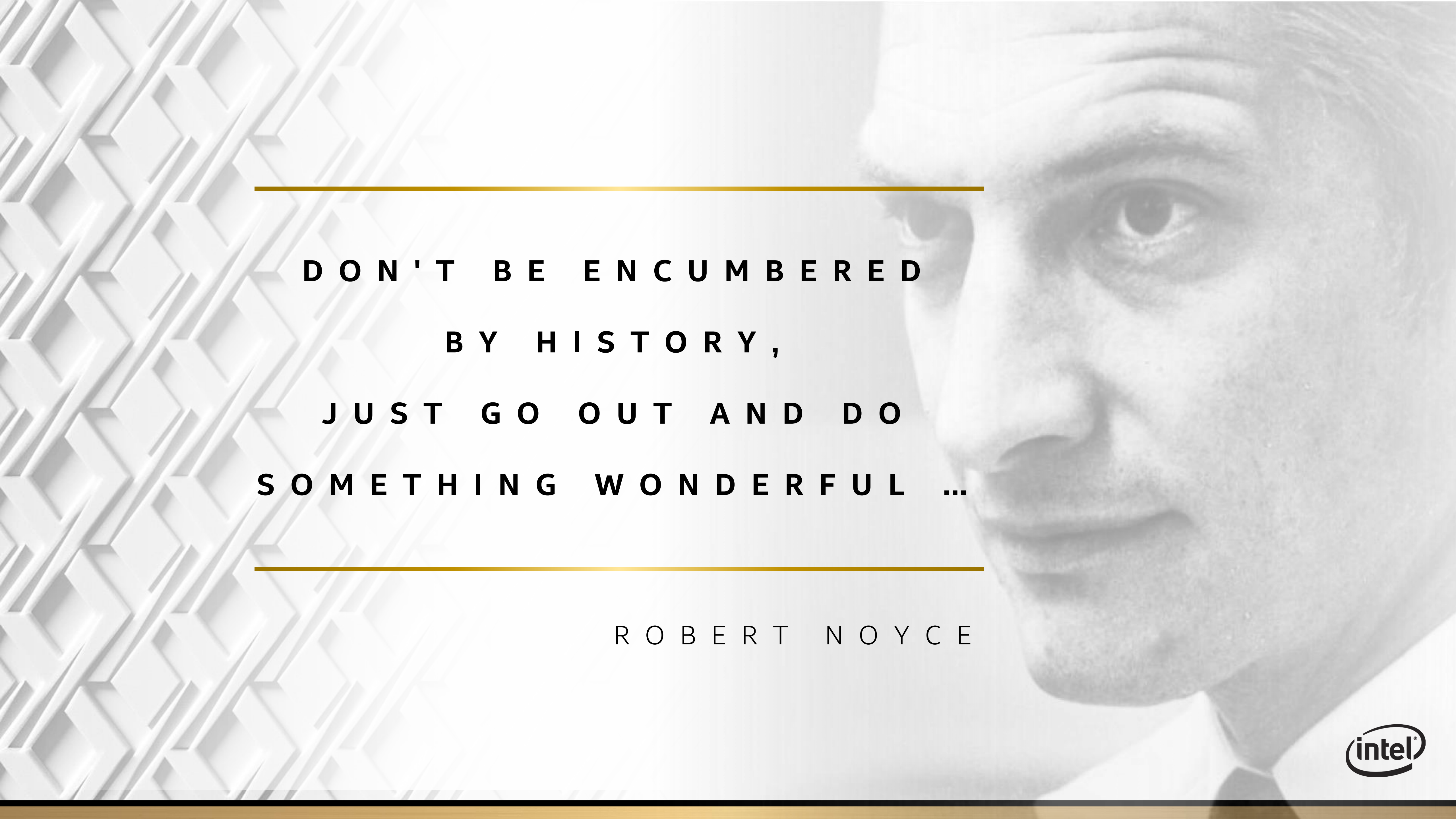
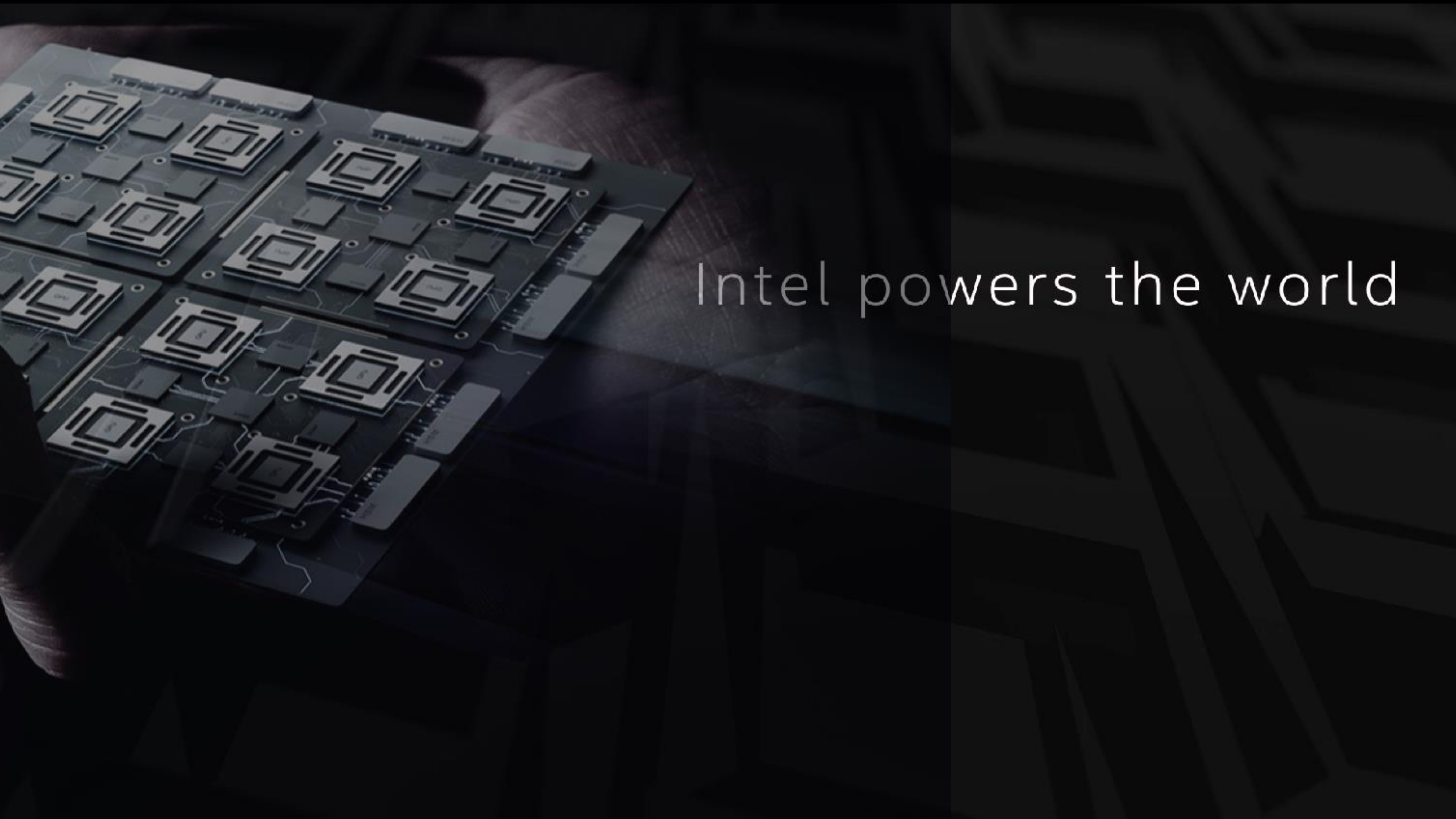




D O N ' T B E E N C U M B E R E D
B Y H I S T O R Y ,
J U S T G O O U T A N D D O
S O M E T H I N G W O N D E R F U L ...

R O B E R T N O Y C E



A close-up, low-angle shot of a hand holding a dark-colored circuit board. The board is populated with several square Intel processors, each with a distinct square pad in the center. The background is dark and out of focus, showing the keys of a laptop keyboard. The overall lighting is dim, creating a professional and technical atmosphere.

Intel powers the world



EXASCALE FOR EVERYONE

R A J A K O D U R I



1ST ERA IN HPC

VERTICALLY INTEGRATED SYSTEMS

PROPRIETARY HARDWARE
AND SOFTWARE

PARAGON
143 GFLOP

DELTA
8.1 GFLOP

OF HPC SYSTEMS

10,000,000's

1,000,000's

100,000's

1000's

10's

1970

1980

1990

2000

2010

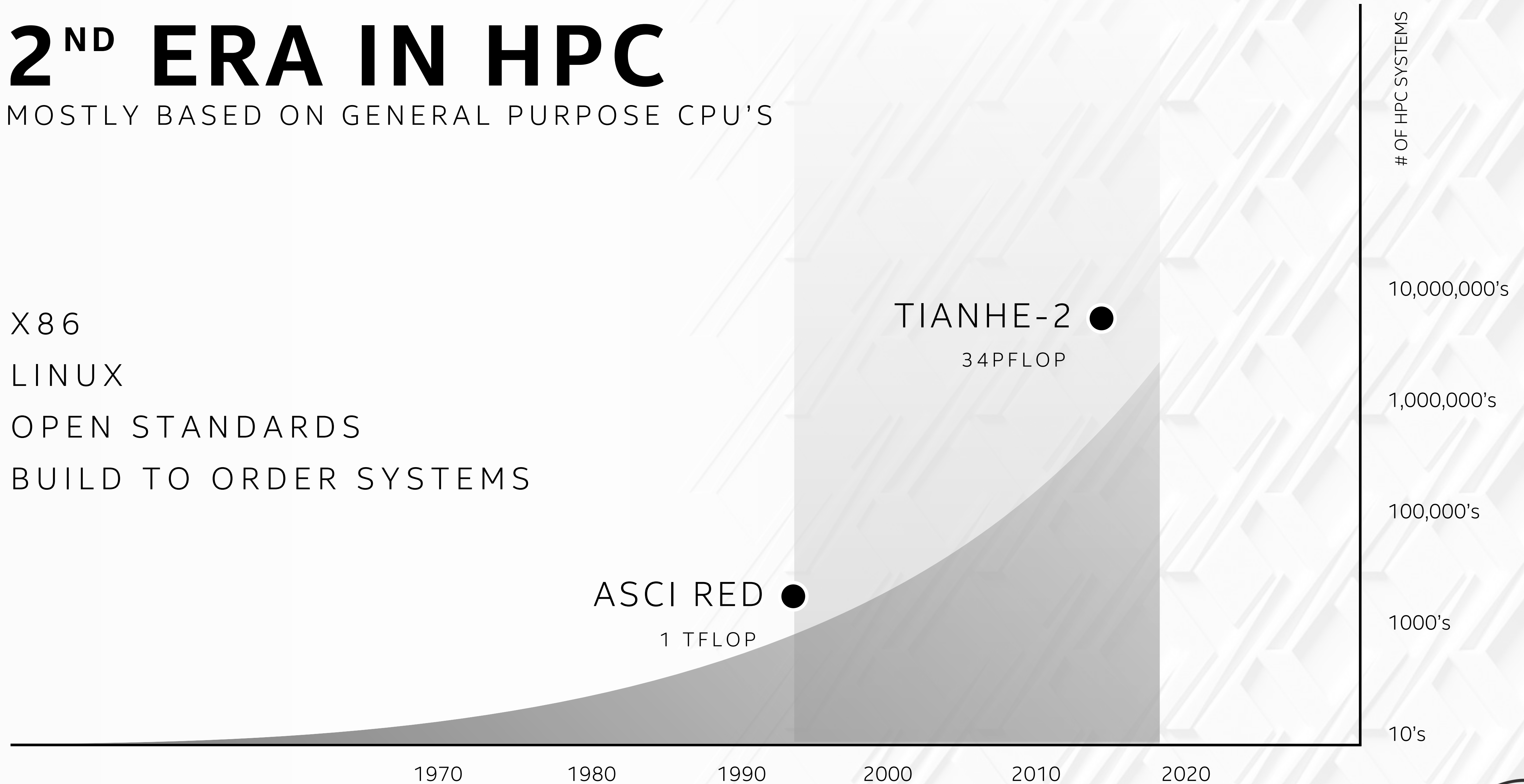
2020



2ND ERA IN HPC

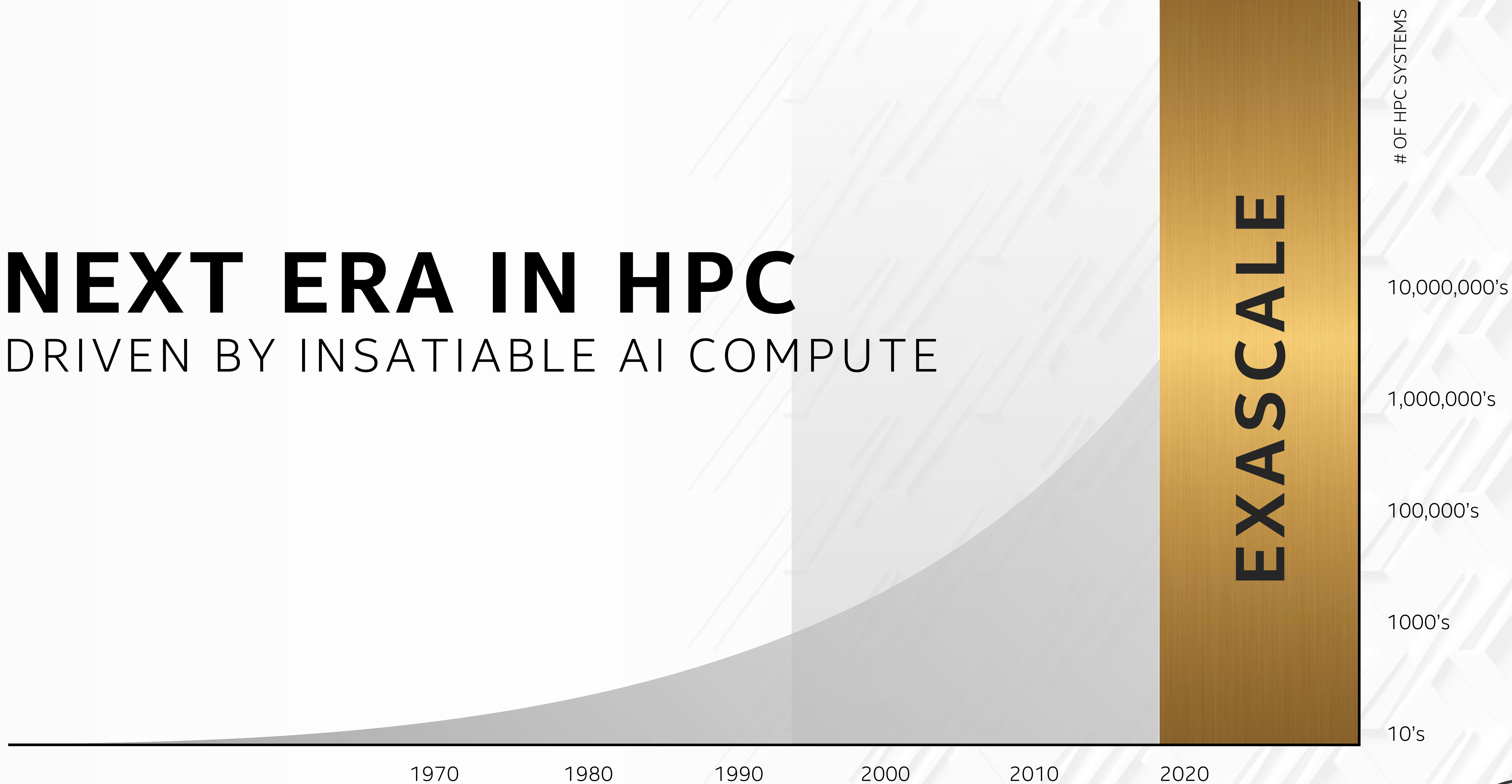
MOSTLY BASED ON GENERAL PURPOSE CPU'S

X86
LINUX
OPEN STANDARDS
BUILD TO ORDER SYSTEMS



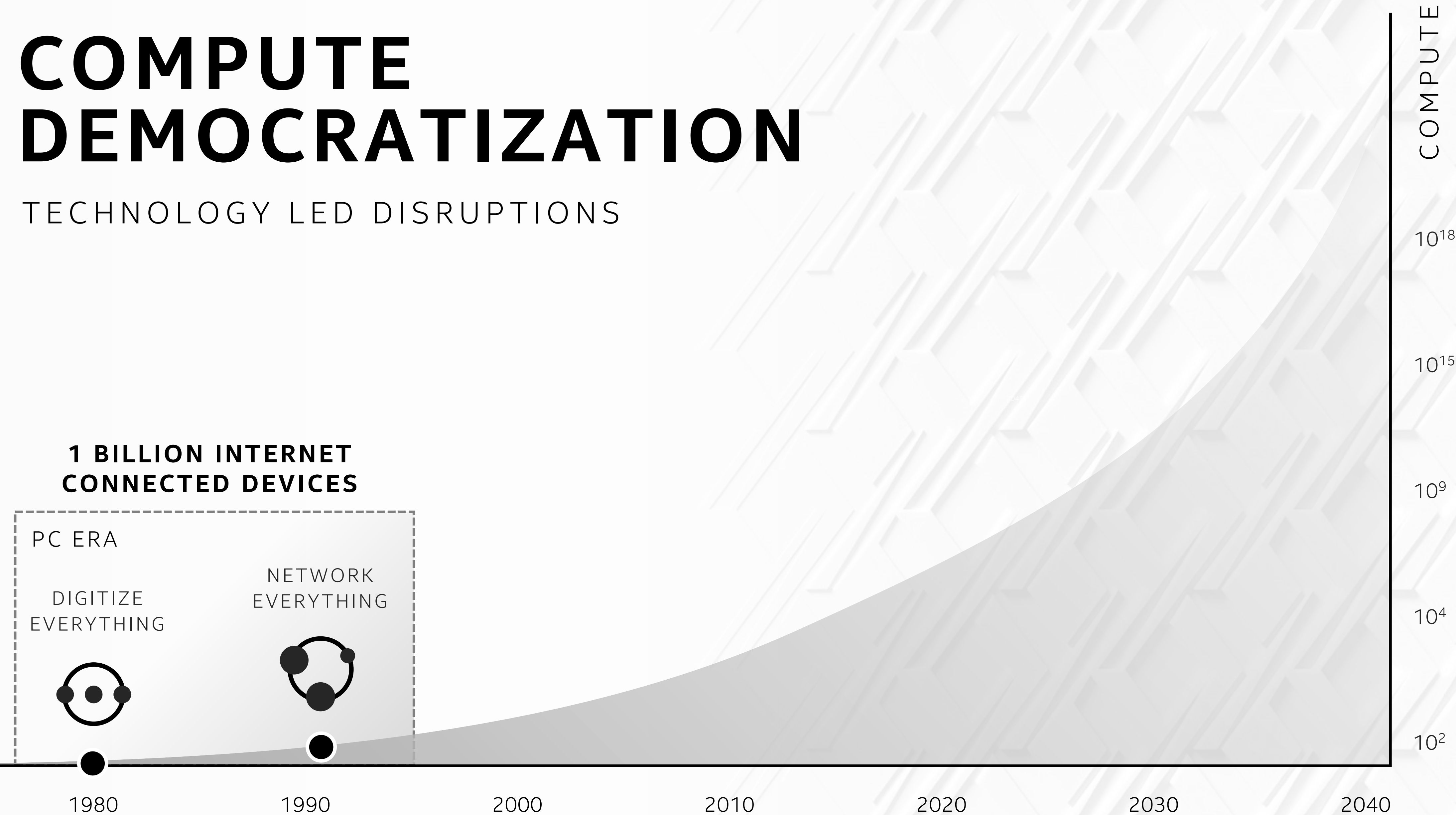
NEXT ERA IN HPC

DRIVEN BY INSATIABLE AI COMPUTE



COMPUTE DEMOCRATIZATION

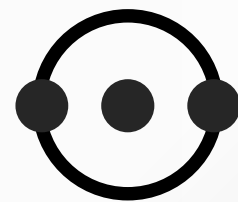
TECHNOLOGY LED DISRUPTIONS



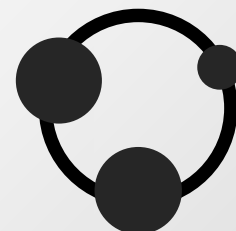
**1 BILLION INTERNET
CONNECTED DEVICES**

PC ERA

DIGITIZE
EVERYTHING

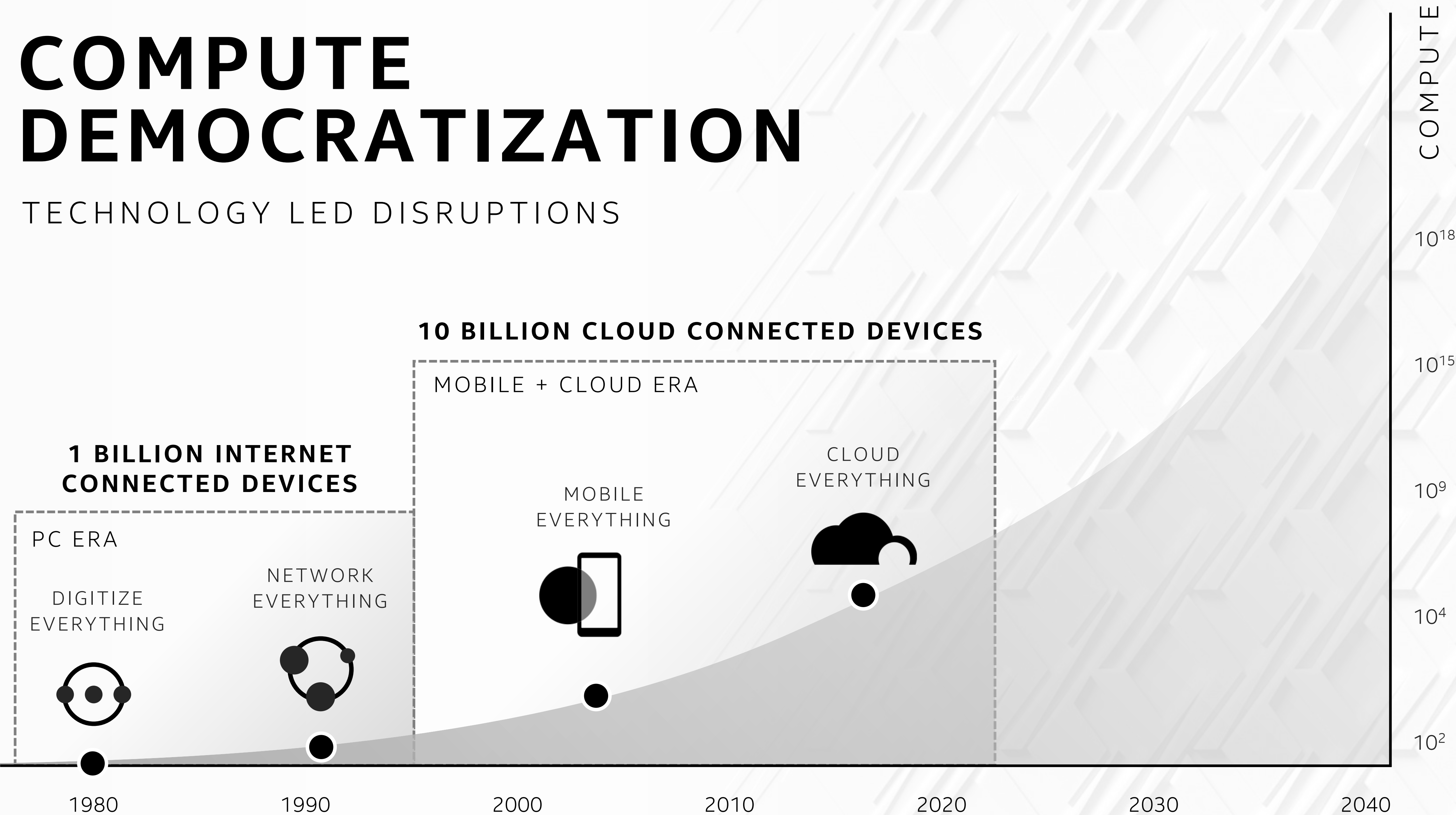


NETWORK
EVERYTHING



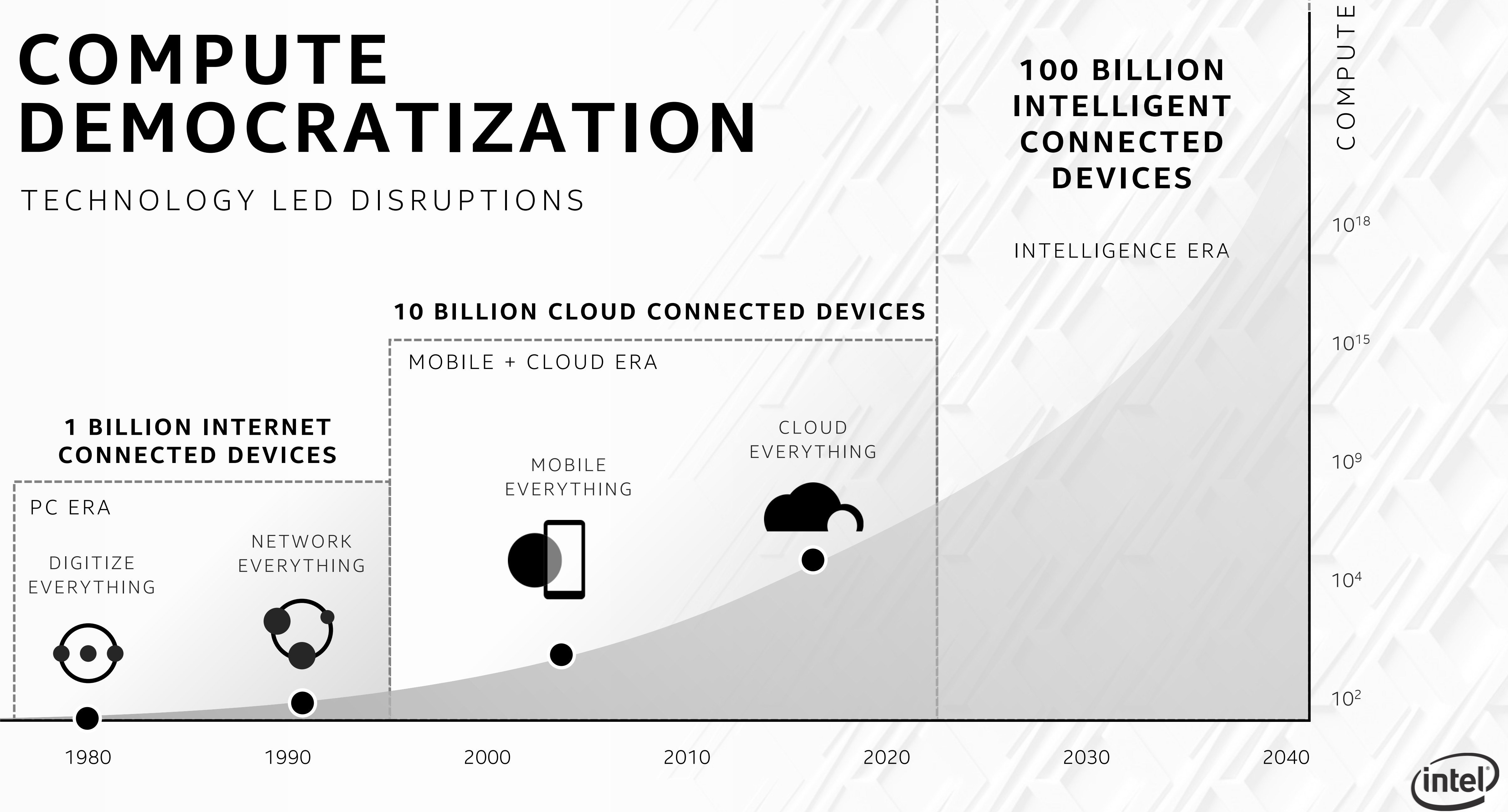
COMPUTE DEMOCRATIZATION

TECHNOLOGY LED DISRUPTIONS



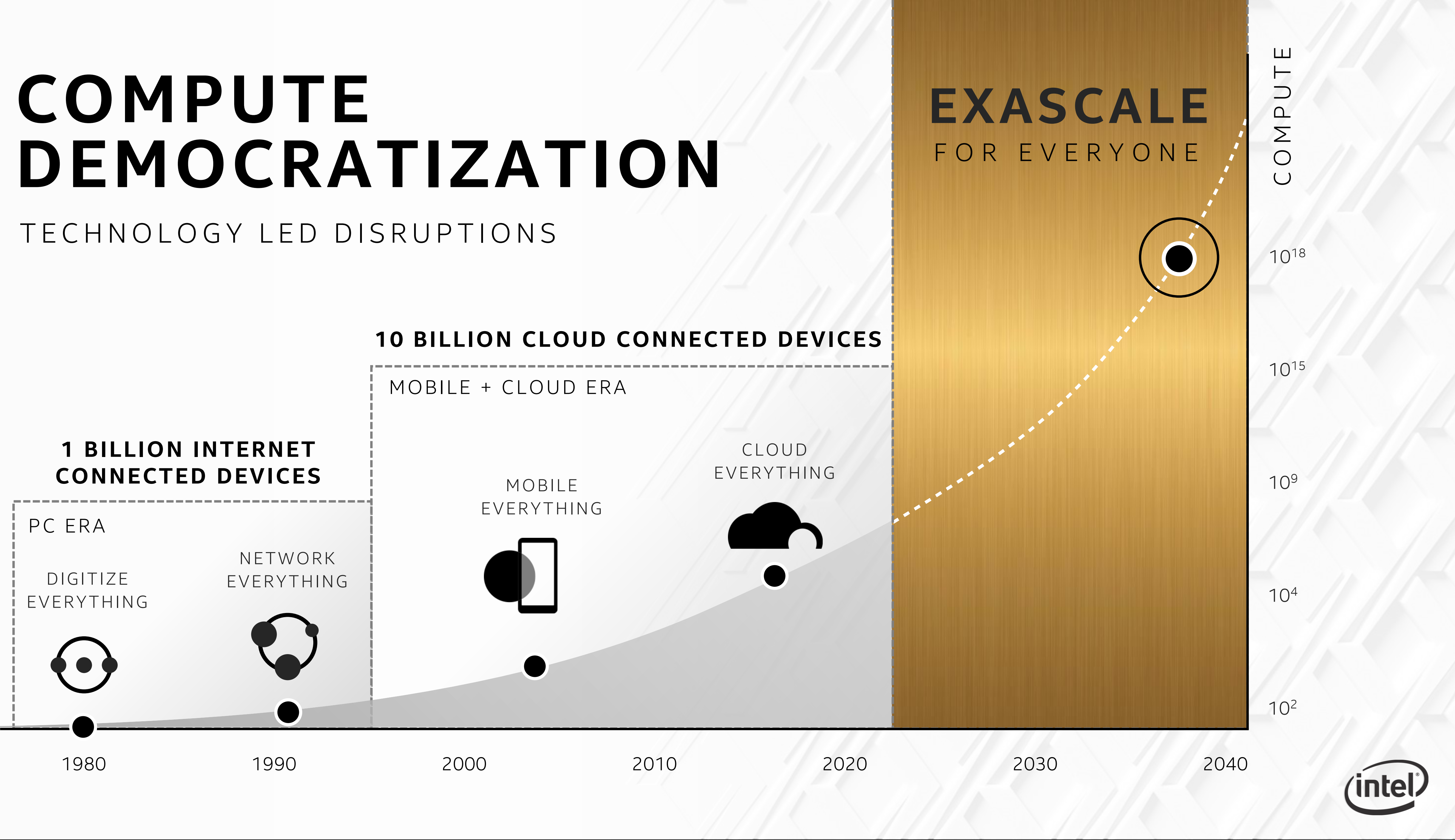
COMPUTE DEMOCRATIZATION

TECHNOLOGY LED DISRUPTIONS



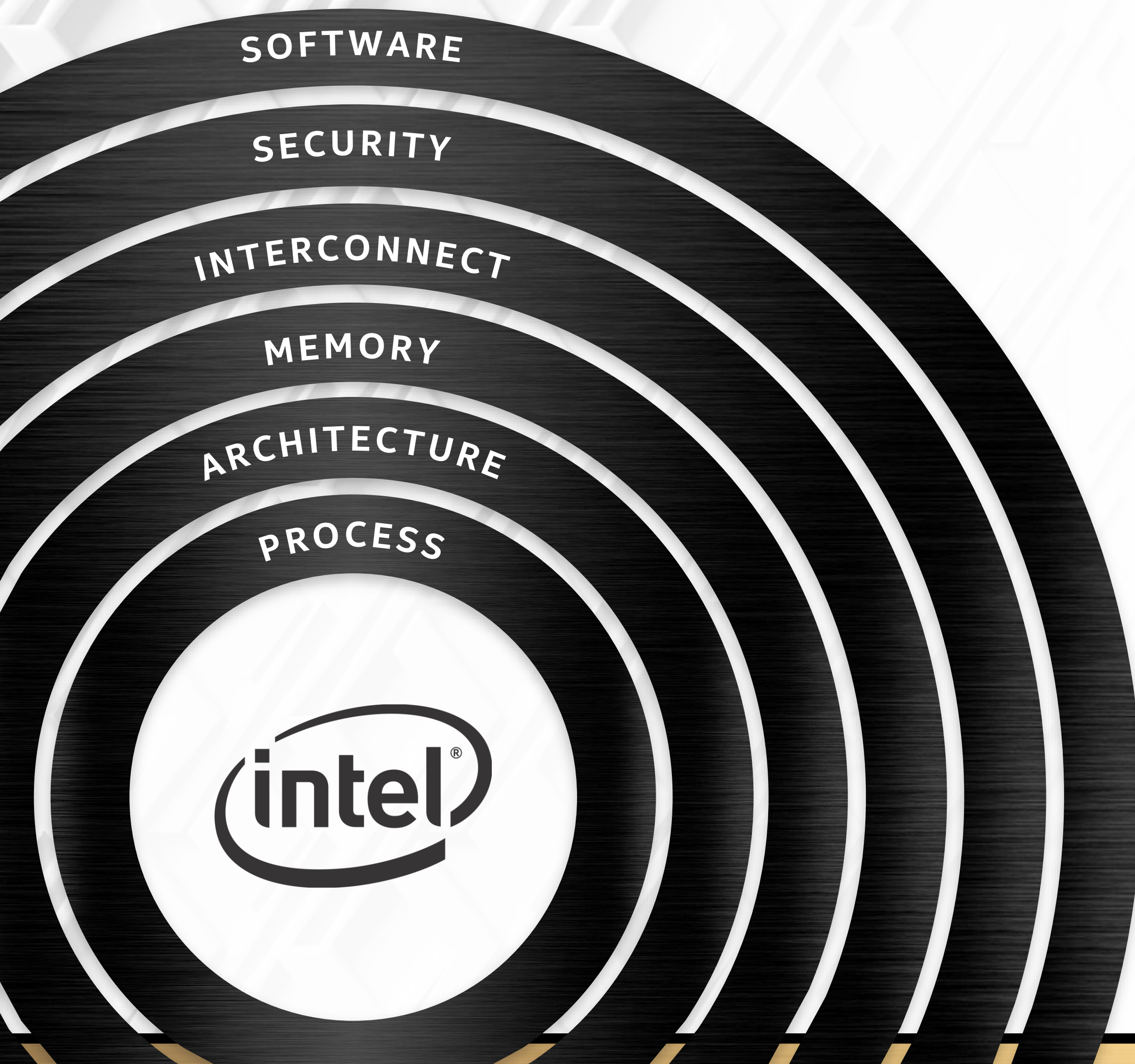
COMPUTE DEMOCRATIZATION

TECHNOLOGY LED DISRUPTIONS



X^e **EXASCALE FOR EVERYONE**

INTEL ARCHITECTURE DAY DEC'18



Intel will deliver a diverse mix of **Scalar, Vector, Matrix and Spatial Architectures** designed with state of the art **Process technology**, fed by disruptive **Memory** hierarchies, integrated into systems with advanced **Packaging**, deployed at hyperscale with lightspeed **Interconnect** links, unified by a **single Software** abstraction, with benchmark defining Security features

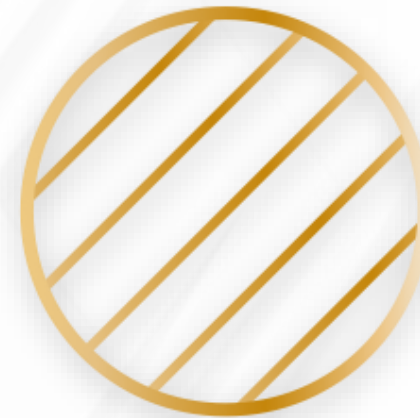


HETEROGENEOUS TAXONOMY

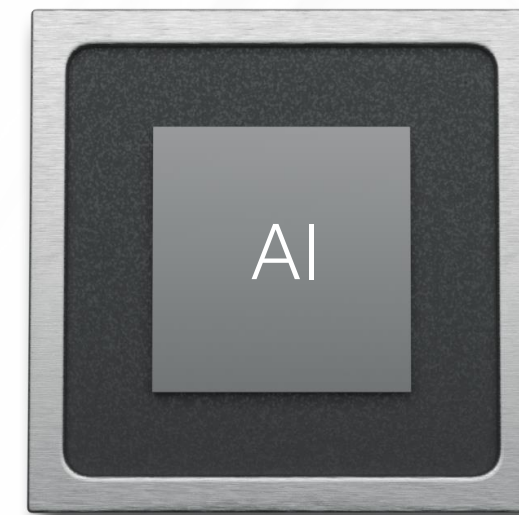
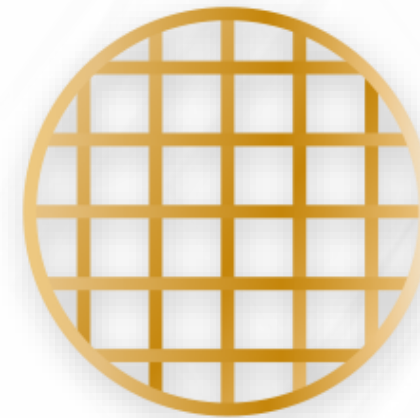
SCALAR



VECTOR



MATRIX



SPATIAL



**ARCHITECTURE IMPACT =
PERFORMANCE X GENERALITY**

ARCHITECTURE IMPACT

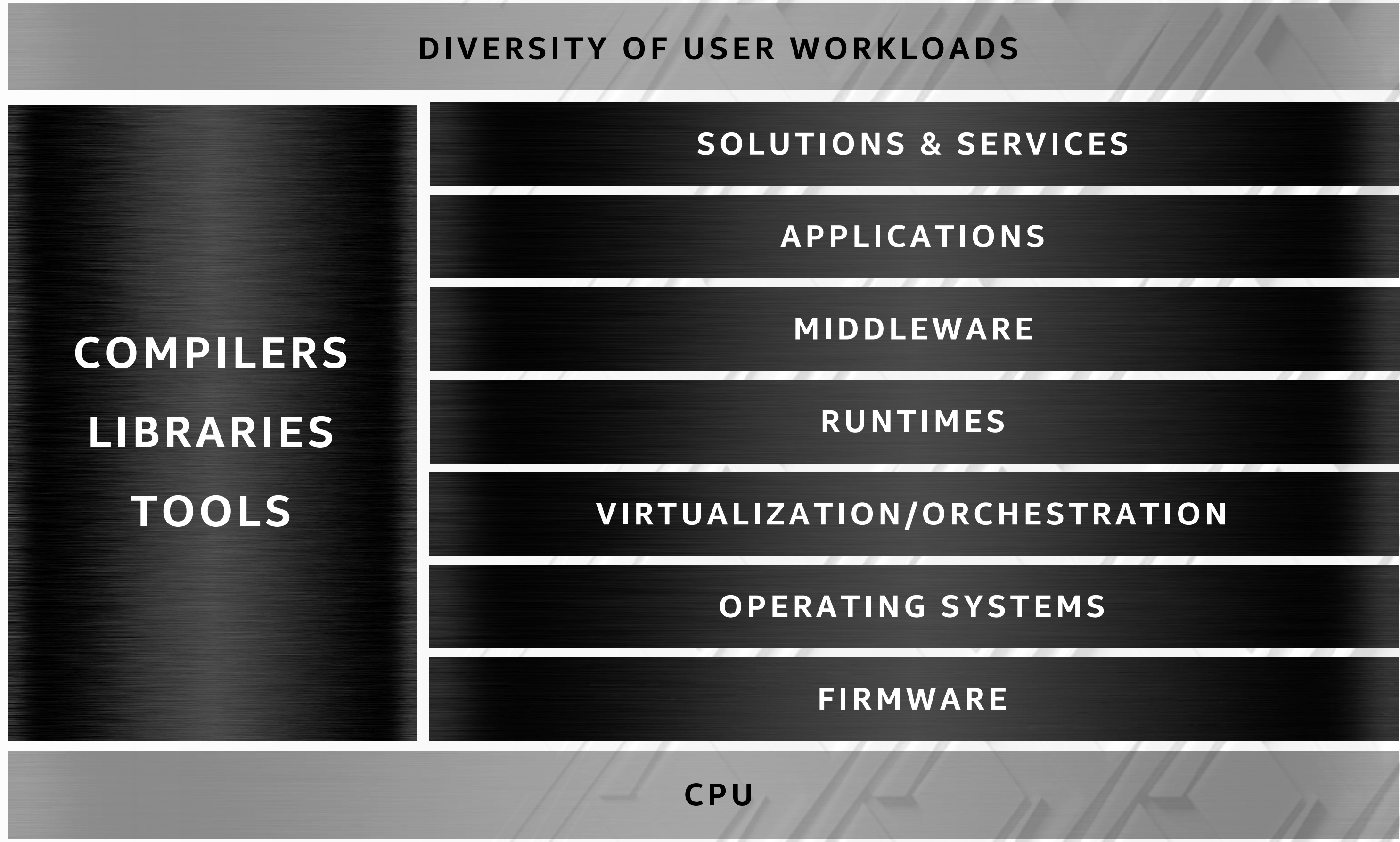


GENERALITY $\propto$ SOFTWARE STACK SCALE



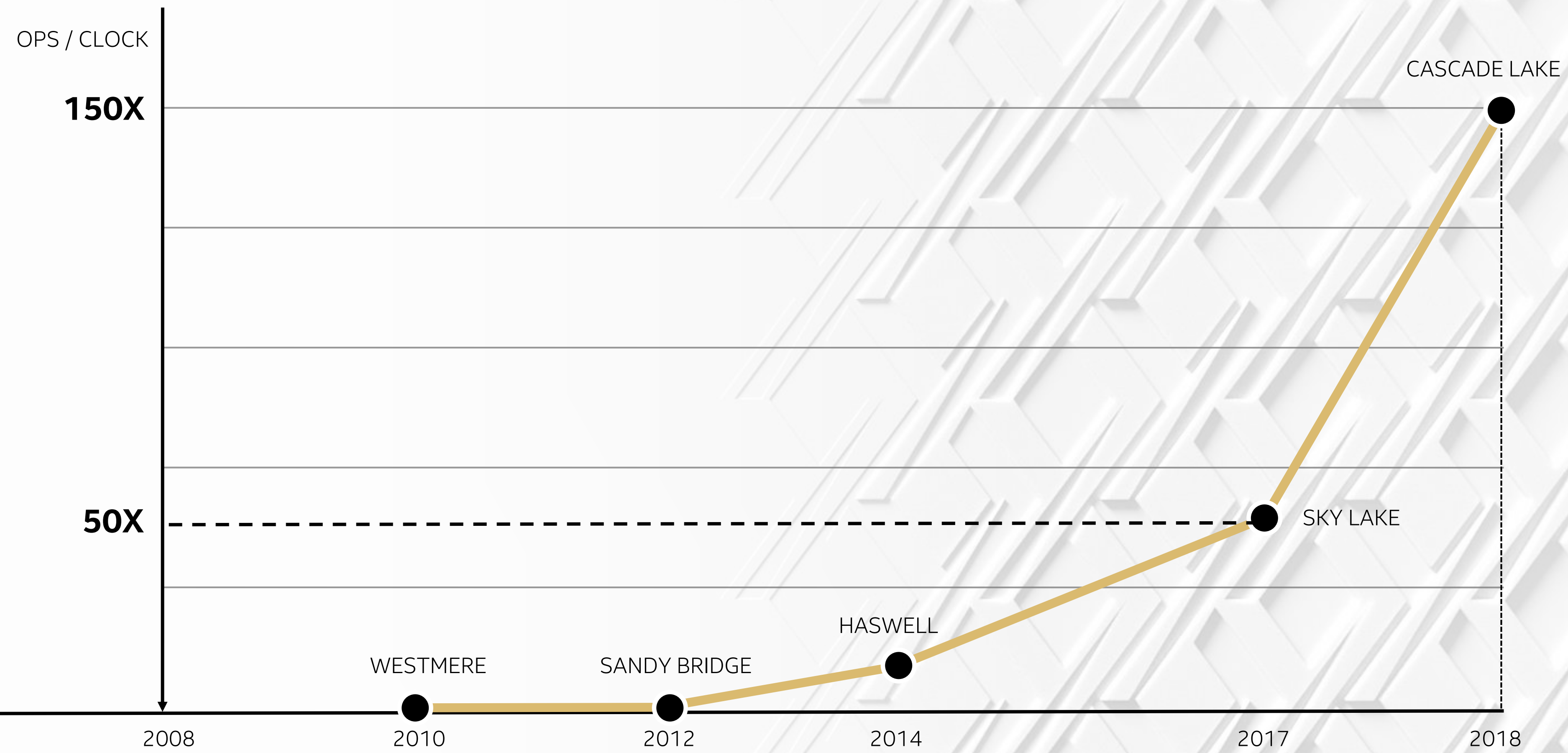
GENERALITY $\propto$ SOFTWARE STACK SCALE

20M
DEVELOPERS



GENERALITY $\propto$ 1/ARCHITECTURE HETEROGENEITY

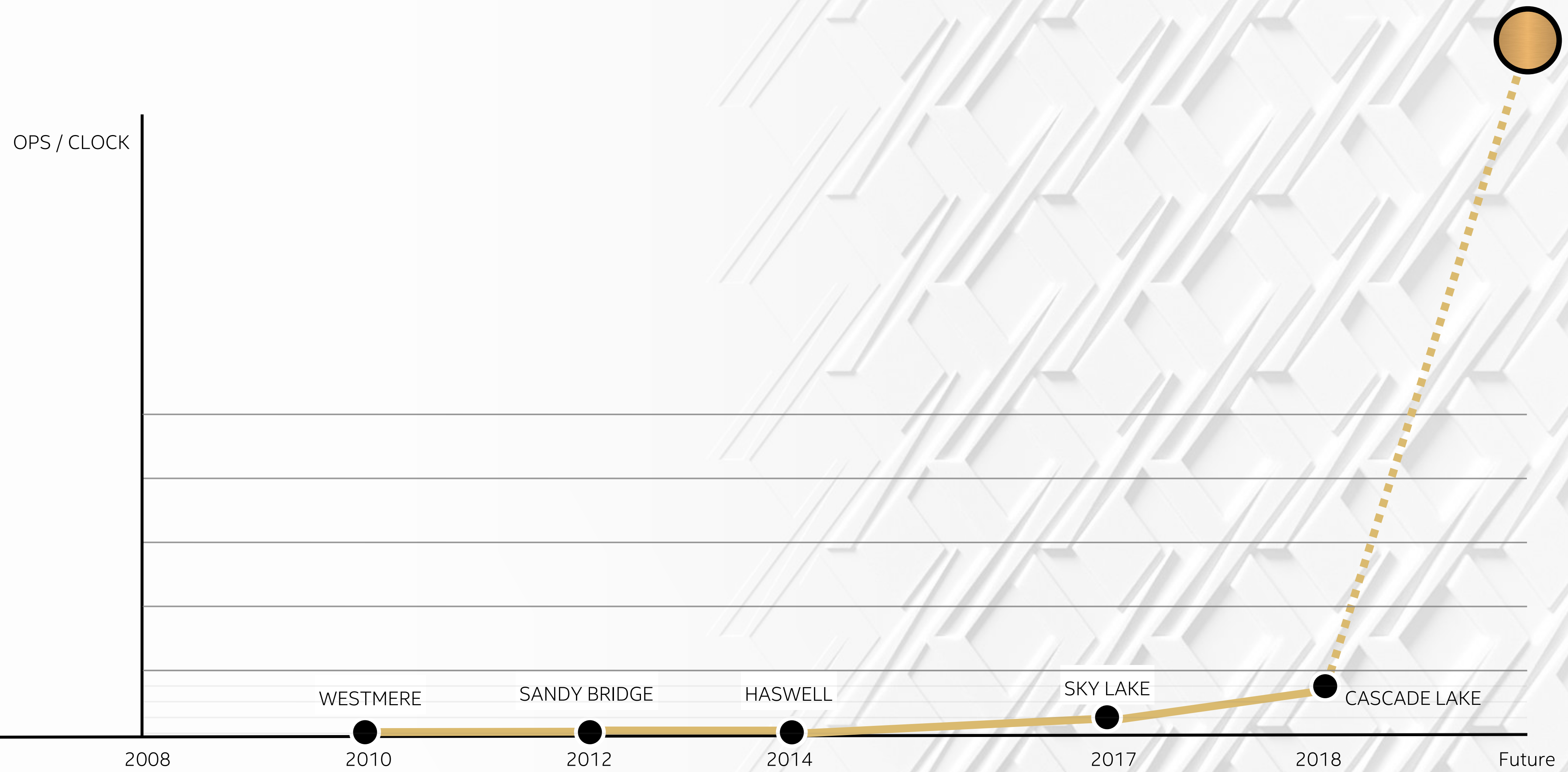
HETEROGENEOUS MATH IN CPU



Source: Results have been calculated based on internal Intel analysis and are provided for informational purposes only



HETEROGENEOUS MATH IN CPU



Graph not drawn to scale

Source: Intel Projections



CPU IMPACT WITH ISA EXTENSIONS AND SOFTWARE



CPU IMPACT WITH ISA EXTENSIONS AND SOFTWARE



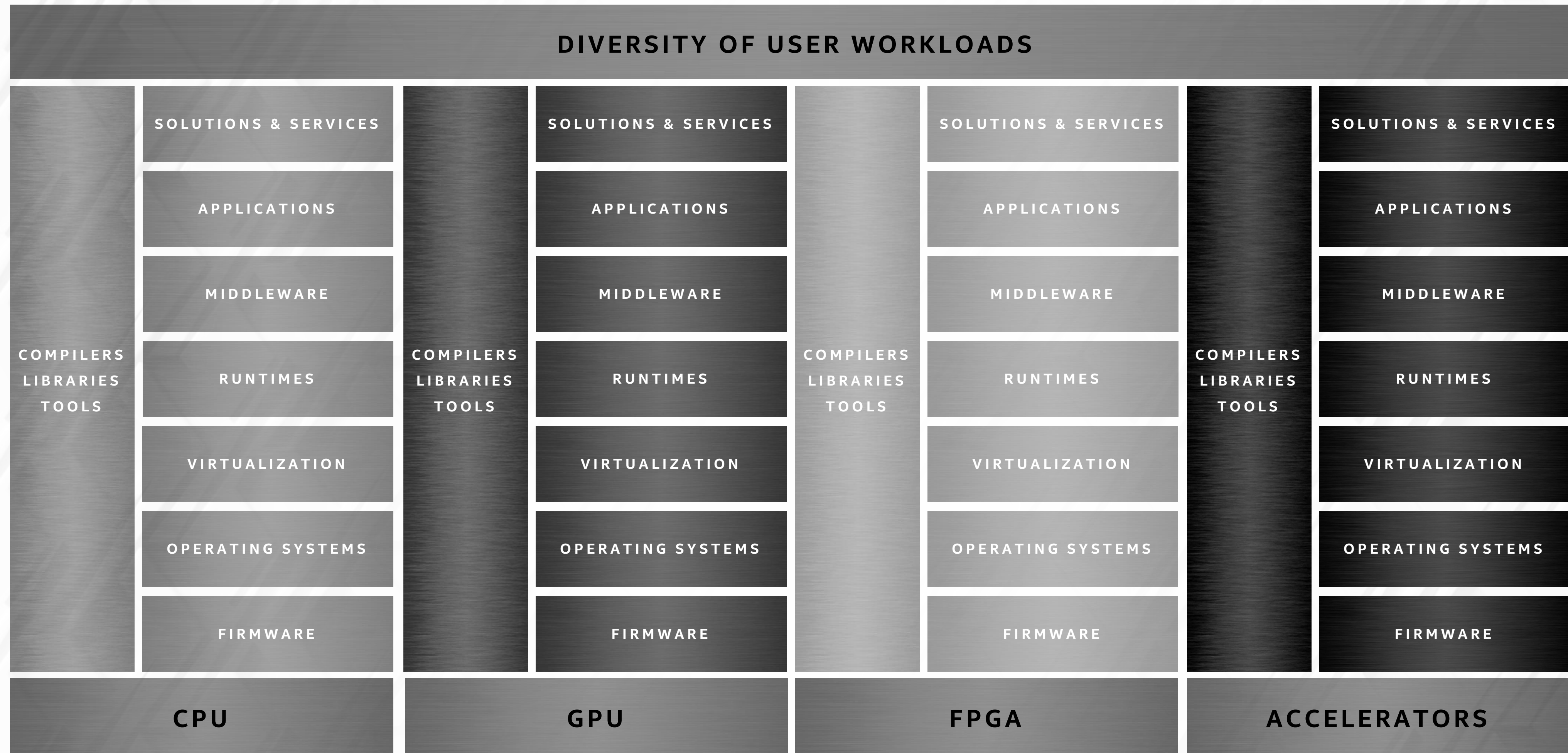
DISCRETE GPUS

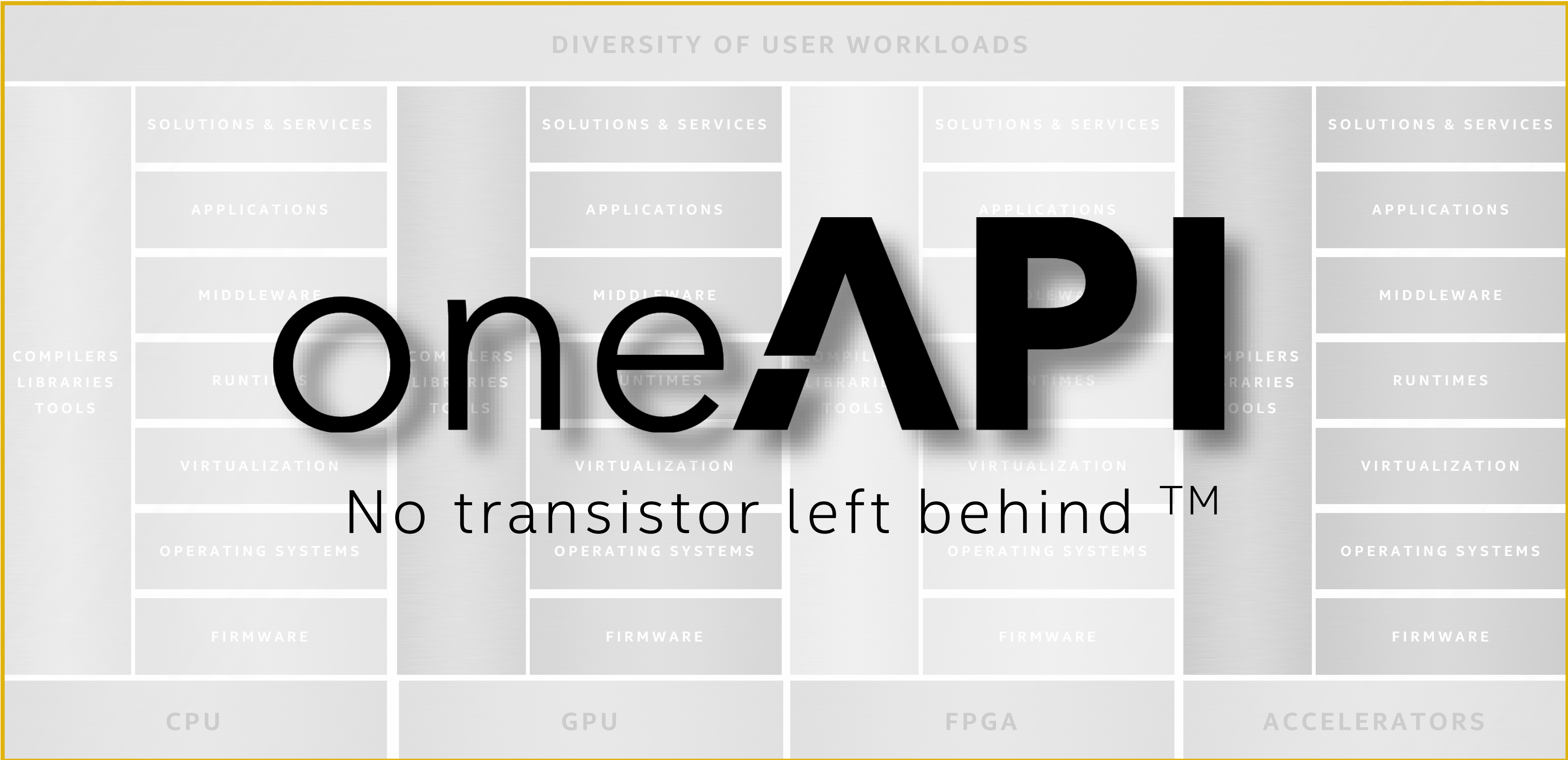


DISCRETE GPU IMPACT



UNSCALABLE HETEROGENOUS SOFTWARE





oneAPI GOALS

DIVERSITY OF USER WORKLOADS

OPEN

SCALAR – VECTOR – MATRIX – SPATIAL

oneAPI GOALS

DIVERSITY OF USER WORKLOADS

OPEN

SIMPLE
&
SCALABLE

SCALAR – VECTOR – MATRIX – SPATIAL

oneAPI GOALS

DIVERSITY OF USER WORKLOADS

OPEN

SIMPLE
&
SCALABLE

NO
DEVELOPER
LEFT BEHIND

SCALAR – VECTOR – MATRIX – SPATIAL

oneAPI STACK

DIVERSITY OF USER WORKLOADS

DIRECT
PROGRAMMING
LANGUAGES

DOMAIN SPECIFIC
LIBRARIES

MIGRATION
TOOLS

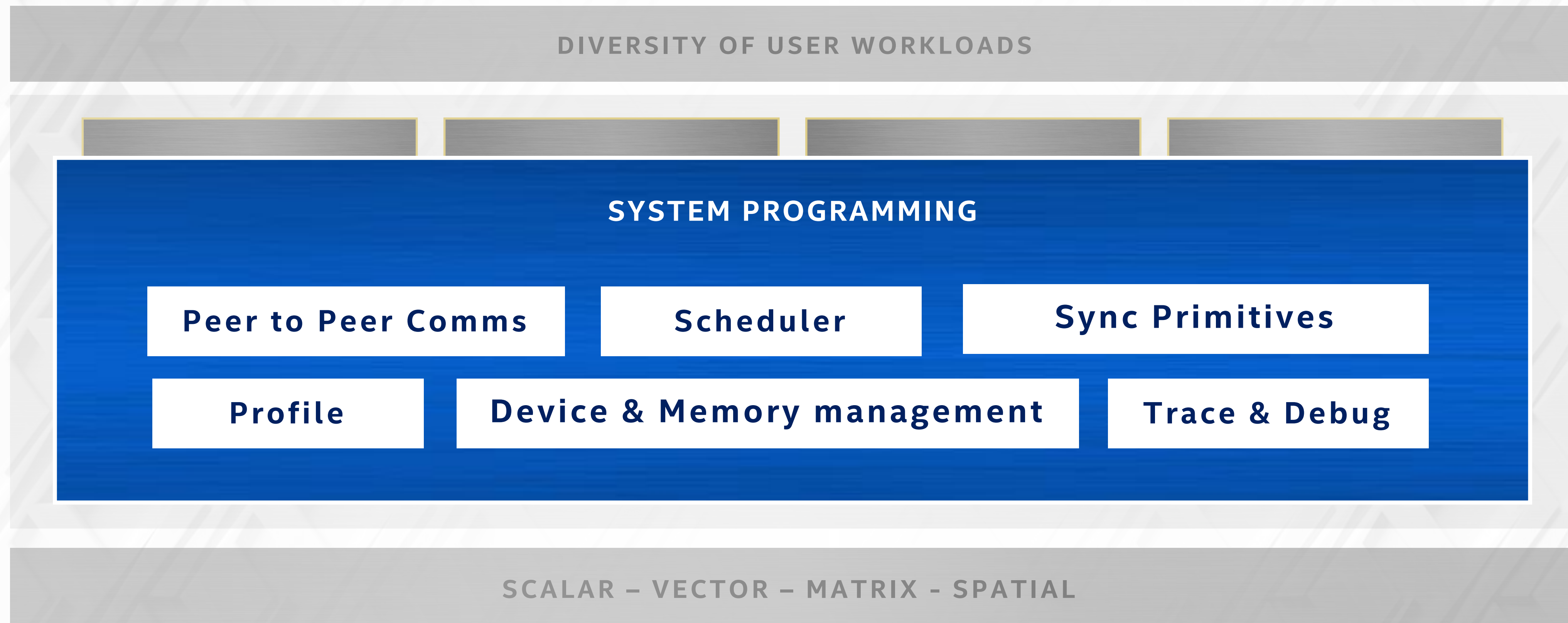
ANALYSIS &
DEBUG TOOLS

SYSTEM PROGRAMMING

SCALAR – VECTOR – MATRIX – SPATIAL

oneAPI

SYSTEM PROGRAMMING



oneAPI STACK

DIVERSITY OF USER WORKLOADS

DIRECT
PROGRAMMING
LANGUAGES

DOMAIN SPECIFIC
LIBRARIES

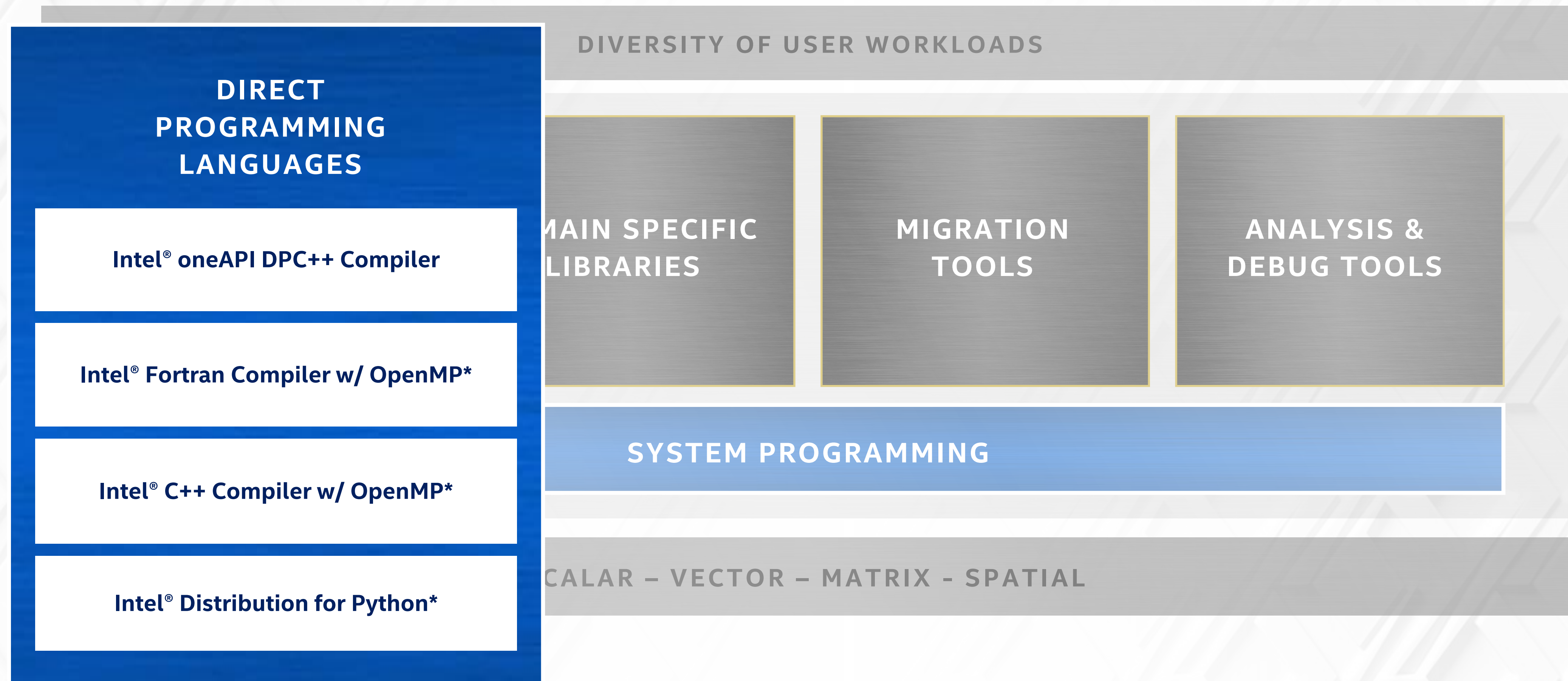
MIGRATION
TOOLS

ANALYSIS &
DEBUG TOOLS

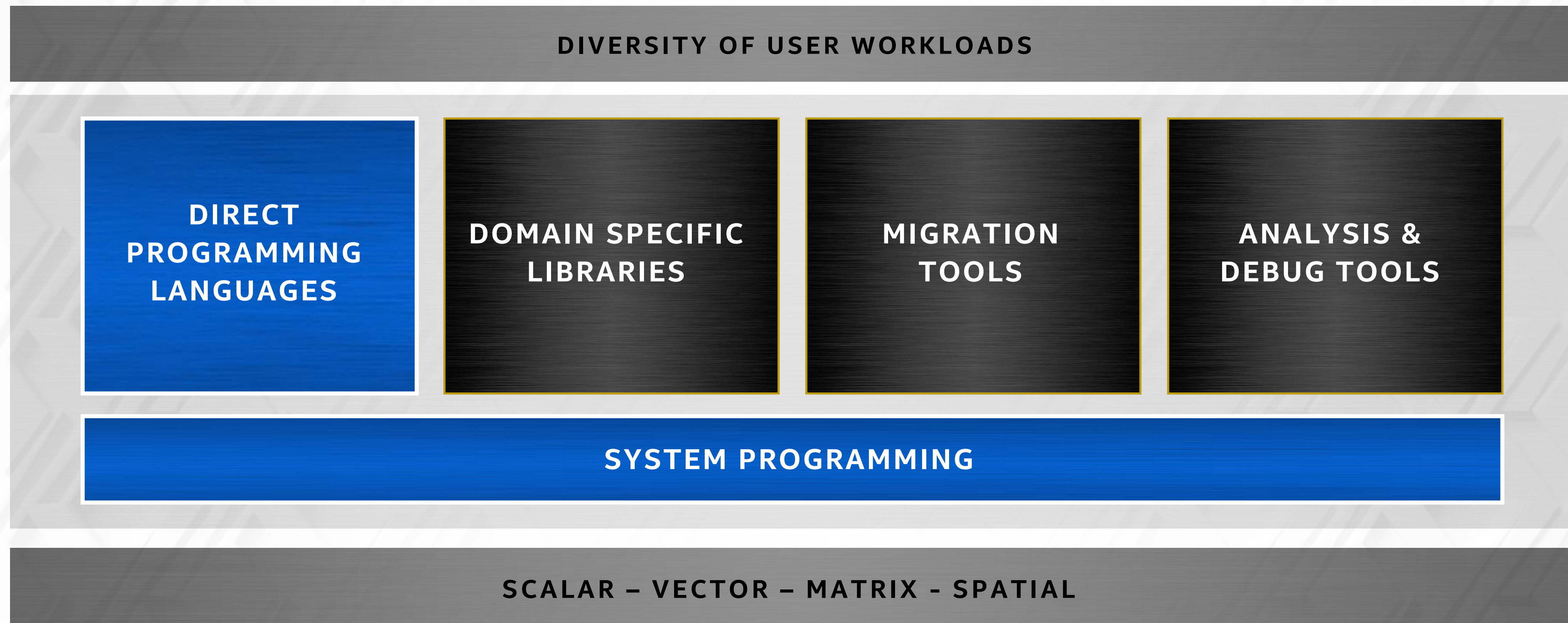
SYSTEM PROGRAMMING

SCALAR – VECTOR – MATRIX – SPATIAL

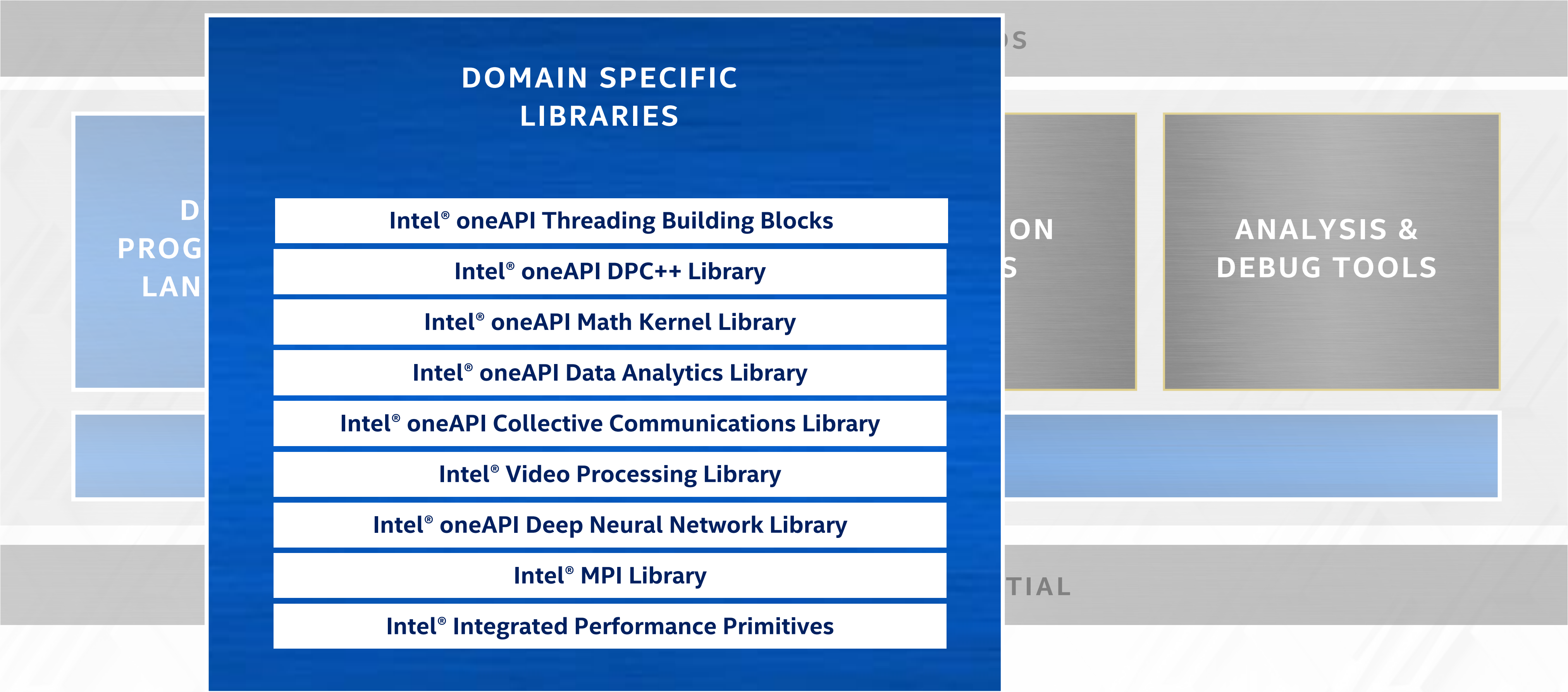
oneAPI DIRECT PROGRAMMING



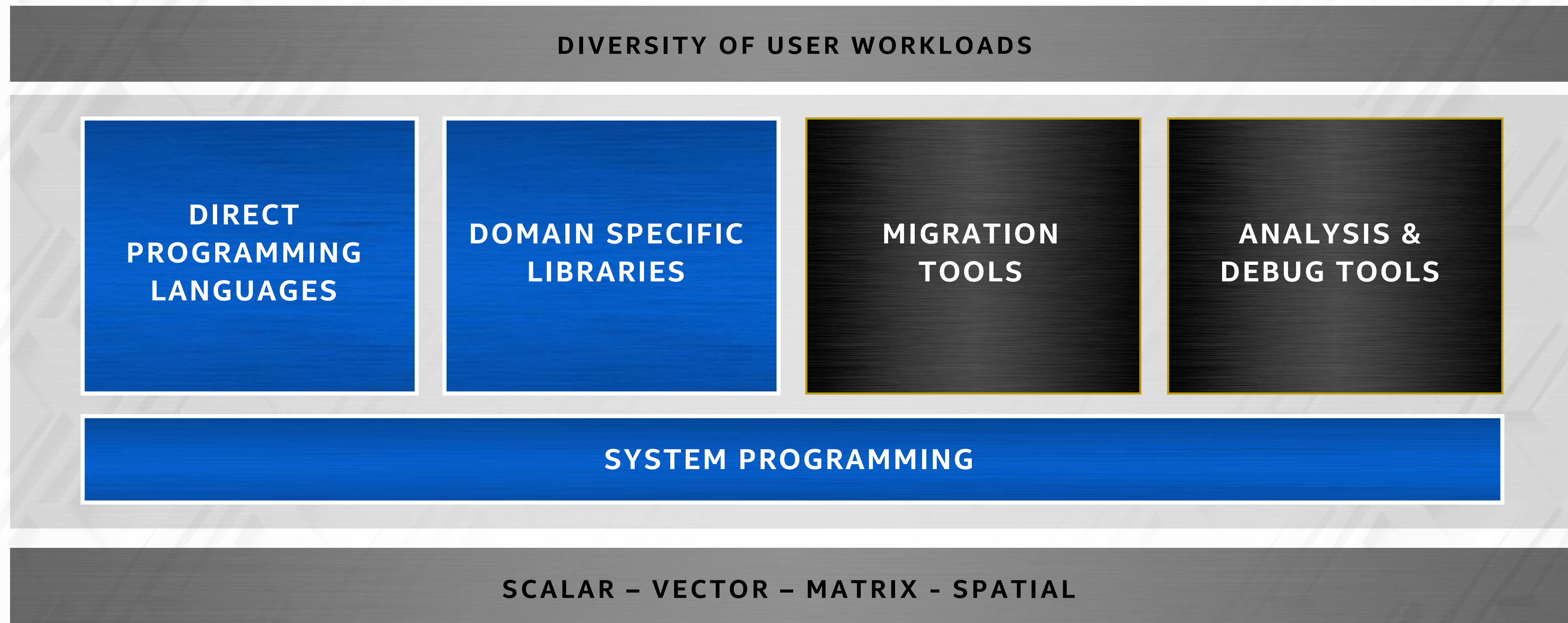
oneAPI STACK



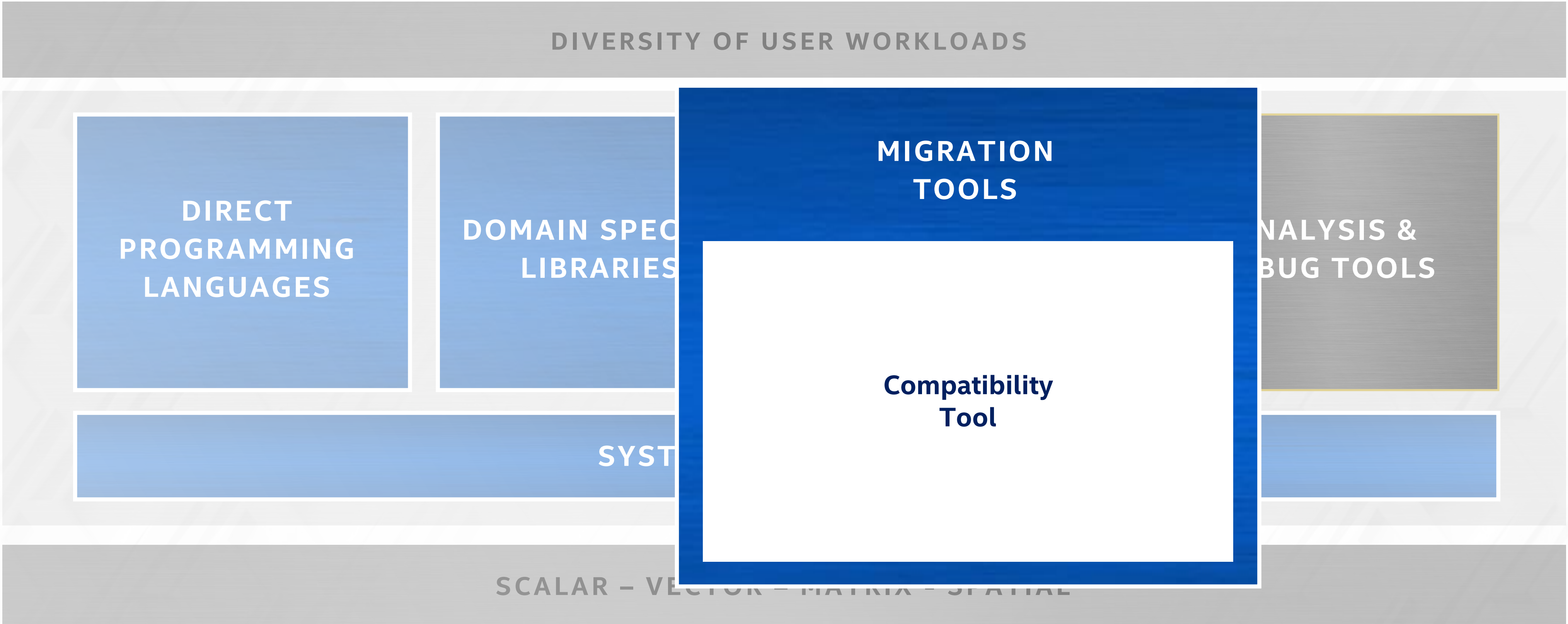
oneAPI LIBRARIES



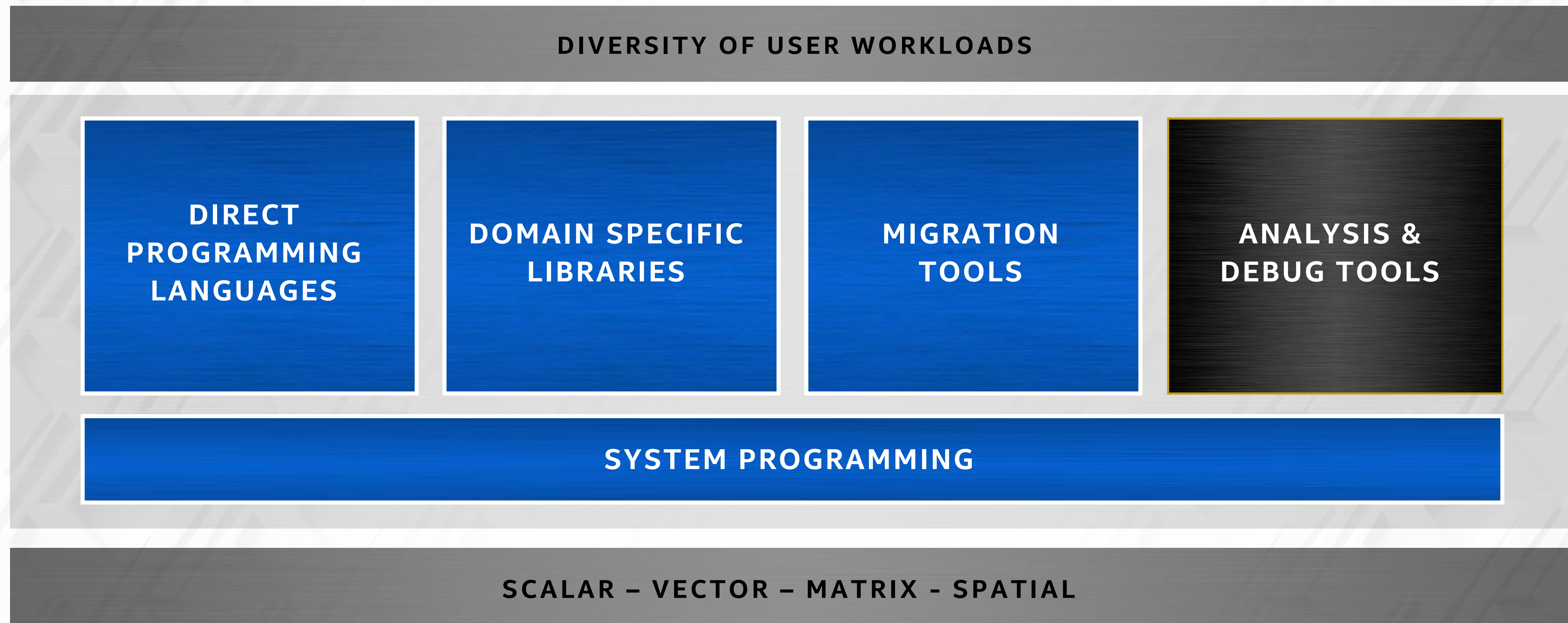
oneAPI STACK



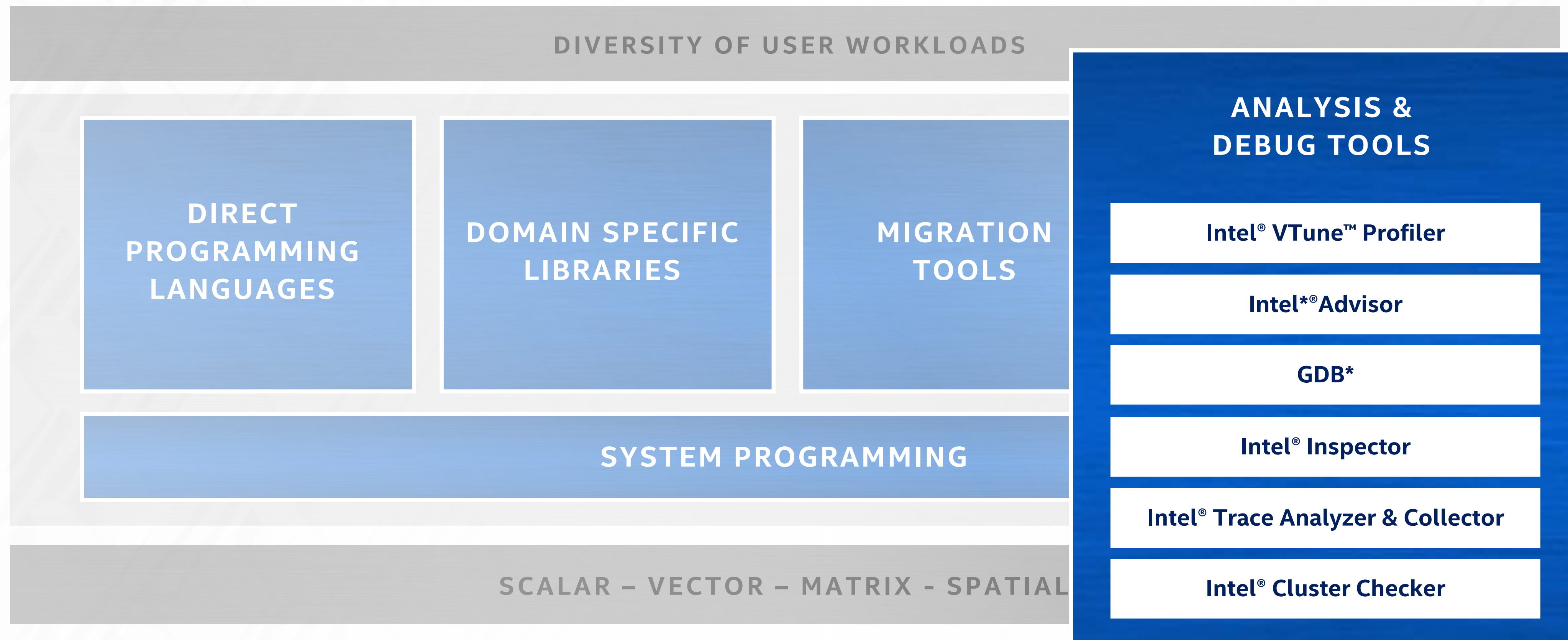
oneAPI TOOLS



oneAPI STACK



oneAPI TOOLS



oneAPI STACK

DIVERSITY OF USER WORKLOADS

DIRECT PROGRAMMING LANGUAGES

Intel® oneAPI DPC++ Compiler

Intel® Fortran Compiler w/
OpenMP*

Intel® C++ Compiler w/ OpenMP*

Intel® Distribution for Python*

DOMAIN SPECIFIC LIBRARIES

Intel® oneAPI Threading Building Blocks

Intel® oneAPI DPC++ Library

Intel® oneAPI Math Kernel Library

Intel® oneAPI Data Analytics Library

Intel® oneAPI Collective Communications Library

Intel® Video Processing Library

Intel® oneAPI Deep Neural Network Library

Intel® MPI Library

Intel® Integrated Performance Primitives

MIGRATION TOOLS

Compatibility
Tool

ANALYSIS & DEBUG TOOLS

Intel® VTune™ Profiler

Intel® Advisor

GDB*

Intel® Inspector

Intel® Trace Analyzer & Collector

Intel® Cluster Checker

SYSTEM PROGRAMMING

Peer to Peer
Comms

Scheduler

Profile

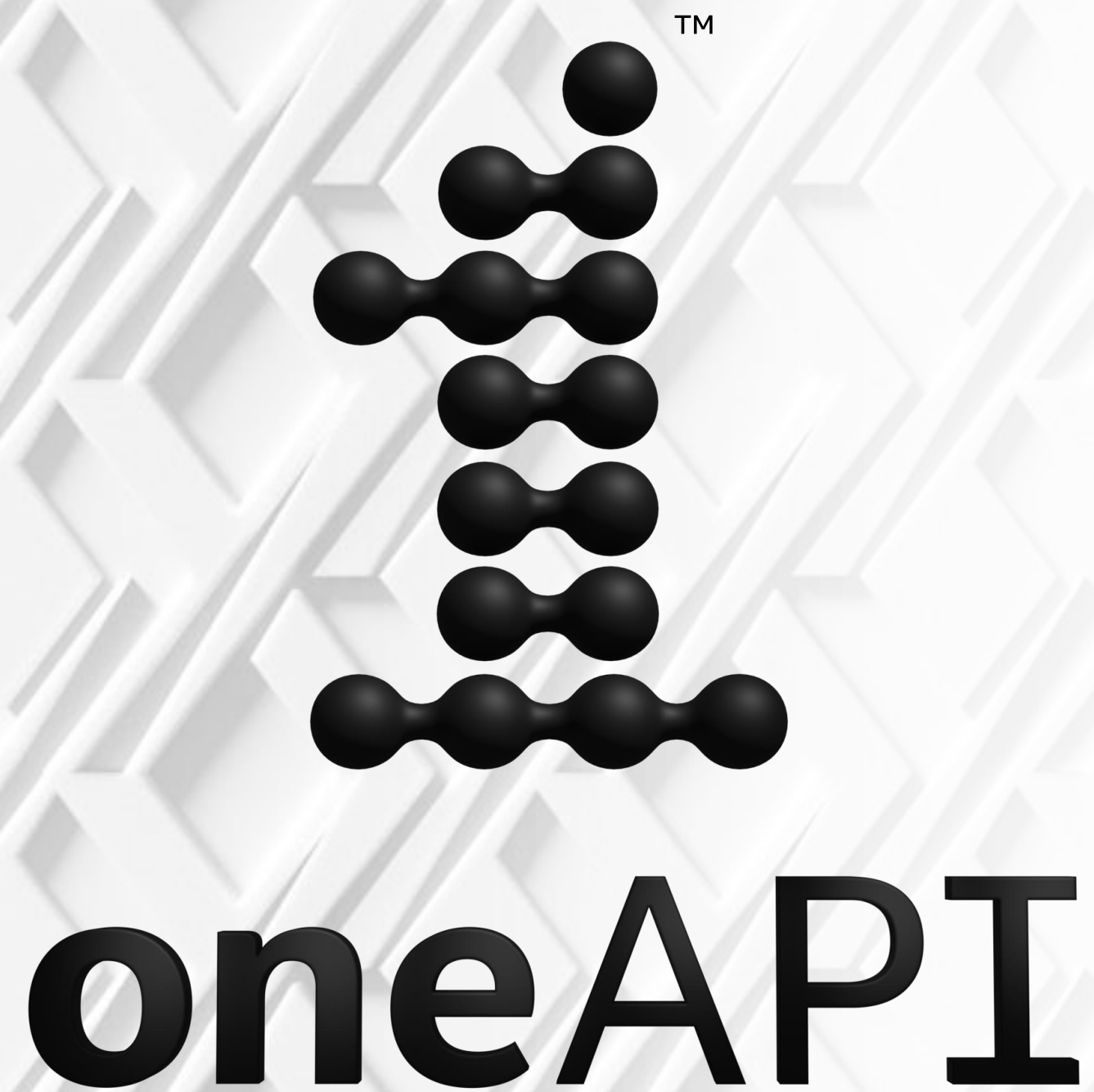
Sync Primitives

Device & Memory
management

Trace & Debug

SCALAR – VECTOR – MATRIX – SPATIAL





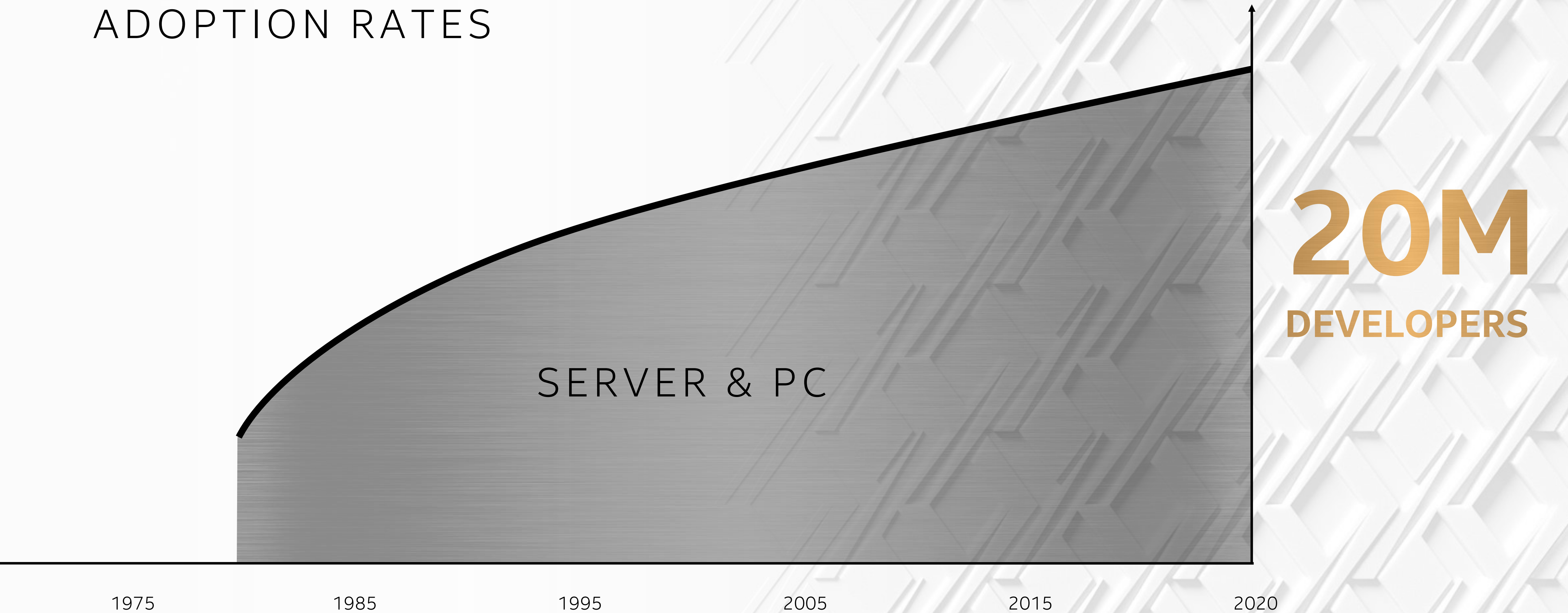
**PUBLIC BETA
AVAILABLE
TODAY!**

GO TO [SOFTWARE.INTEL.COM/ONEAPI](https://software.intel.com/oneapi)



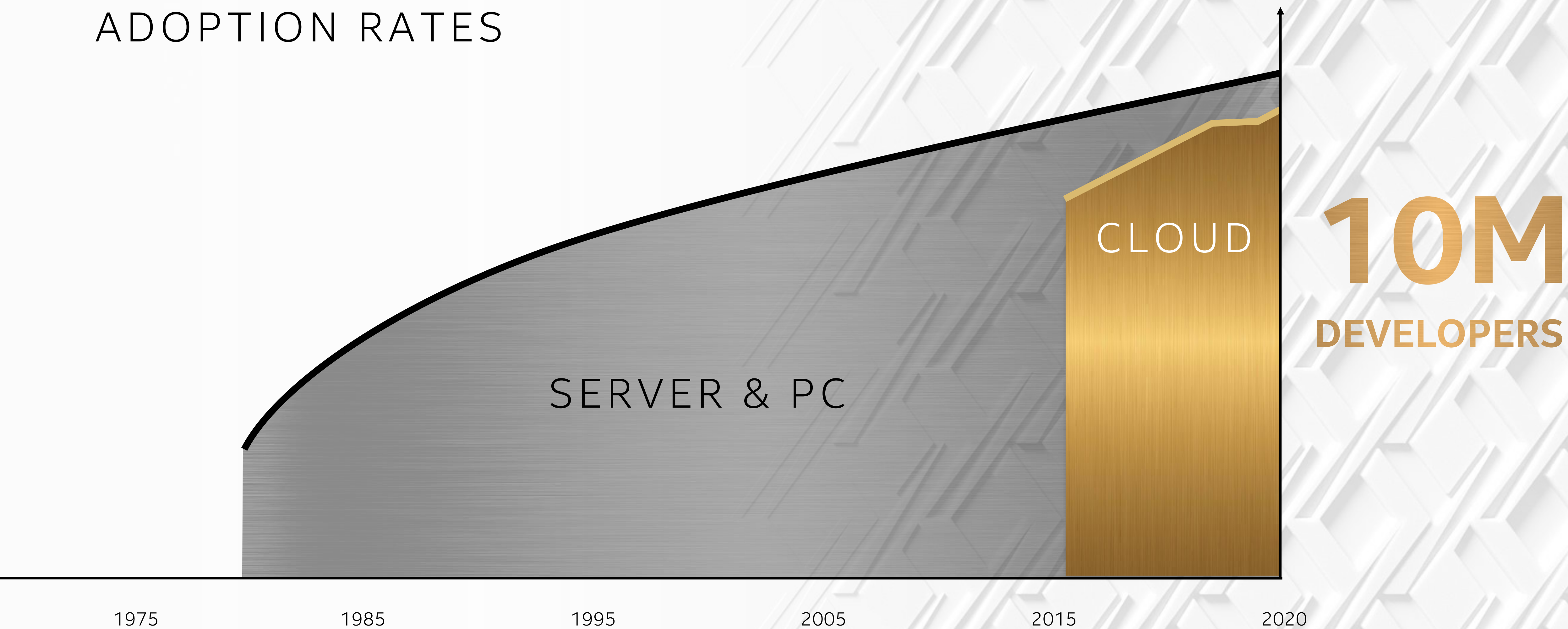
TRADITIONAL

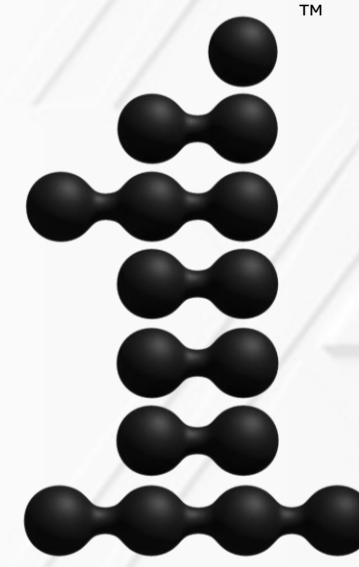
ADOPTION RATES



CLOUD

ADOPTION RATES





oneAPI



DevCloud

A V A I L A B L E T O D A Y

NO HARDWARE ACQUISITION

NO INSTALLATION

NO DOWNLOADS

**NO SET-UP AND
CONFIGURATION**





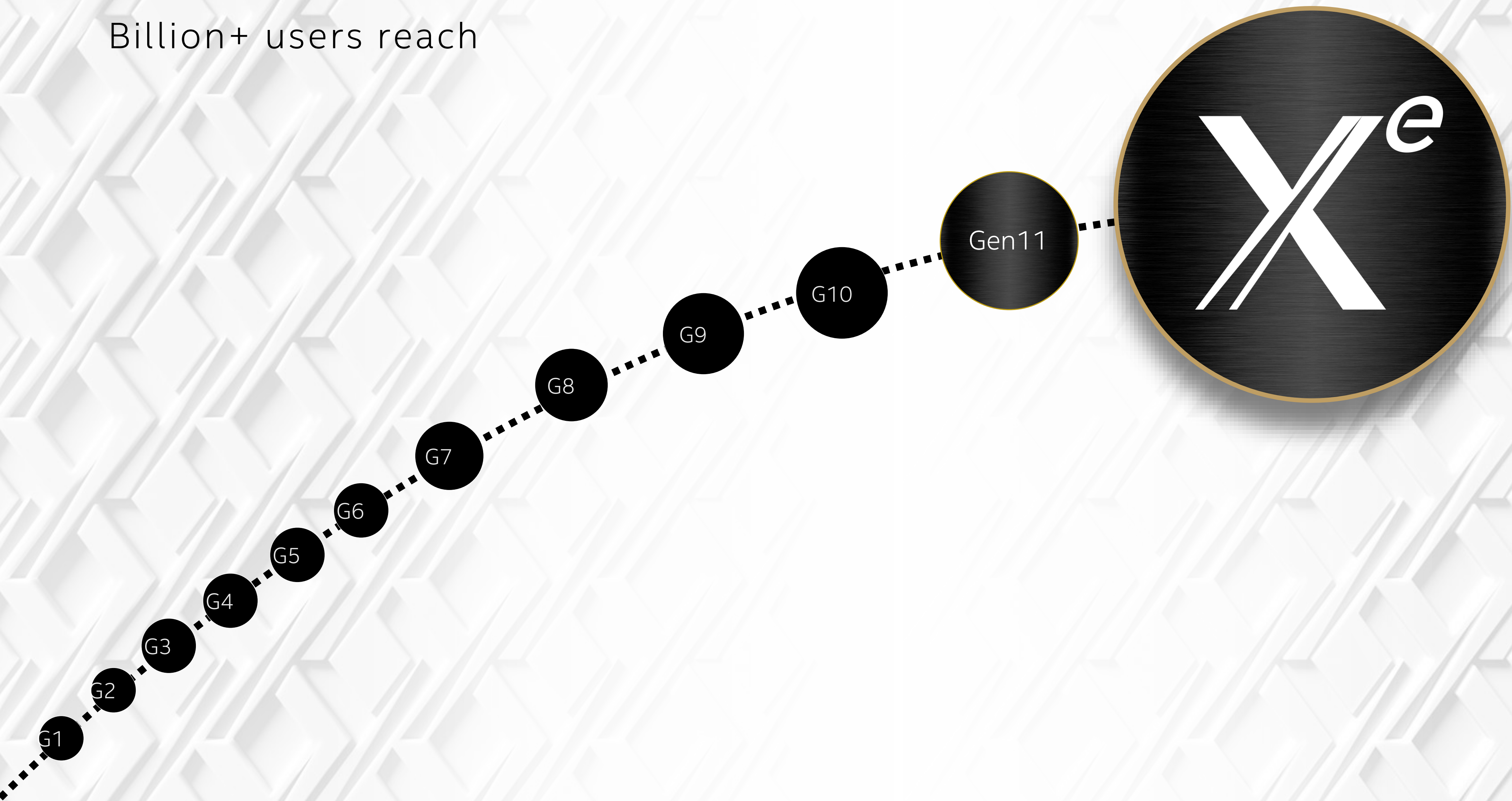
INTEL GPU

ARCHITECTURE
ROADMAP



INTEL GPU IMPACT

Billion+ users reach



X^e ARCHITECTURE

TWO MICRO ARCHITECTURES

ONE GPU
ARCHITECTURE

DATACENTER / AI

ENTHUSIAST

MID-RANGE

INTEGRATED + ENTRY

PETAFLIPS

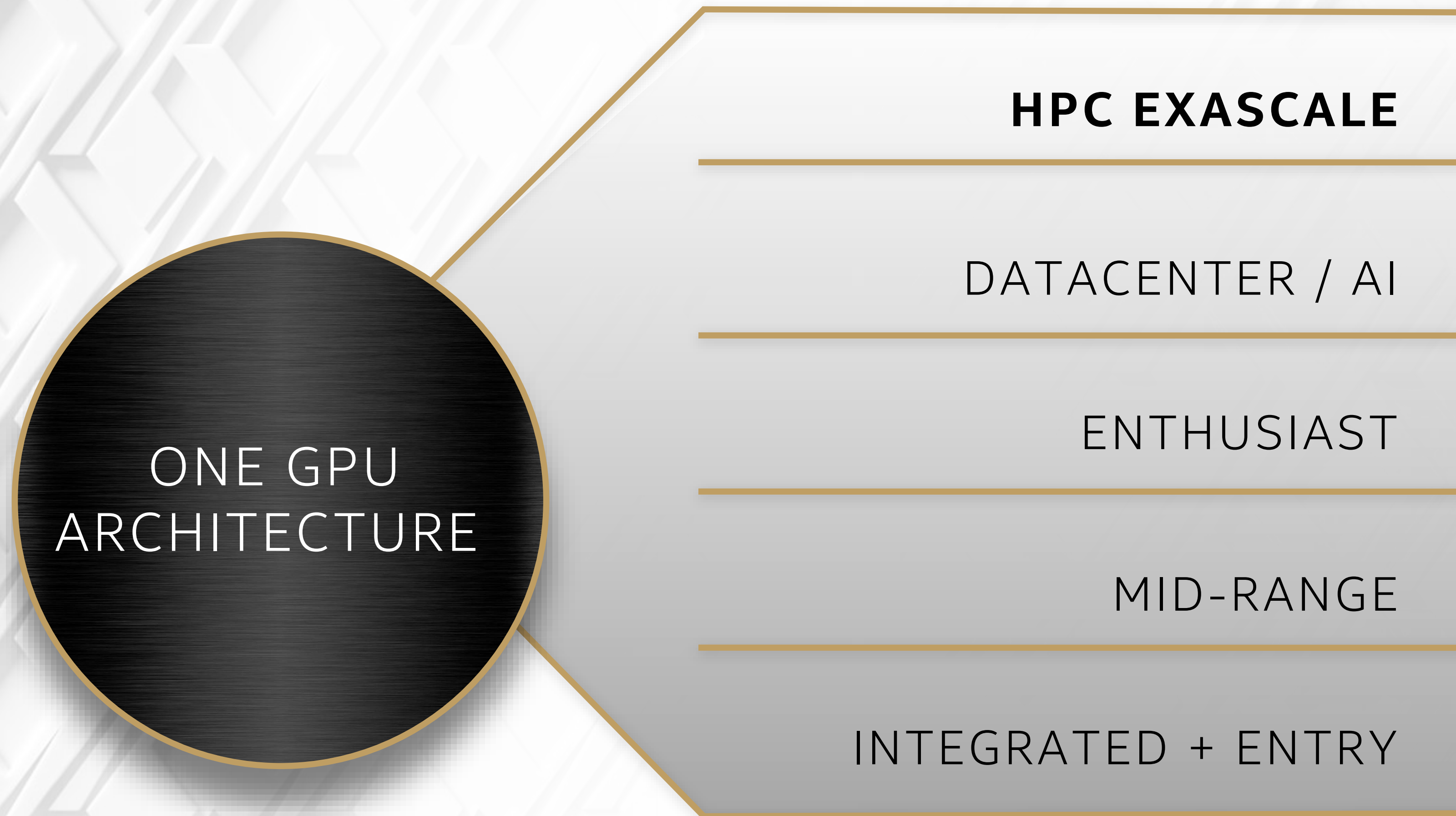


TERAFLOPS



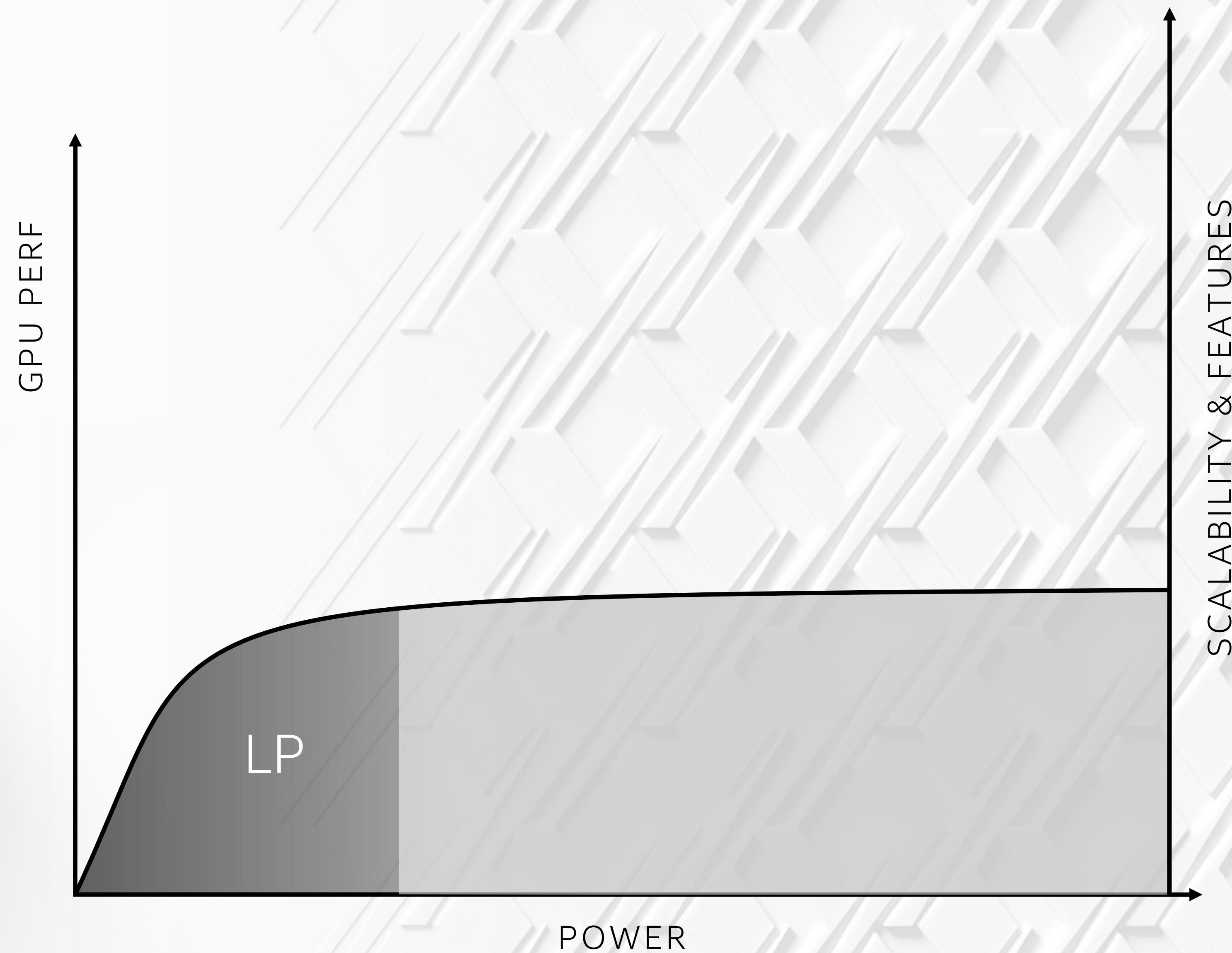
X^e ARCHITECTURE

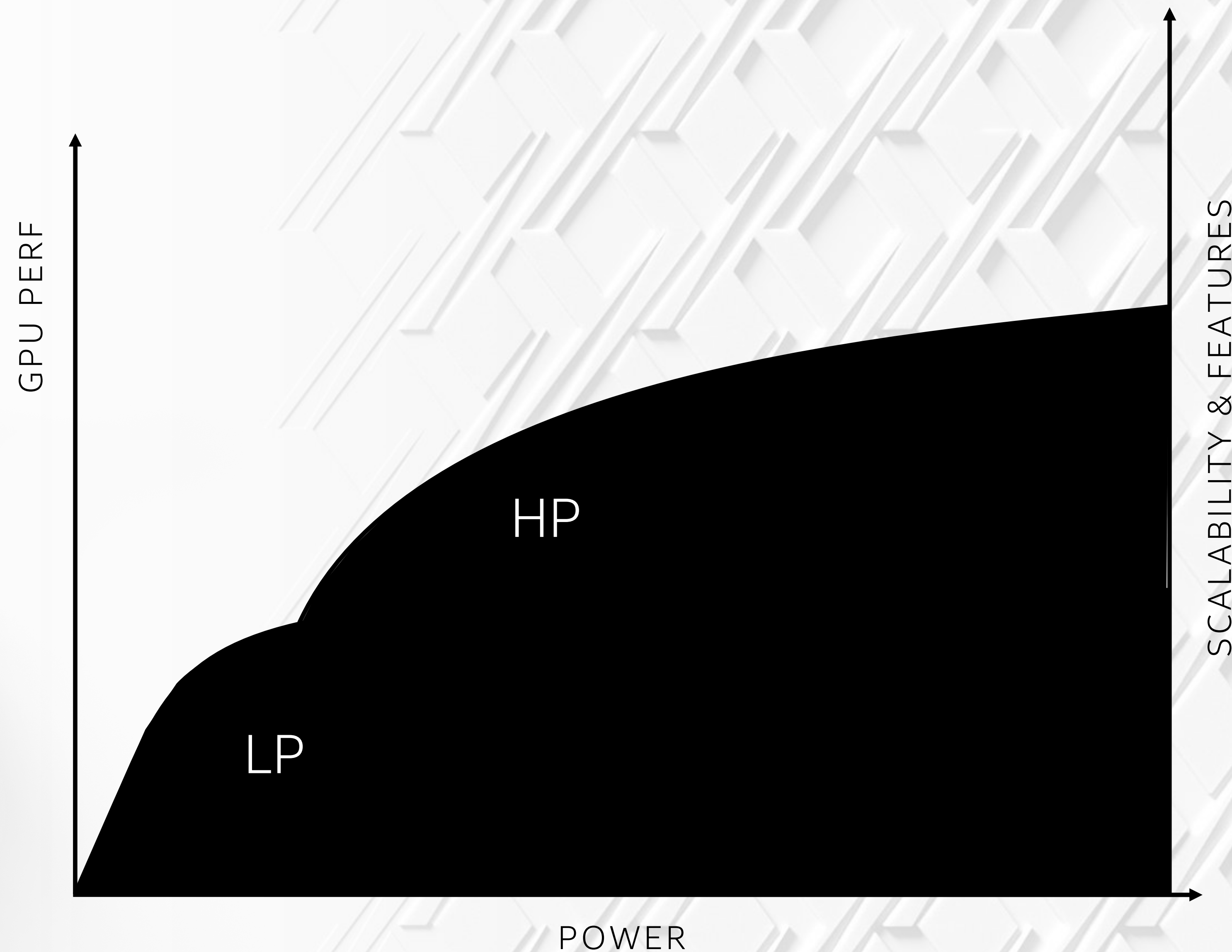
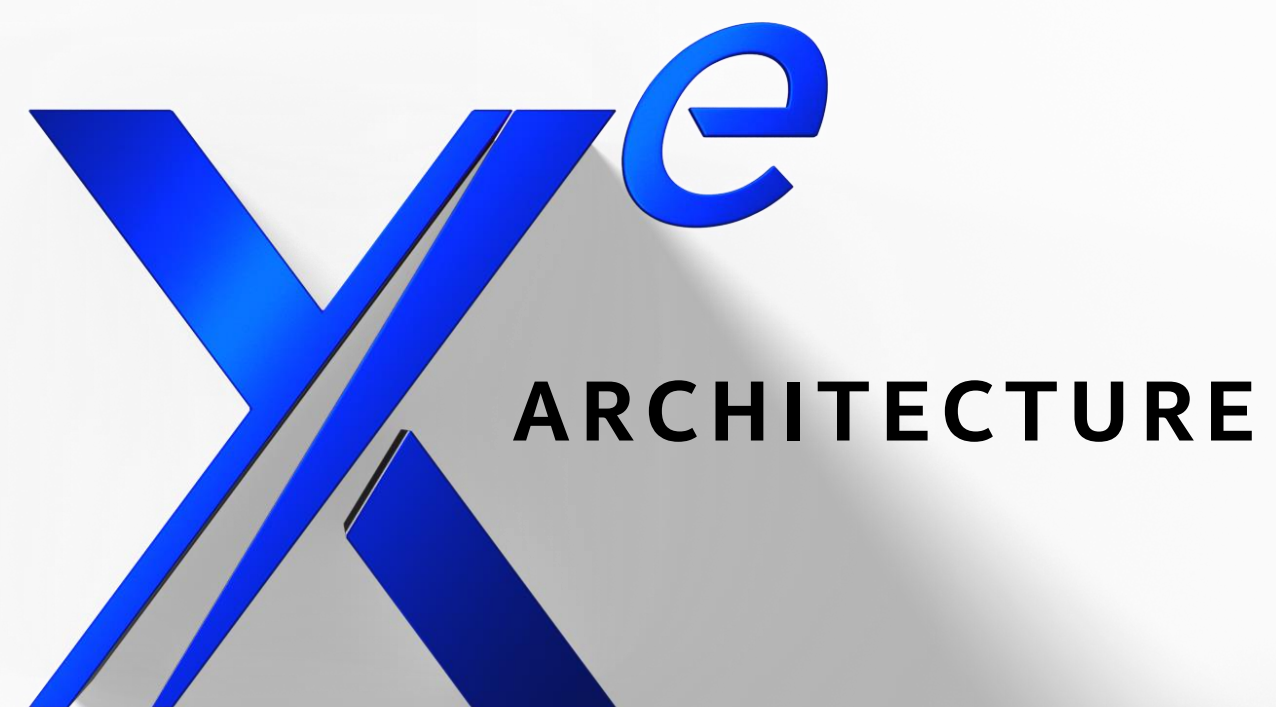
FROM TERA FLOPS TO EXASCALE

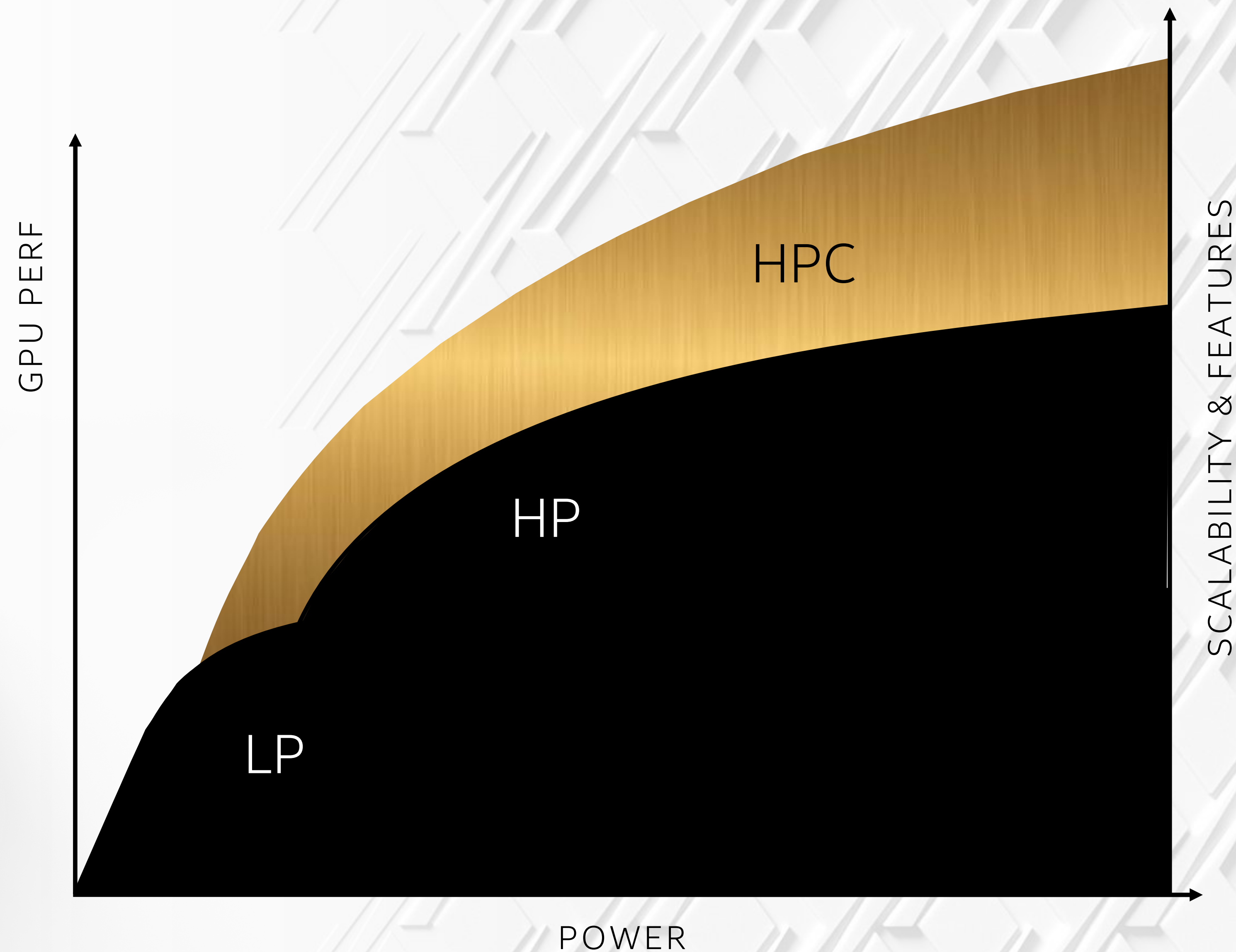
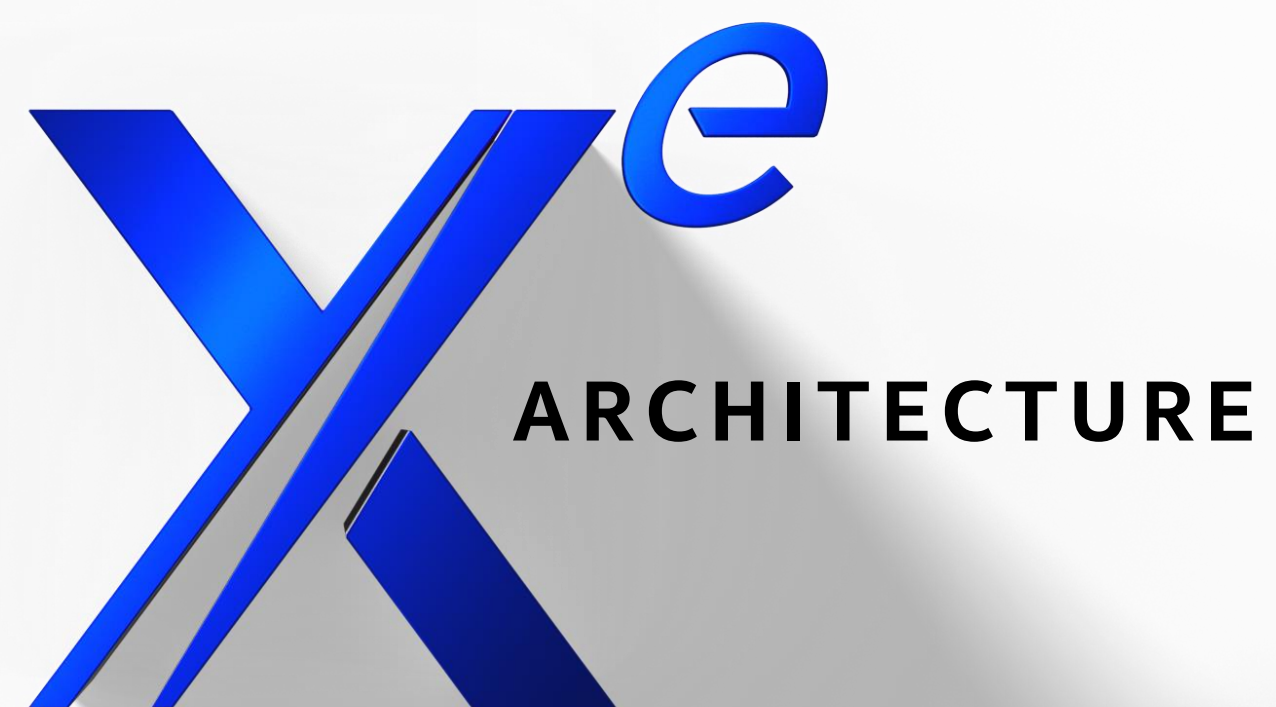


TERAFLOPS











HPC FEATURES

COMPUTE

MEMORY



HPC FEATURES

COMPUTE

SCALABILITY

AI PERFORMANCE

HPC PERFORMANCE

MEMORY

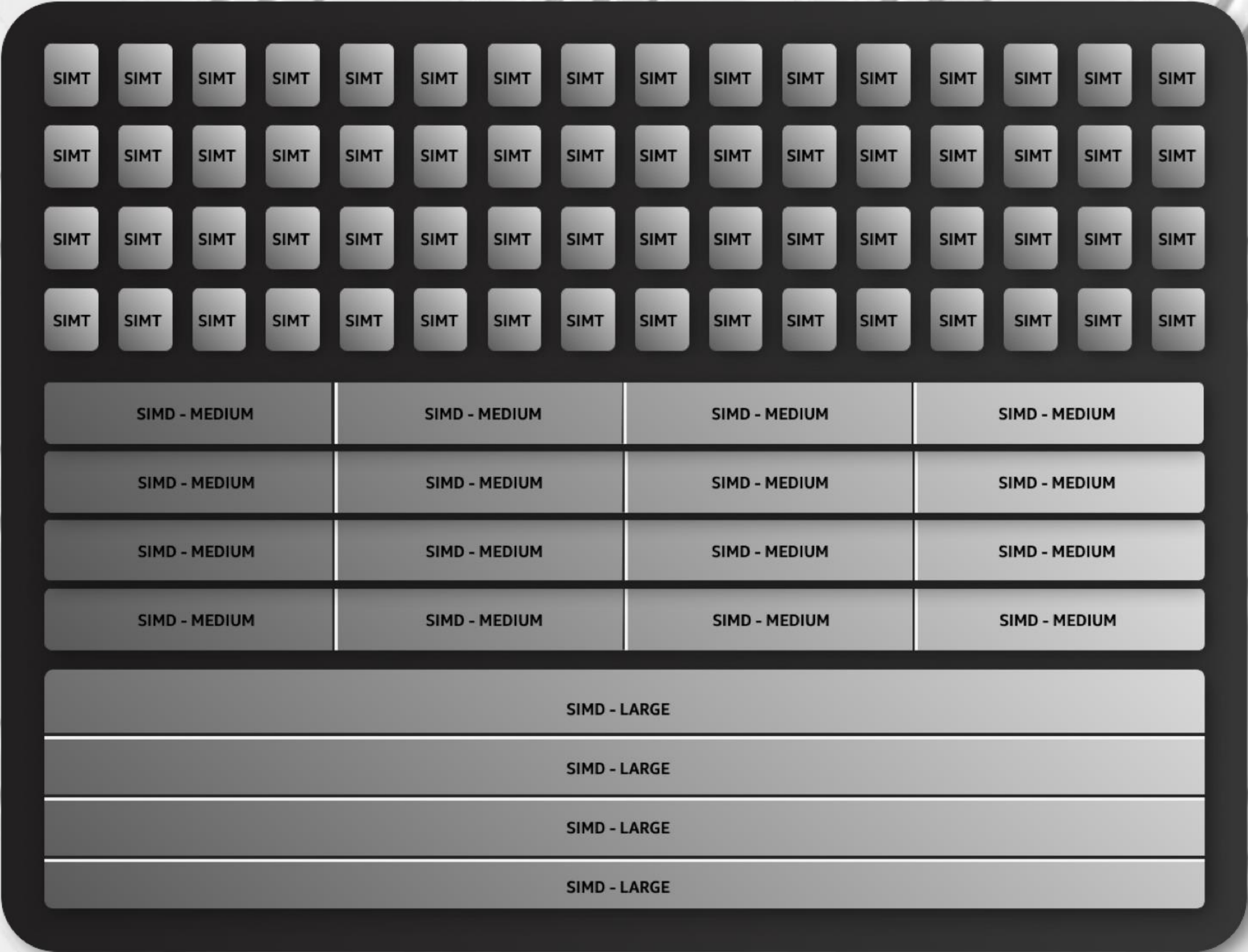
VARIABLE VECTOR WIDTH



SIMT (GPU STYLE)
SIMD (CPU STYLE)
SIMT + SIMD (MAX PERF)

COMPUTE

SCALABILITY



VARIABLE VECTOR WIDTH



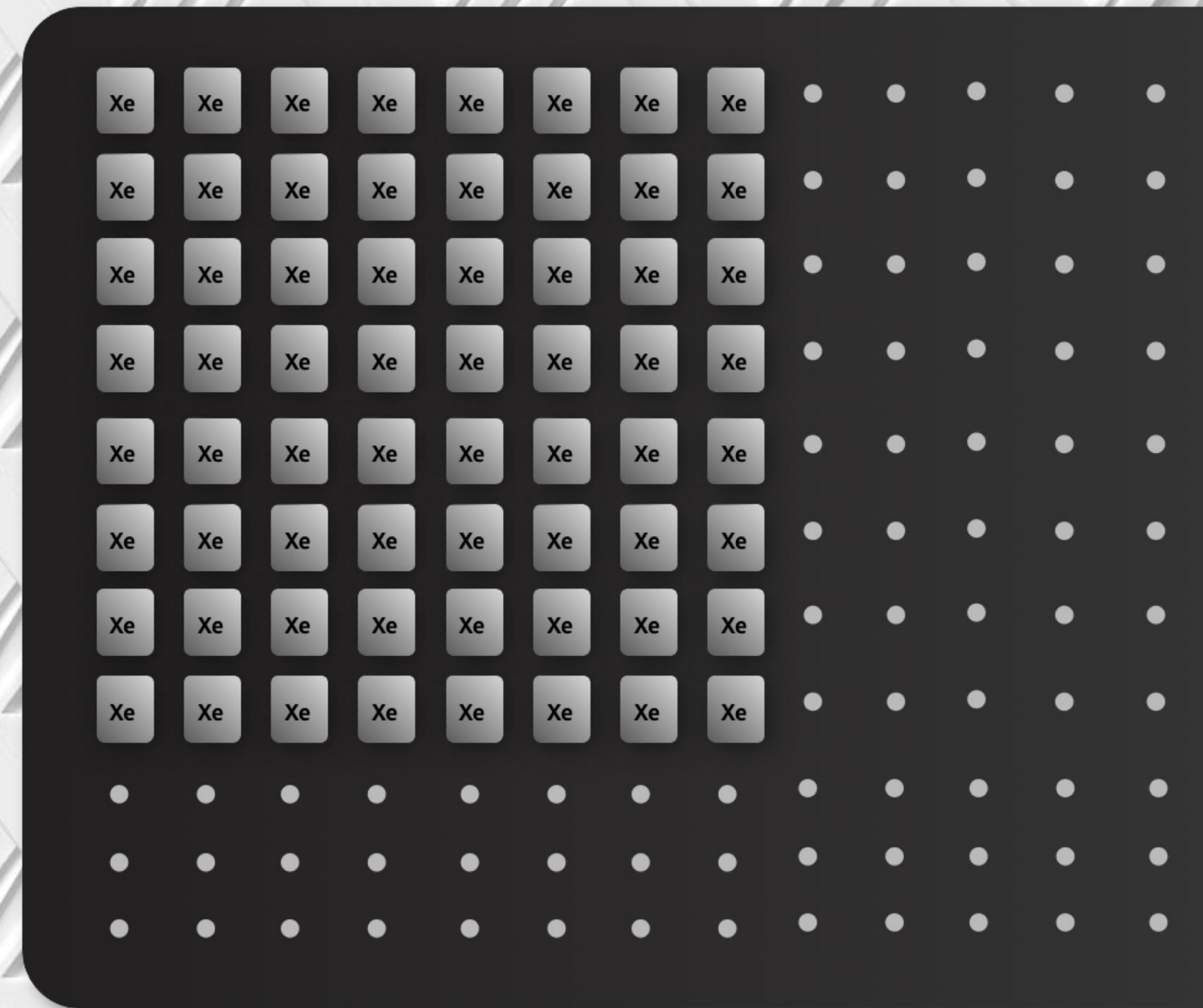
SIMT (GPU STYLE)
SIMD (CPU STYLE)
SIMT + SIMD (MAX PERF)

COMPUTE

SCALABILITY



SCALABLE TO THOUSANDS
OF EXECUTION UNITS



COMPUTE

SCALABILITY



NEW DATA PARALLEL
MATRIX ENGINE

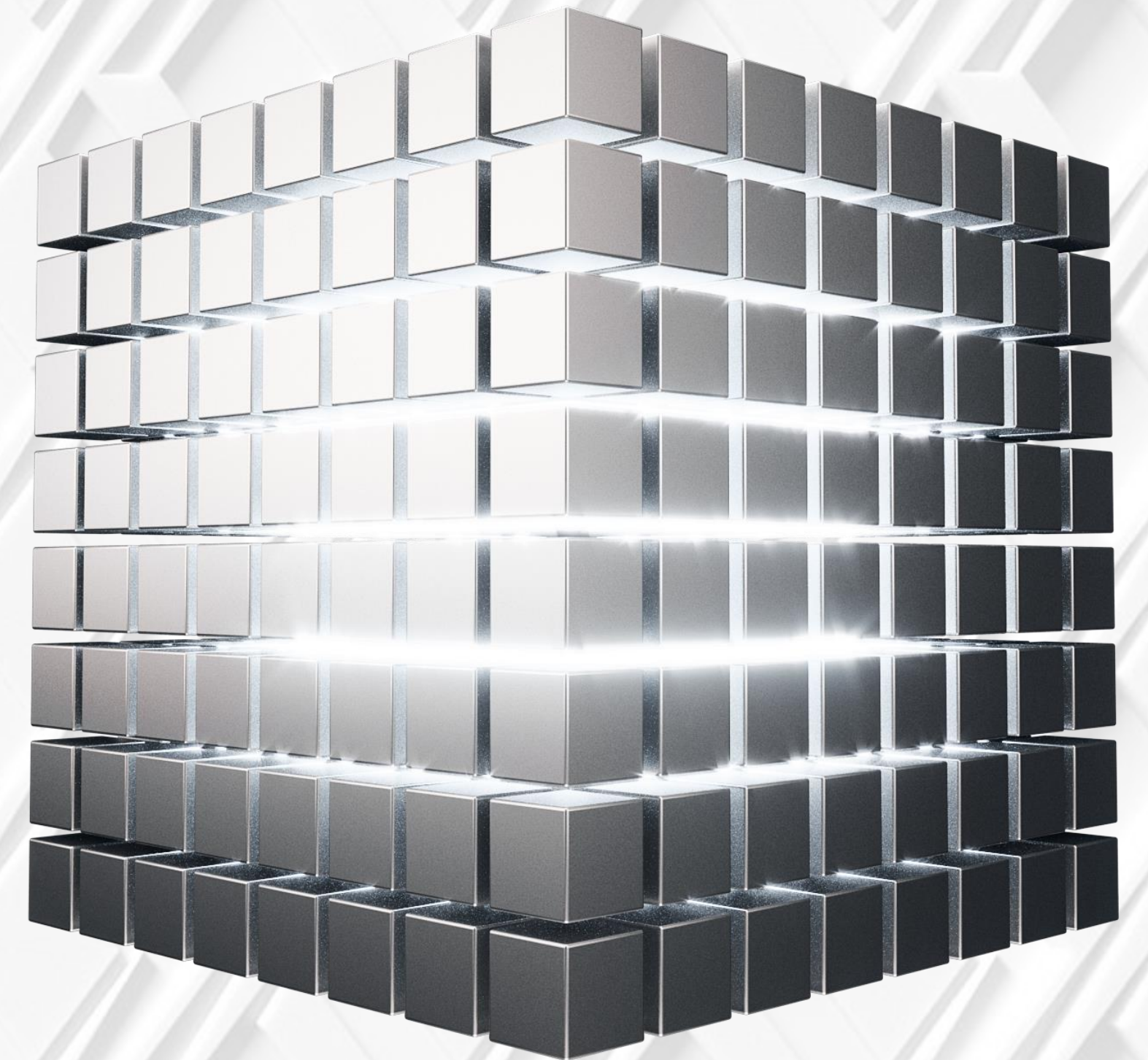


AI DATA TYPES SUPPORTED
INT8, BF16, FP16
UPTO 32X VECTOR RATE

COMPUTE

SCALABILITY

Source: Intel projections



40X INCREASE IN DPFP PER EU

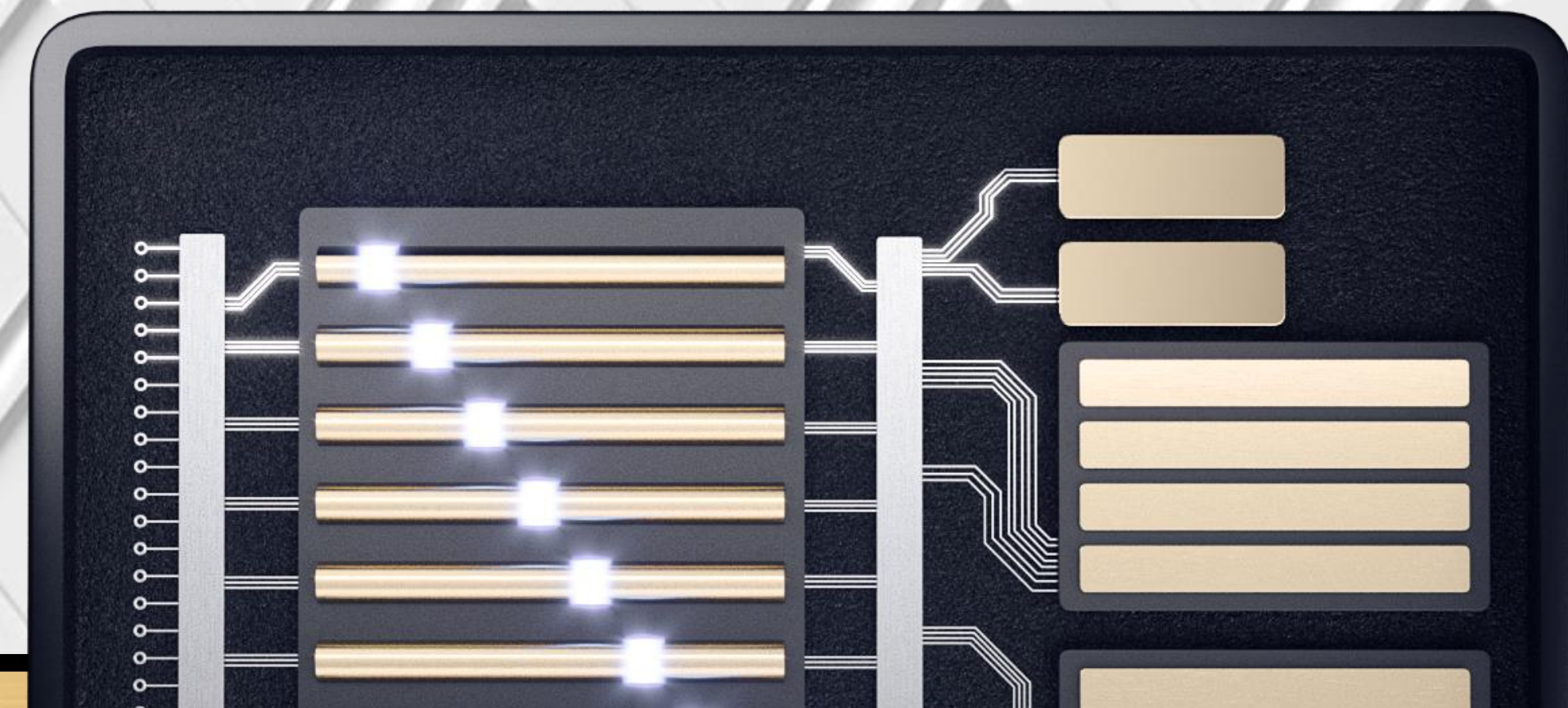
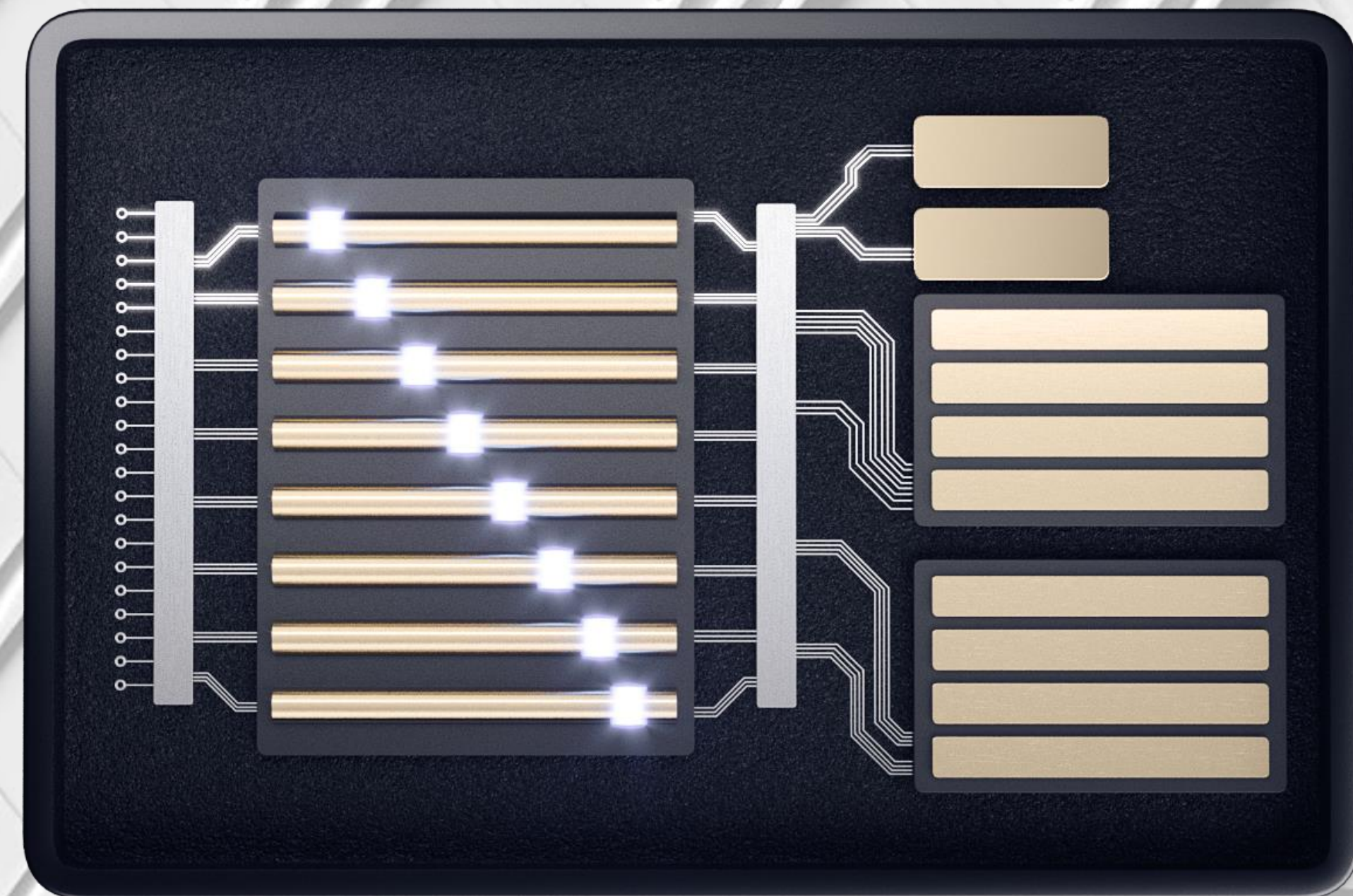
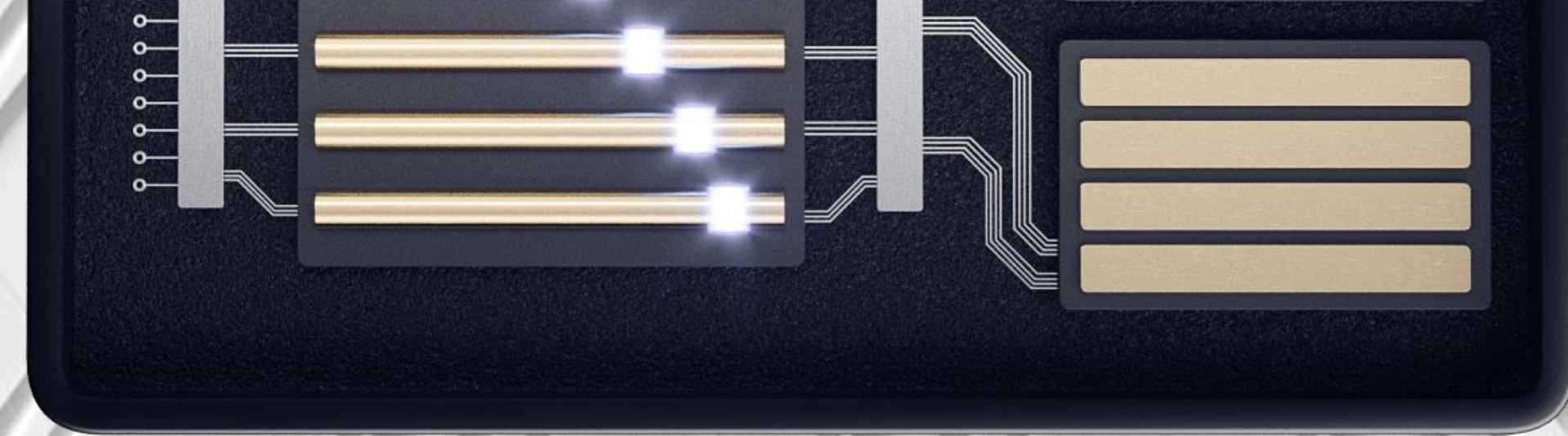


DOUBLE PRECISION
FLOATING POINT
SUPPORT

COMPUTE

HPC PERFORMANCE

Source: Intel projections





HPC FEATURES

COMPUTE

MEMORY

SCALABILITY

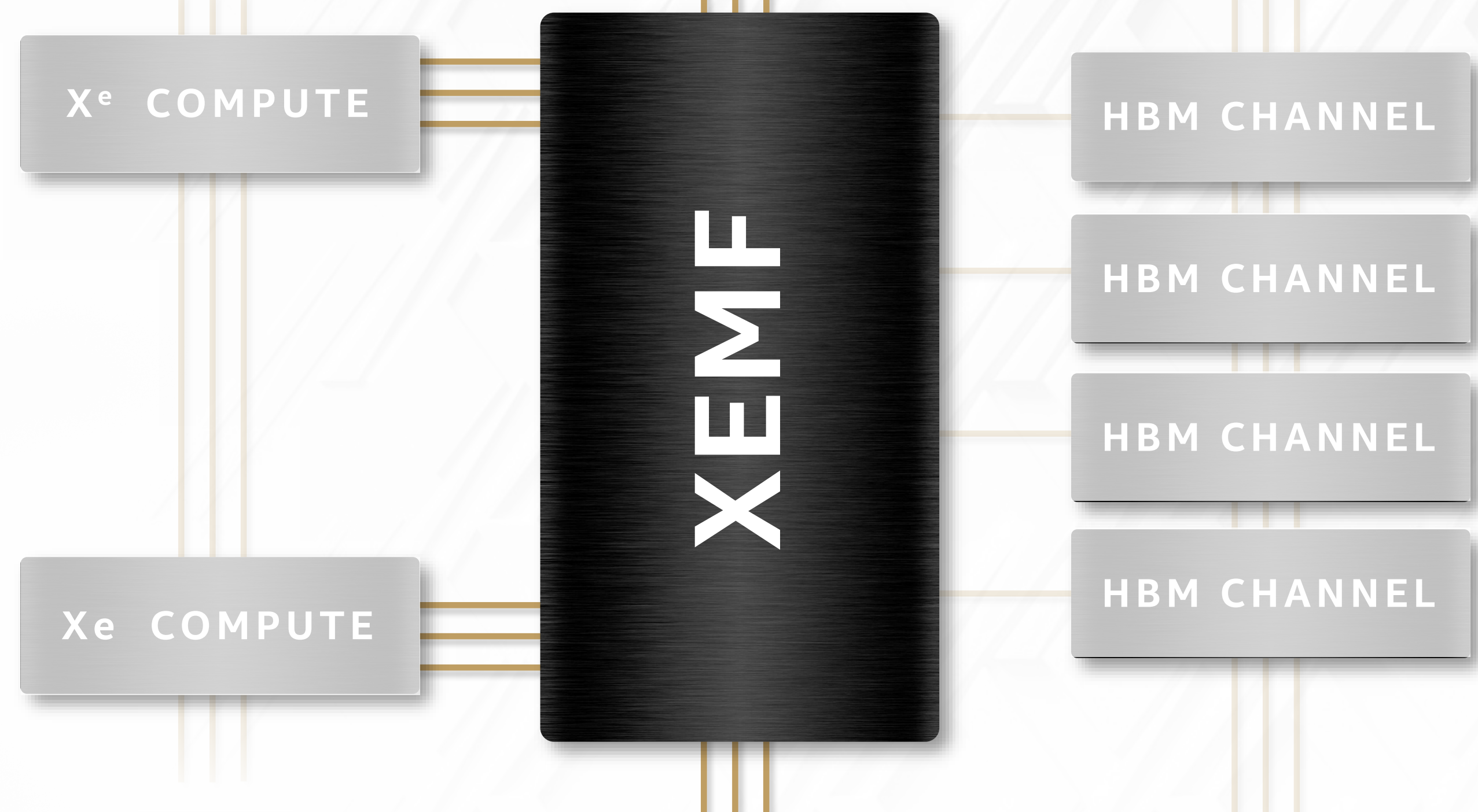
BANDWIDTH

UNIFIED MEMORY

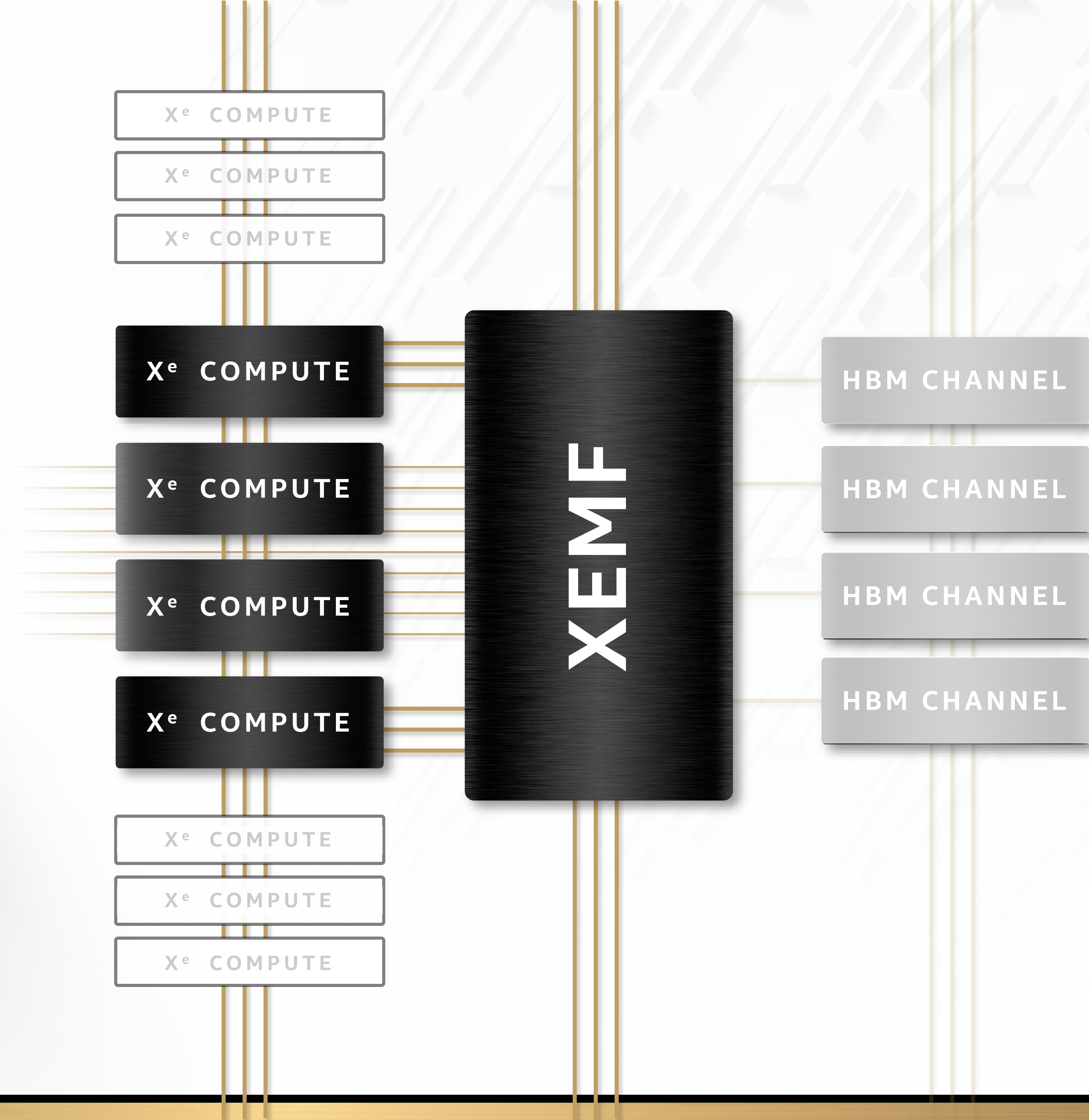


SCALABLE MEMORY FABRIC

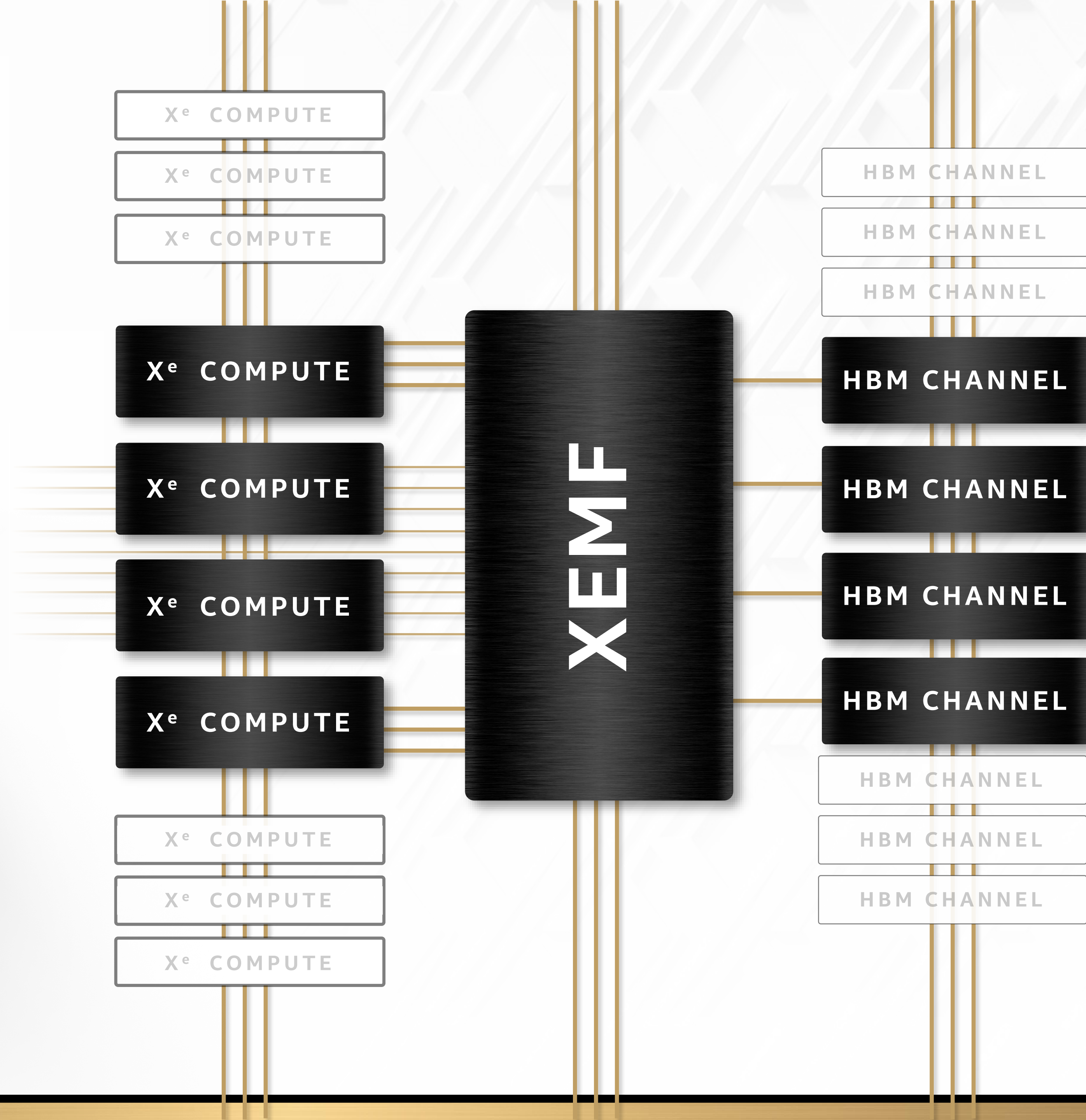
X^e
X HPC



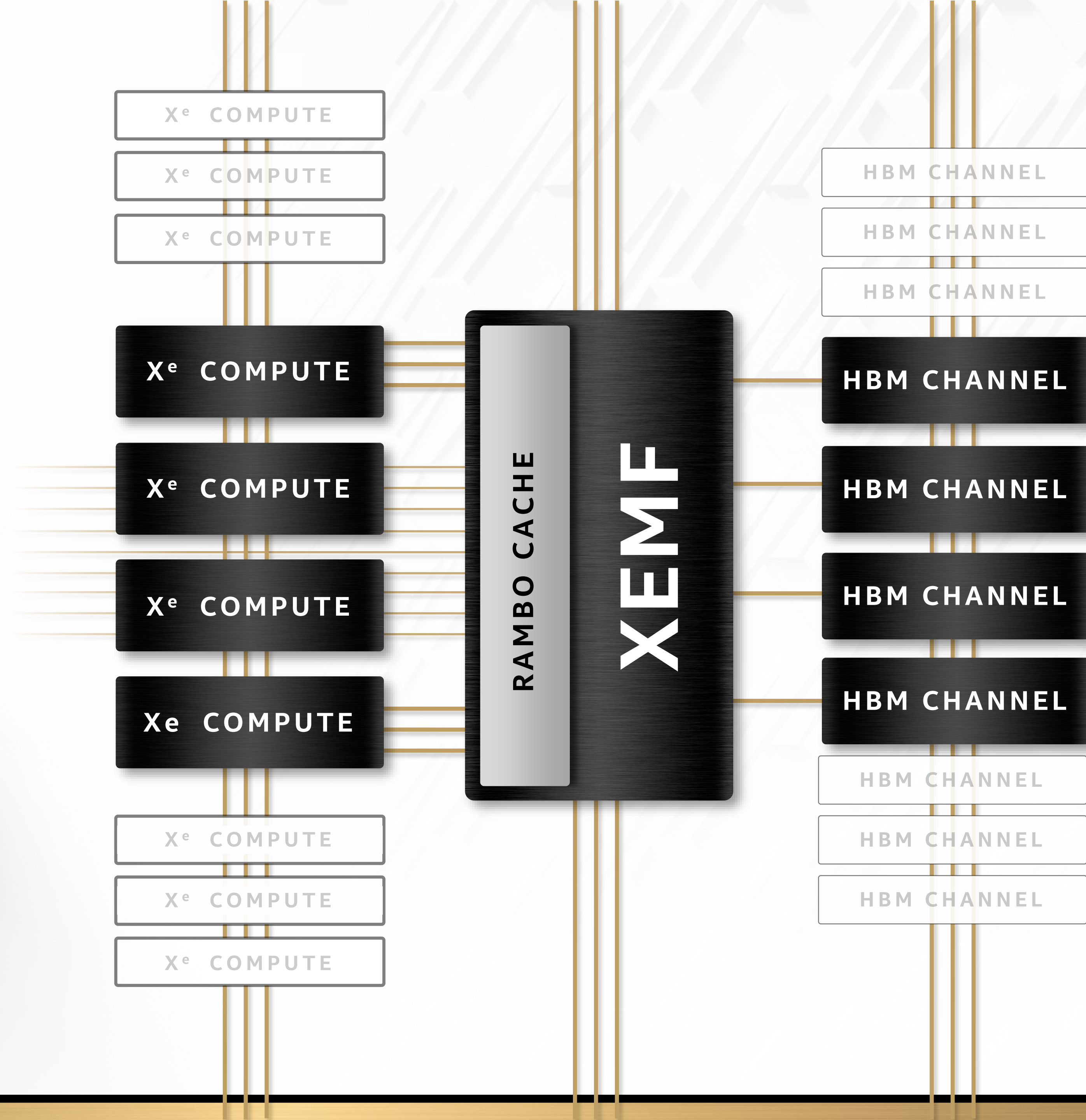
SCALABLE MEMORY FABRIC



SCALABLE MEMORY FABRIC



SCALABLE MEMORY FABRIC



RAMBO
CACHE

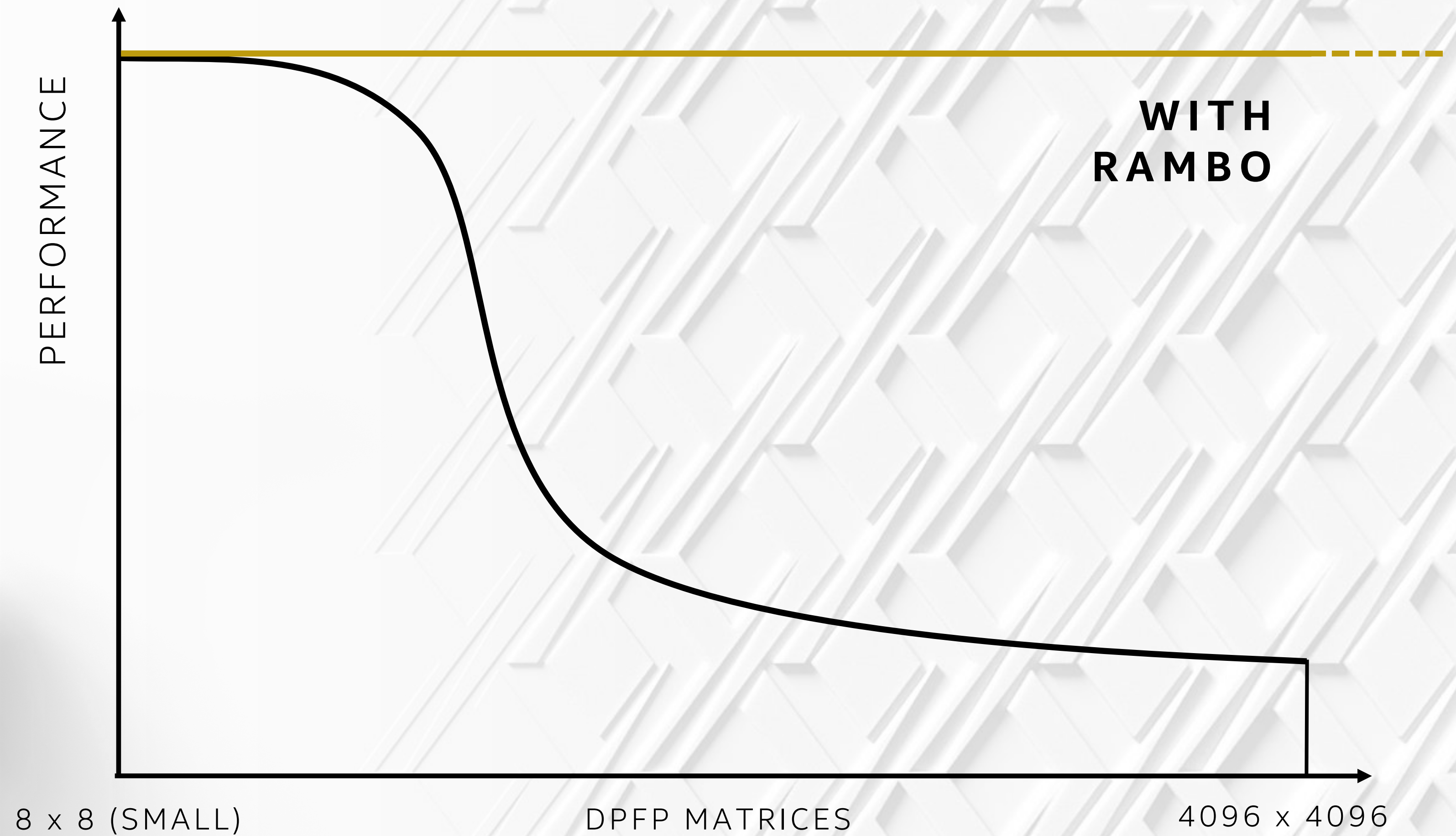
X^e
HPC



RAMBO

CACHE

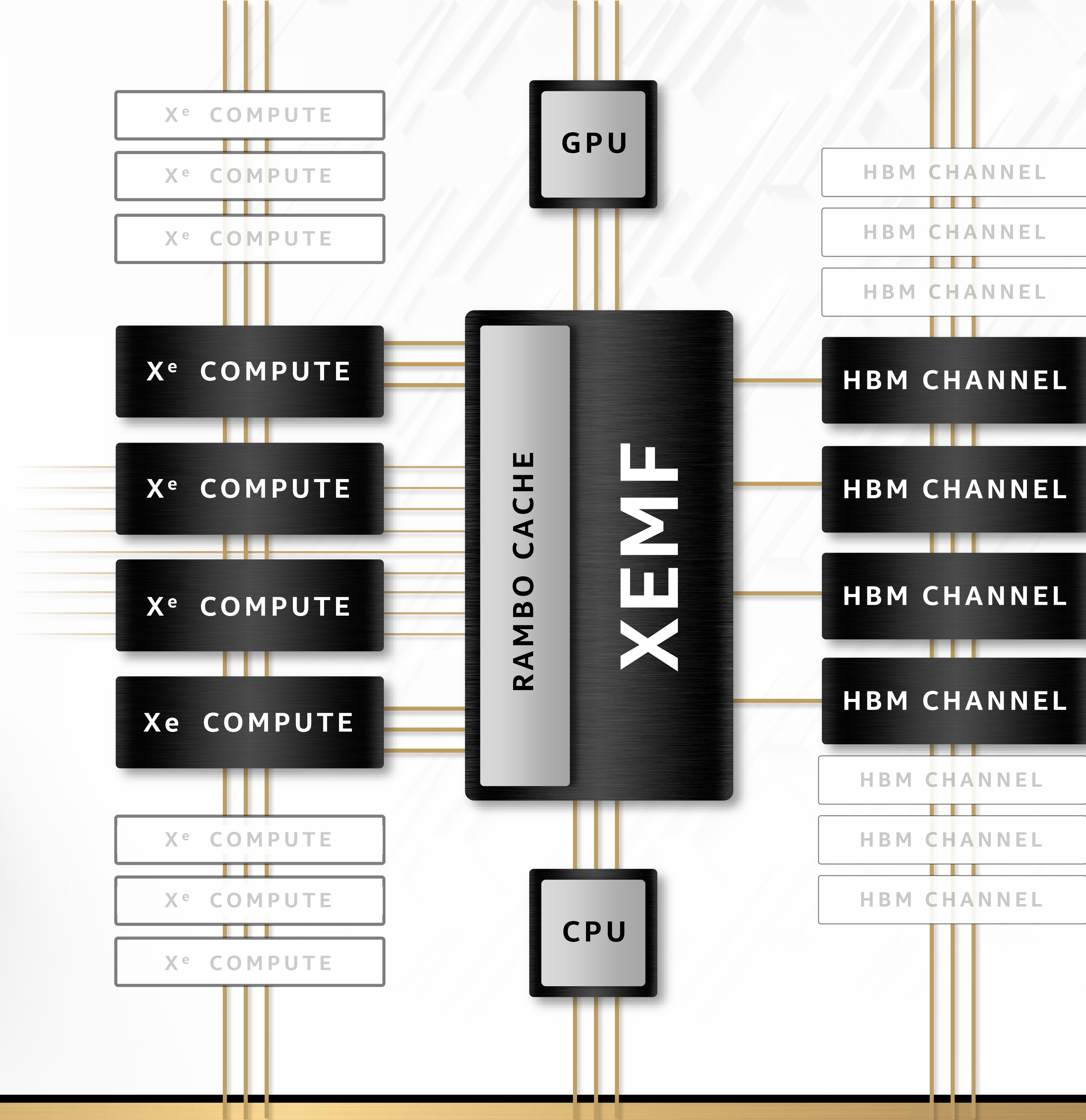
Xe
HPC



Source: Intel projections



UNIFIED COHERENT MEMORY TO CPU AND OTHER GPUS





HPC FEATURES

COMPUTE

SCALABILITY

AI PERFORMANCE

HPC PERFORMANCE

MEMORY

SCALABILITY

BANDWIDTH

UNIFIED MEMORY



PATH TO EXASCALE GPU

EXASCALE CHALLENGES



COMPUTE DENSITY

MEMORY

CONNECTIVITY

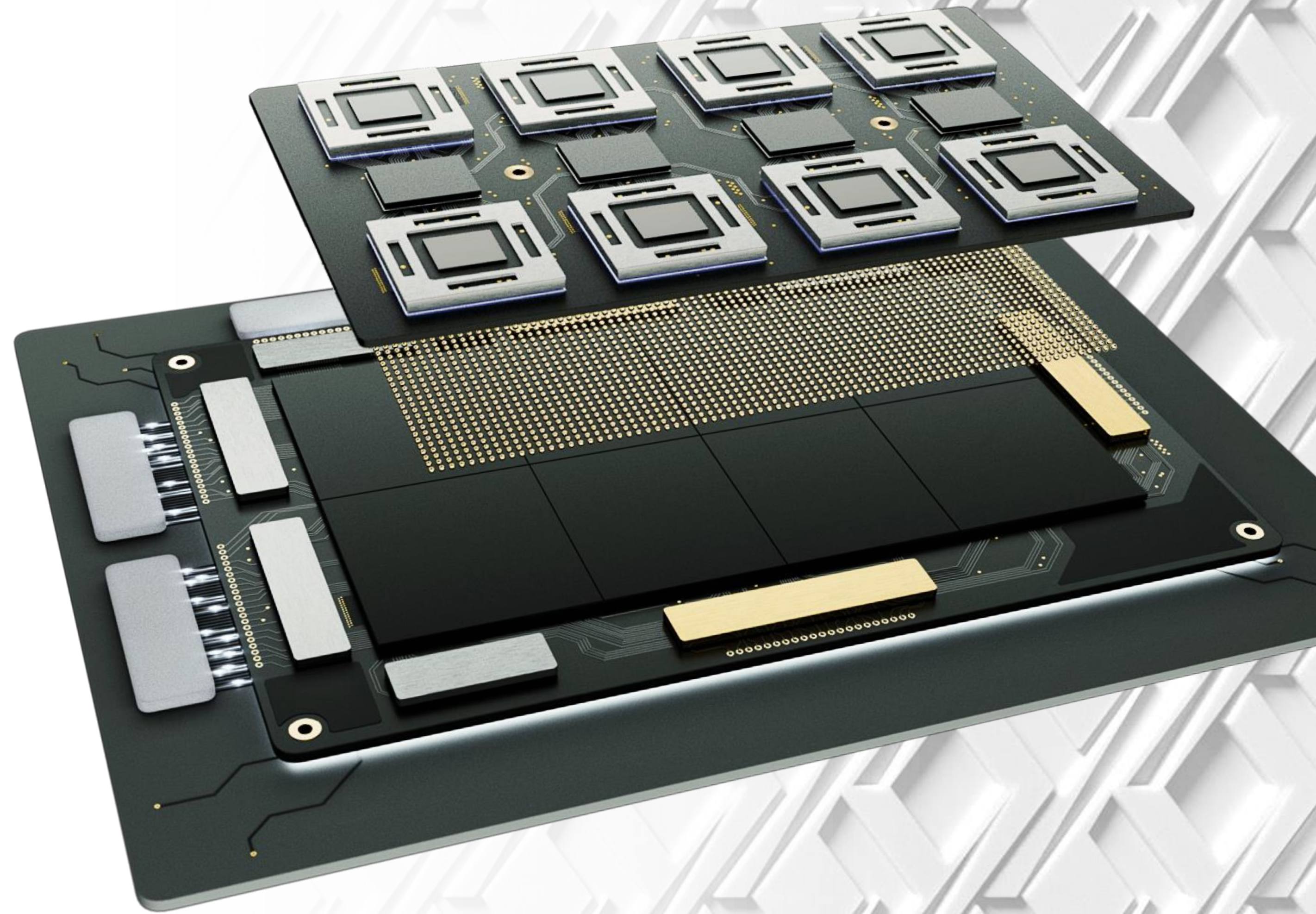
RELIABILITY

TECHNOLOGY FOR EXASCALE

NEED HUGE LEAP IN PERF/WATT
AND PERF/MM²

INTEL NEXT GEN 7nm
PROCESS &
FOVEROS™ PACKAGING

COMPUTE DENSITY

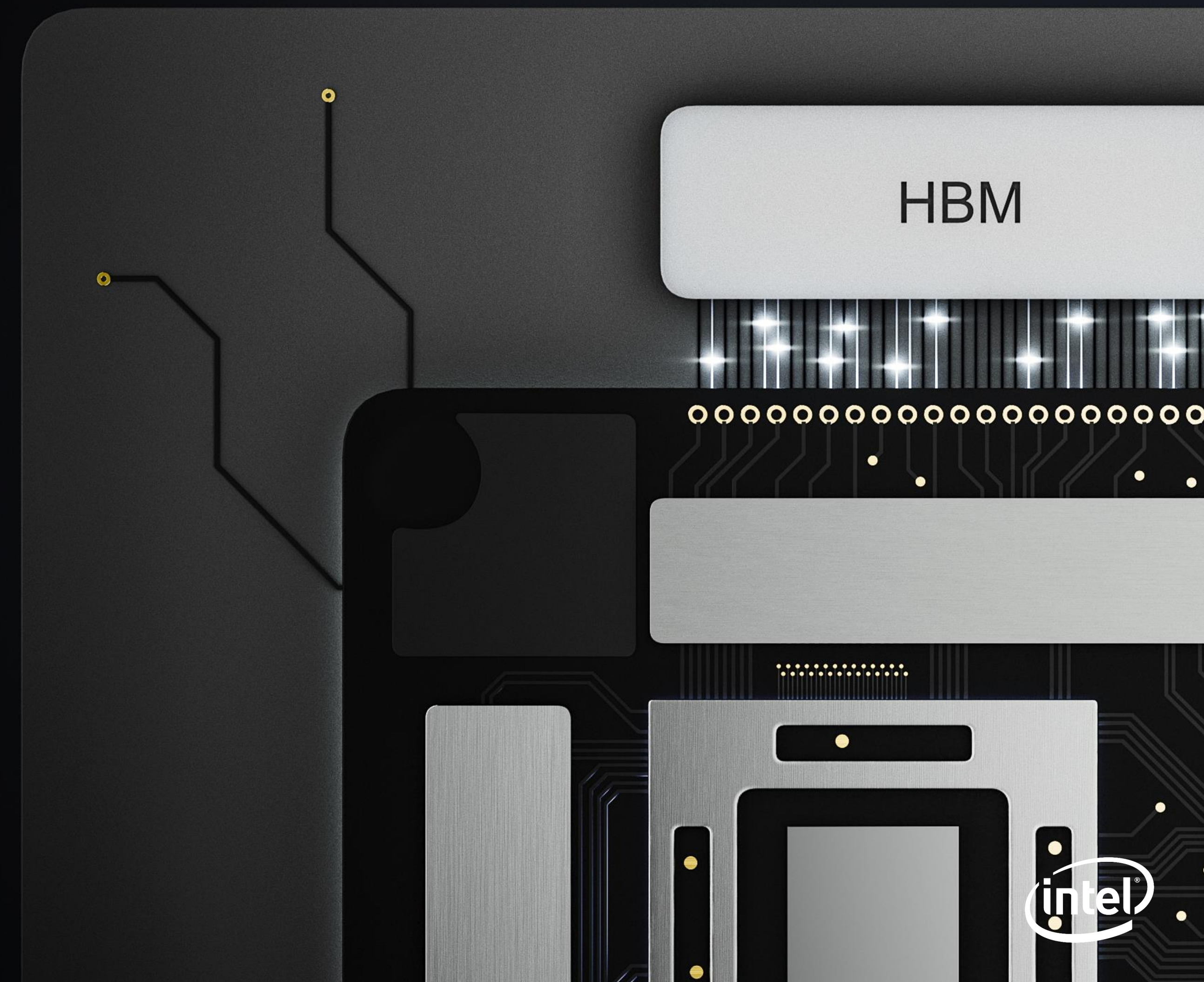


MEMORY FOR EXASCALE

NEED HUGE LEAP IN
BANDWIDTH/WATT
& FOOTPRINT/MM²

EMIB FOR HBM
&
FOVEROS™ FOR RAMBO CACHE

MEMORY



CONNECTIVITY FOR EXASCALE

X^e LINK

SCALE OUT TO MANY
GPUS/NODE
UNIFIED MEMORY
CXL BASED

CONNECTIVITY



RELIABILITY FOR EXASCALE

XEON™ CLASS RAS

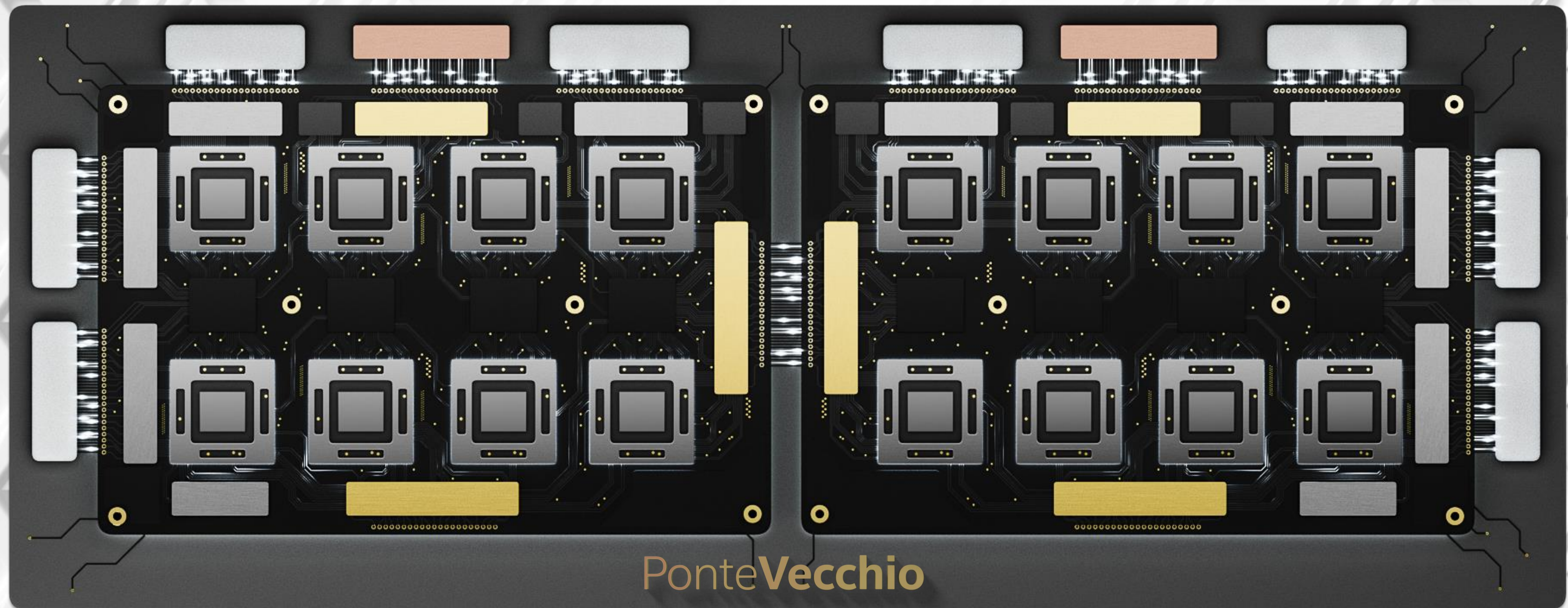
IN-FIELD REPAIR

ECC, PARITY ACROSS ALL
MEMORY AND CACHES

CONNECTIVITY



EXASCALE GPU



PonteVecchio

X^e
HPC



Aurora

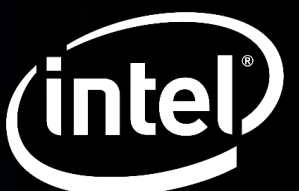


Aurora

US FIRST EXASCALE COMPUTER

RICK STEVENS

ASSOCIATE DIRECTOR FOR COMPUTING, ENVIRONMENT
AND LIFE SCIENCES AT ARGONNE NATIONAL
LABORATORY



AURORA – BRINGING IT ALL TOGETHER

LEADERSHIP PERFORMANCE

For HPC, Data Analytics, AI

UNIFIED MEMORY ARCHITECTURE

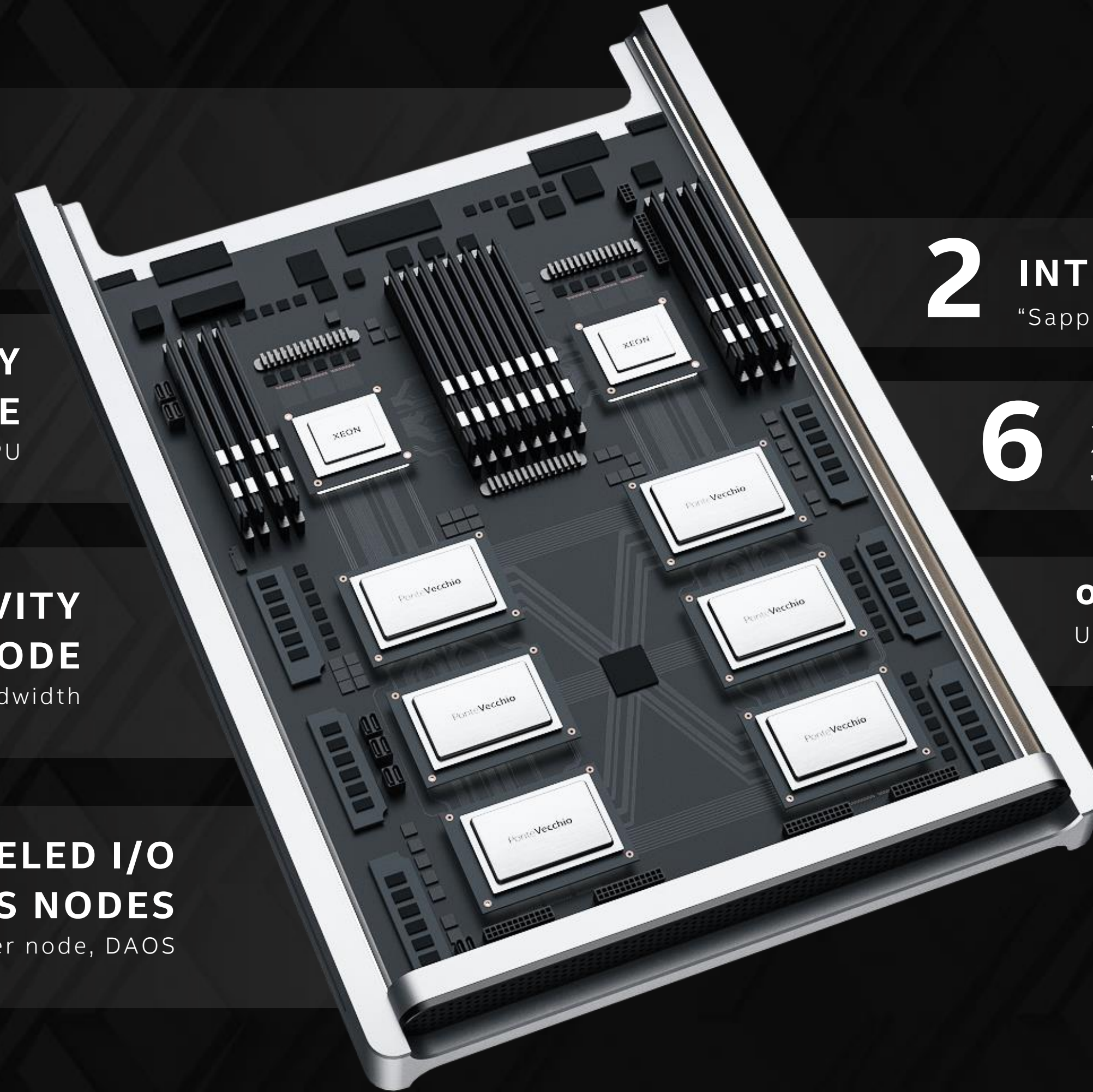
Across CPU & GPU

ALL-TO-ALL CONNECTIVITY WITHIN NODE

Low latency, high bandwidth

UNPARALLELED I/O SCALABILITY ACROSS NODES

8 fabric endpoints per node, DAOS

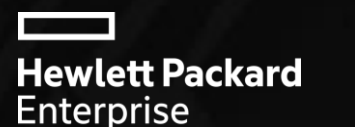


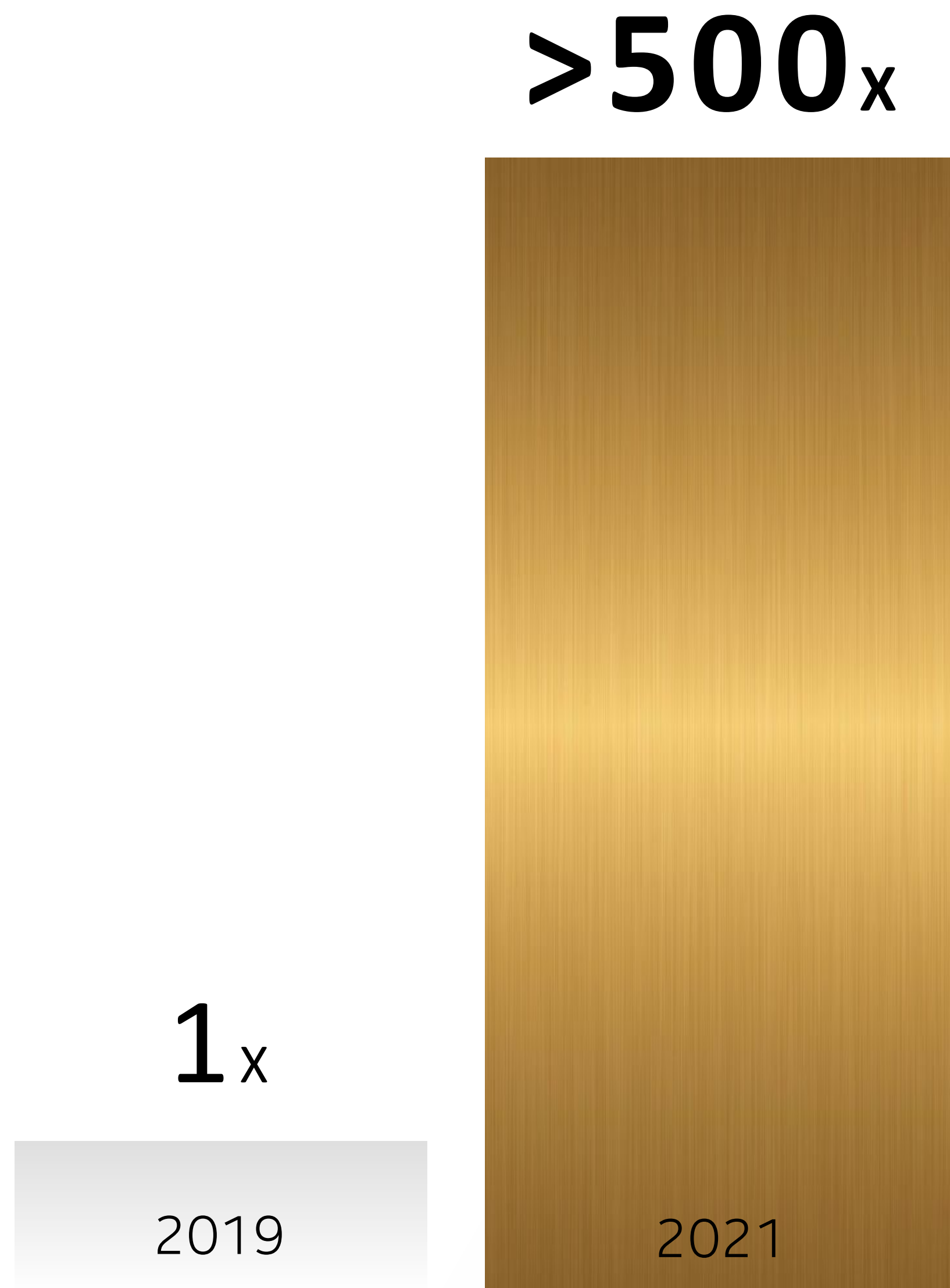
2 INTEL XEON™ SCALABLE PROCESSORS
“Sapphire Rapids”

6 X^e ARCHITECTURE BASED GPUS
“Ponte Vecchio”

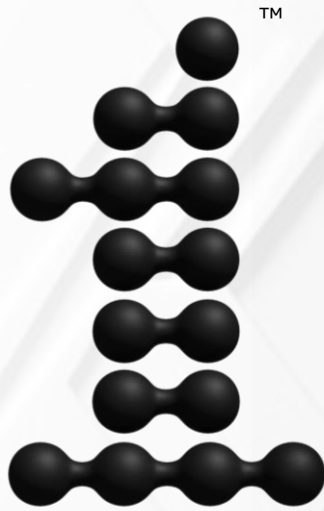
oneAPI
Unified programming model

DELIVERED IN 2021

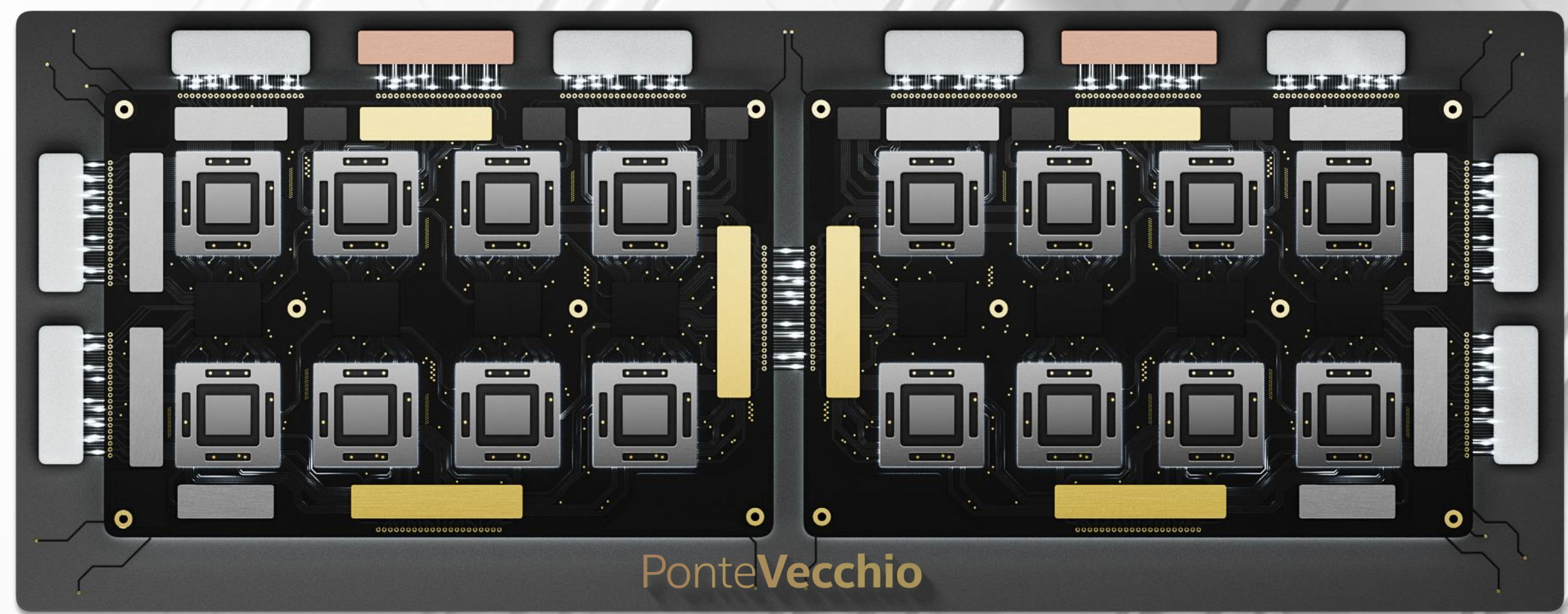




PERFORMANCE INCREASE PER NODE


oneAPI

X^e
HPC



Source: Intel projections





THANK YOU!

R A J A K O D U R I



Notices & Disclaimers

Intel® technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. No computer system can be absolutely secure. Check with your system manufacturer or retailer or learn more at [intel.com](https://www.intel.com).

Performance results are based on testing as of the dates shown in configurations and may not reflect all publicly available security updates. See configuration disclosure for details. No product or component can be absolutely secure.

Software and workloads used in performance tests may have been optimized for performance only on Intel® microprocessors. Performance tests, such as SYSmark* and MobileMark*, are measured using specific computer systems, components, software, operations, and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products. For more information, go to www.intel.com/benchmarks.

Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel® microprocessors. These optimizations include SSE2 and SSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel.

Microprocessor-dependent optimizations in this product are intended for use with Intel® microprocessors. Certain optimizations not specific to Intel® microarchitecture are reserved for Intel® microprocessors. Please refer to the applicable product user and reference guides for more information regarding the specific instruction sets covered by this notice. Notice revision #20110804.

Cost reduction scenarios described are intended as examples of how a given Intel®-based product, in the specified circumstances and configurations, may affect future costs and provide cost savings. Circumstances will vary. Intel does not guarantee any costs or cost reduction.

No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document.

Intel does not control or audit third-party benchmark data or the websites referenced in this document. You should visit the referenced website and confirm whether referenced data are accurate.

© Intel Corporation. Intel, Xeon, Optane, AVX, and DL Boost are trademarks of Intel Corporation in the U.S. and/or other countries.

*Other names and brands may be claimed as property of others.



Business Forecast: Statements in this document that refer to Intel's plans and expectations for the quarter, the year, and the future, are forward-looking statements that involve a number of risks and uncertainties. A detailed discussion of the factors that could affect Intel's results and plans is included in Intel's SEC filings, including the annual report on Form 10-K.

