Cache Size vs. Throughput (tokens/sec)