

Scaling Speedup of BitNET INT8xINT1 on A100

