



THE A.I. PARADOX

Biggest Opportunities and Challenges of our Lifetimes

Gary Dickerson

PRESIDENT AND CEO
APPLIED MATERIALS, INC.

A.I. Design Forum | July 9, 2019

EXTERNAL USE

PHOTO : https://commons.wikimedia.org/wiki/File:San_Francisco_Bay_Bridge_Western_Span_at_night.jpg

Forward-Looking Statements

This presentation contains forward-looking statements, including those regarding anticipated growth and trends in Applied's businesses and markets, industry outlooks, technology transitions, and other statements that are not historical facts. These statements are subject to risks and uncertainties that could cause actual results to differ materially from those expressed or implied by such statements and are not guarantees of future performance. Information concerning these risks and uncertainties is contained in Applied's most recent Form 10-Q and 8-K filings with the SEC. All forward-looking statements are based on management's current estimates, projections and assumptions, and Applied assumes no obligation to update them.

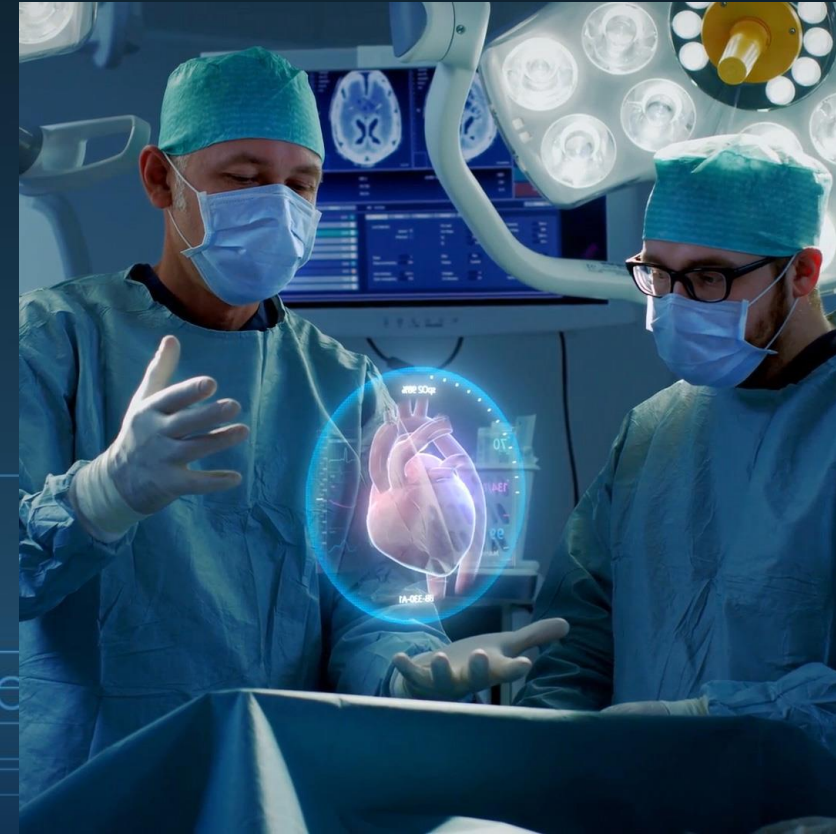
AGRICULTURE



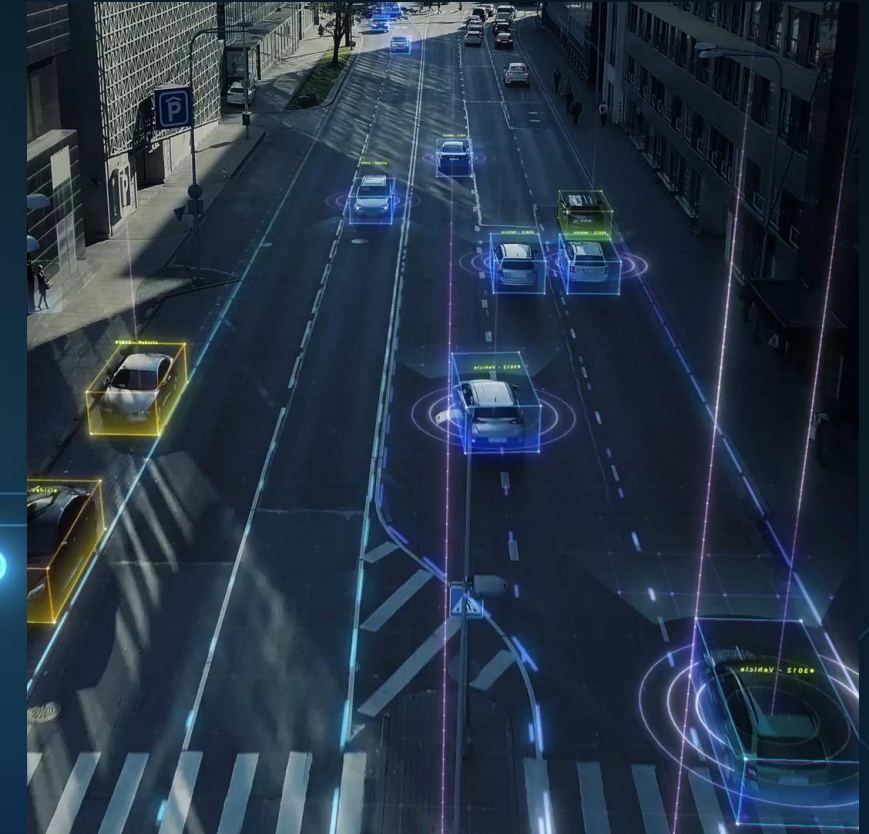
EDUCATION



HEALTHCARE



TRANSPORTATION



“In the future, **every business** will be in the **technology business**”



**RATE OF
A.I. ADOPTION
DETERMINED BY**

Speed of innovation

Society's acceptance

Security and safety

Sustainability



EDGE DEVICE
Compute + Storage

COMMUNICATIONS



CLOUD DATACENTER
Compute + Storage

Compute and store at
EDGE vs. CLOUD
determined by 3 “laws”

1. Law of **the land**
2. Law of **physics**
3. Law of **economics**

500B

TO

1T

Connected devices by 2030

SOURCE: CISCO

Connected devices by 2030

SOURCE: ARM



1 GB
PER DAY

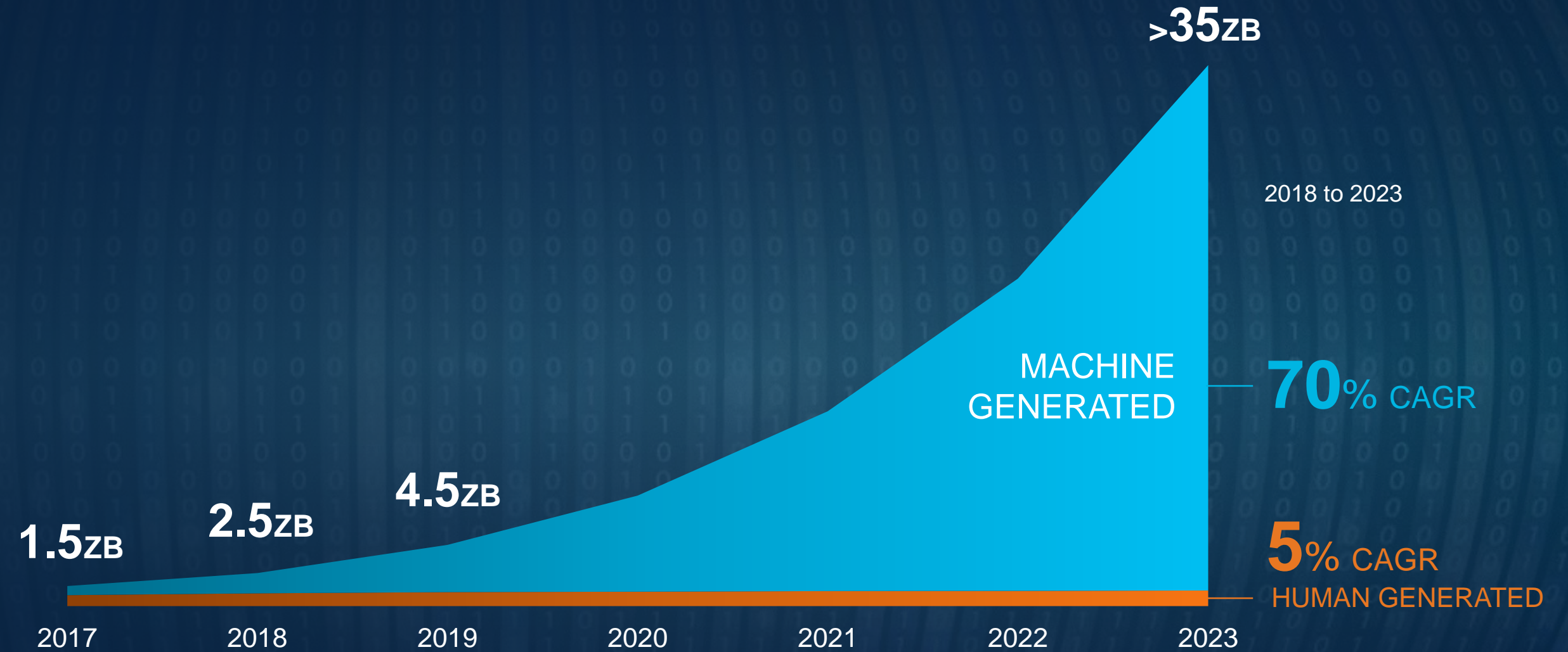


4,000 GB
PER DAY



Source: Intel, Waymo

EXPLOSION OF DATA GENERATION



SOURCE: Applied Materials model based on forecasts published by Cisco, Intel, Western Digital

GENERAL PURPOSE COMPUTING
(GOOD AT MANY THINGS)

Multi-level on-die cache architecture

Deep instruction pipeline,
out-of-order execution, branch
prediction, pre-fetching, ...

64 bit or higher precision data paths

A.I. COMPUTING
(CUSTOMIZED FOR SPECIALIZED TASKS)

High Speed Memories

To improve memory latency
and memory to logic bandwidth

Extreme Parallelism

Simple operations on multiple
chunks of data in parallel

Lower Precision

Training and inference at lower floating point
(16 bit) and integer (8 bit) precision

New Computing Approaches Needed to Unlock Value of Data Explosion

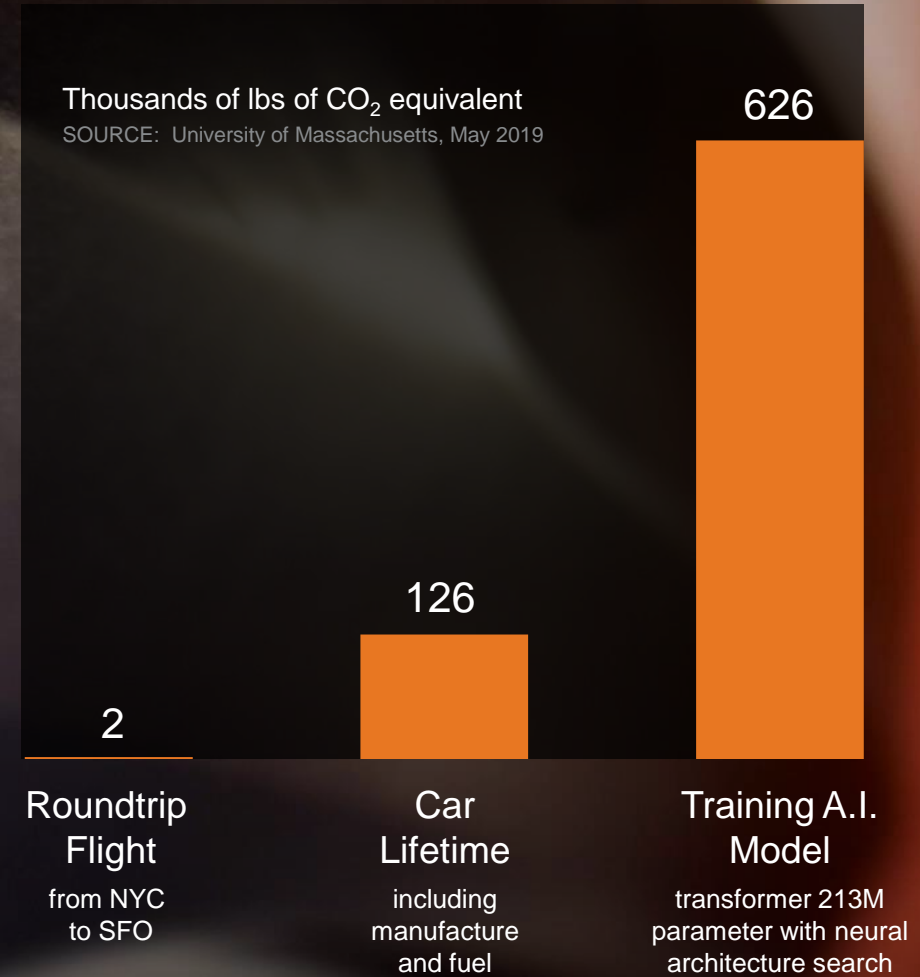
FASTEST GROWING A.I. WORKLOADS ARE VISUAL



	VISUAL	AUDIO	TEXT / NUMERIC
Examples	IP / Security cameras Autonomous vehicles Robotics	Voice assistants Wearables Call Centers	Search optimization Advertising Business intelligence
Neural Network (NN) Type	Convolutional NN (CNN)	Recurrent NN (RNN) or CNN	Multilayer Perceptron (MLP)
% of total NN use today	10%	25%	65%
% of total NN use 2025	45%	15%	40%

“Training a single A.I. model can emit as much carbon as five cars in their lifetimes”

MIT TECHNOLOGY REVIEW, JUNE 2019



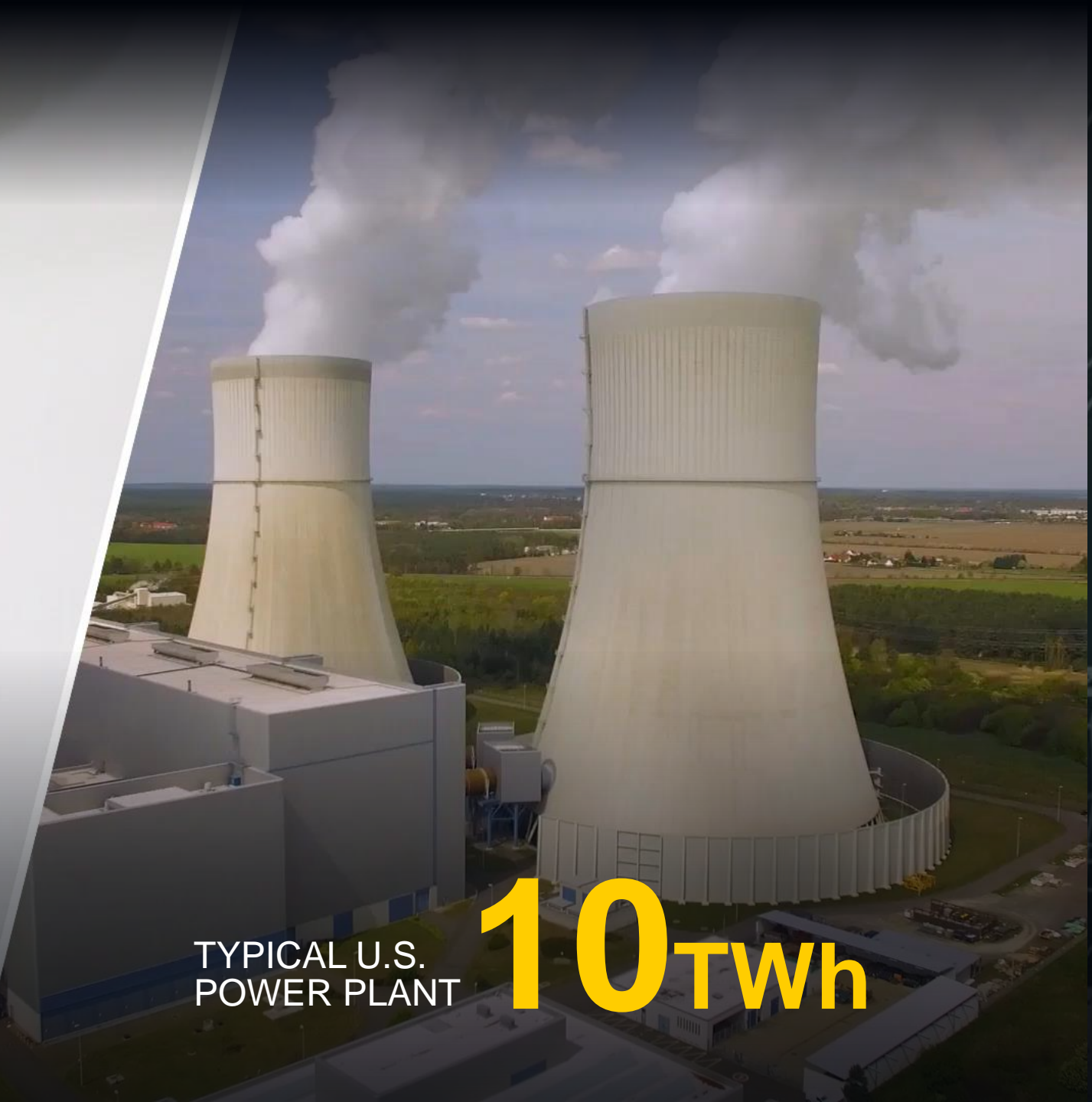
Energy Consumption Could Become a Key Constraint to Growth of A.I.

SHEER NUMBER
OF EDGE DEVICES
CAN CHANGE
WORLD'S ENERGY
EQUATION



5 to 8 WATTS
PER UNIT x
500M UNITS

12TWh

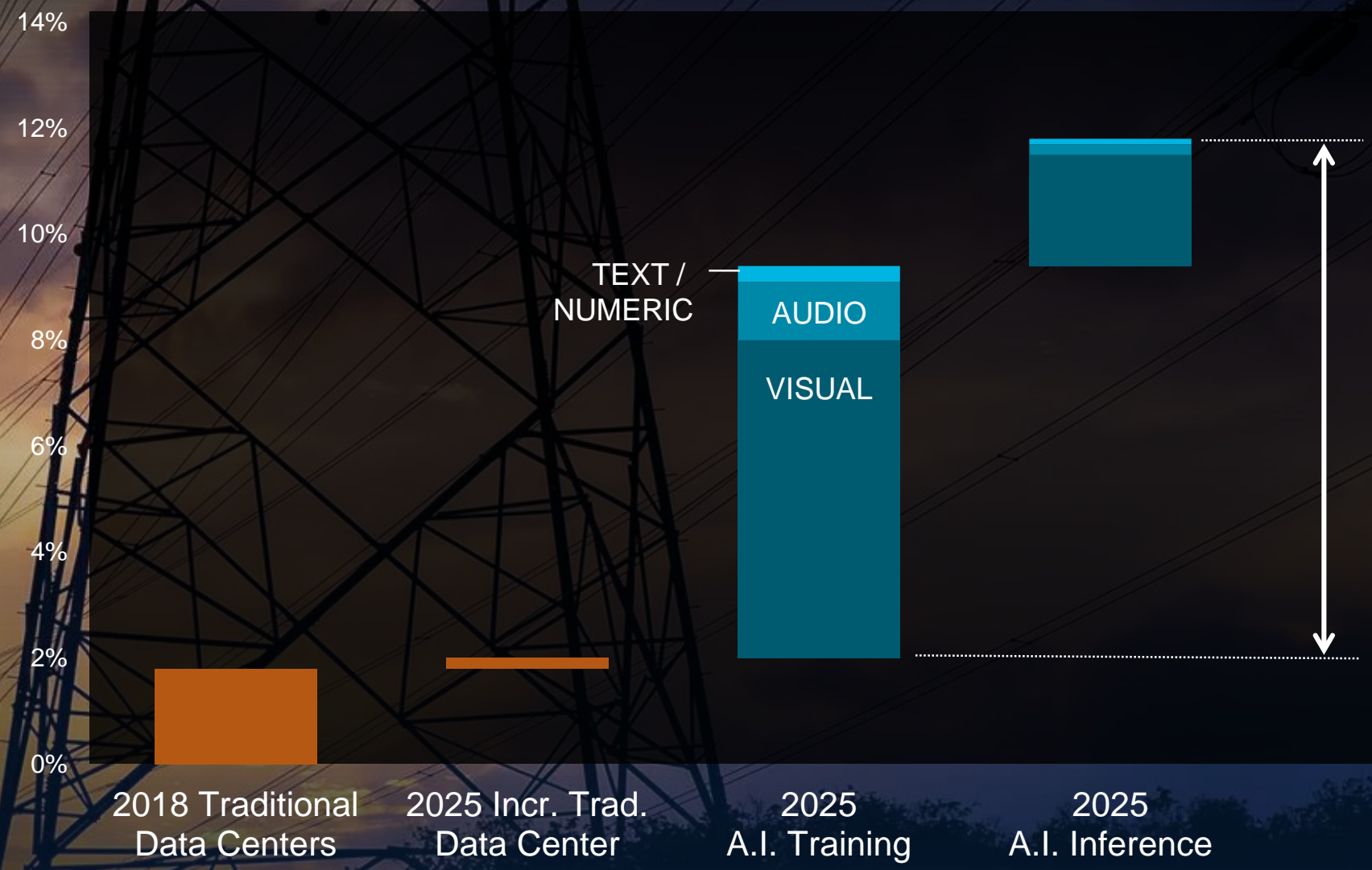


TYPICAL U.S.
POWER PLANT

10TWh

IMPACT OF SCALING THE CLOUD IS EVEN MORE SIGNIFICANT

Power Consumption of Data Centers as % of Global Electricity Generation



A.I. data centers could consume more than 10% of global electricity supply by 2025

IF BUILT USING TECHNOLOGY
from 1980s...

14 Meters

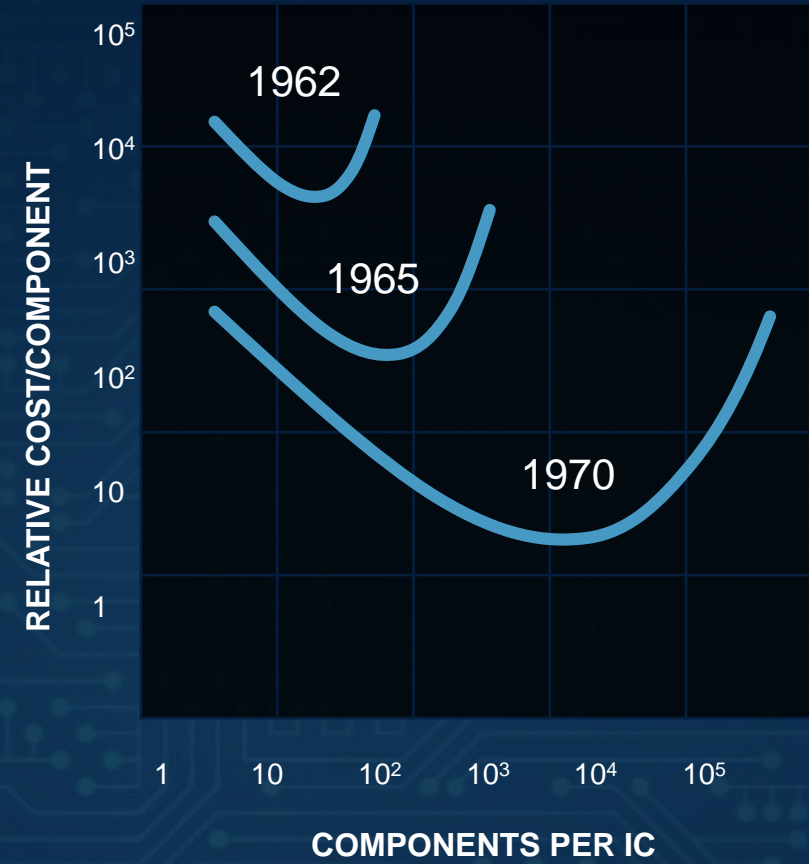
600kW

\$100M

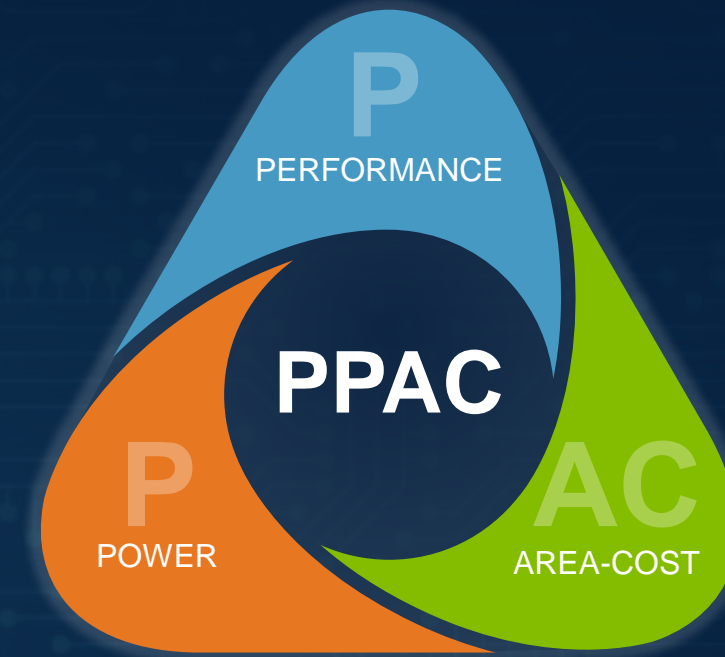


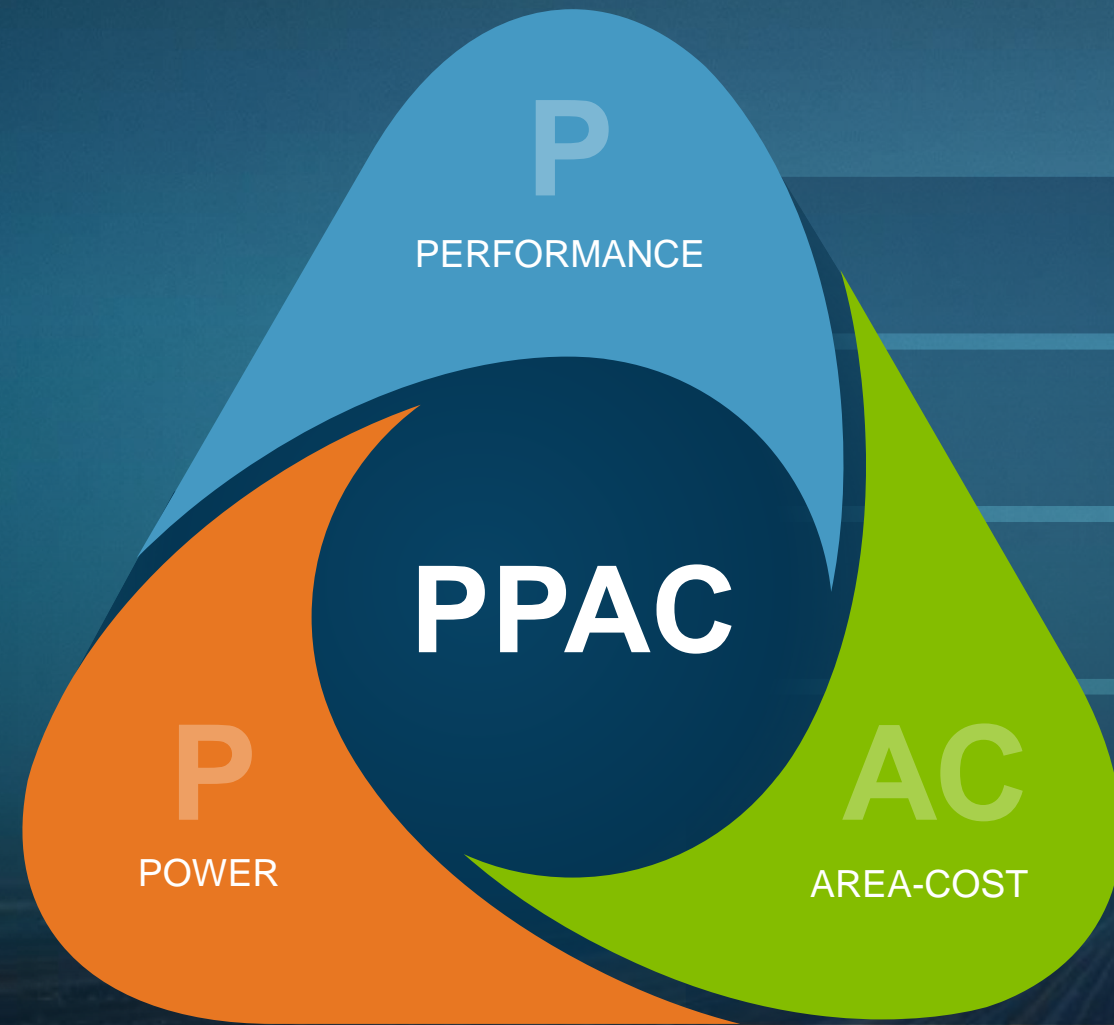
FOR **40** YEARS

MOORE'S LAW



IMPROVEMENTS IN...





ENABLED BY

New architectures

New structures / 3D

New materials

New ways to shrink

Advanced packaging

A New Playbook for Semiconductor Design and Manufacturing is Needed

**Major advances at edge
and in cloud needed**
to realize potential of A.I.
without impacting world's
energy equation

TRAINING

CLOUD

NEEDED VS. CURRENT SOLUTION (GPU)

75x
throughput and
1,000x
performance per watt

EDGE

Limited impact on roadmap

INFERENCE

NEEDED VS. CURRENT SOLUTION (CPU)

50x
better latency and
500x
throughput

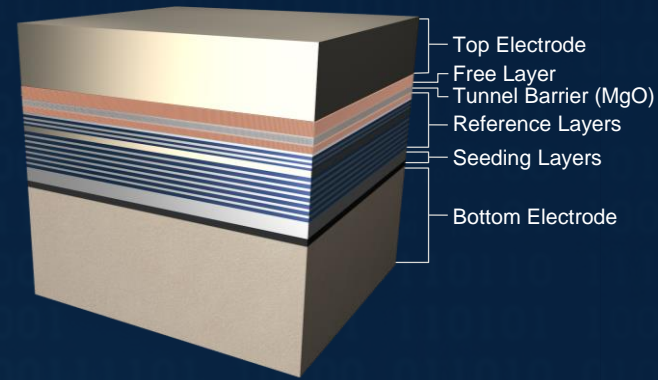
NEEDED VS. CURRENT SOLUTION (CPU)

50x
better latency and
75x
throughput

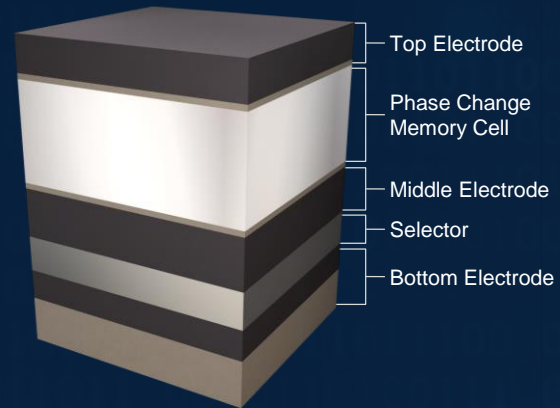
	INITIAL DEPLOYMENT	CLOUD	EDGE
Accelerators GPU, TPU, ASICs, FPGAs	Now	✓	Autos
Near Memory DDR, SRAM, HBM, NAND, SCM	Now to 2 years	✓	✓
New Memory MRAM, ReRAM, PCRAM, FeRAM	Now to 5 years	✓ ←	✓
In-Memory Compute Analog, ReRAM, PCRAM	2 to 5 years	✓ ←	✓
Novel HPC Quantum, Synaptic	5 to 10 years	✓ →	✓

Promising Pipeline of New Hardware Innovations

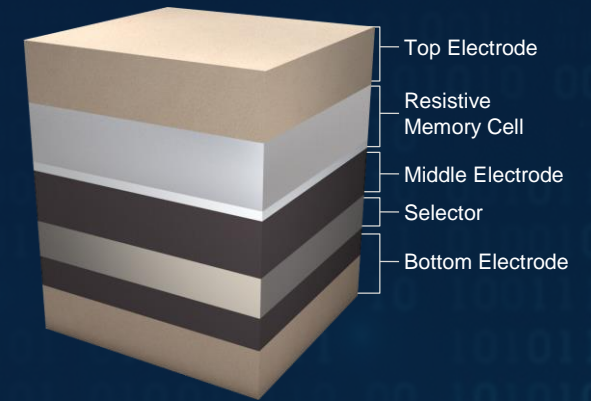
ENABLING ROADMAP REQUIRES NEW APPROACHES TO MANUFACTURING



MRAM
Magnetic RAM



PCRAM
Phase Change RAM



ReRAM
Resistive RAM

INITIAL DEPLOYMENT	Edge	Cloud
KEY ATTRIBUTES	<ul style="list-style-type: none"> ✓ Extremely fast (as fast as L3 cache) ✓ High density (one transistor per cell) ✓ Non-volatile 	<ul style="list-style-type: none"> ✓ Faster than NAND ✓ Lower cost than DRAM ✓ Non-volatile ✓ Ideal for storage class memory
ENABLED BY	New materials, 3D structures, Sub atomic-level precision films	

Endura[®] Clover[™] MRAM PVD platform

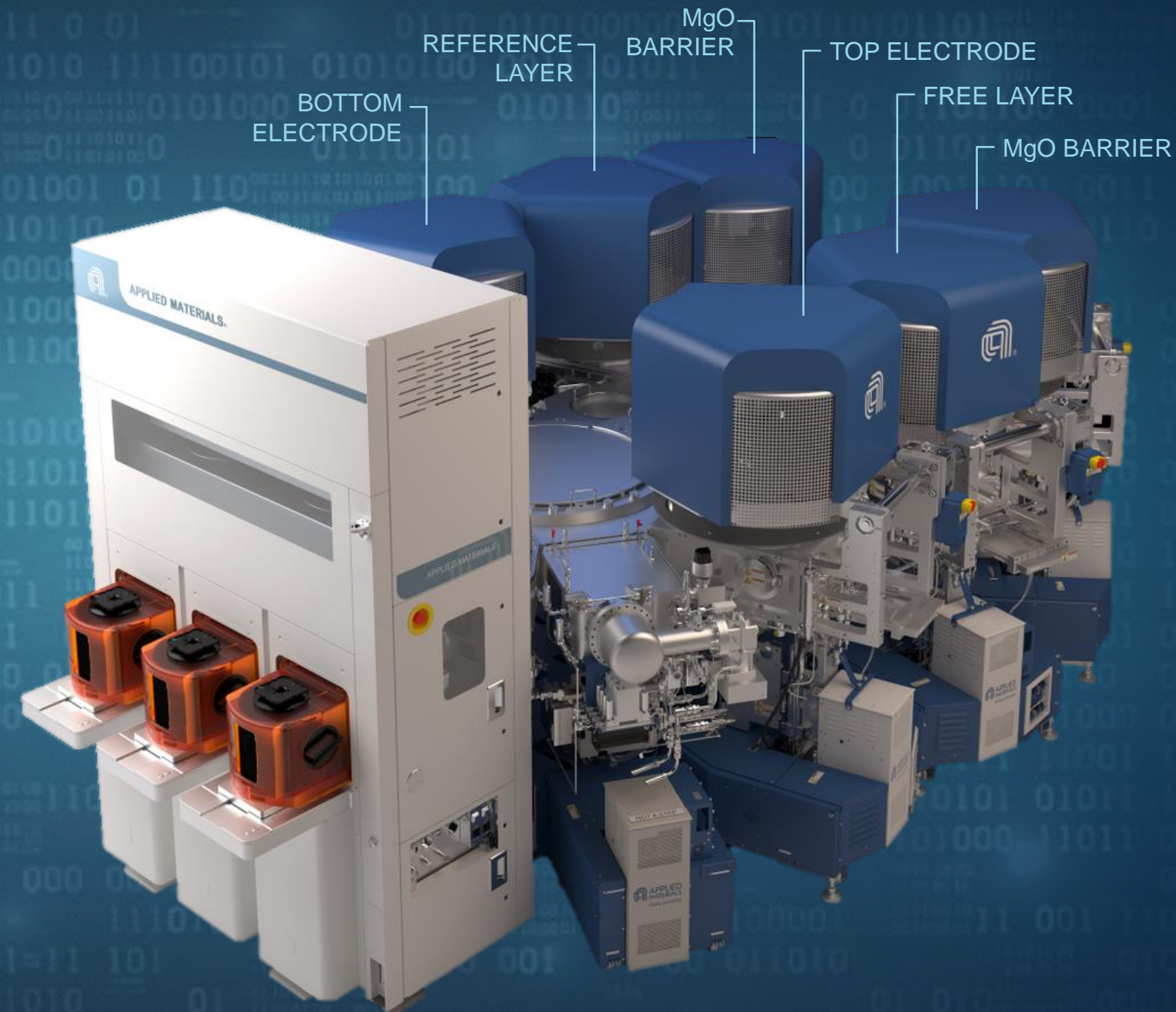
Integrated Materials Solution

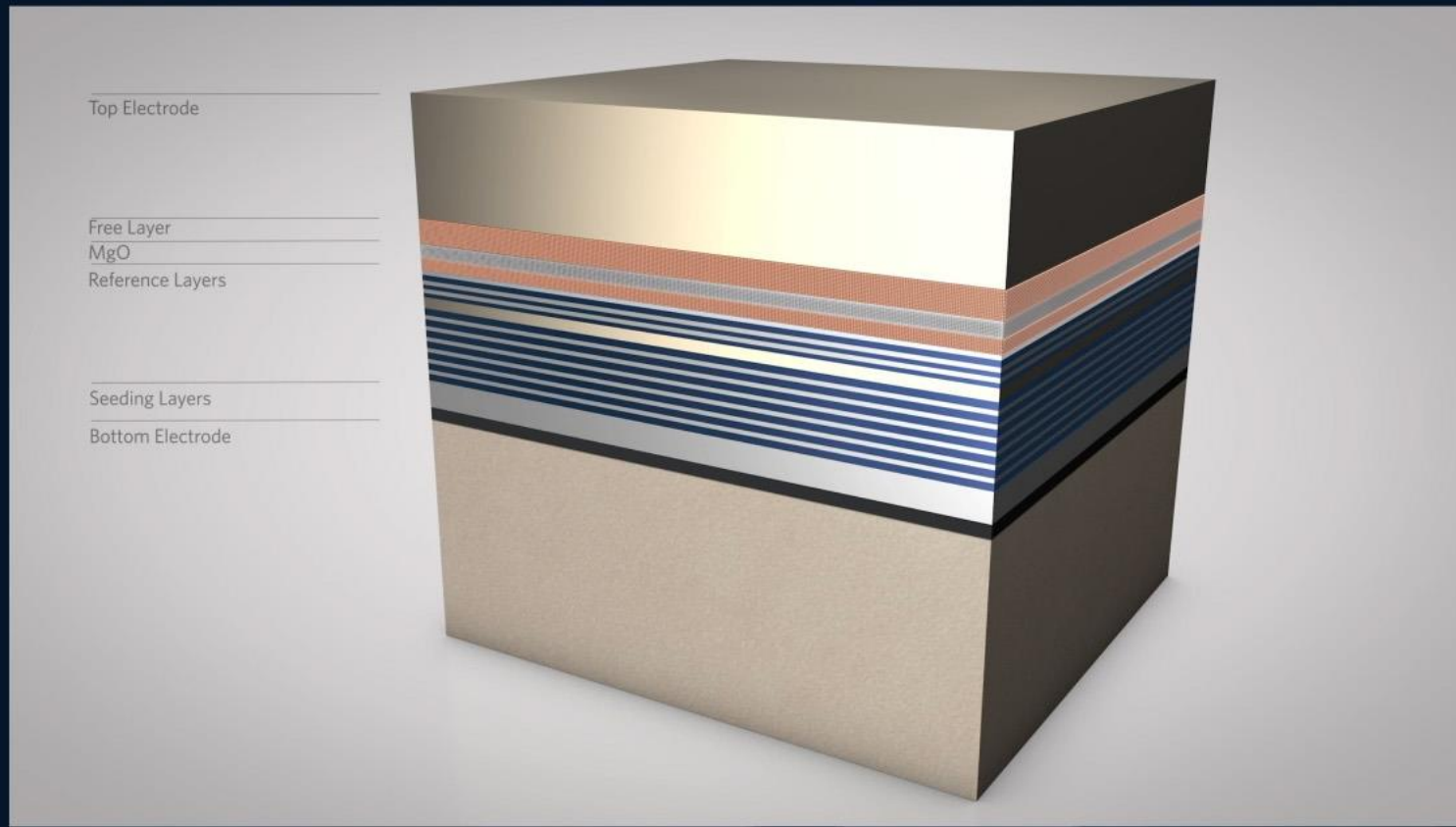
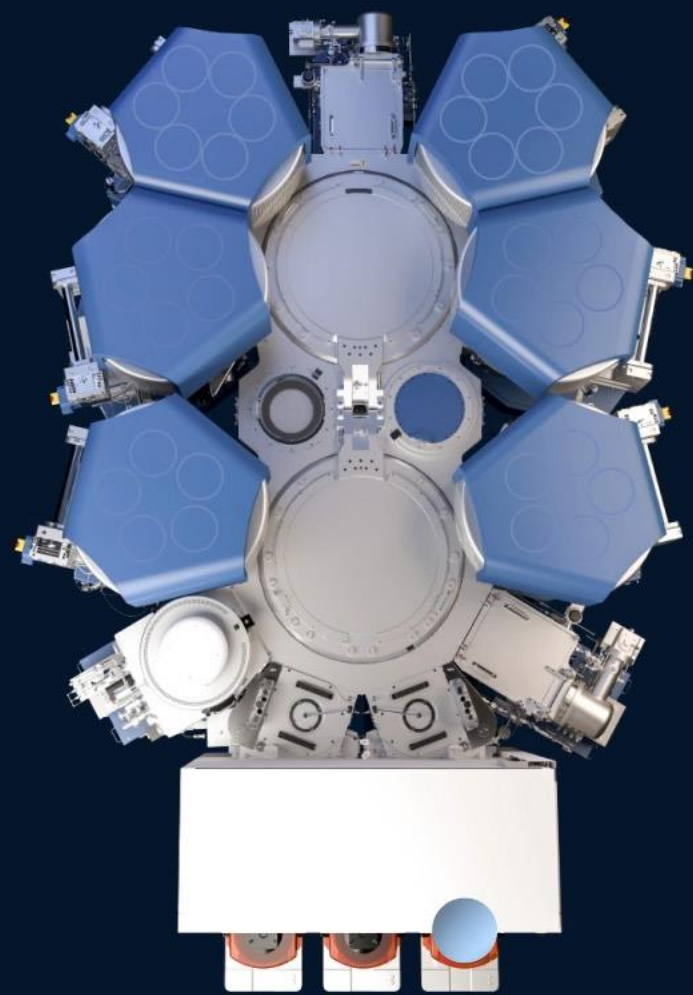
9 chambers

Up to 5 materials per chamber

On-board metrology

“Fab inside a fab”

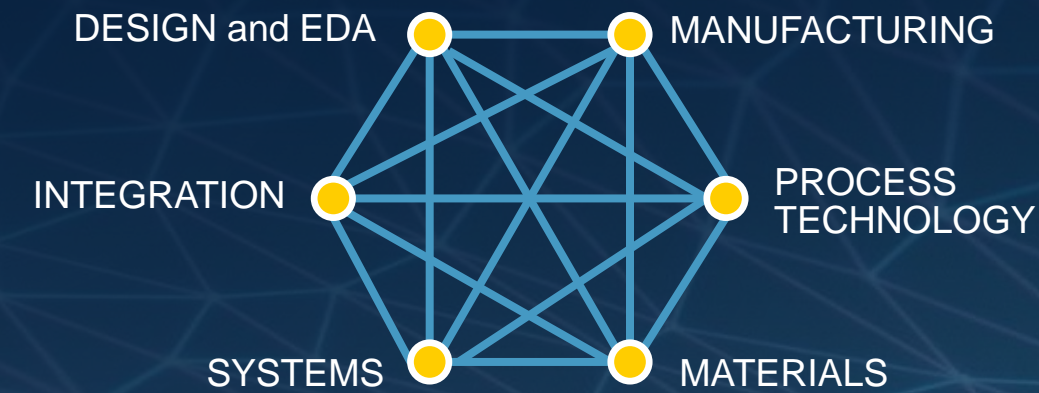




“Von Neumann” mindset vs. “Neuromorphic” mindset



TODAY: **Serial** / compartmentalized interaction between key parts of ecosystem



OPPORTUNITY: **Parallel** development to get powerful tools to designers faster

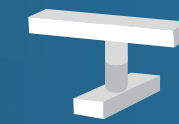
Ecosystem Collaboration is Key to Accelerating A.I. Adoption



MATERIALS



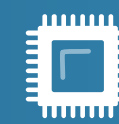
PROCESS
TECHNOLOGY



DEVICES &
STRUCTURES



DESIGN



ARCHITECTURE



ALGORITHMS



SYSTEMS

MAYDAN TECHNOLOGY CENTER

State of the art 300mm
process technology lab

ADVANCED PACKAGING LAB

Fully integrated 300mm advanced
wafer level packaging lab

META CENTER

opening fall 2019

'Lab to fab' accelerator

ADVANCED MATERIALS LAB

Foundational materials
engineering R&D

Applied's Collaborative R&D Platform
to Accelerate **Materials to Systems 10x**

A.I. GRAND CHALLENGES

IMPROVE COMPUTE
PERFORMANCE / WATT

1,000x

ACCELERATE
MATERIALS TO SYSTEMS

10x