Rahul Vigneswaran K
CS23MTECH02002

4 4.5

2) Most significant Bits (MSBs) of weights and bias are not as fault tolerant as the LSB. If the calculation duration exceeds to ~~read~~ retention time because of partial SET, then an error is induced. Error in the LSB can be tolerated as they are by definition Least Significant Bits. The error because of faulty MSB will be accumulated and will lead to bad accuracy.

2

One way to get away by using partial SET on MSB is by proactively track and invoking the partial SET to each bit based on their end of retention time. This would cause unnecessary tracking & refresh overhead. ~~One thing I can think is, it is observed in~~

3) ~~The~~ → ~~addon~~ Addional overhead caused by tracking each layer proactively in the layer aware SET policy. The implementation details of the tuners, etc are not clear.

→ In the buffered marching based levelling policy, they mention that the write is done in a marching sequential fashion but the physical memory is not always allocated exactly in that fashion. The reference & address are arranged in different places. So this marching fashion is questionable.

2

→ The reproducibity of this work is poor. The benchmarking of this is done in an inhouse accelerator according to the authors. This makes it very difficult to reproduce or verify or extend their work.

1) The write bypassing may not be effective for several reasons :

→ The predictor model of when to bypass would cause additional overhead. This would nullify any extra boost we get by sparsity exploitation.

→ Irrespective of the write bypass, we would anyway have to access the data and write them in local low caches for operation and this transfer of data would reduce to time gain.

4

0.5