

BERT-TORCH-FP32 REPORT (STC)

Graph Compilation Results		
Compile Duration	Compile Precision	Subgraph Coverage
463.07309	fp16	99.76

Performance Results			
BS	AVG_Latency	P99_Latency	QPS
2	19.81	21.69	807.89

Accuracy Results (OPEN_SQUAD)			
Data Percent	100	Mean Diff	0.02692
Exact Match	78.99742		
F1 Score	86.14194		

