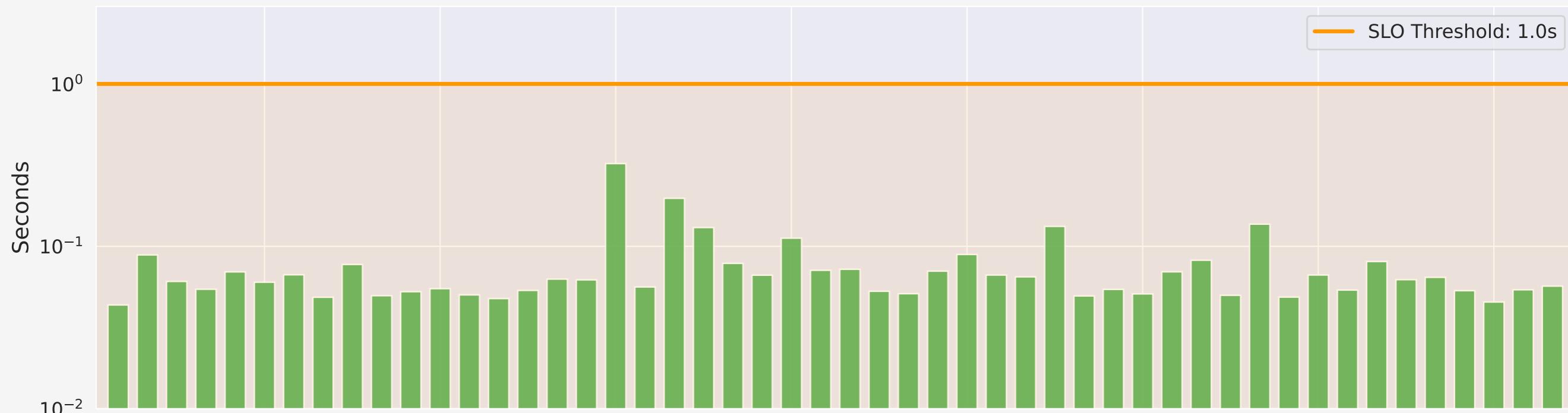
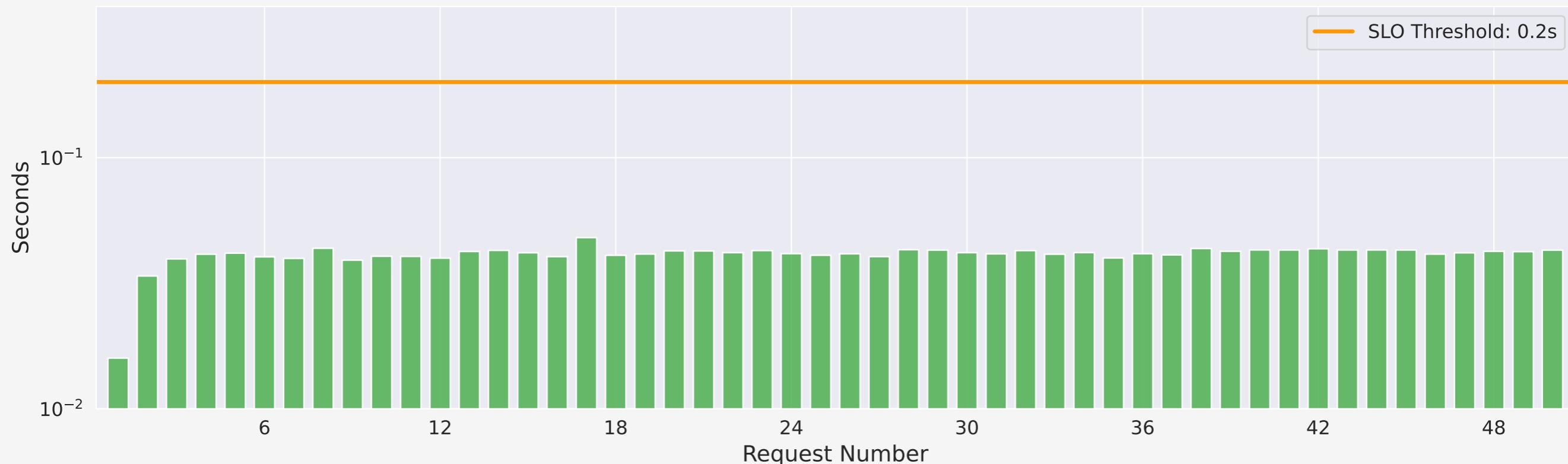


Chatbot Performance Metrics

Time to First Token (TTFT) (Logarithmic Scale)



Time Per Output Token (TPOT) (Logarithmic Scale)



Summary Statistics:

TTFT - Avg: 0.07s, Max: 0.32s, SLO Compliance: 100.0%
TPOT - Avg: 0.041s, Max: 0.048s, SLO Compliance: 100.0%