

Ollama Phi Model

Load Test Report

Generated: 2025-12-09 07:09:53

Executive Summary

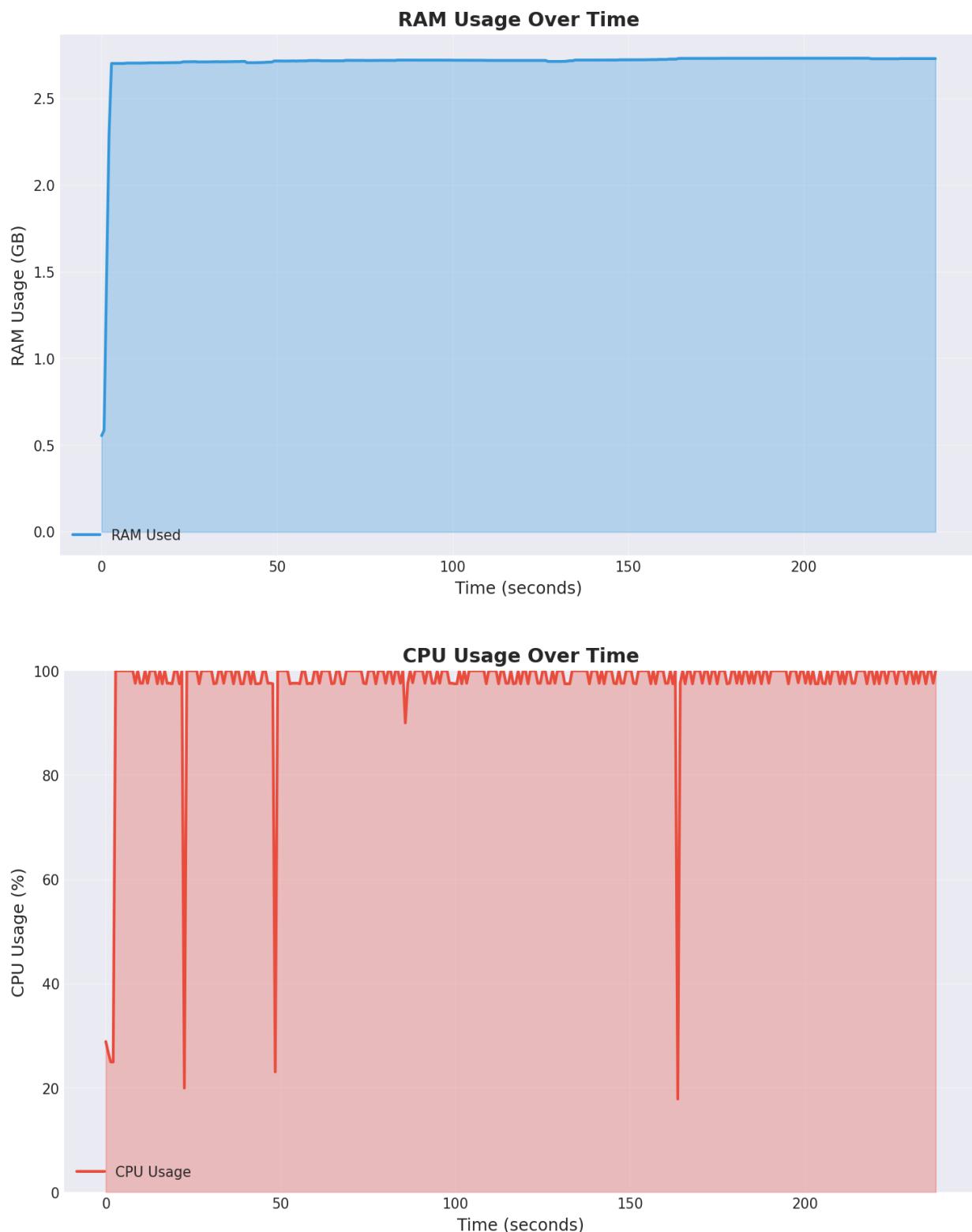
■■ CRITICAL: CPU usage exceeded 90%

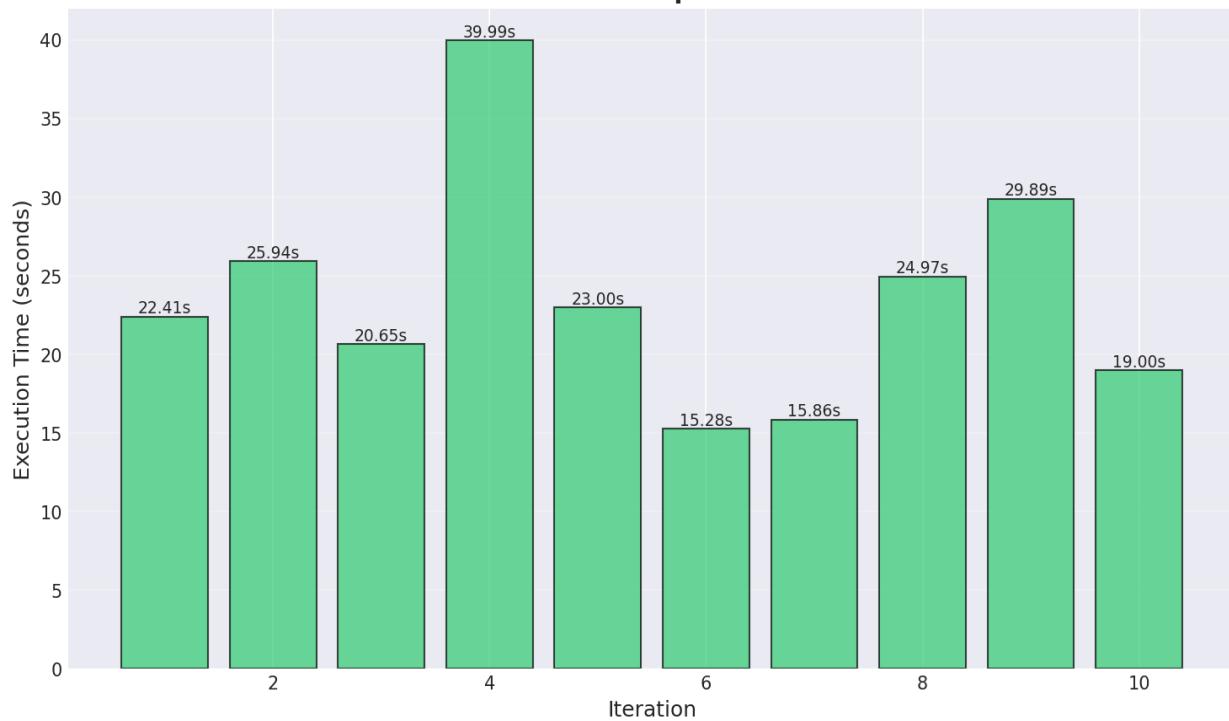
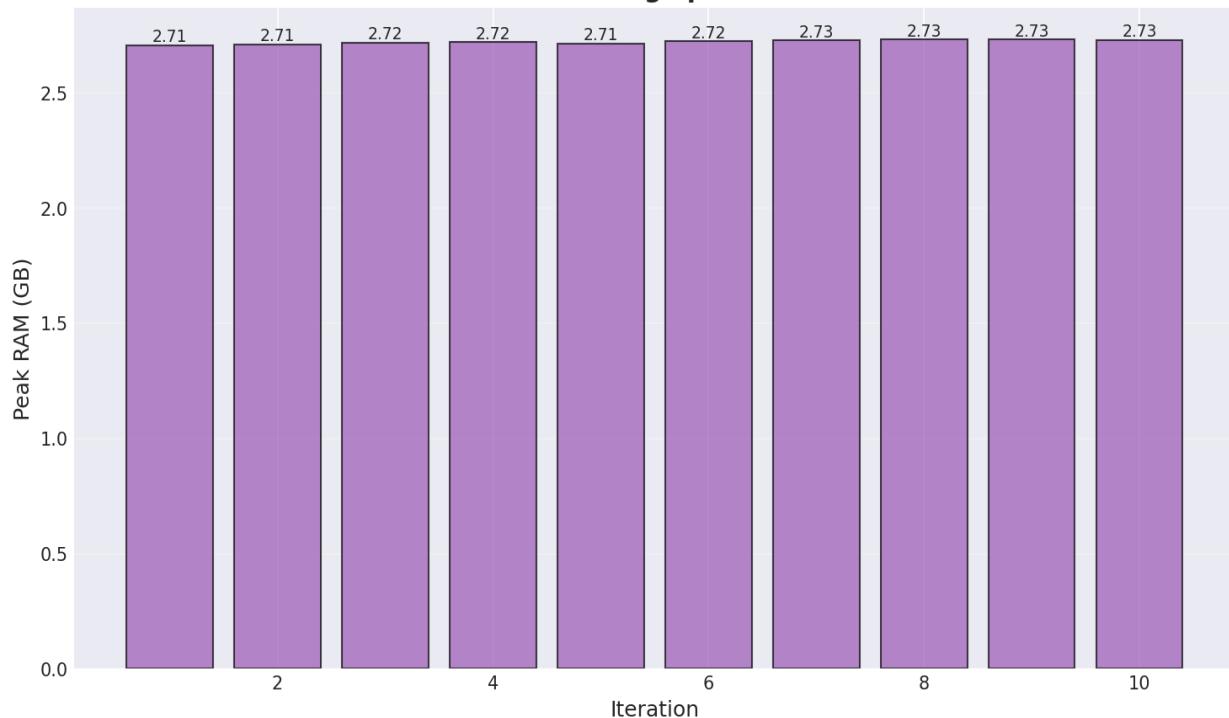
Metric	Value
Total Requests	10
Concurrent Requests	1
Model Name	phi
Success Rate	100.0%
Throughput	0.04 req/s
Avg Execution Time	23.70s
Min Execution Time	15.28s
Max Execution Time	39.99s
Avg RAM Usage	2.72 GB
Peak RAM Usage	2.73 GB (8.7%)
Avg CPU Usage	100.0%
Peak CPU Usage	100.0%

System Information

Component	Specification
CPU	x86_64
CPU Cores	4 cores / 4 threads
CPU Frequency	1575.4225000000001 MHz
Total RAM	31.29 GB
Platform	posix

Performance Metrics



Execution Time per Iteration**Peak RAM Usage per Iteration**

Iteration Details

Iteration	Time (s)	Peak RAM (GB)	Peak CPU (%)	Response Length
1	22.41	2.71	100.0	585
2	25.94	2.71	100.0	839
3	20.65	2.72	100.0	713
4	39.99	2.72	100.0	1265
5	23.00	2.71	100.0	717
6	15.28	2.72	100.0	525
7	15.86	2.73	100.0	619
8	24.97	2.73	100.0	829
9	29.89	2.73	100.0	1066
10	19.00	2.73	100.0	635