4-GPU_ivf_quantizer Average Latency vs. Batch Size with dim=960, nb=100000 Average Latency (µs) Batch Size (nq)