gradients/bert/encoder/layer_3/attention/self/key