The benchmark of flops
dist-attn-fwd with varlen_full mask