The benchmark of flops
attn-bwd with varlen_causal mask