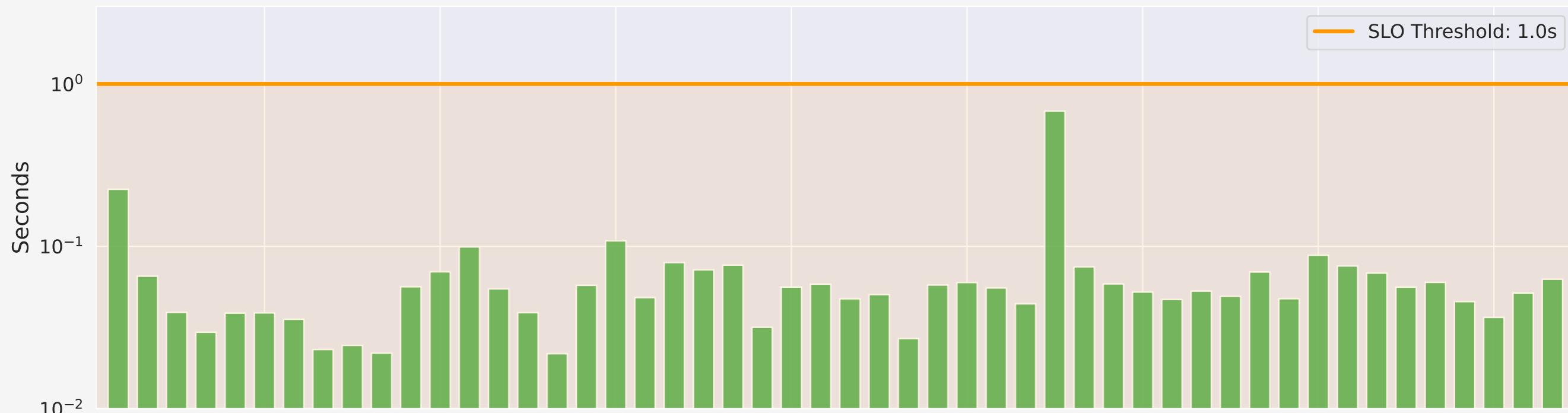
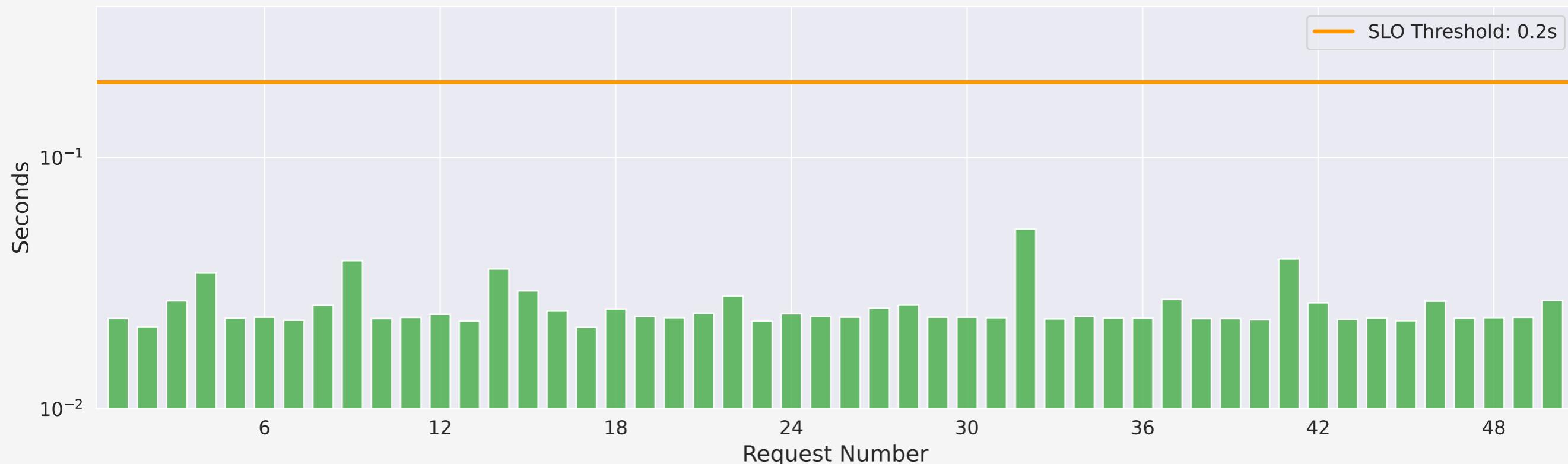


Chatbot Performance Metrics

Time to First Token (TTFT) (Logarithmic Scale)



Time Per Output Token (TPOT) (Logarithmic Scale)



Summary Statistics:

TTFT - Avg: 0.07s, Max: 0.68s, SLO Compliance: 100.0%

TPOT - Avg: 0.026s, Max: 0.052s, SLO Compliance: 100.0%