

Latencies vs CPU request with Batch sizes

