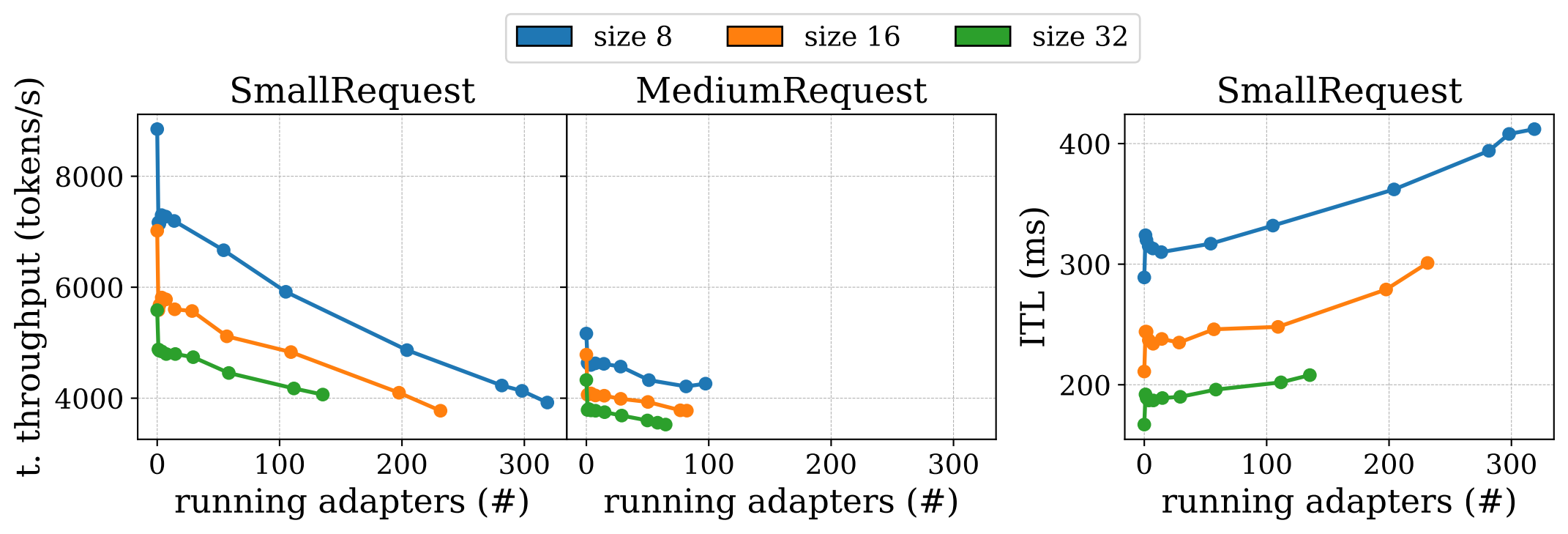
size 8
size 16
size 32
SmallRequest
MediumRequest
SmallRequest
t. throughput (tokens/s)
8000
6000
4000
0
100
200
300
running adapters (#)
0
100
200
300
running adapters (#)
ITL (ms)
400
300
200
0
100
200
300
running adapters (#)