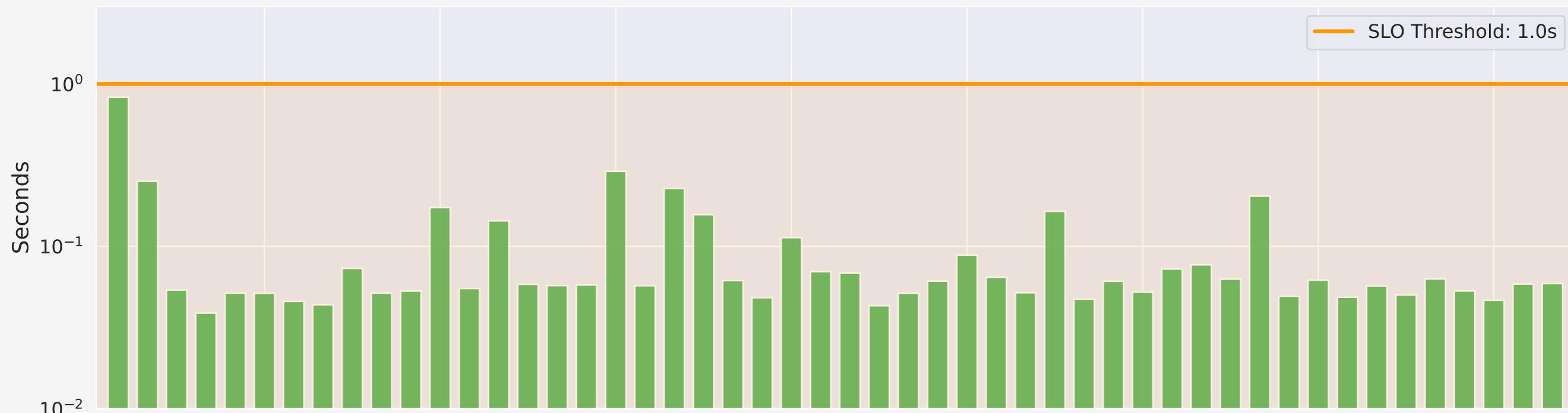
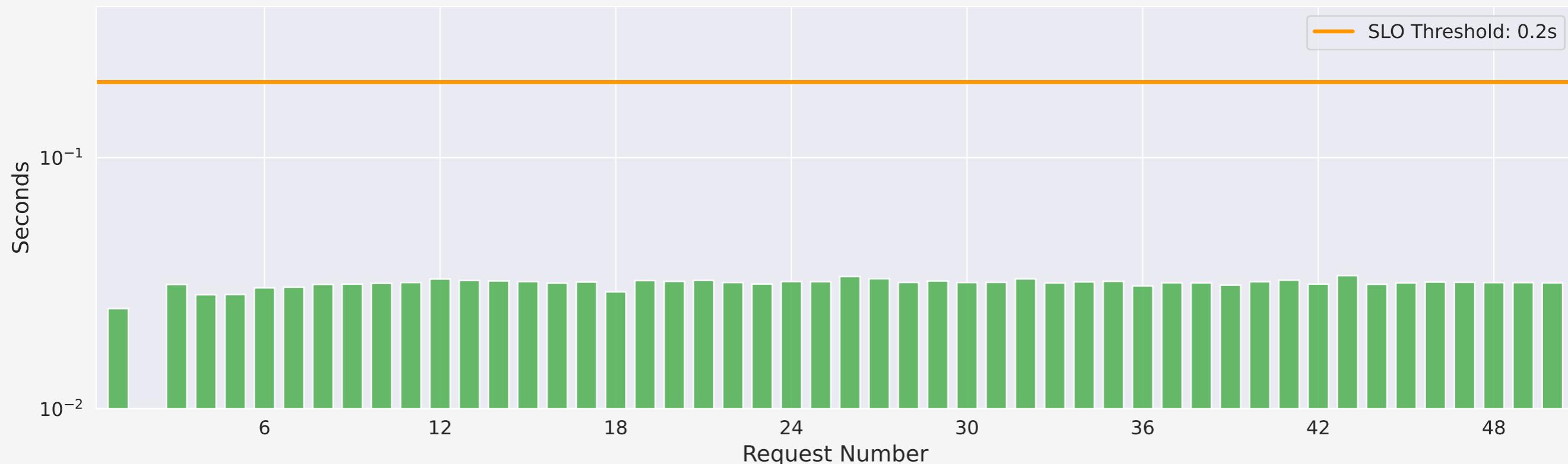


Chatbot Performance Metrics

Time to First Token (TTFT) (Logarithmic Scale)



Time Per Output Token (TPOT) (Logarithmic Scale)



Summary Statistics:

TTFT - Avg: 0.10s, Max: 0.83s, SLO Compliance: 100.0%

TPOT - Avg: 0.031s, Max: 0.034s, SLO Compliance: 100.0%