SDC 19 | September 23-26, 2019
Santa Clara, CA

Computational Storage
Architectural Discussion

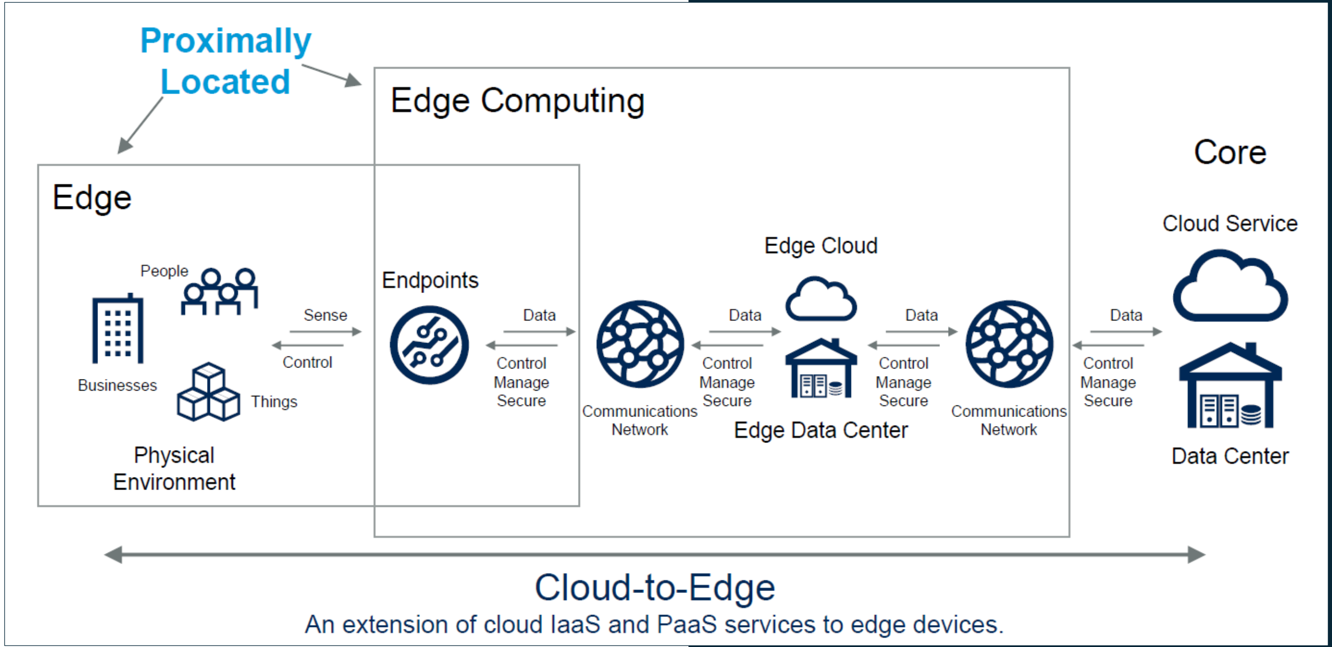
Scott Shadley
NGD Systems
Co-Chair Computational Storage TWG

Today's Learning Opportunities (TLO).

- EDGE needs CSD in M.2, EDSFF
- Architecture Our Way – CSD with PCSS
- AI – ML – CSD – the Overlap
- Hadoop & DB – CSD – Growth, Scale

Data, Data, Data. But Don't Take Our Word For it.

What Are You Doing with Your Data Today?

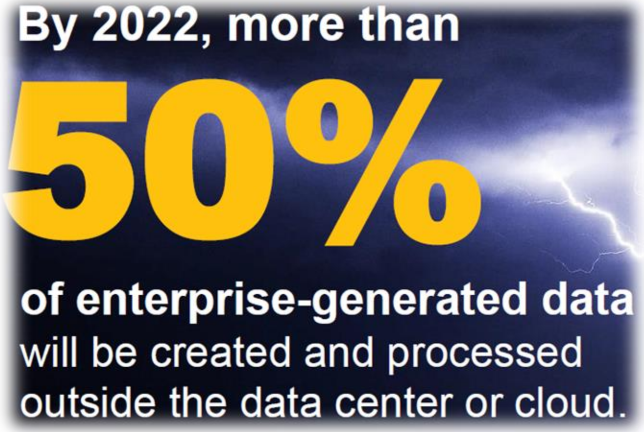


Source: Gartner - Bittman

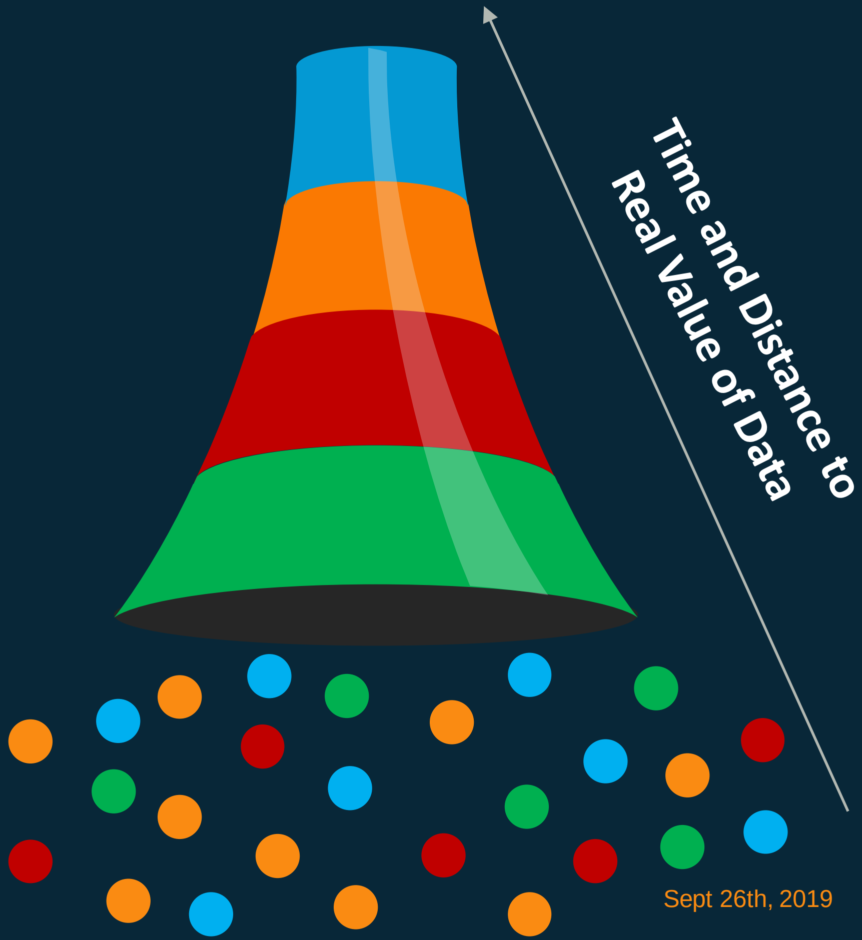
It's No Longer Black and White.

What is Driving Our Data Analytics Issues?

Weeding through the Noise at the Edge



Source: Gartner - Bittman



From Edge Sensors to Centralized Cloud

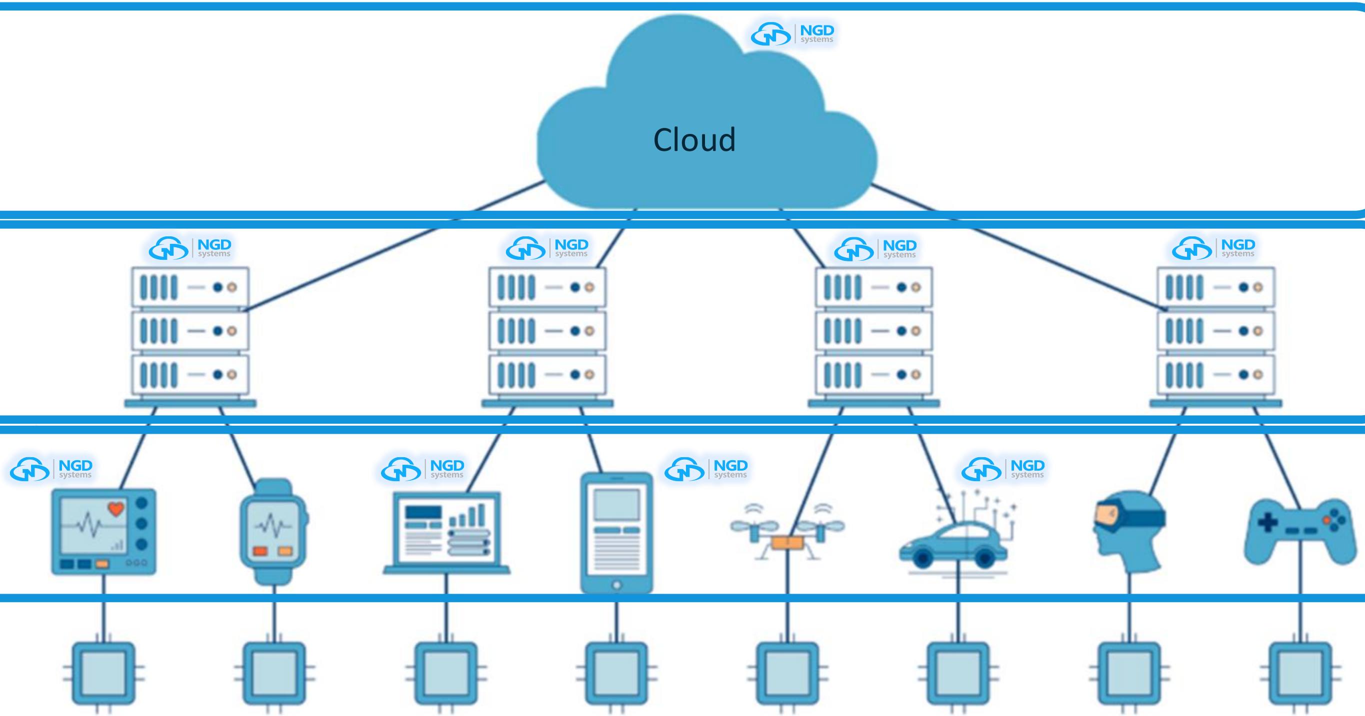
Computational Storage opportunities exist throughout the distributed compute environment

Centralized Cloud
Farthest from the network edge
High density of compute, storage and network resources

Infrastructure
Small distributed data centers
Low roundtrip latency 5 – 10ms

Edge Devices
Real-time data processing

Edge Sensors &Chip
Data Collection & Origination

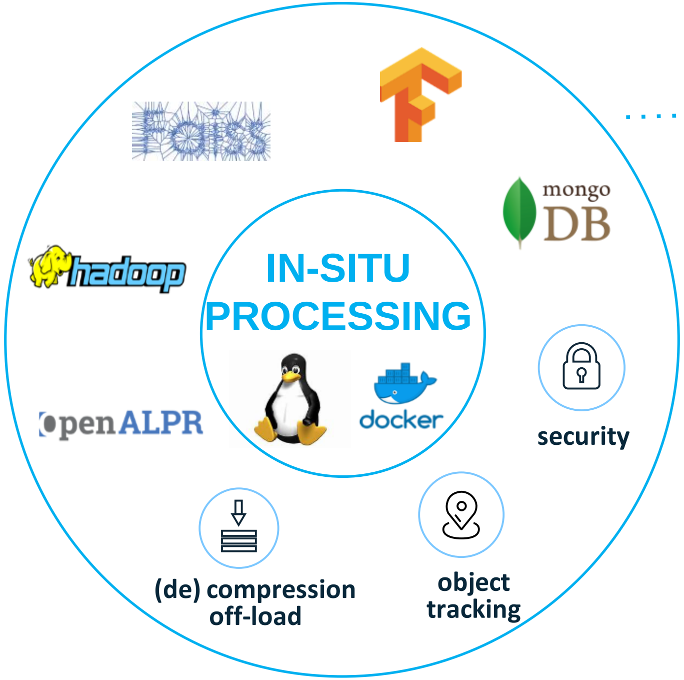


Innovative Computational Storage Uses.

FACIAL RECOGNITION

MACHINE LEARNING

HPC



CONTENT DELIVERY

SMART CITIES

AUTONOMOUS DRIVING

Highest Capacity, Lowest Power.

Industry leading W/TB

Industry's Only 16-Channel M.2

Industry's Largest Capacity U.2

8TB/8W
M.2

32TB/12W

16TB/12W
E1.S

Form Factor	Capacity (TB)	MAX Power (W)
M.2 22110	Up to 8	8
EDSFF E1.S	Up to 16	12
EDSFF E1.L	Up to 32	12
U.2 15mm	Up to 32	12
AiC FHTQL	Up to 64	15

Today's Learning Opportunities (TLO).

- EDGE needs CSD in Compact Form Factor
- Architecture One Way – CSD with PCSS
- AI – ML – CSD – the Overlap
- Hadoop & DB – CSD – Growth, Scale

Complete Solution and Disruptive Technology.

1st FULLY INTEGRATED COMPUTATIONAL STORAGE SOLUTION

SoC Controller

Industry's First 14nm

Management

- Modular firmware
- Efficient algorithm
- Flash characterization

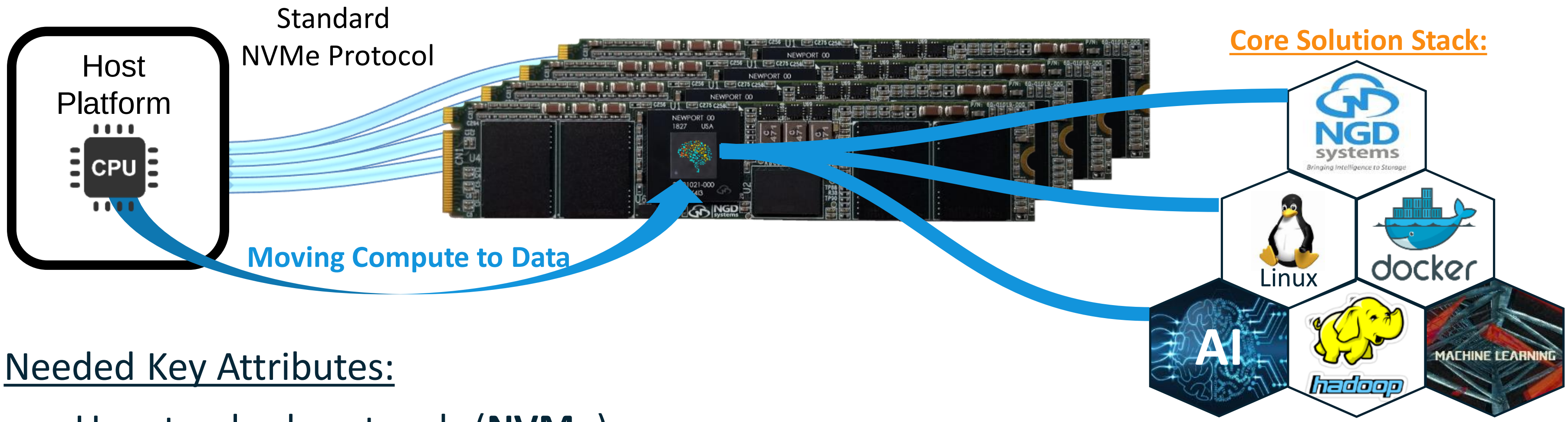
Optimized Hardware

"In-Situ Processing" Computational Storage Stack

- Full fledged on drive OS — Linux
- Light virtualization — docker
- Quad-core 64-bit application processor — ARM
- Hardware acceleration

The Scalable, ASIC-based Computational Storage Drive.

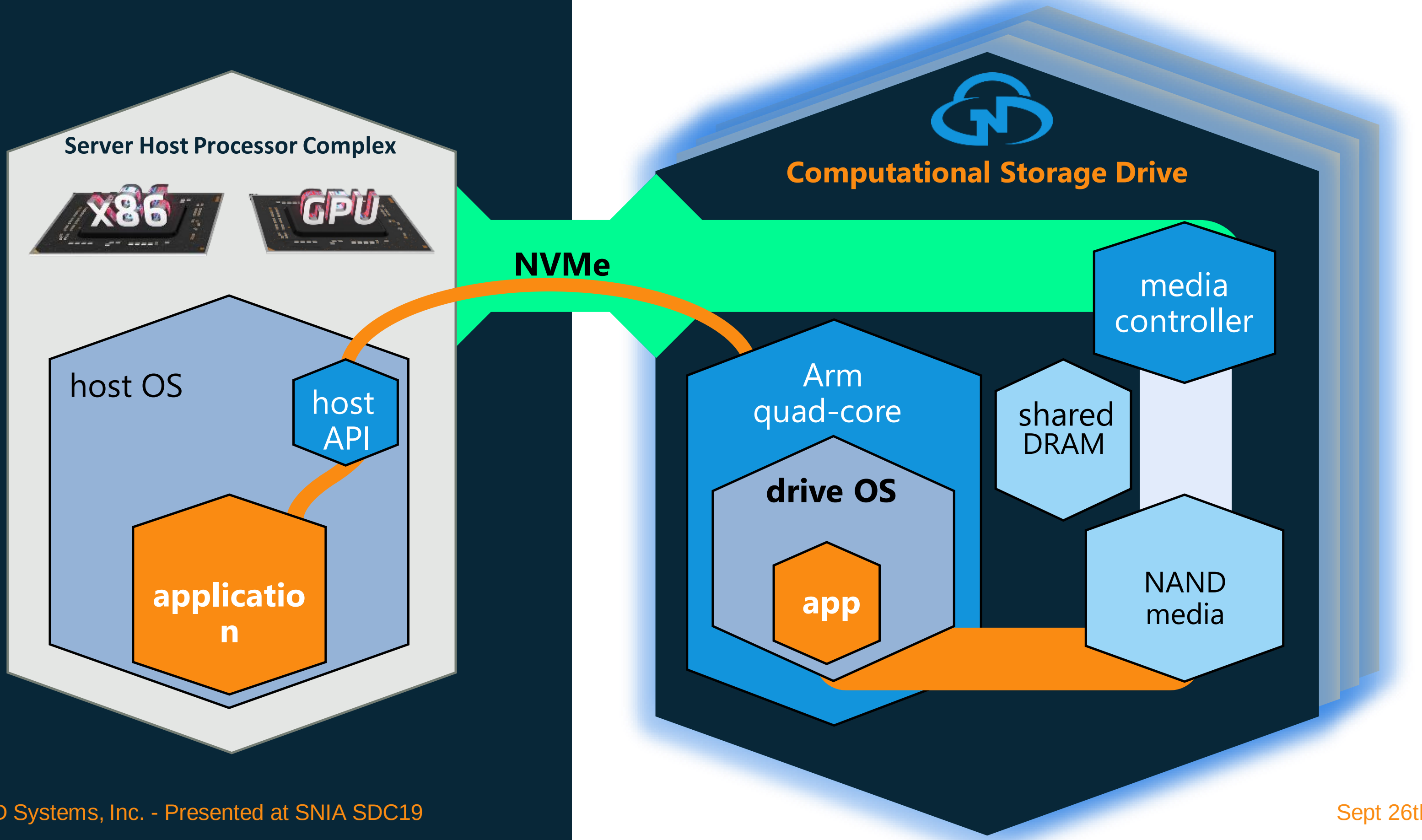
An enterprise class device capable of processing workloads in storage at the source



Needed Key Attributes:

- Use standard protocols (**NVMe**)
- Minimize data movement (Faster Response, Lower **W/TB**)
- Improve (**TB/in^3**) with maximize (Customer **TCO**)

The Data Lives on Storage. Why Not Work on it There?



Today's Learning Opportunities (TLO).

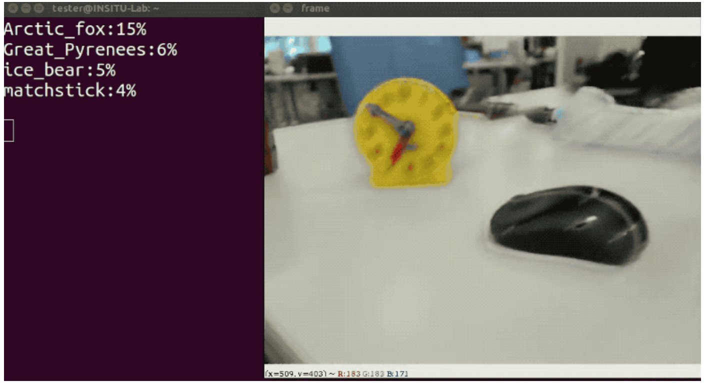
- EDGE needs CSD in Compact Form Factor
- Architecture One Way – CSD with PCSS
- AI – ML – CSD – the Overlap
- Hadoop & DB – CSD – Growth, Scale

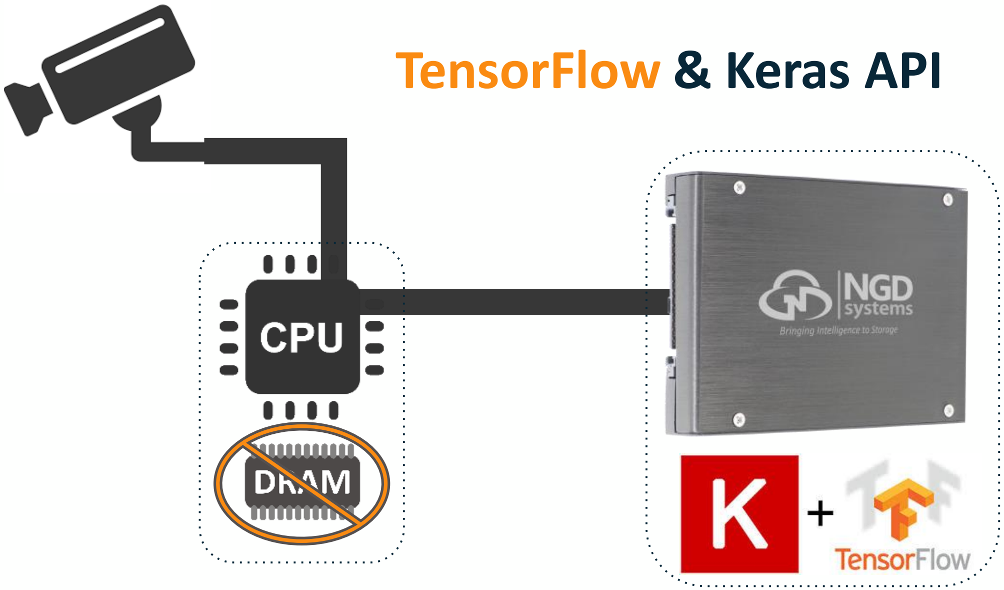
NGD systems

Using Computational Storage Drives for ML.

MobileNet Object Classification

TensorFlow & Keras API





No Host Interaction Required

The Leader in Computational Storage Solutions

Weightless Neural Networks Used for Object Tracking.

ESANN 2015 proceedings, European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning. Bruges (Belgium), 22-24 April 2015, i6doc.com publ., ISBN 978-287587014-8. Available from http://www.i6doc.com/en/.

A WiSARD-based multi-term memory framework for online tracking of objects

Daniel N. do Nascimento[1], Rafael L. de Carvalho[1,3], Félix Mora-Camino[4], Priscila V. M. Lima[2], Felipe M. G. França[1] *

1 – COPPE, 2 – NCE, Universidade Federal do Rio de Janeiro, BRAZIL

3 - Universidade Federal do Tocantins, UFT, BRAZIL

4 - Ecole Nationale de l'Aviation Civile - Laboratoire d'Automatique, FRANCE

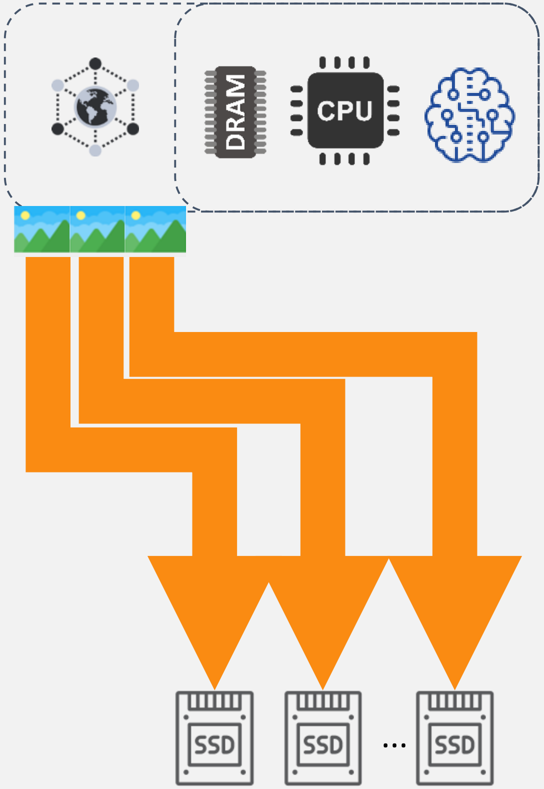
Abstract. In this paper it is proposed a generic object tracker with real-time performance. The proposed tracker is inspired on the hierarchical short-term and medium-term memories for which patterns are stored as discriminators of a WiSARD weightless neural network. This approach is evaluated through benchmark video sequences published by Babenko et al. Experiments show that the WiSARD-based approach outperforms most of the previous results in the literature, with respect to the same dataset.

NGD systems

Moving Beyond Traditional Models.

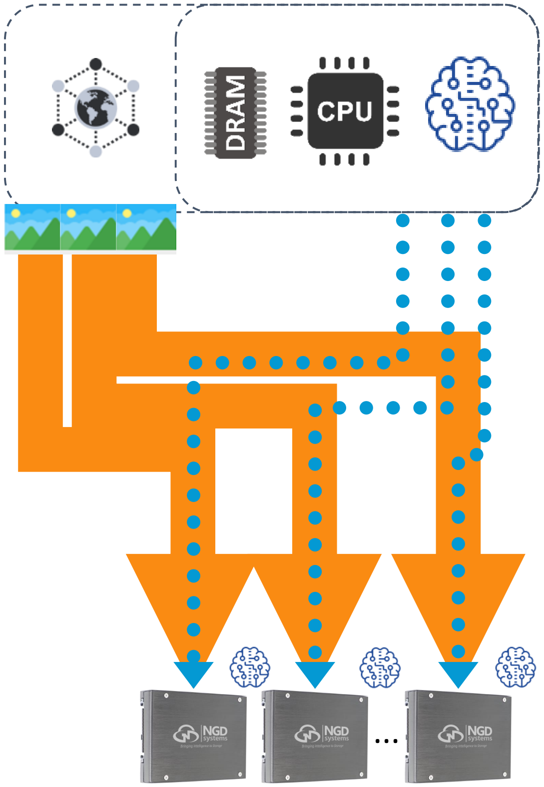
- Parallel & distributed Training in Computational Storage
- Federated/Transfer Learning
- Reduce data transfers by sending sparse model updates

ML Training with Traditional Approach.

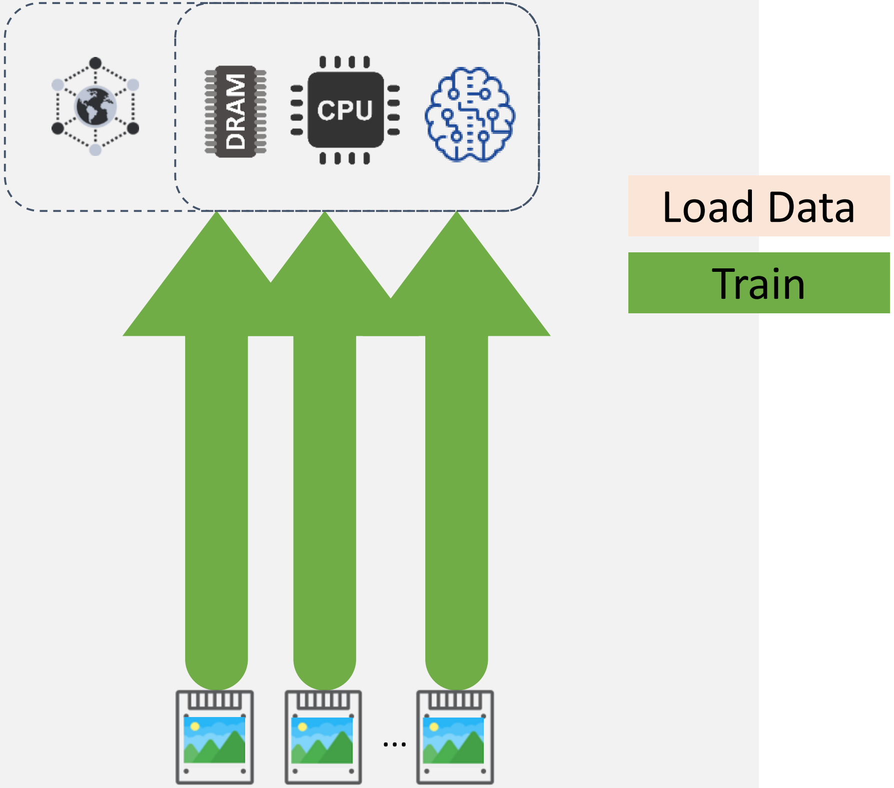


Load Data

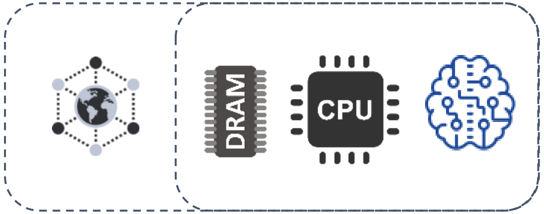
ML Training with Computational Storage.



ML Training with Traditional Approach.

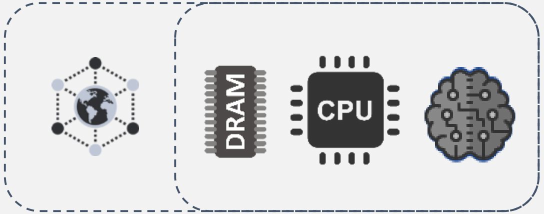


ML Training with Computational Storage.



- No data movement
- No host CPU needed
- Distributed training

ML Training with Traditional Approach.



- Host CPU still needed
- No Parallelism

Load Data

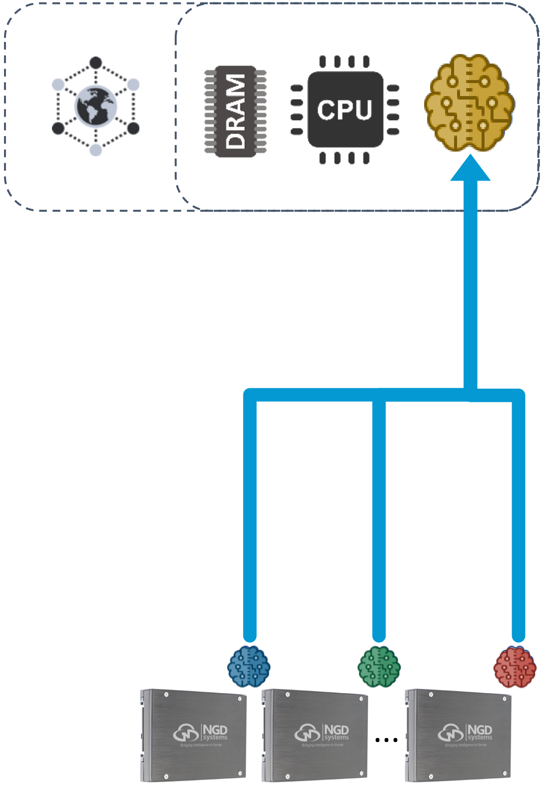
Train

Evaluate

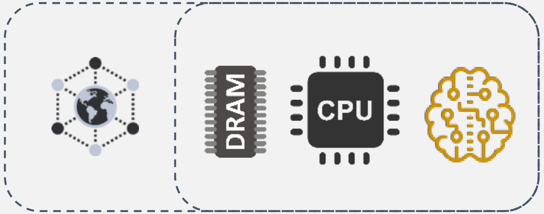
Update

ML Training with Computational Storage.

NGD systems



ML Training with Traditional Storage.



Repeat Steps

Train

Evaluate

Update

Load Data

Train

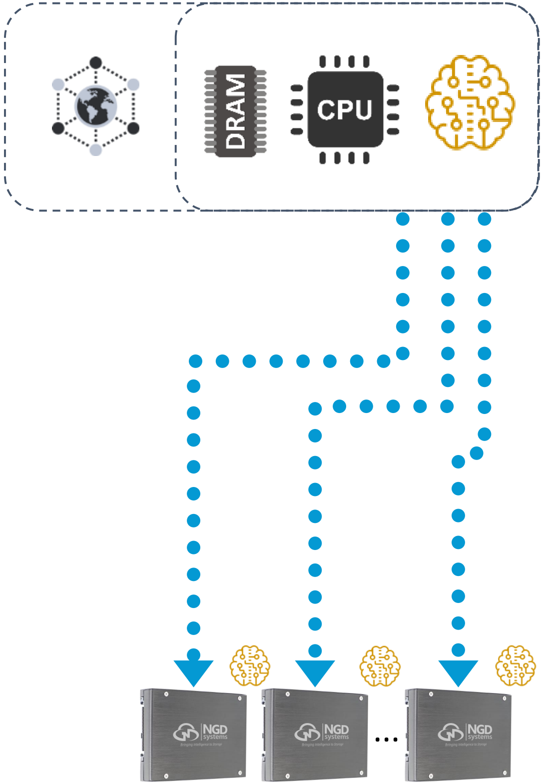
Evaluate

Update

Deploy

ML Training with Computational Storage.

NGD systems



Federated/Transfer Learning.

MNIST DATASET

60,000 samples

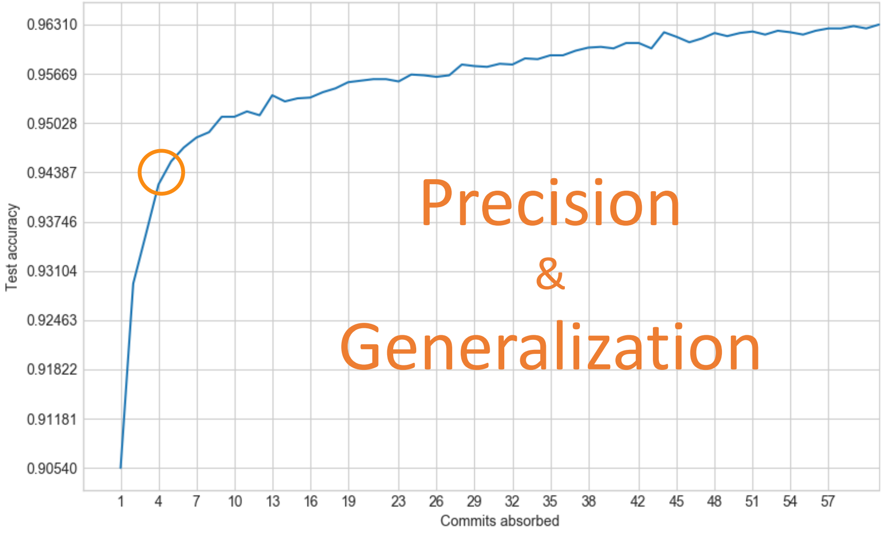
From the training set

61 updates

Model updates transferred

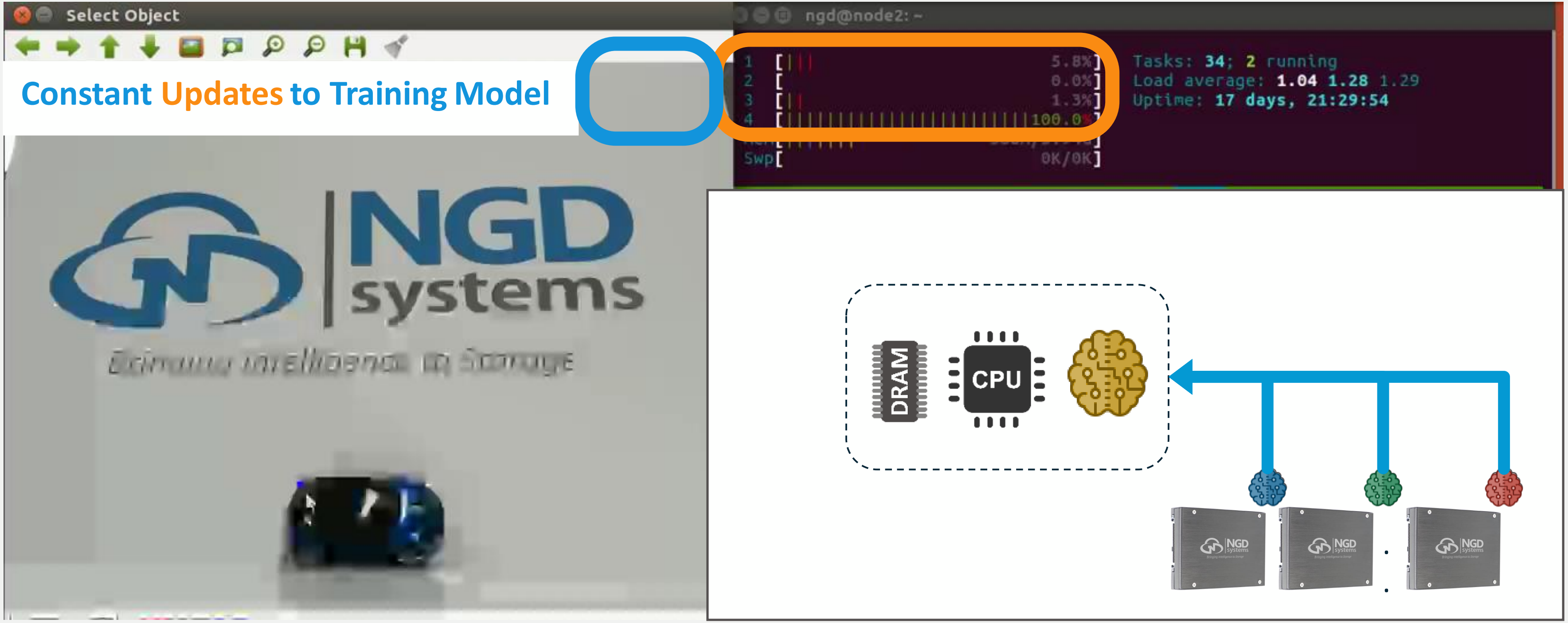
94% accuracy

With only 4 partial model updates



Computational Storage Cores Running WiSARD.

The Leader in Computational Storage Solutions

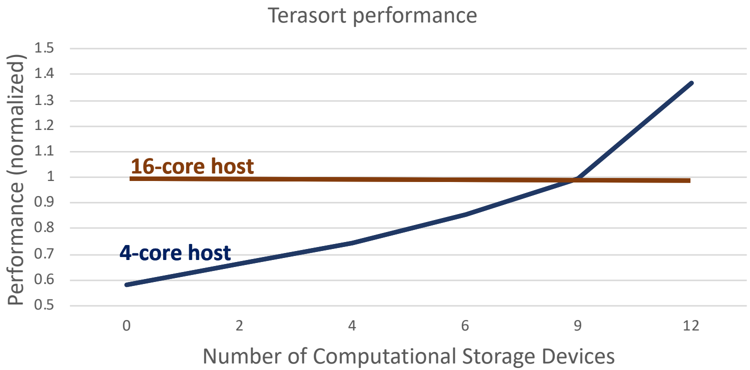


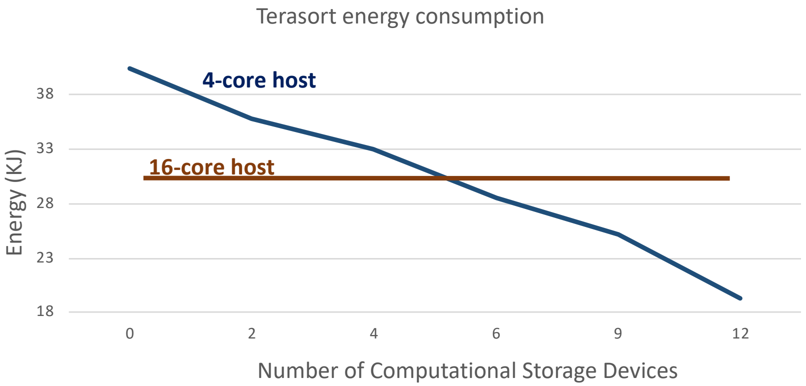
Constant Updates to Training Model

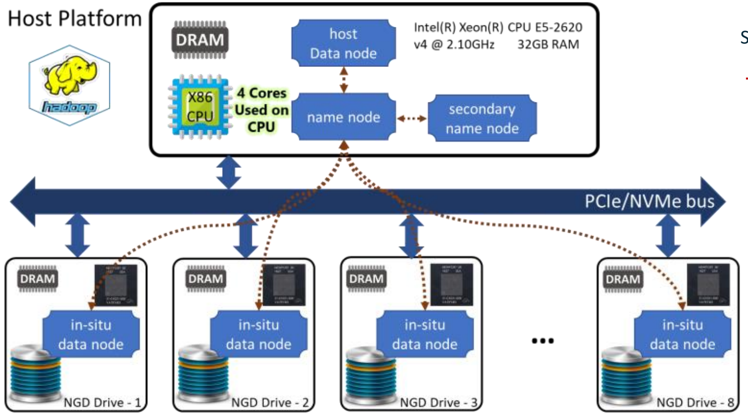
Today's Learning Opportunities (TLO).

- EDGE needs CSD in Compact Form Factor
- Architecture One Way – CSD with PCSS
- AI – ML – CSD – the Overlap
- Hadoop & DB – CSD – Growth, Scale

Amplifying TCO for Hadoop



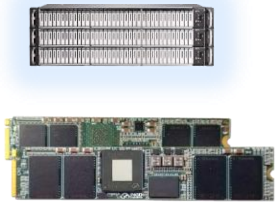




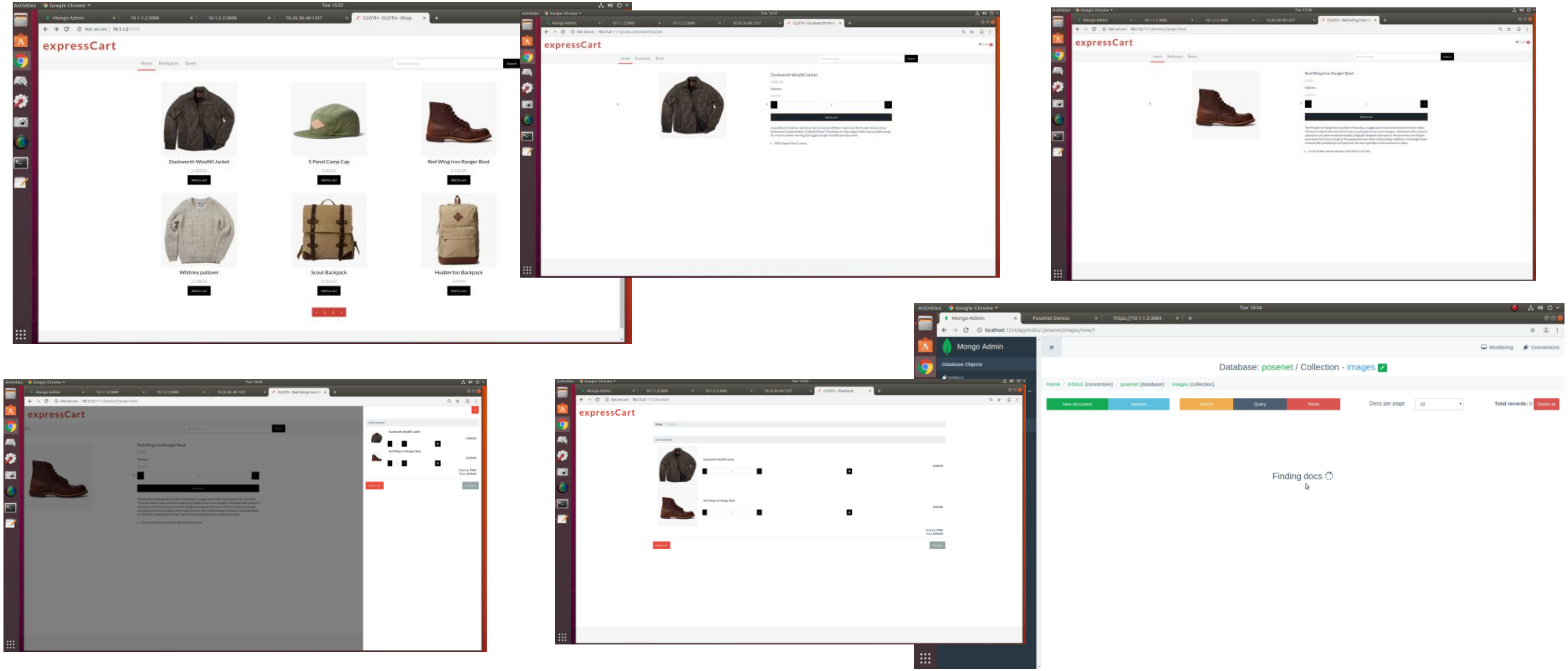
Datanode Config:
Single E5-2620v4, 32GB DRAM, 12*8 TB SAS HDD
18U Total Density in 18U = 864TB

@ Scale
Saves Power!
Saves Space!
Saves Time!

Datanode Config:
Single E5-2620v4, 32GB DRAM, 36*8TB NVMe
3U Total Density in 3U = 864TB
432 Additional Drive Cores



Using MongoDB within Computational Storage.



Scalable Computational Storage.

A New Storage Paradigm is Here

- **The "New Cloud" needs the Distributed Edge**
  - There is no longer just a 'central' storage location
- **Edge data growth challenges HW platforms**
  - Innovative form factors and high capacity for the Edge
- **In-Situ Processing brings ML closer to data**
  - Exploit data locality and enable distributed processing

SDC19

September 23-26, 2019
Santa Clara, CA

Computational Storage

Scott.Shadley@NGDSystems.com
www.NGDSystems.com

TW: @SMShadley
@NGDSystems