Memory use per step over the first 4 iterations

Batch of 48 sequences of max length 40 - Tesla M60 (8Gb)