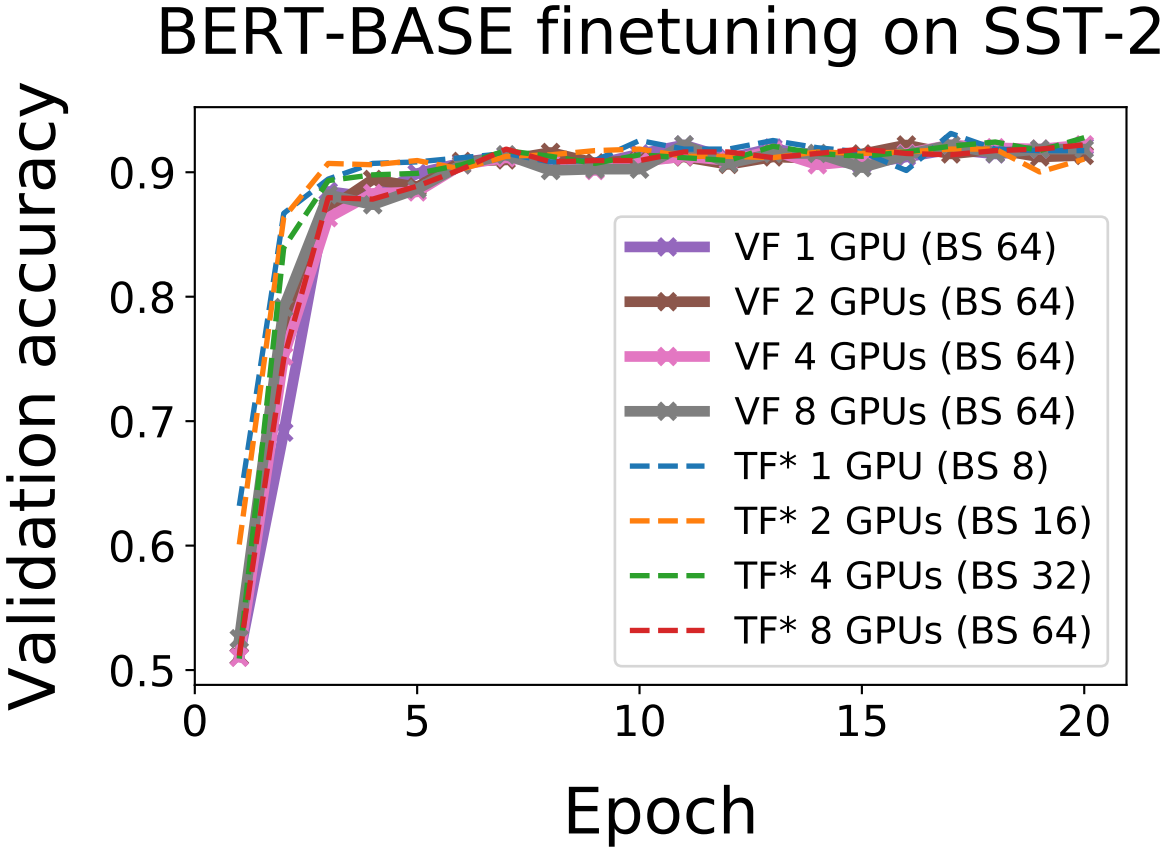
BERT-BASE finetuning on SST-2
Validation accuracy
0.9
0.8
0.7
0.6
0.5
0
5
10
15
20
Epoch
VF 1 GPU (BS 64)
VF 2 GPUs (BS 64)
VF 4 GPUs (BS 64)
VF 8 GPUs (BS 64)
TF* 1 GPU (BS 8)
TF* 2 GPUs (BS 16)
TF* 4 GPUs (BS 32)
TF* 8 GPUs (BS 64)