

Step Time vs Number of Finetuning Forward Tokens
Max Tokens per Batch: 256

