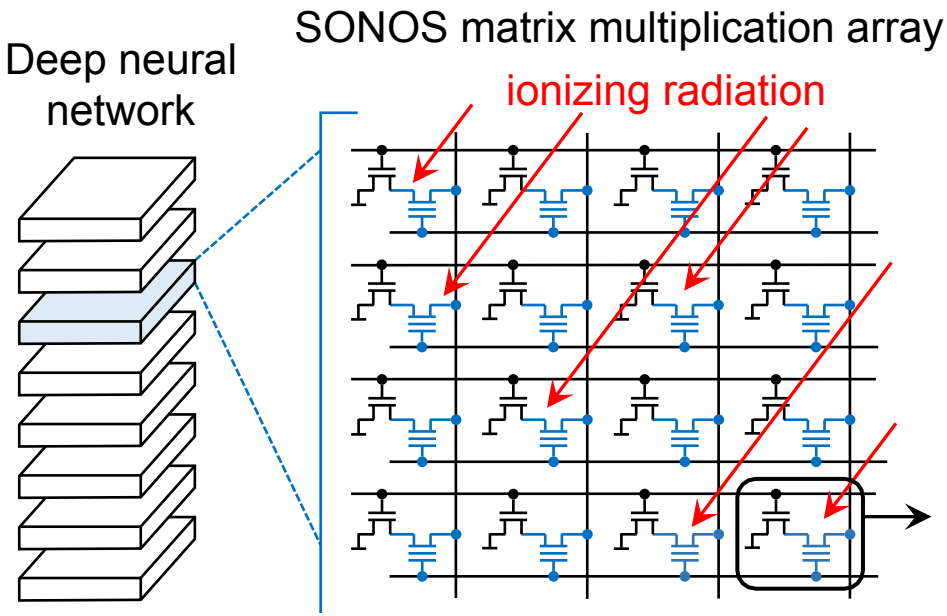
# Introduction

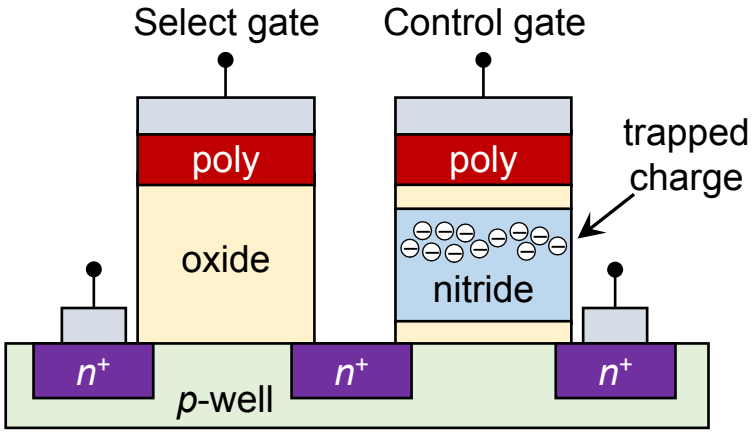
SAND2020-11757C

- Non-volatile memory arrays can be used for ***analog in situ matrix-vector multiplication***: faster & more energy efficient than digital processors
  - Key application: acceleration of deep neural network inference

- **SONOS** (silicon-oxide-nitride-oxide-silicon) memory is attractive because:
  - Programmable to many current levels [1]
  - In large-scale commercial production
  - Low read noise

***How sensitive are SONOS-based inference accelerators to total ionizing dose (TID)?***

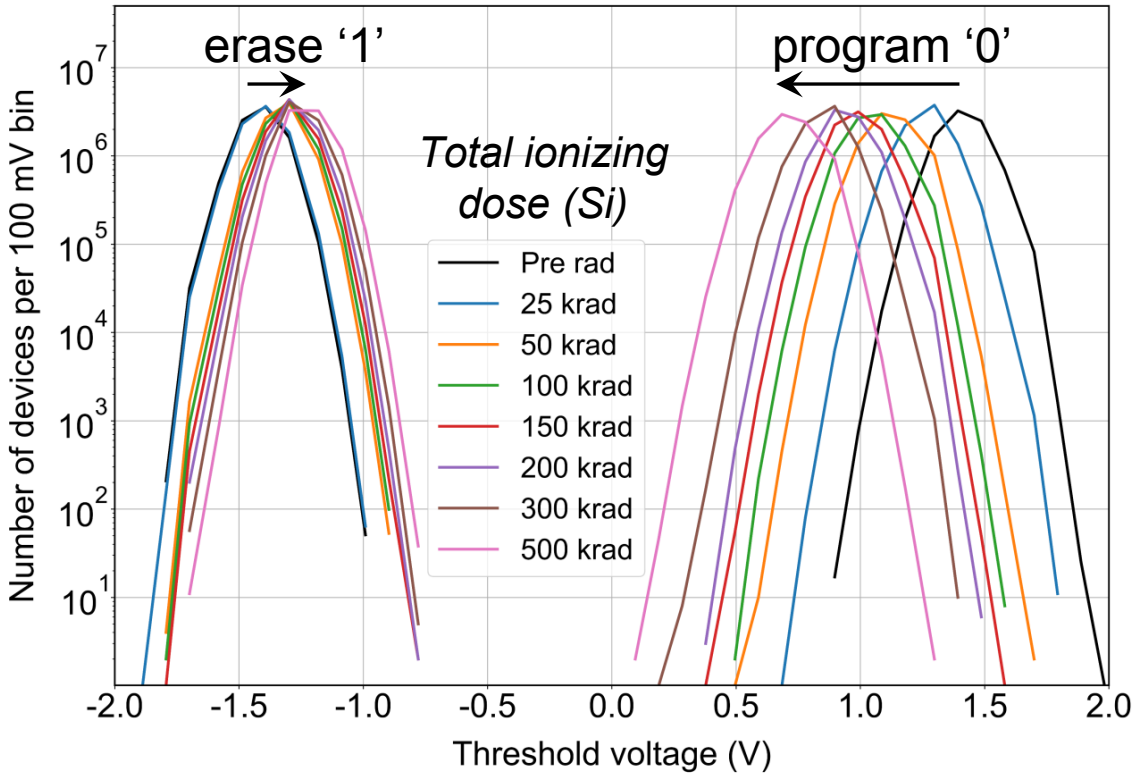



Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

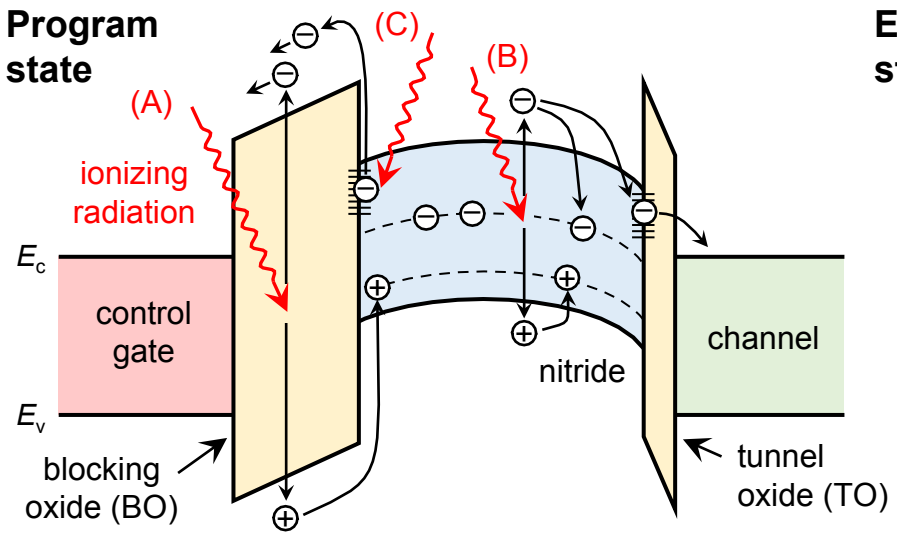
# TID experiments on digital SONOS memory

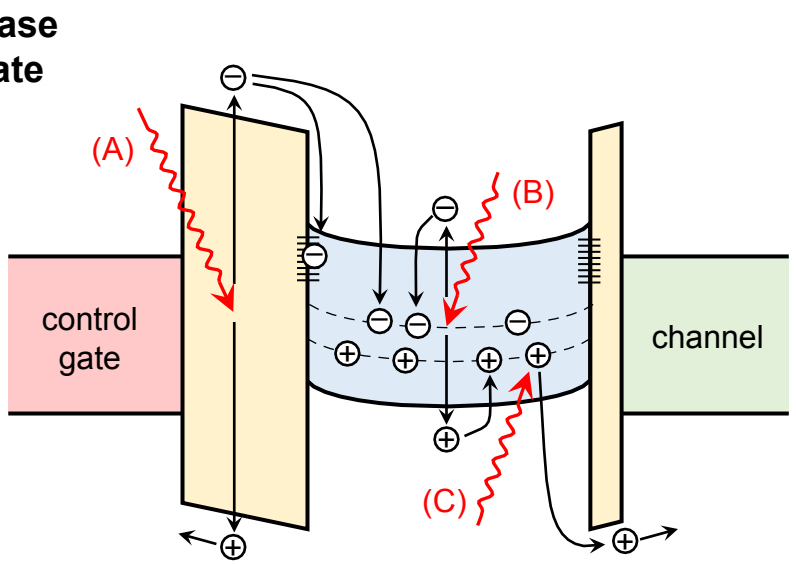
- **4 Mbit of 40nm SONOS memory** (2 Mbit ‘0’, 2 Mbit ‘1’) was exposed to ionizing radiation from a Co-60 source, while connected to DC power
- The threshold voltage $V_T$ distributions of all 4 Mbit was profiled following each step of a sequence of dose steps at 300K
- All bits function reliably as digital memory up to the maximum TID of 500 krad(Si)



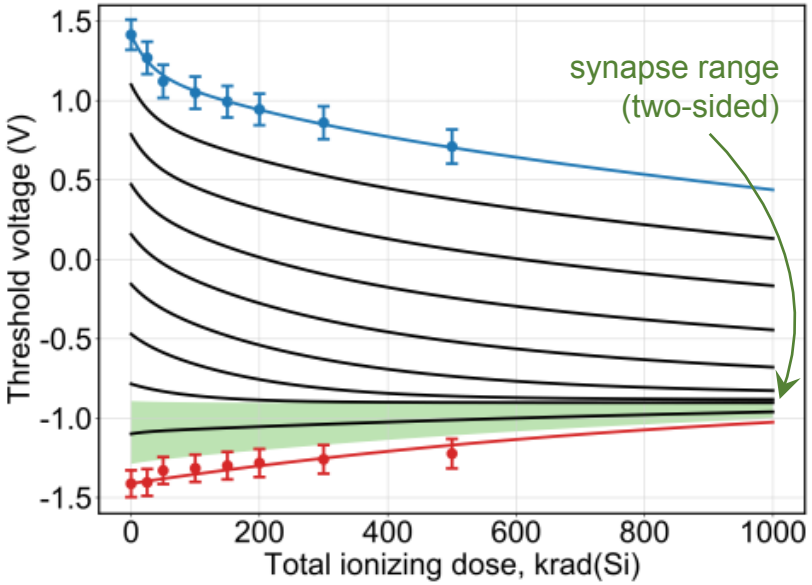
# Modeling carrier response to TID (I)

- We formulated a rate-equation based model to predict how intermediate SONOS device states (between program and erase) decay upon ionizing radiation exposure
- The following effects were considered:
  A. Charge generation in oxide and injection into nitride
  B. Charge generation in nitride, followed by re-trapping/escape
  C. Radiation-assisted emission of bulk and interface trapped charge out of the storage layer



## Modeling carrier response to TID (II)



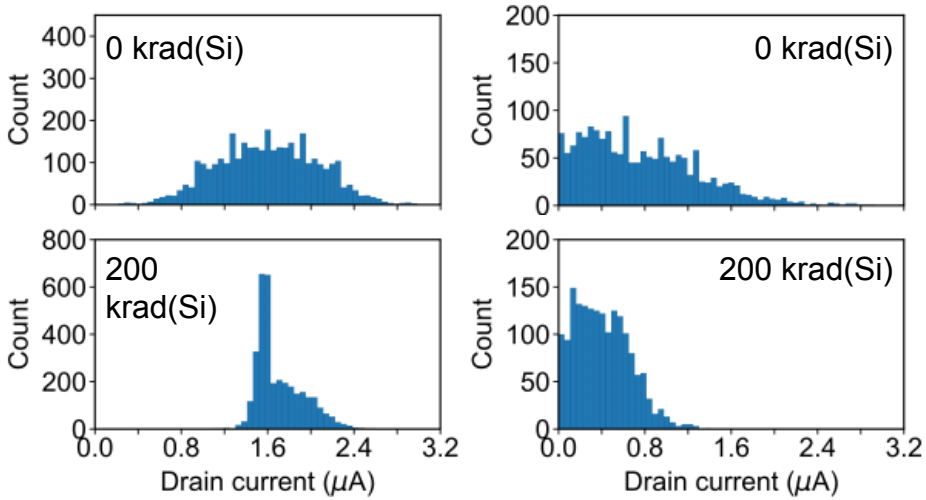

- The model was fit to the measured data on digital memory devices
- The neutral point $V_T$ = –0.91V is insensitive to TID. A band of $V_T$ states near the neutral point are used to represent the synaptic weights of a neural network
- $V_T$ values are converted to drain current $I_{DS}$ using measured I-V curves; devices operate in subthreshold regime

## Decay of SONOS synaptic weights

- TID model was integrated with **CrossSim**: a device-level *in situ* matrix multiplication simulator [2]. Radiation effects were considered in the SONOS devices only
- Positive/negative weights encoded as the difference of two SONOS currents using one of two bias schemes:

<u>Two-sided mapping</u>: neutral point set to 1.6 μA (midpoint)

<u>One-sided mapping</u>: neutral point set to 0.0 μA

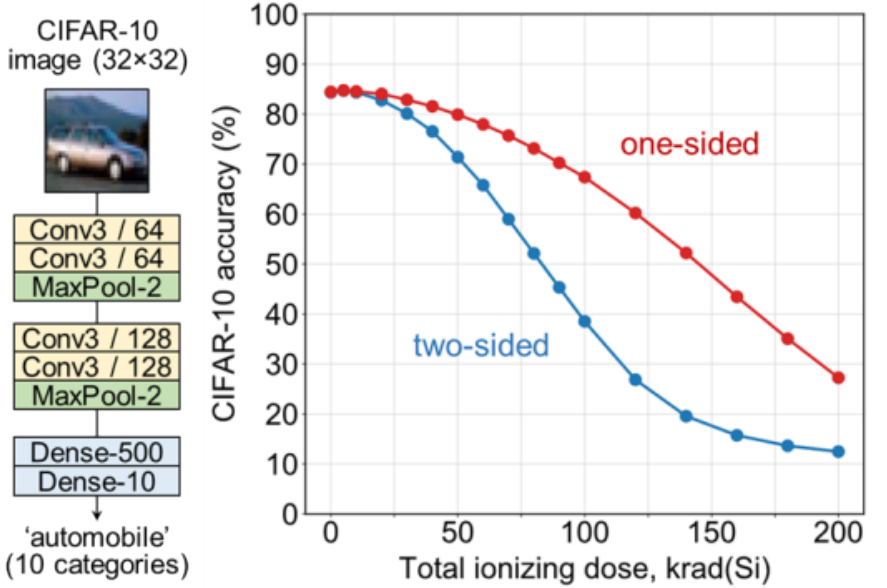



**Synaptic current distribution in layer 1 of CIFAR-10 CNN**

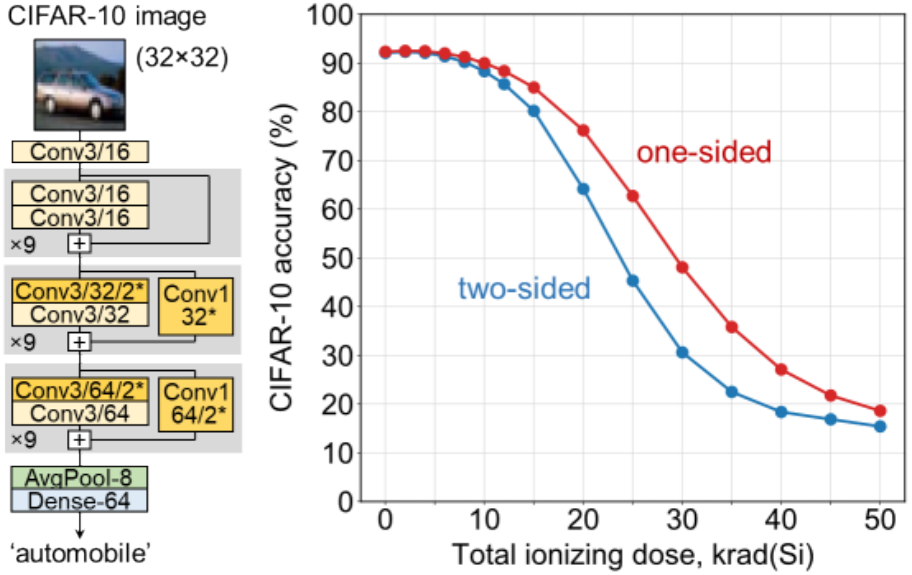
# TID response of SONOS inference accelerator: CIFAR-10

- The SONOS accelerator was evaluated on two convolutional neural networks (CNNs) for the CIFAR-10 image classification task
  - A six-layer plain CNN
  - ResNet-56v1: a state-of-the-art deep residual network with 56 layers [3]
- The one-sided bias scheme is more resilient than the two-sided bias scheme
- The deeper network is substantially more sensitive to TID effects, due to the multiplicative effect of TID-induced weight decay from layer to layer

**6-layer CNN for CIFAR-10**
4.36M weights, 100.4M ops

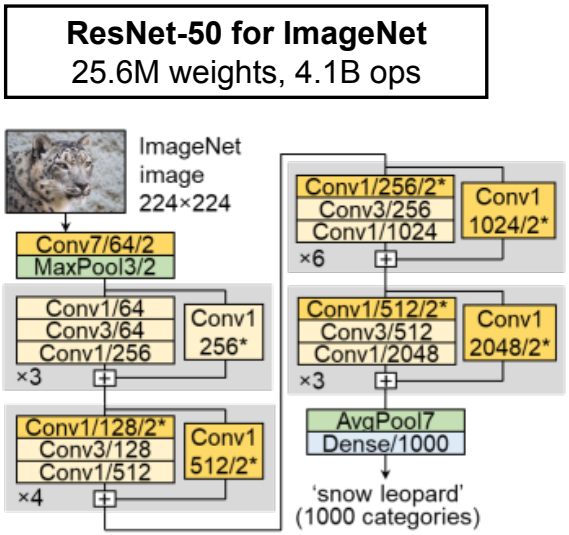


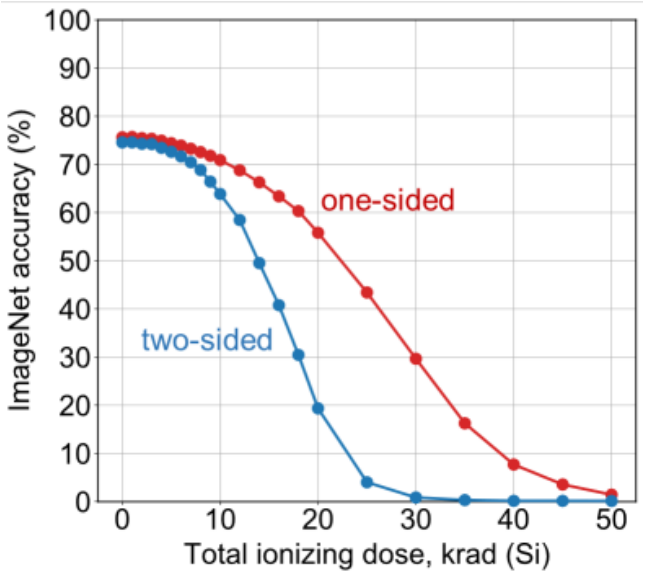

**ResNet-56v1 for CIFAR-10**
0.85M weights, 126.3M ops



# TID response of SONOS inference accelerator: ImageNet

- ResNet-50: a state-of-the-art deep CNN on the much more difficult ImageNet dataset [3]
- TID sensitivity is not substantially greater than the similarly deep ResNet-56v1, despite the large difference in dataset complexity.




## Conclusion

Considering the effects of radiation on SONOS devices alone, the evaluated neural networks have TID tolerances from 10 krad(Si) to 100 krad(Si), depending on the network topology, especially the depth; this level of resilience may be suitable for deployment at geosynchronous orbit.

## References


[1] V. Agrawal et al, *IEEE IMW,* 2020

[2] S. Agarwal et al, *Symp. VLSI,* 2017

[3] K. He et al, *IEEE CVPR*, 2016



U.S. DEPARTMENT OF ENERGY

Sandia National Laboratories

National Nuclear Security Administration

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.