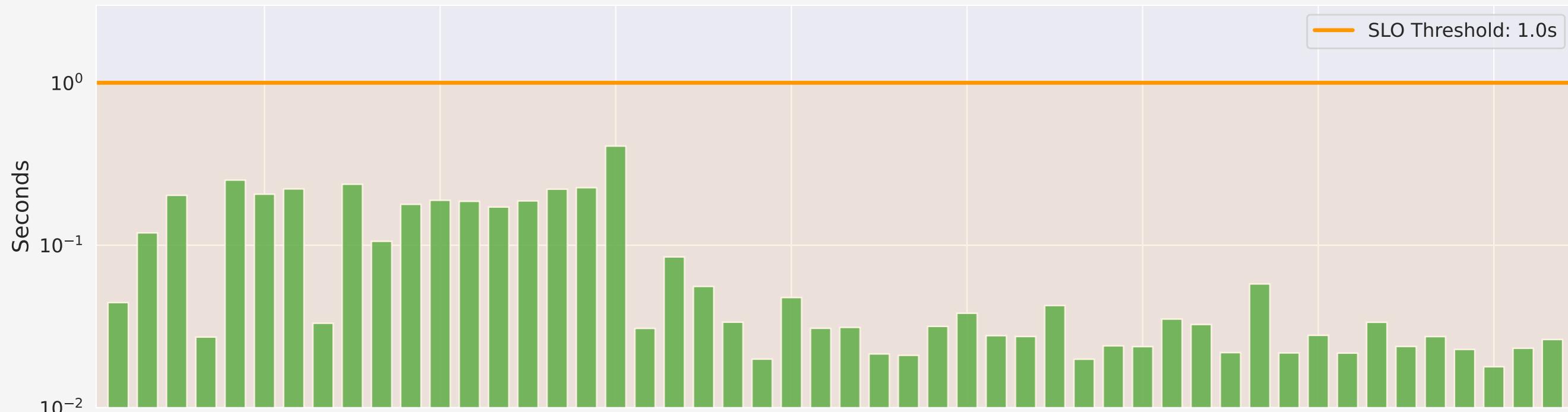
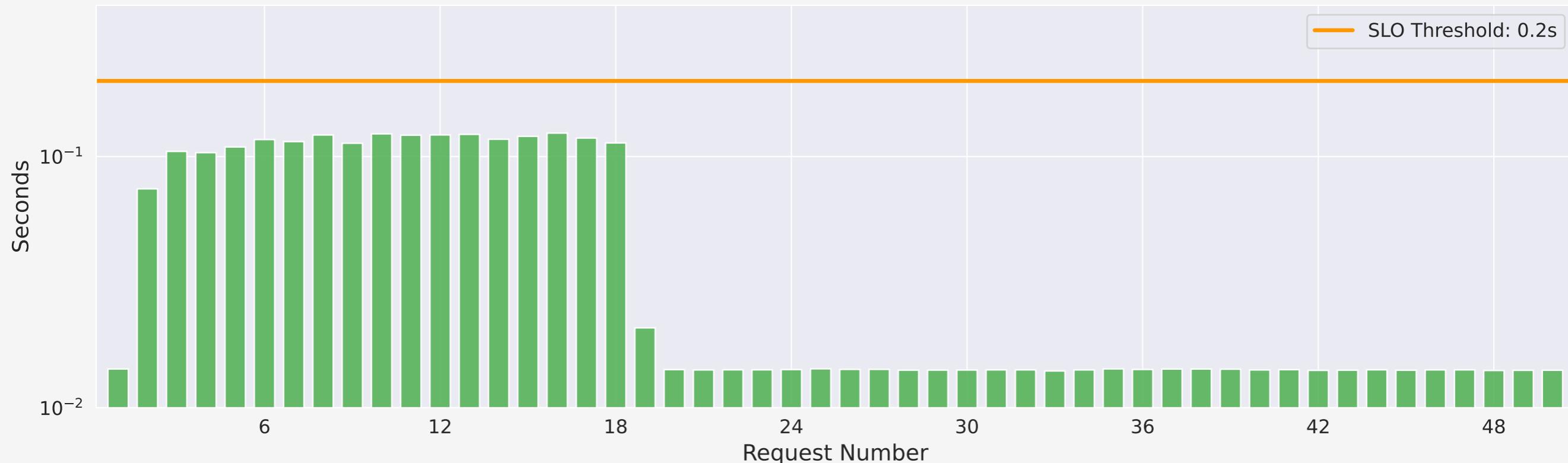


Chatbot Performance Metrics

Time to First Token (TTFT) (Logarithmic Scale)



Time Per Output Token (TPOT) (Logarithmic Scale)



Summary Statistics:

TTFT - Avg: 0.08s, Max: 0.41s, SLO Compliance: 100.0%

TPOT - Avg: 0.048s, Max: 0.124s, SLO Compliance: 100.0%