

Inference latency for norm

