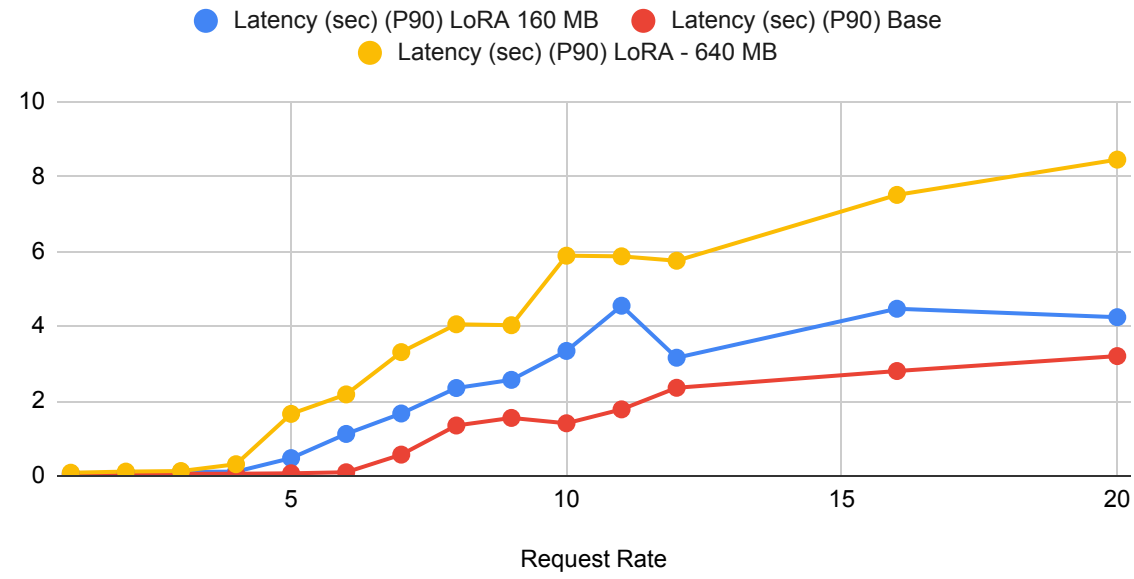


Latency (sec) (P90) LoRA 160 MB, Latency (sec) (P90) Base and Latency (sec) (P90) LoRA - 640 MB



Throughput LoRA (output tokens per sec), Throughput Base (output tokens per sec) and Throughput LoRA 640 (output to...

