

Agent Normalized Latency (token)

chunk_prefill=512, prompt=300

