Commutativ

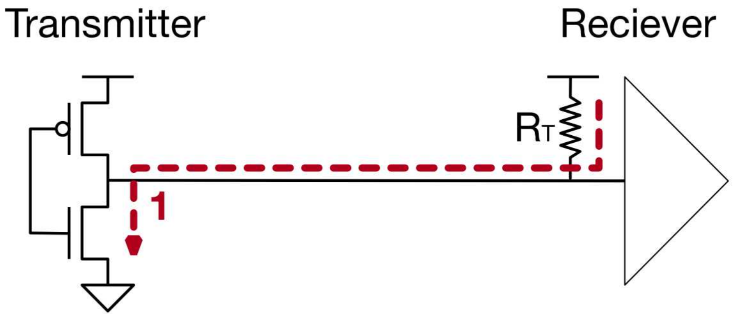
This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

Reduce Data Movement Energy on Sparse Inference Worklo

SAND2020-5473C

Ben Feinberg[*], Benjamin C. Heyman[†], Darya Mikhailenko[†], Ryan Wong[†], An C. Ho[†], Engin Ipek[†]

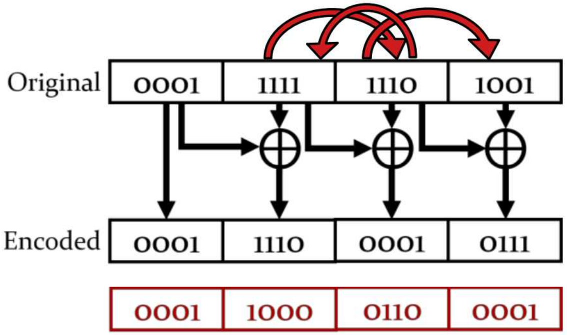
[*] Sandia National Laboratories, [†] University of Rochester



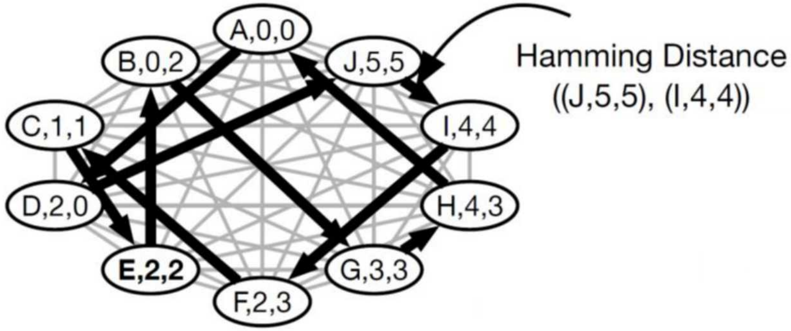
Reducing DRAM IO energy can be modeled as reducing transmitted 1s

Prior Work: Base+XOR Encoding

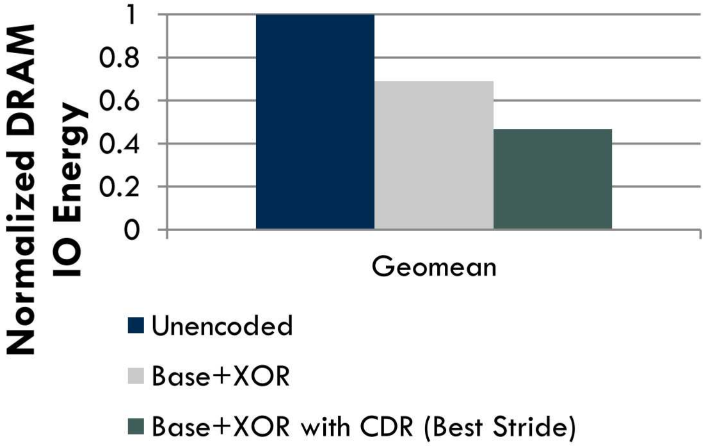
This Work



Key Idea: Reorder data to reduce the number of transmitted 1s



Model data reordering as an instance of the Traveling Salesman Problem



Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

22.4% reduction in DRAM IO energy over Base+XOR Encoding