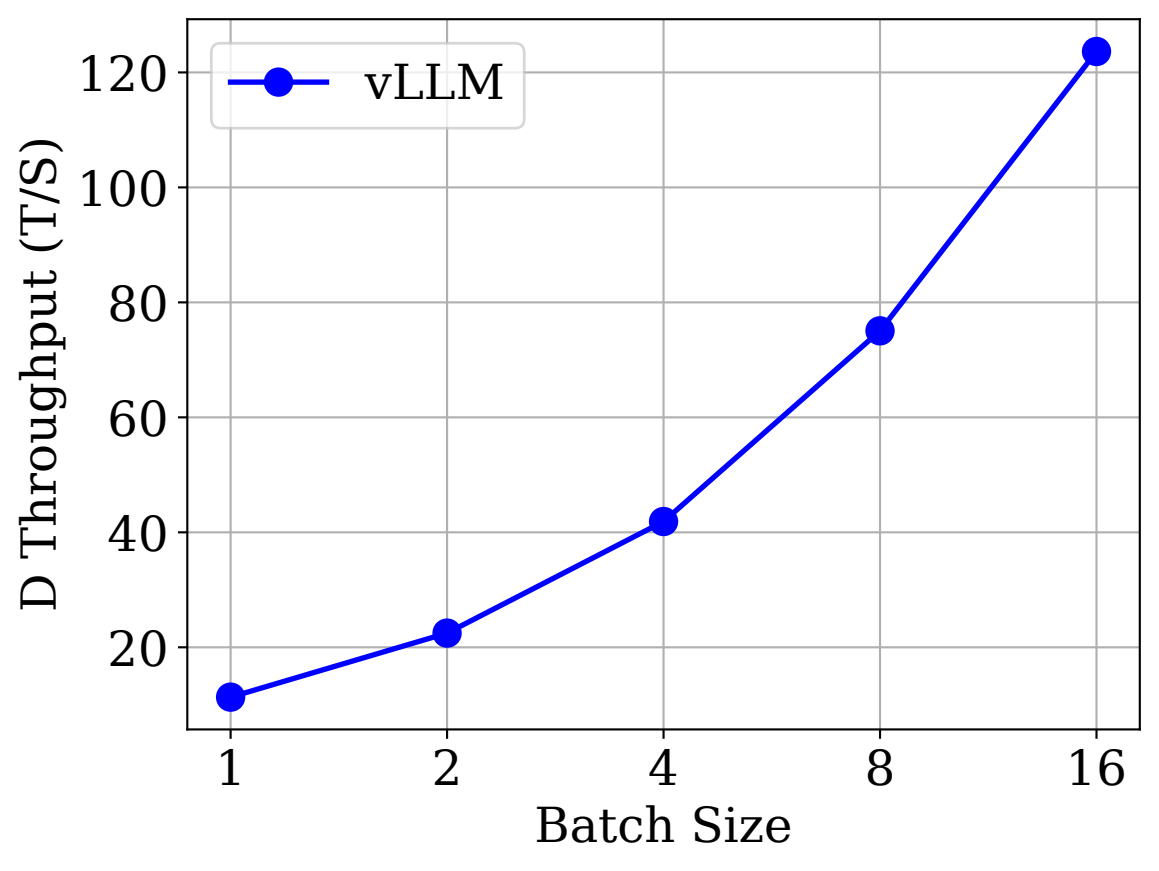
vLLM
D Throughput (T/S)
120
100
80
60
40
20
1
2
4
8
16
Batch Size