SAND2016-11460D

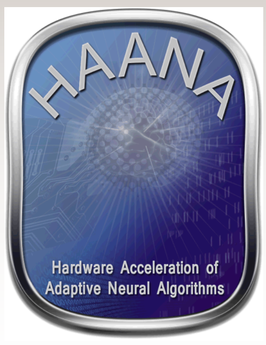
Exceptional service in the national interest

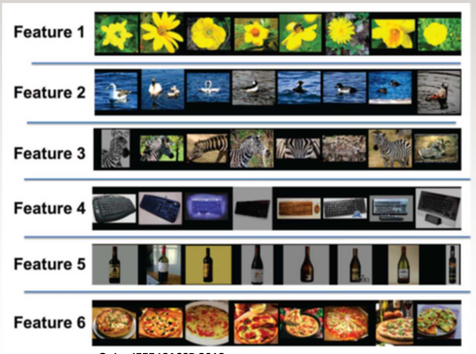
Sandia National Laboratories

Evaluating Resistive Memory Devices for Neuromorphic Computing Using Ultrashort Voltage Pulses

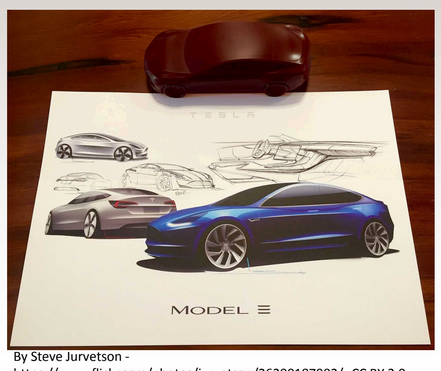
Jacobs-Gedrim, Robin[1]; Kotula, Paul; Goeke, Ronald; Mook, Bill; Finnegan, Patrick Sean; Smith, Carl Lee; Agarwal, Sapan; Gastian, Loren; Van Benthem, Mark; Jungjohann, Katie; Marinella, Matthew; James, Conrad D;

Hardware Based Deep Learning





Q. Le, IEEE ICASSP 2013



By Steve Jurvetson - https://www.flickr.com/photos/jurvetson/26200187902/, CC BY 2.0,

Cybersecurity – Image Recognition – Self Driving Vehicles

Deep Learning provides solutions but currently requires a supercomputer and MWs of power…
Hardware accelerator using analog devices may offer 10^5-10^6 times power reduction – on a single chip!

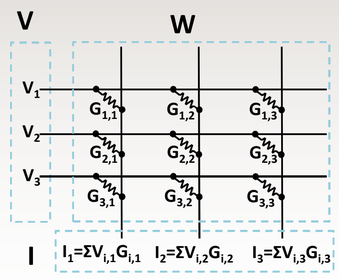
Matrix Vector Multiply

Mathematical

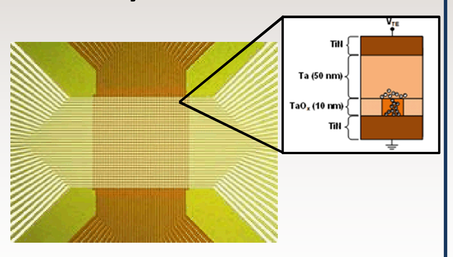
$$V^T W = I$$

$$\begin{bmatrix} V_1 & V_2 & V_3 \end{bmatrix} \begin{bmatrix} W_{1,1} & W_{1,2} & W_{1,3} \\ W_{2,1} & W_{2,2} & W_{2,3} \\ W_{3,1} & W_{3,2} & W_{3,3} \end{bmatrix} = \begin{bmatrix} I_1=\Sigma V_{i,1}W_{i,1} & I_2=\Sigma V_{i,2}W_{i,2} & I_3=\Sigma V_{i,3}W_{i,3} \end{bmatrix}$$

Electrical



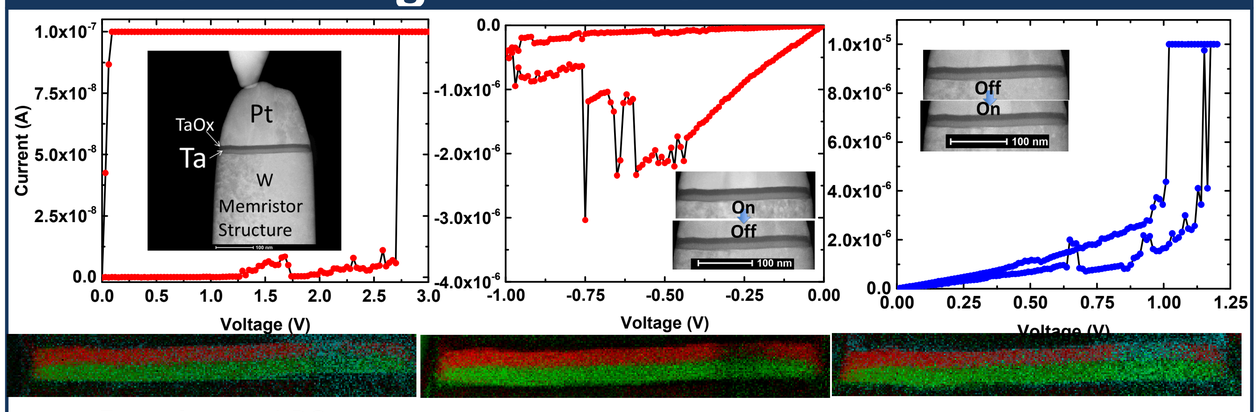
Physical



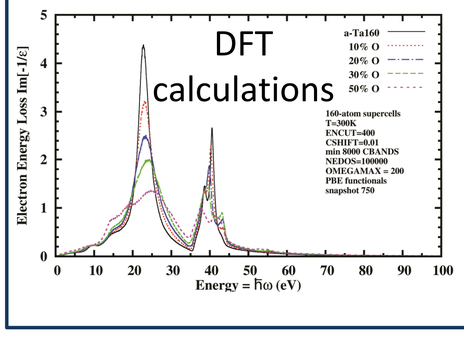
- Matrix vector multiply is most computationally expensive process in deep learning
- Kirchhoff's laws allow for matrix vector multiply on crossbar of resistors
- Can implement programmable array of resistors using analog metal oxide resistive switching devices

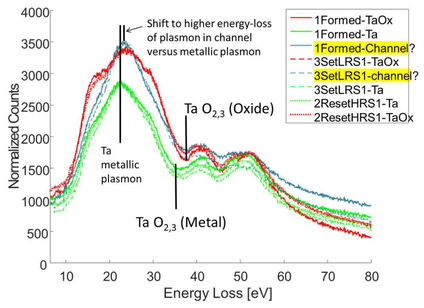
"Let physics do the computation"

Switching Mechanism: *In-Situ* TEM



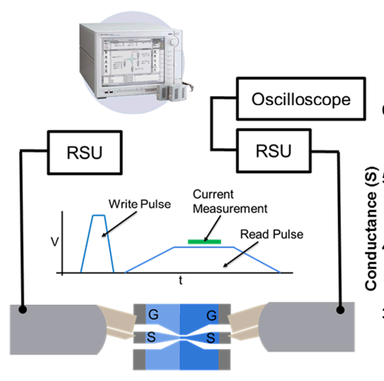
- **Forming creates 50 nm filament (cyan).**
- **RESET changes the TaOx in the filament less metallic state (no cyan)**
- **SET returns the filament to conductive state (cyan)**



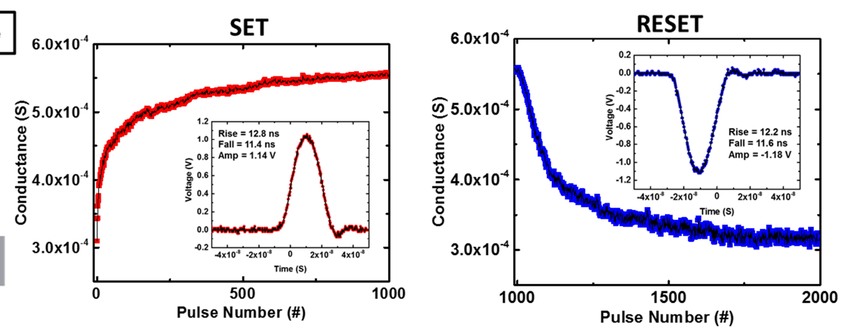


In filamentary region, Ta O2,3 edge is consistent with oxidized Ta but there is a sharp metallic plasmon feature shifted to slightly higher energy loss consistent with DFT calculations for more metallic oxides

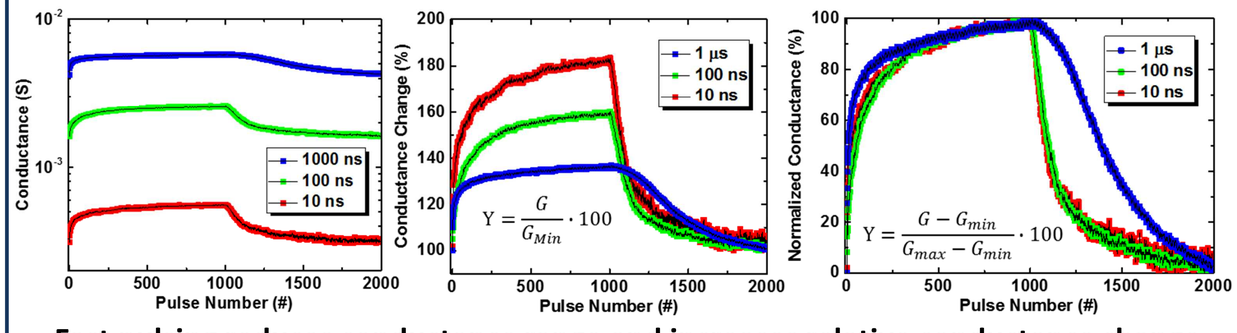
Ultrashort Voltage Pulses: Analog Operation



Fast pulsing allows analog weights to be written

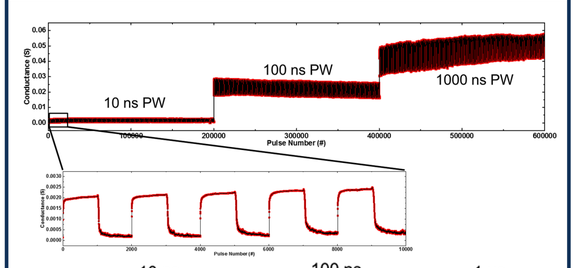


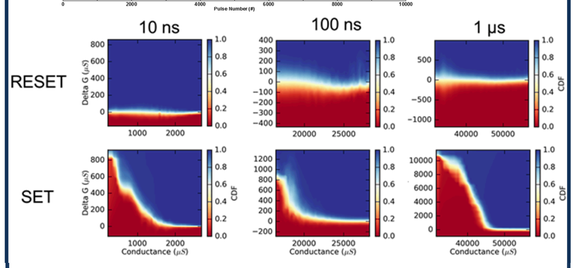
Effect of Pulse Width and Edge Time



Fast pulsing reduces conductance range and increases relative conductance change

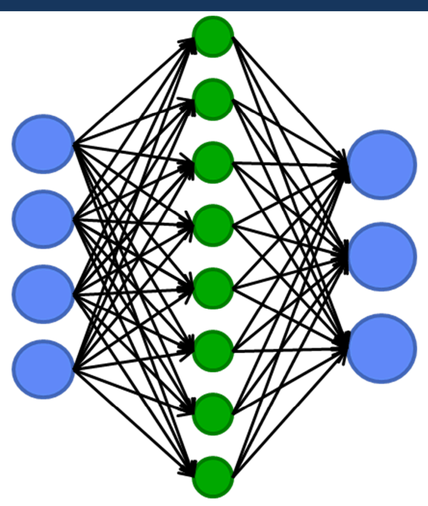
$CDF\ of\ \Delta G/G$





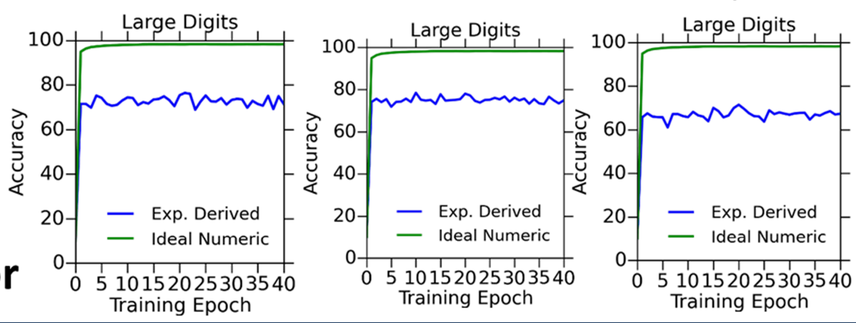
Stable cycling at short pulse lengths

MNIST Data Set Training Results on CrossSim



Backpropagation of Error

Data set	# Training Examples	# Test Examples	Network Size
MNIST Large Digits	60,000	10,000	784×300×10



- **Conductance switch in these TaOx devices is based on a metallic filamentary conduction mechanism**
- **Relative conductance change increased with shorter Pulse Width / Edge Time**
- **Ultrafast pulses may be employed to lower conductance switching range without reducing overall conductance change**
  - **Ultrafast pulses down to 100s of ps are possible to implement in-silico!**
- **Major progress has been made toward "fab friendly" devices suitable for image recognition, but still a long way to go before achieving numerical equivalent**

U.S. DEPARTMENT OF ENERGY
National Nuclear Security Administration

This work was supported by Sandia National Laboratories' Laboratory Directed Research and Development (LDRD) Program. Sandia National Laboratories is a multi-mission laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

[1]Contact Information –*rbjaco@sandia.gov*