



The background features a high-resolution image of the Earth from space, showing the Americas. A vertical line divides the image. To the right of this line, a complex, multi-colored digital grid of glowing lines (in shades of blue, green, and orange) is overlaid on the Earth, representing a digital or data landscape. The overall theme is sustainable computing and technology.

arm

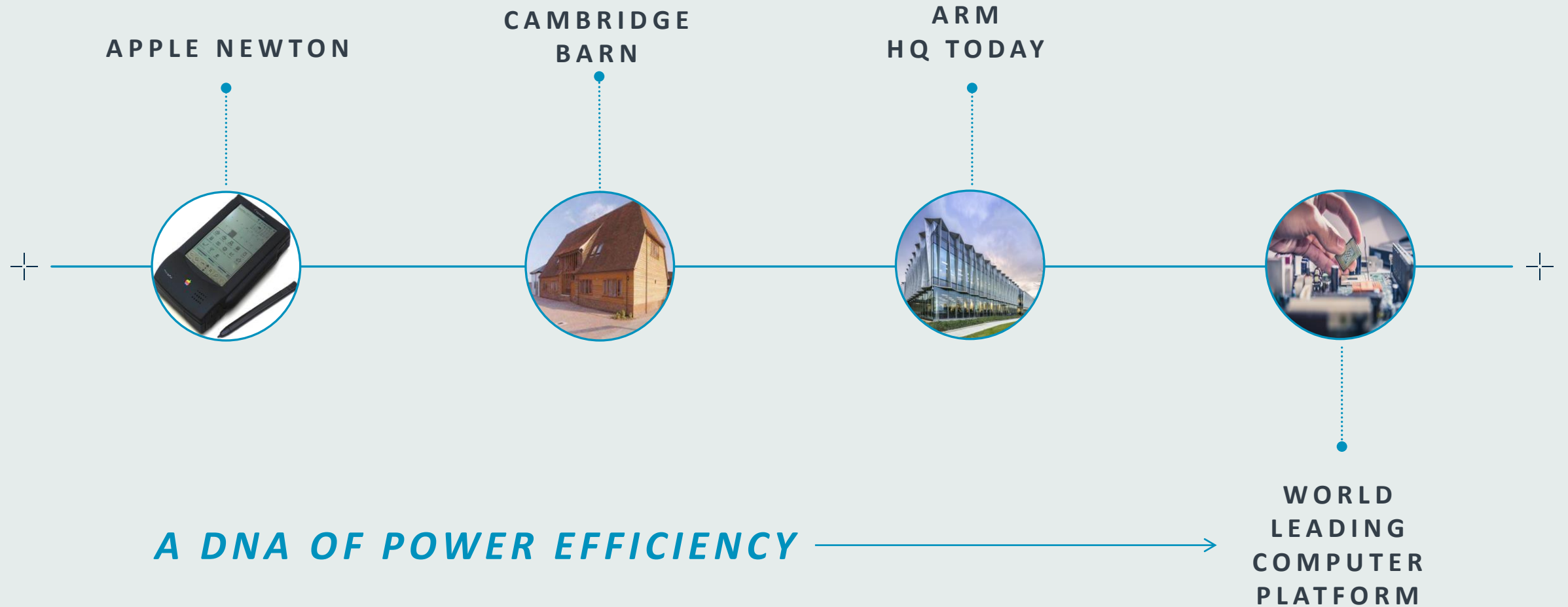
The Future of Sustainable Computing

Mohamed Awad

SVP & GM of Infrastructure, Arm

 @awadmo  [linkedin.com/in/moawad](https://www.linkedin.com/in/moawad)

30+ Years of Engineering Innovation



Arm Technology is Defining the Future of Computing

A microprocessor design and software platform company

250 Billion

Arm-based chips shipped to date

30 Billion

Arm-based chips shipped in 2022

650+

Active licensees, growing by 50+ every year

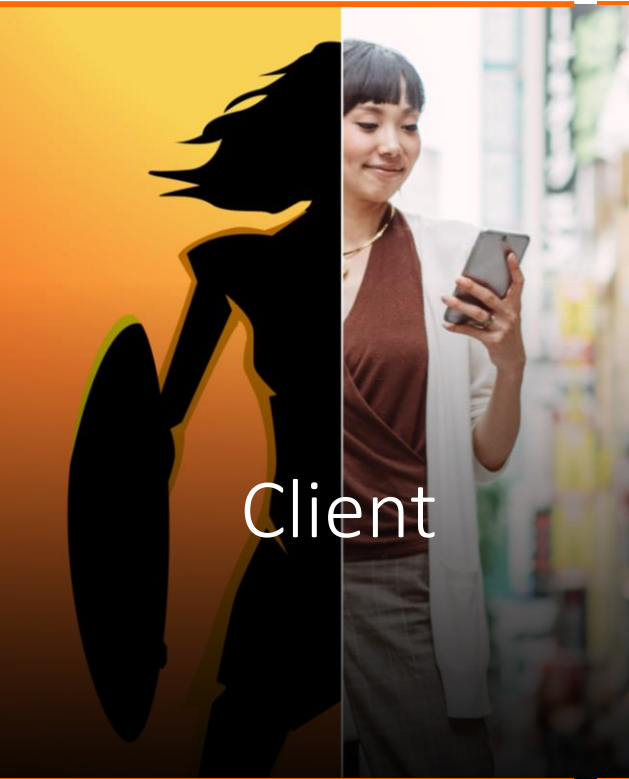


70%

of the world's population use products and services powered by Arm technology.



Arm has Diversified and Delivers for Key Markets



Total Compute

Embedded and Edge AI

Autonomous & Safety

Deliver, Store & Analyze

The Pressure on Infrastructure



CONNECTIVITY AND DATA



14.75B
IoT CONNECTIONS



1B
5G SUBSCRIBERS



120 ZB
DATA GENERATED

As of 2023



SCIENCE AND POLITICS



Moore's Law
ENDING

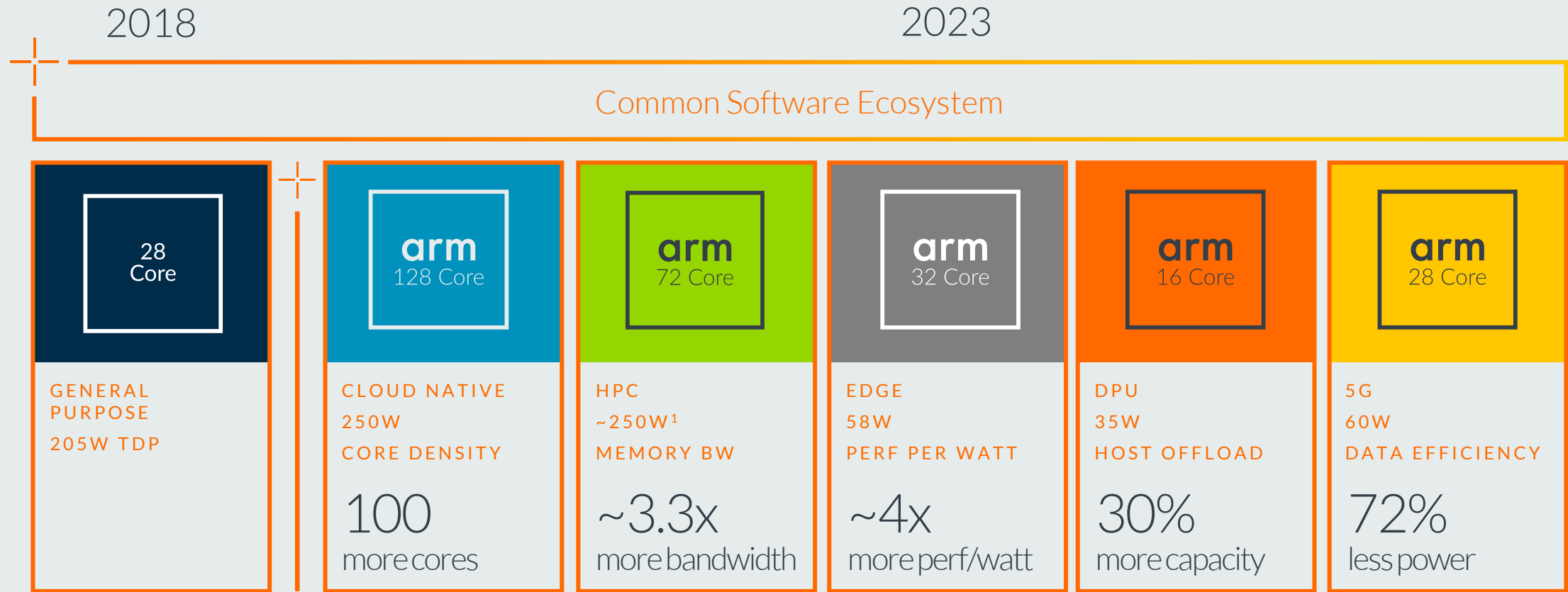


Sustainability
NET 0 BY 2050



Politics
LOCAL DEMANDS

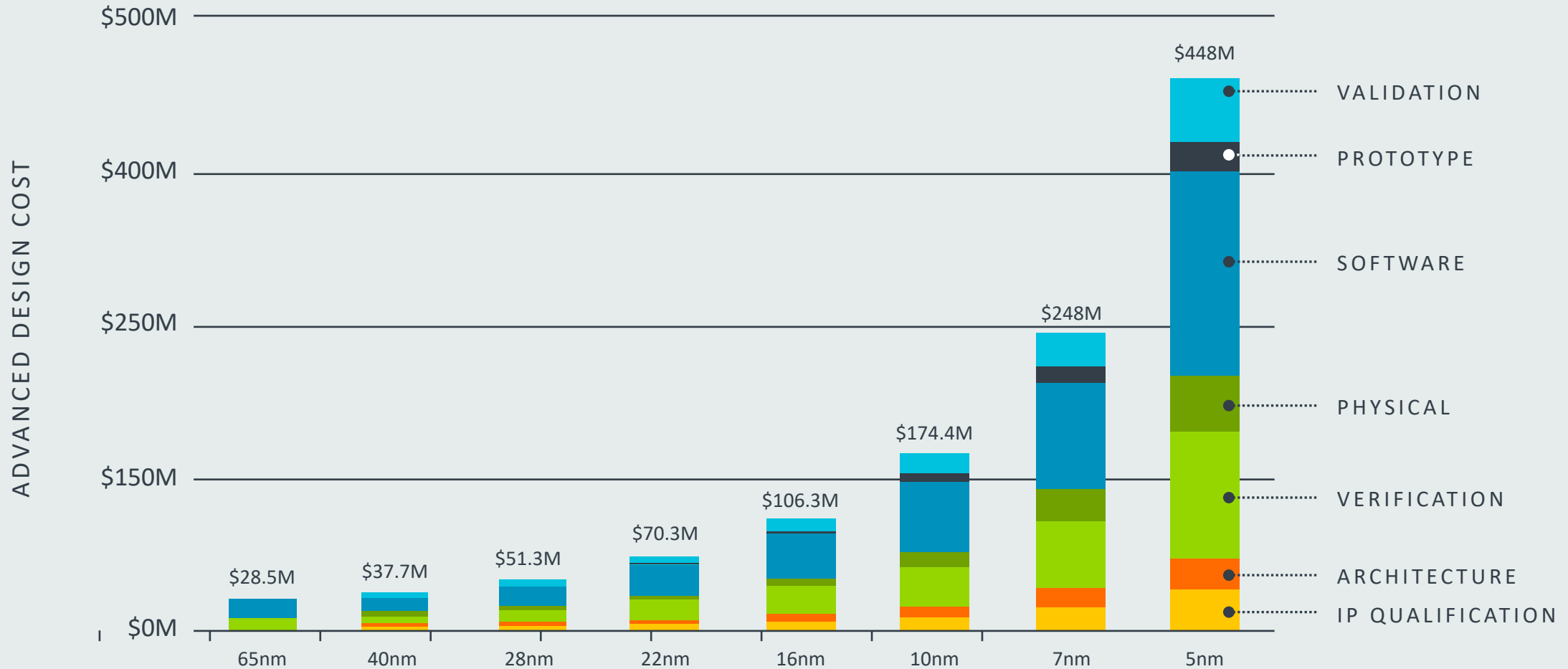
The Shift to Specialized Processing



Source: <http://www.specbench.org/cpu2017/results/res2018q2/cpu2017-20180612-06879.html>

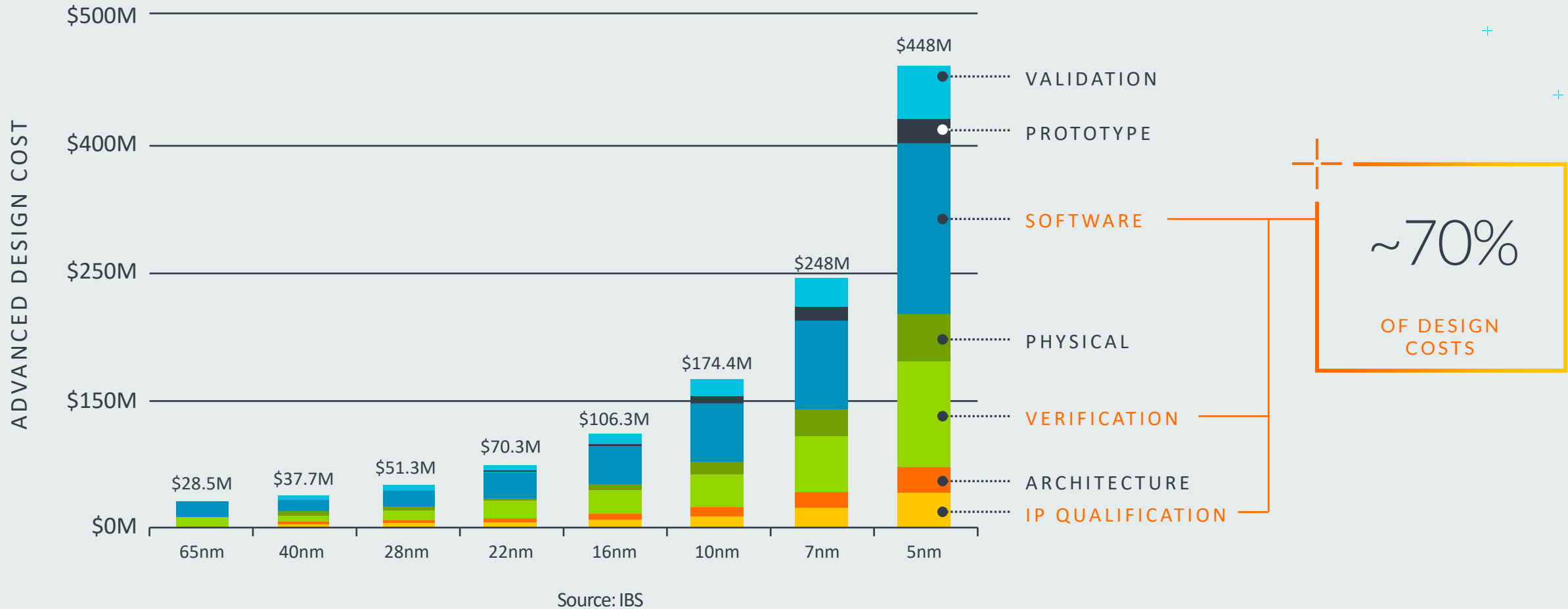
1. Per die, including LPDDR5 memory sub-system

Greater Demand Means Greater Costs & Complexity



Source: IBS

Greater Demand Means Greater Costs & Complexity





TECH LEADERSHIP



INNOVATION



EFFICIENCY

+

+

+

The Future of
Computing
Infrastructure



TECH LEADERSHIP



INNOVATION



EFFICIENCY

+
+
+
The Future of
Computing
Infrastructure

Technology Leadership: From CPU to Systems

+ CPU



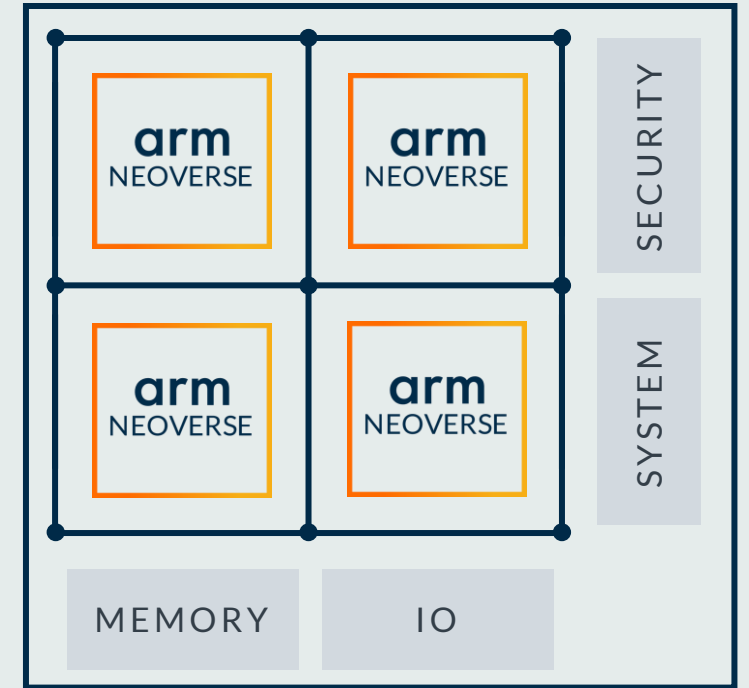
PERFORMANCE
PER WATT

+ INTERCONNECT



SCALE OUT CORE
CONNECTIVITY

+ COMPUTE SUB SYSTEMS



INTEGRATED &
VALIDATED
SYSTEMS

Technology Leadership: From CPU to Systems

+ CPU



PERFORMANCE
PER WATT

1.8x

More Vector
Workloads*

2x

Better Floating-
point Performance*



Scaling for high single-
thread performance



Built on Arm Scalable Vector
Extension (SVE) foundation

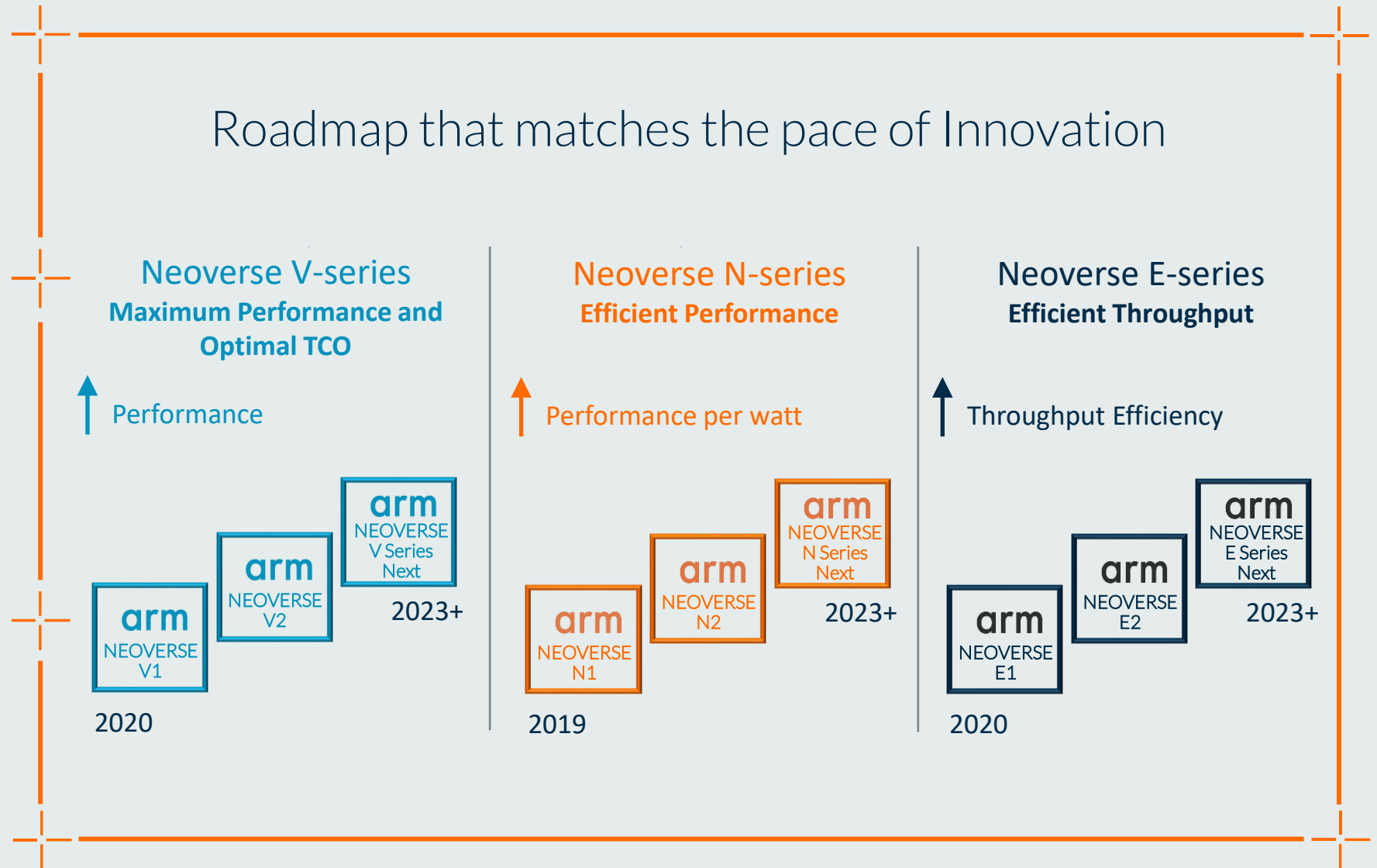
Technology Leadership: From CPU to Systems

+ CPU

Roadmap that matches the pace of Innovation

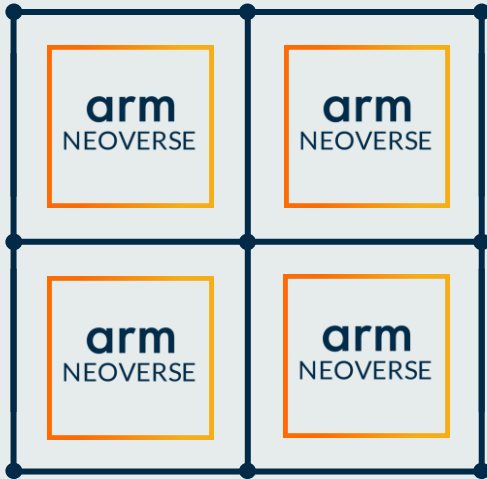


PERFORMANCE
PER WATT



Technology Leadership: From CPU to Systems

INTERCONNECT



SCALE OUT CORE
CONNECTIVITY

Custom core counts increase power efficiency



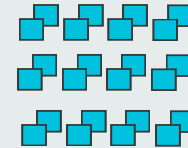
8-16 cores



20-35W Power



12-36 cores

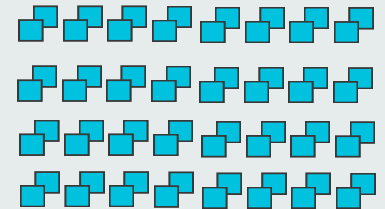


30-80W Power



High Performance
Computing

32-256 cores



80-350W Power

Arm Neoverse Core Scalability

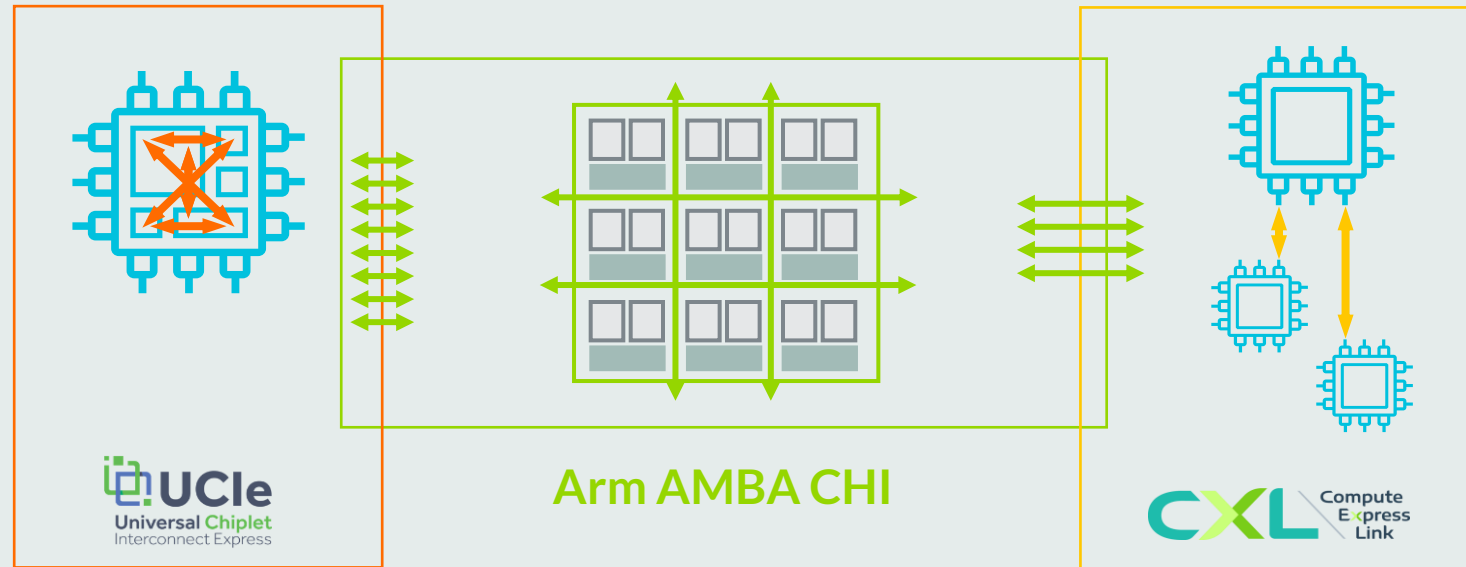
Technology Leadership: From CPU to Systems

+ INTERCONNECT



SCALE OUT CORE
CONNECTIVITY

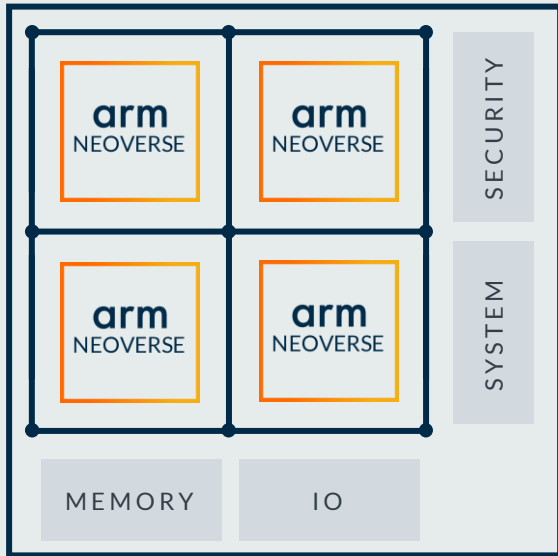
High performance and standard interconnect
across die and chip boundaries



Supports up to 4 Terabits/sec bandwidth

Technology Leadership: From CPU to Systems

COMPUTE SUB SYSTEMS



INTEGRATED & VALIDATED SYSTEMS

Integrated

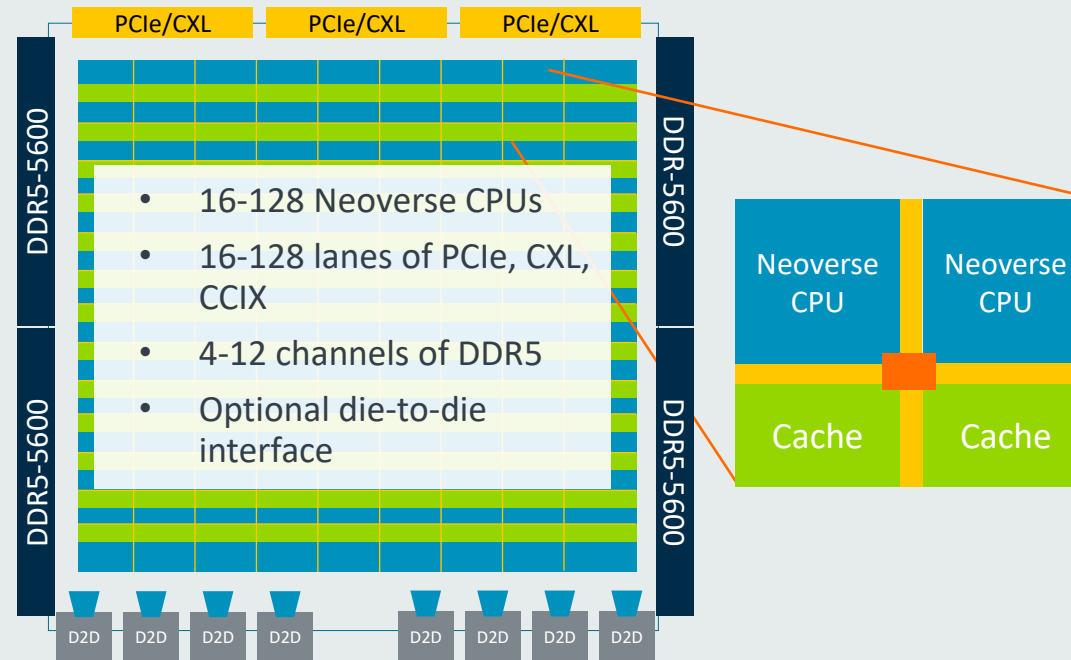
COMPUTE, INTERCONNECT, CACHE, PERIPHERALS, CONTROLLERS, INTERFACES

Configured

CORES, CACHE, IO

Validated

TO TARGET PROCESS NODE





TECH LEADERSHIP



INNOVATION



EFFICIENCY

+

+

+

The Future of
Computing
Infrastructure

Arm Neoverse Enabling Industry Firsts Innovation

FUJITSU



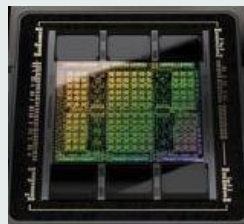
FIRST
1TB/S MEM BW

AMPERE



FIRST
>100 CORES PER CPU

NVIDIA



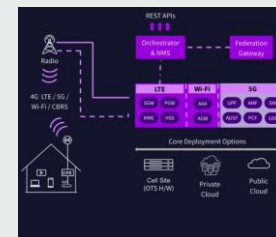
FIRST
LPDDR5X MEM

MARVELL



FIRST
5nm DPUs

CONNECT 5G



FIRST
OPEN SOURCE 5G EPC

FUGAKU



FIRST
>500 PFLOPS/S

AWS



FIRST
DDR5 PCIE GEN5.0

ALIBABA



FIRST
>500 ON SPECINT

NVIDIA



FIRST
DPU WITH >400GB/S

SOLIDRUN



FIRST
SW DEFINED DPU

SUNSEA



FIRST
CLOUD BASE STATION

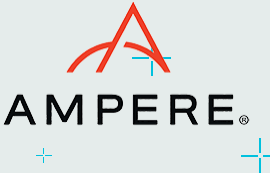
INNOVIUM



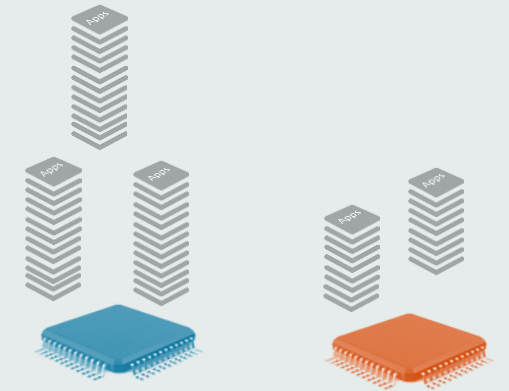
FIRST
SWITCH @ 25.6TBPS

Scalable Efficiency Through High Core Counts

Case Study from Ampere Computing – An Arm Silicon Partner



SoC Level

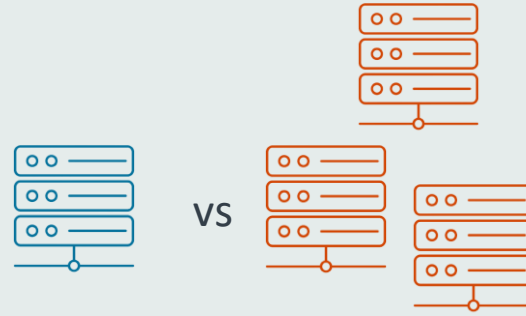


128 Cores

64 Cores

>2x more workloads

Rack Level

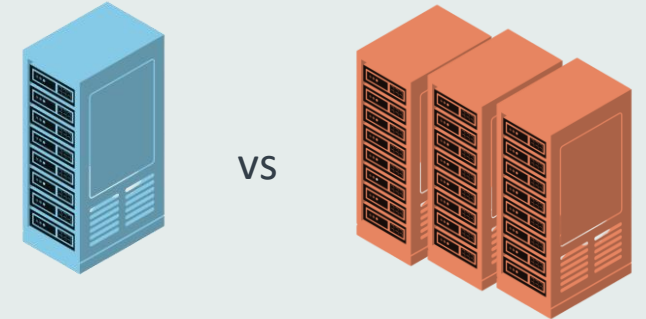


32 CPUs

82 CPUs

2.5x greater perf than legacy

Data Center Level



1 Rack

3 Racks

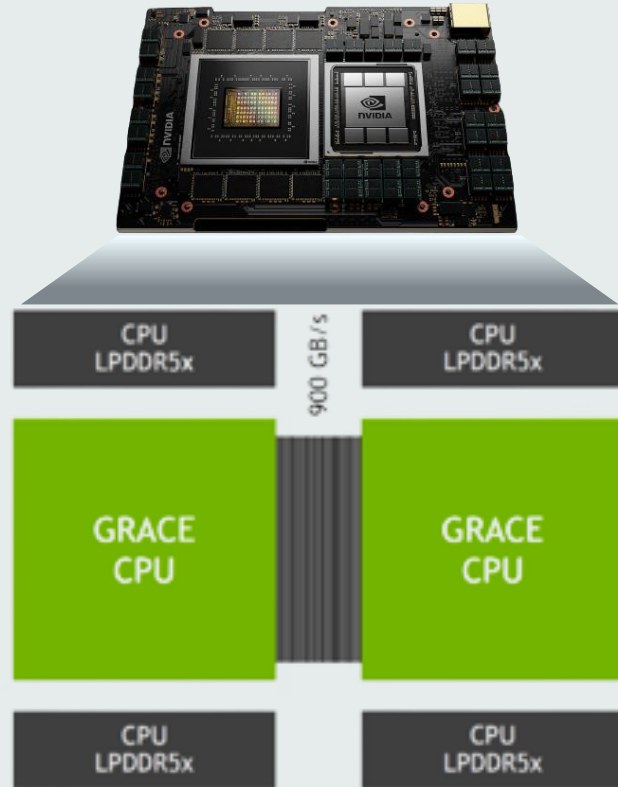
2/3rd less footprint per datacenter

■ Ampere(Arm-based)

■ Legacy architectures

Scalable Efficiency Through SoC Level Innovation

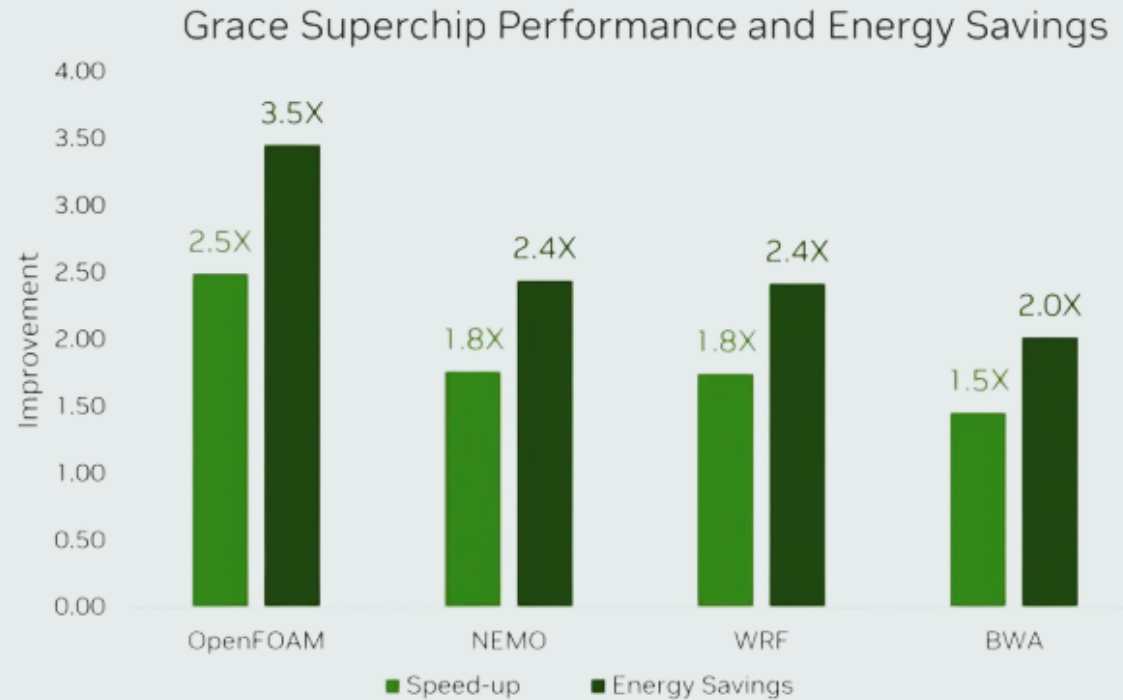
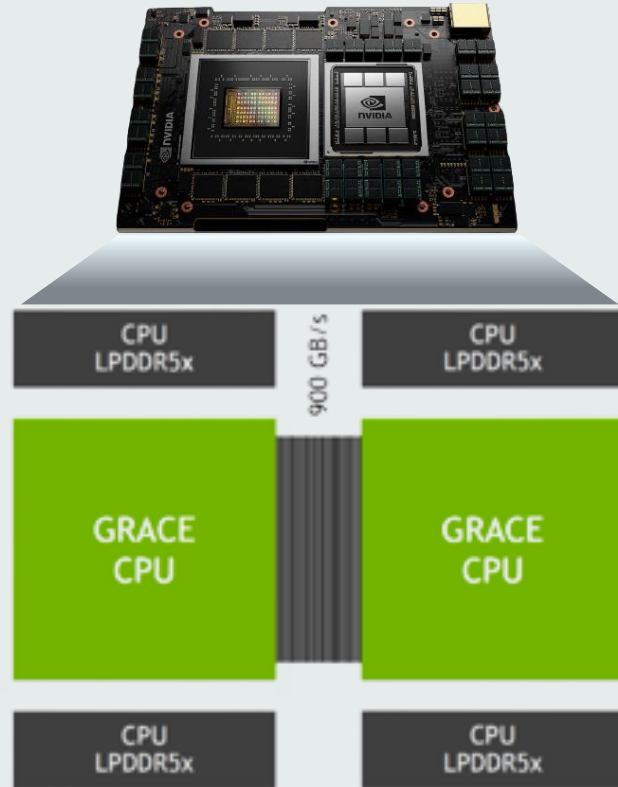
NVIDIA Grace's groundbreaking performance with system level interconnects



- + CPUs and GPUs inside the chips are linked with NVIDIA's NVLink-C2C (Supporting Arm's AMBA CHI)
- + Deliver up to a 10x performance leap in AI's most demanding tasks
- + 144 Arm Neoverse V2 cores with up to 960GB of LPDDR5x memory and 1 TB/sec of memory bandwidth

Scalable Efficiency Through SoC Level Innovation

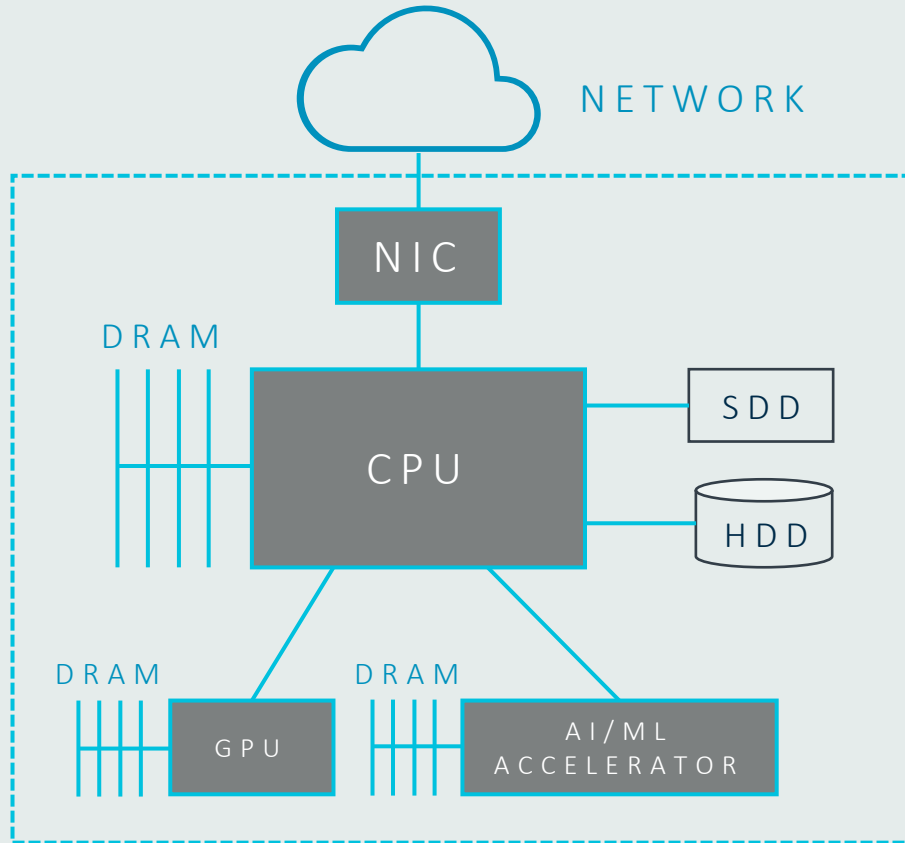
NVIDIA Grace's groundbreaking performance with system level interconnects



DPU's Are the Heart of Data-Centric Computing

CPU's, GPU's, FPGA's and storage become composable assets

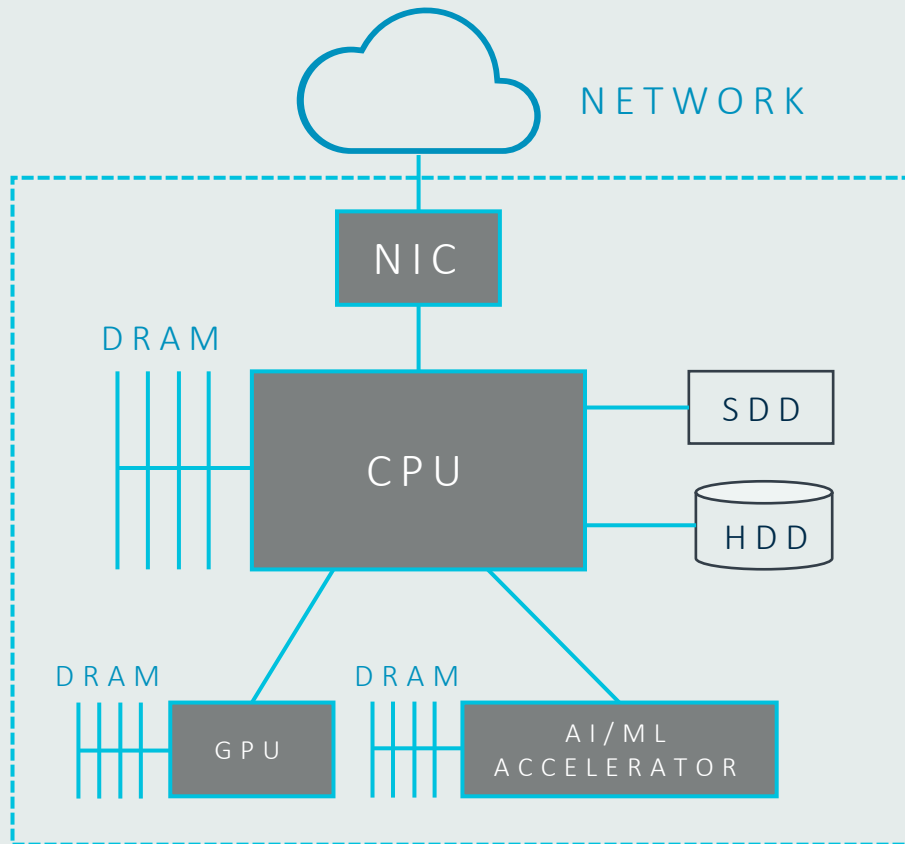
+ COMPUTE CENTRIC ARCHITECTURE



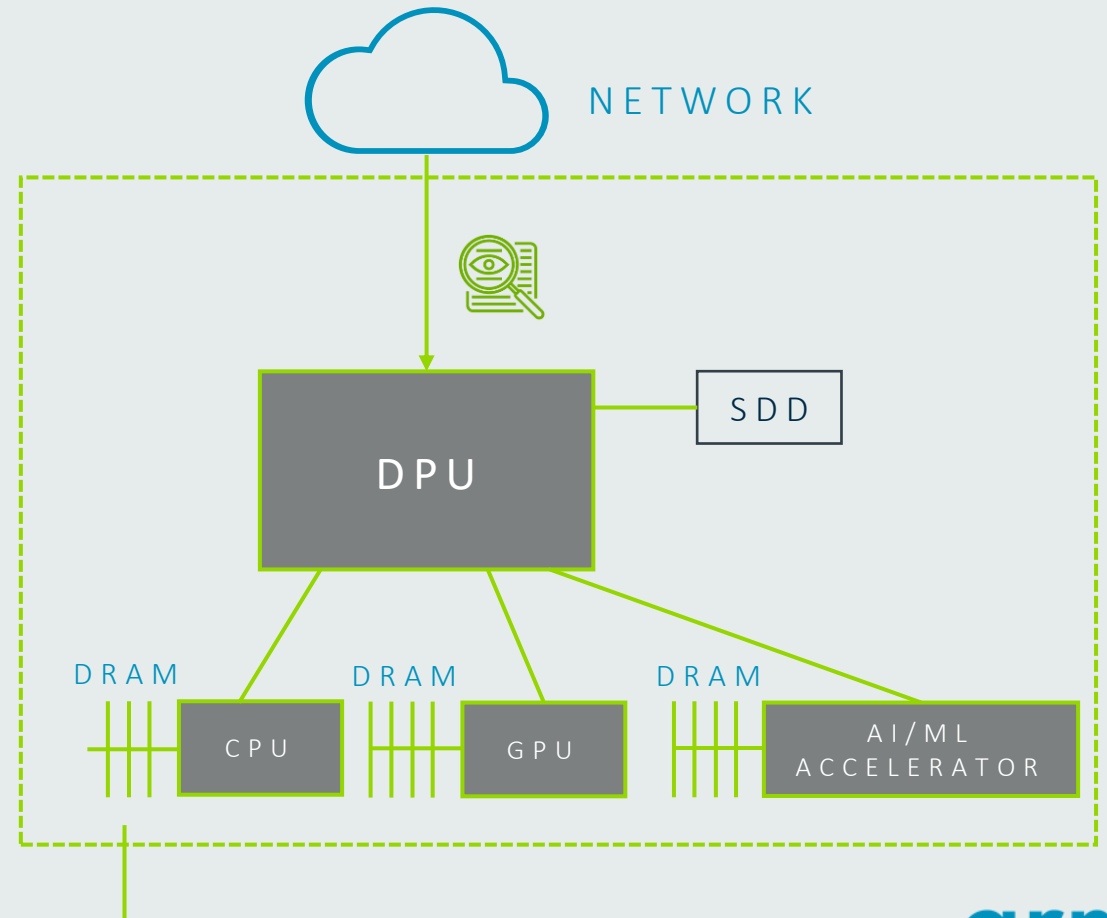
DPU's Are the Heart of Data-Centric Computing

CPU's, GPU's, FPGA's and storage become composable assets

+ COMPUTE CENTRIC ARCHITECTURE



+ DATA CENTRIC ARCHITECTURE

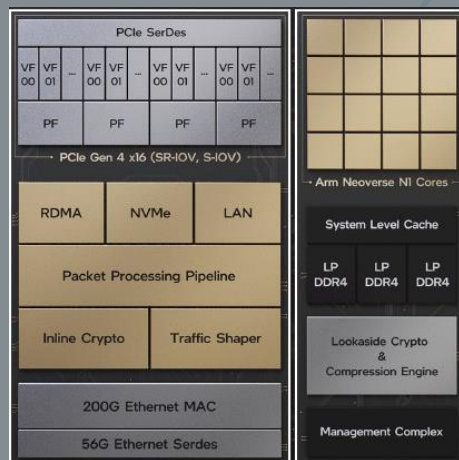


DPU Are the Heart of Data-Centric Computing

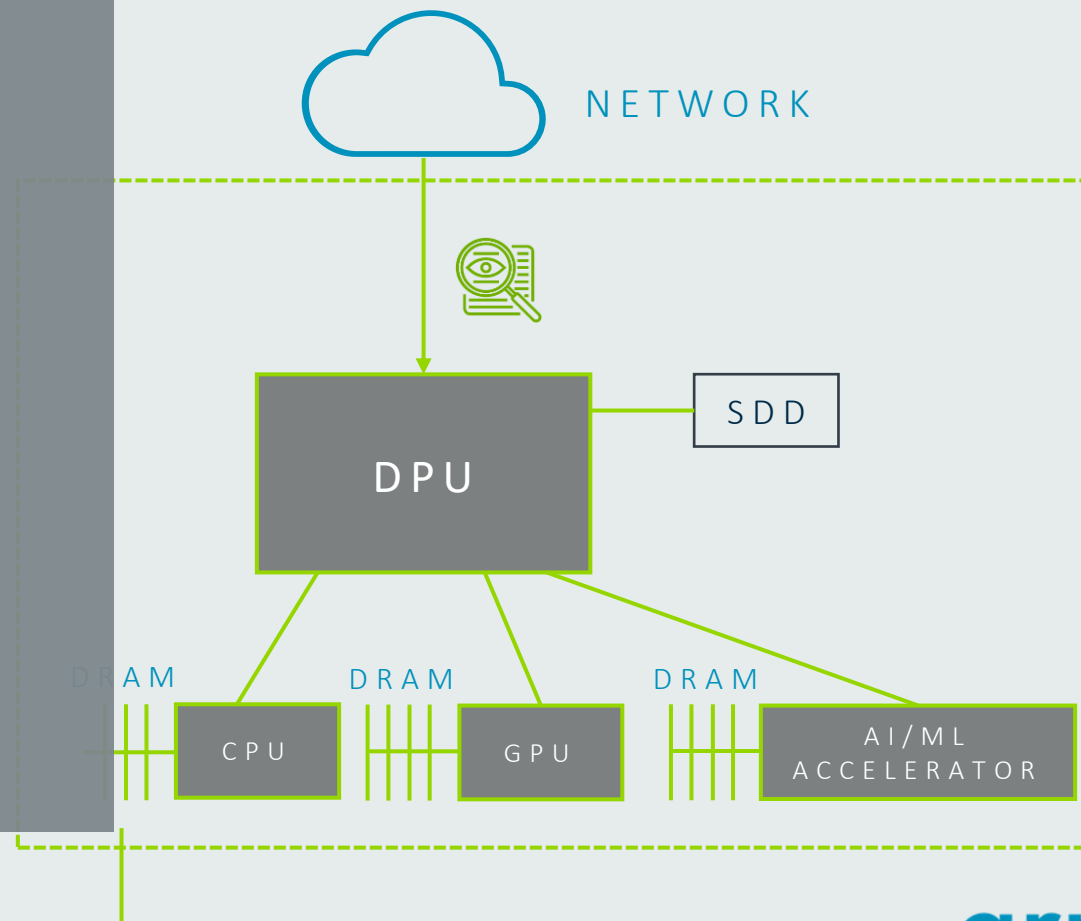
CPU, GPU, FPGA and storage become composable assets

COMPUTE CENTRIC ARCHITECTURE

DATA CENTRIC ARCHITECTURE



Intel Mt Evans DPU deployed by Google utilizes Arm Neoverse cores and can handle data transfers of up to **200 million packets/second**



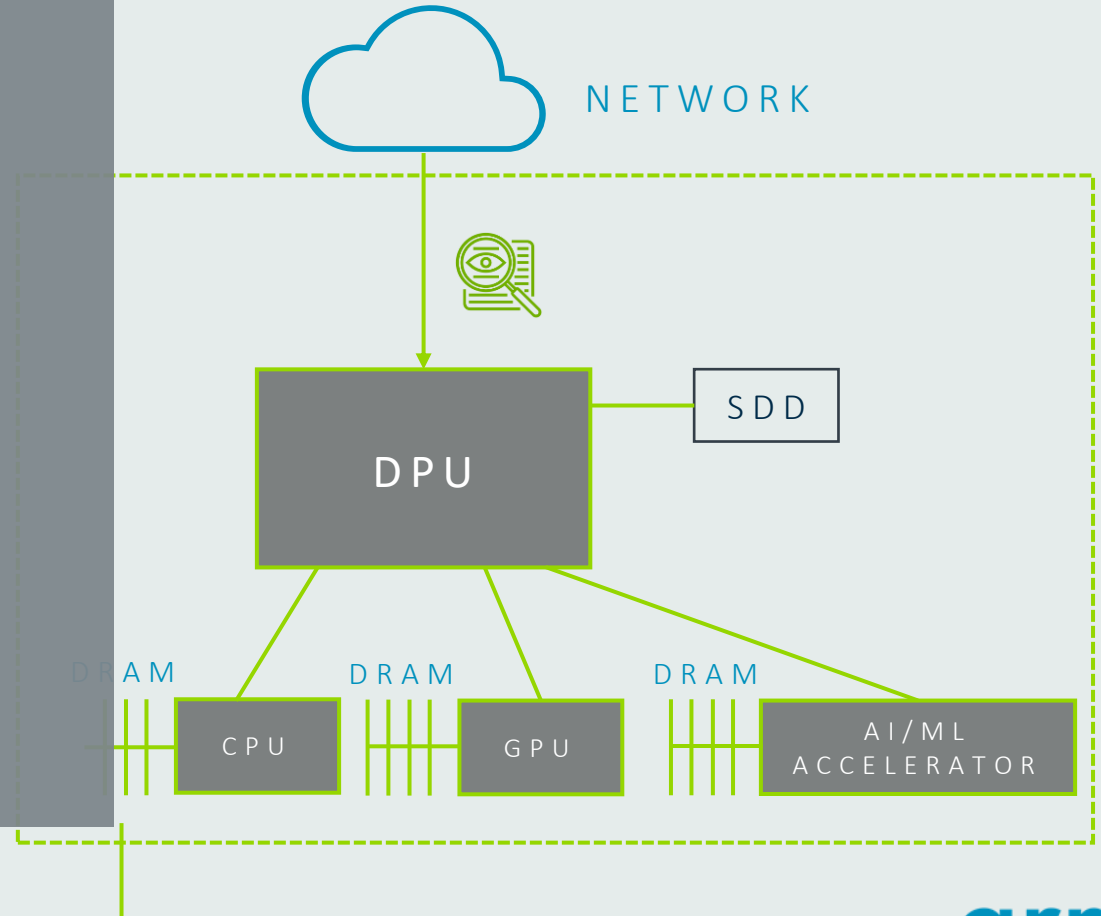
DPU's Are the Heart of Data-Centric Computing

CPU, GPU, FPGA and storage become composable assets



AWS Nitro, powered by Arm Neoverse, delivers 40% better performance/watt

DATA CENTRIC ARCHITECTURE





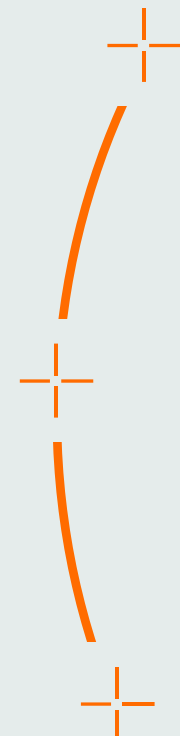
TECH LEADERSHIP



INNOVATION



EFFICIENCY

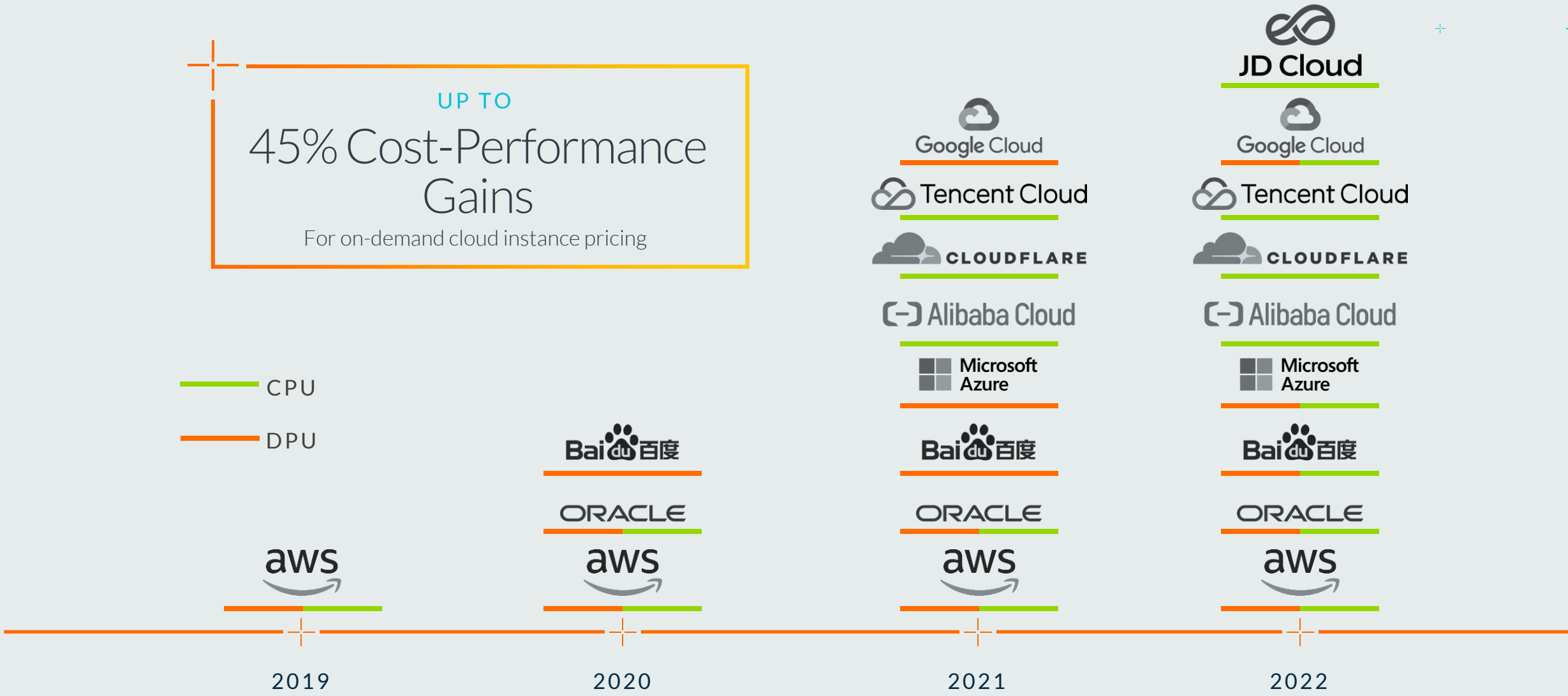


The Future of
Computing
Infrastructure

Arm Neoverse Powered Clouds: Value to the End User

UP TO
45% Cost-Performance Gains
 For on-demand cloud instance pricing

— CPU
 — DPU



Arm Software Ecosystem



Arm Neoverse Sustainability Impact

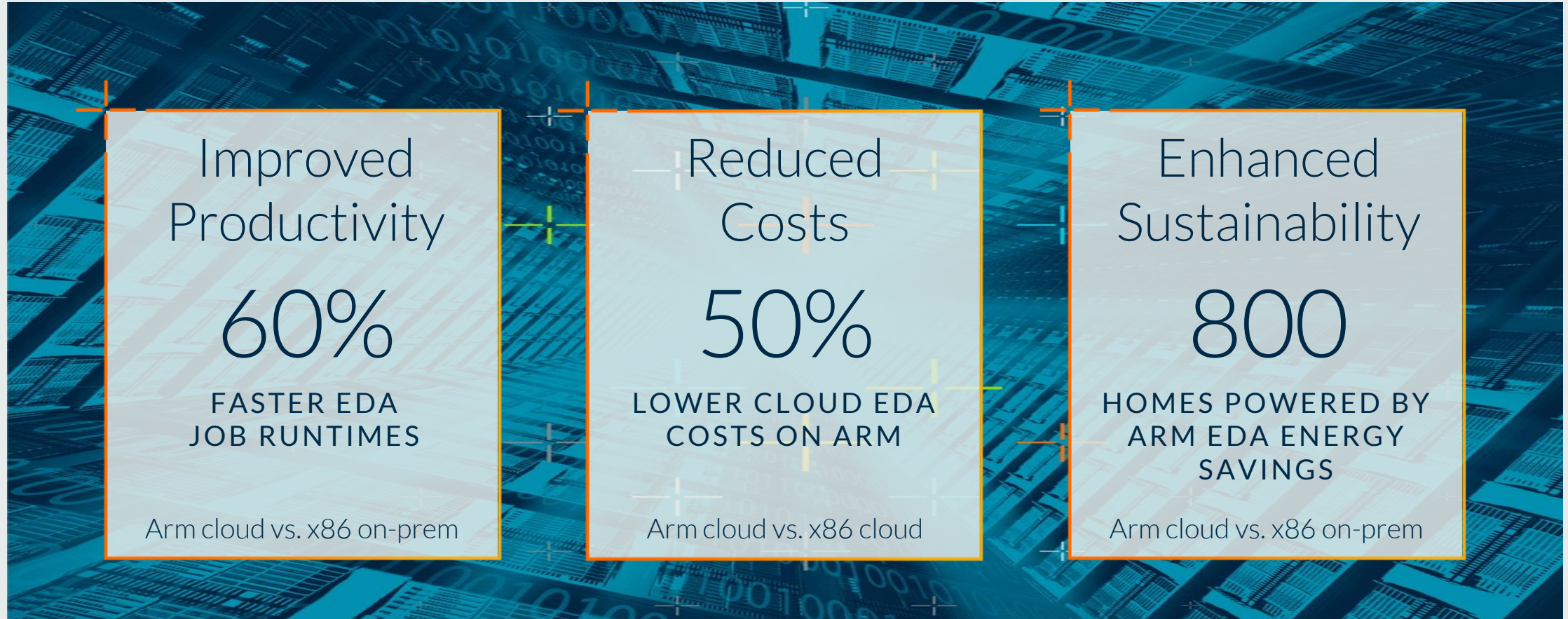
Case Study from Ampere Computing – An Arm Silicon Partner



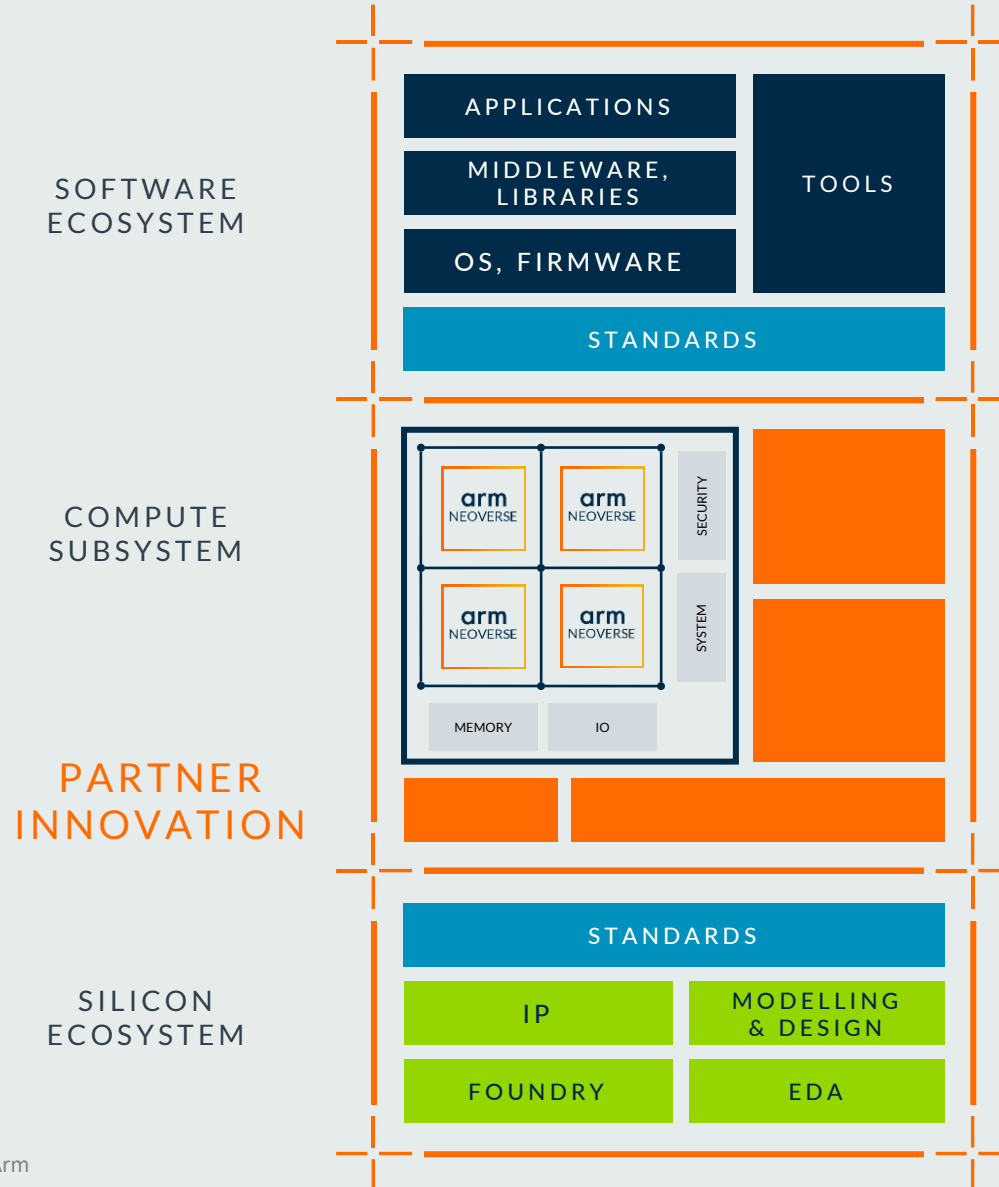
66MW DATA CENTER
(Average)



Running Arm EDA in the Cloud...on Arm



The Future of Computing Infrastructure



Technology Leadership

CPU PERFORMANCE & EFFICIENCY

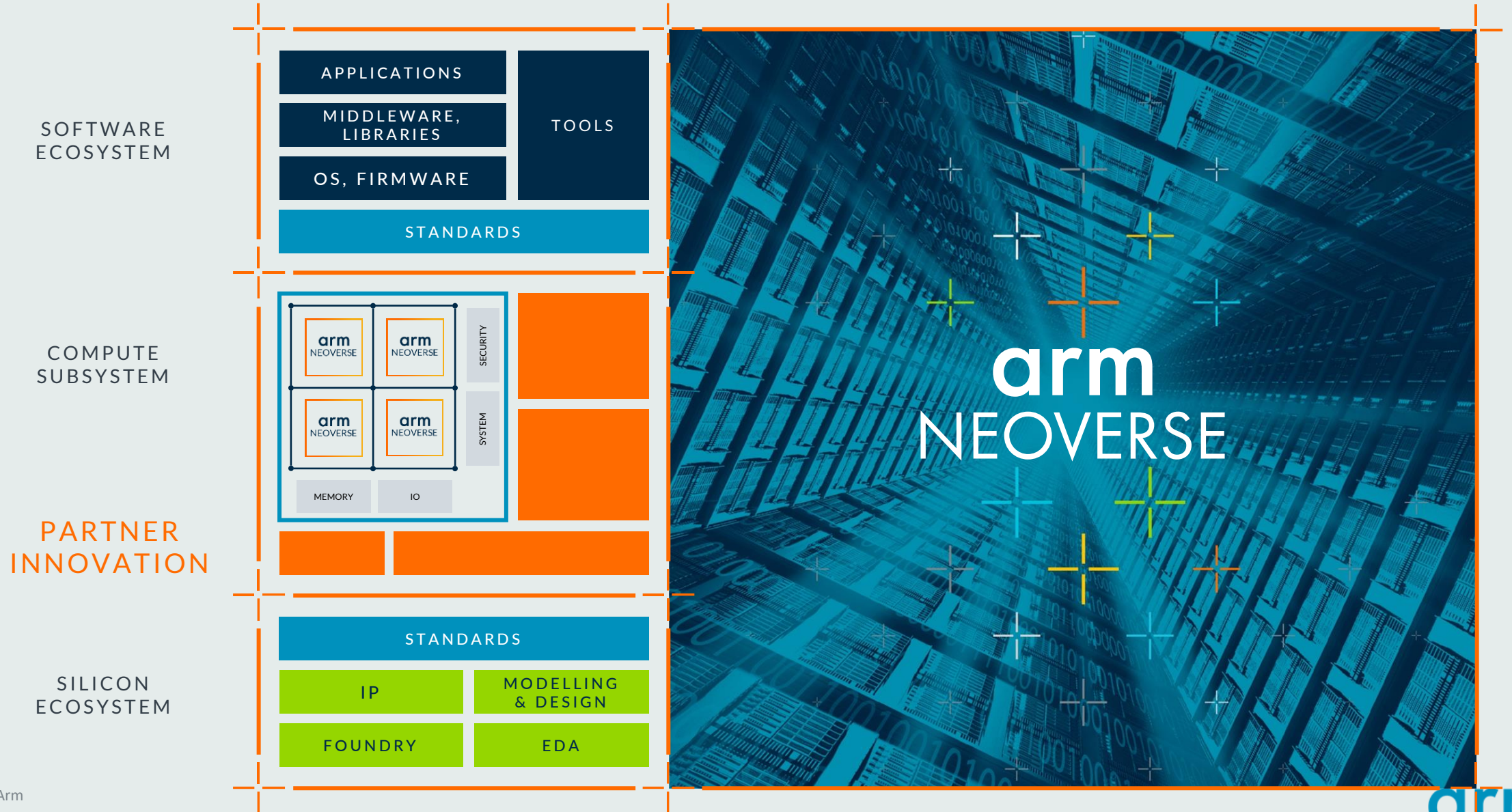
Freedom to Innovate

MAXIMIZING PARTNER LEVERAGE

Scalable Efficiency

SUSTAINABLE COMPUTING AT EVERY POINT

The Future of Computing Infrastructure



arm

Thank You

Danke

Gracias

Grazie

谢谢

ありがとう

Asante

Merci

감사합니다

धन्यवाद

Kiitos

شكرًا

ধন্যবাদ

תודה



The Arm trademarks featured in this presentation are registered trademarks or trademarks of Arm Limited (or its subsidiaries) in the US and/or elsewhere. All rights reserved. All other marks featured may be trademarks of their respective owners.

www.arm.com/company/policies/trademarks