4-GPU-FlatL2-repeated Latency per Iteration with dim=960, nb=100000, nq=1