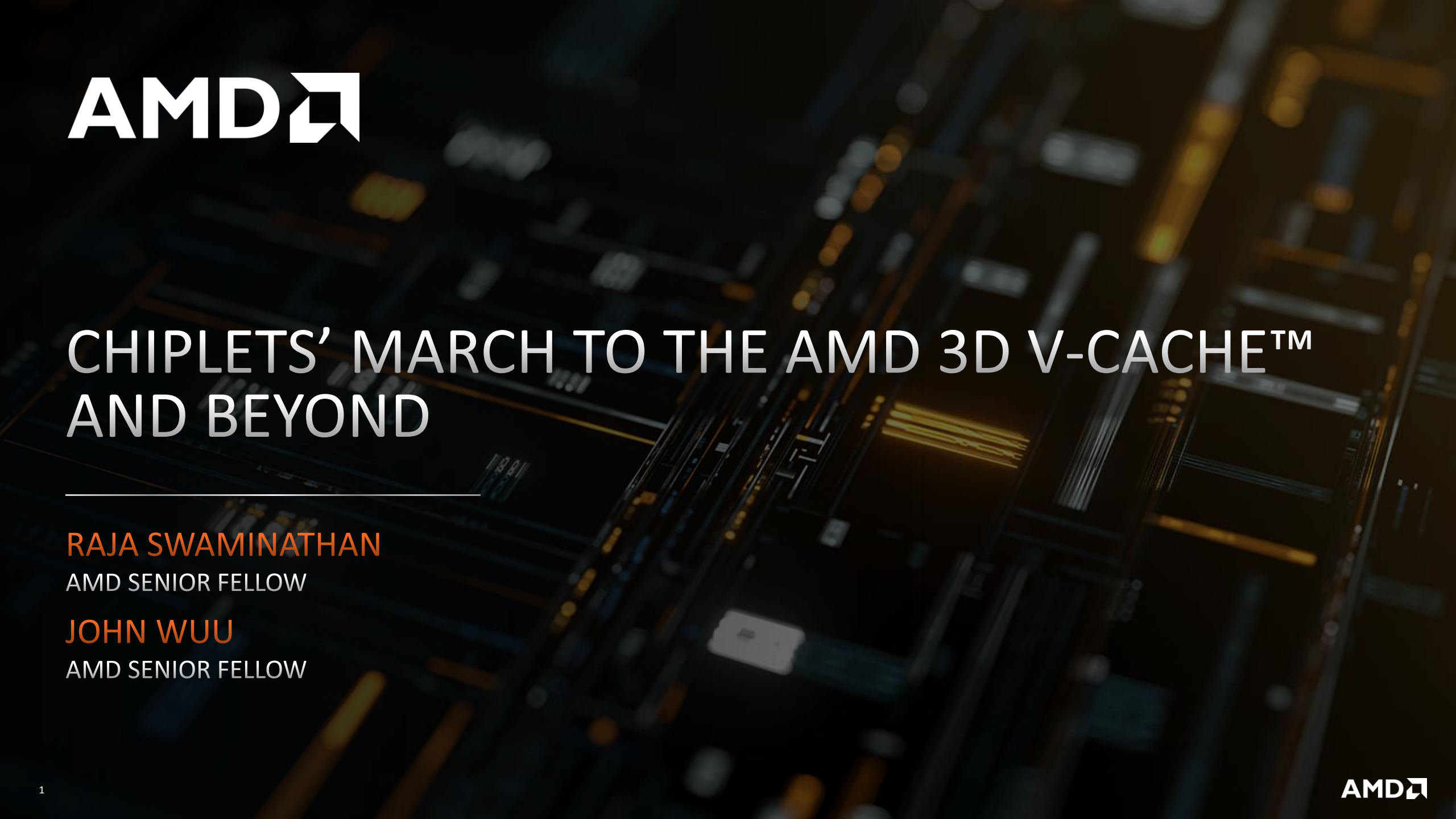


CHIPLETS' MARCH TO THE AMD 3D V-CACHE™ AND BEYOND

RAJA SWAMINATHAN
AMD SENIOR FELLOW

JOHN WUU
AMD SENIOR FELLOW

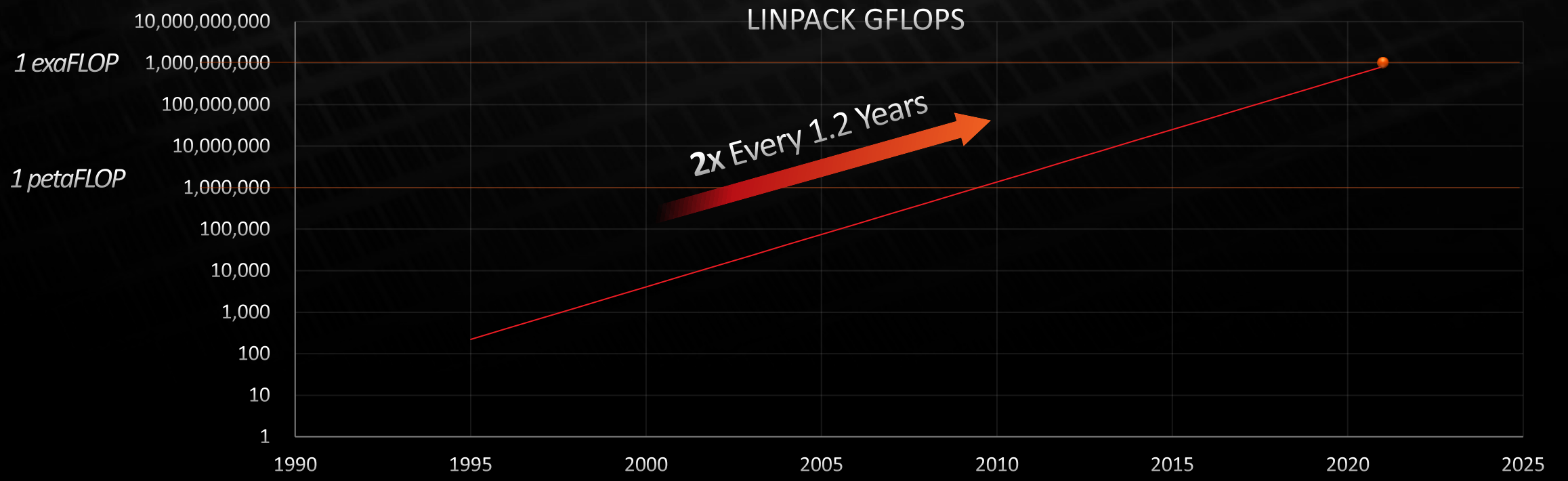
CAUTIONARY STATEMENT

This presentation contains forward-looking statements concerning Advanced Micro Devices, Inc. (AMD) such as the features, functionality, performance, availability, timing and expected benefits of AMD products; TAM for data center, PCs, embedded and gaming; and technology trends, innovation and roadmaps, which are made pursuant to the Safe Harbor provisions of the Private Securities Litigation Reform Act of 1995. Forward-looking statements are commonly identified by words such as "would," "may," "expects," "believes," "plans," "intends," "projects" and other terms with similar meaning. Investors are cautioned that the forward-looking statements in this presentation are based on current beliefs, assumptions and expectations, speak only as of the date of this presentation and involve risks and uncertainties that could cause actual results to differ materially from current expectations. Such statements are subject to certain known and unknown risks and uncertainties, many of which are difficult to predict and generally beyond AMD's control, that could cause actual results and other future events to differ materially from those expressed in, or implied or projected by, the forward-looking information and statements. Investors are urged to review in detail the risks and uncertainties in AMD’s Securities and Exchange Commission filings, including but not limited to AMD’s most recent reports on Forms 10-K and 10-Q.

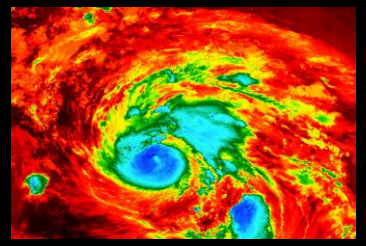
AMD does not assume, and hereby disclaims, any obligation to update forward-looking statements made in this presentation, except as may be required by law.

RELENTLESS DEMAND FOR SCIENTIFIC COMPUTING

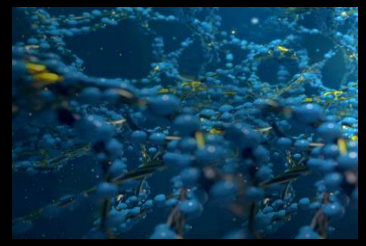
WORLD'S FASTEST SUPERCOMPUTERS



SPACE EXPLORATION

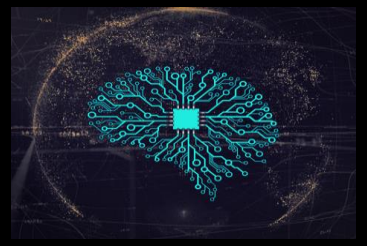


CLIMATE CHANGE

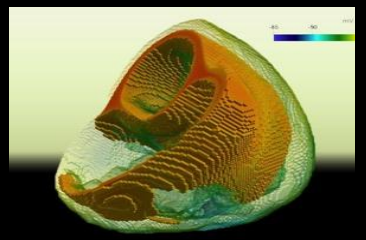


CHEMICAL SCIENCES

ENERGY SOLUTIONS



MACHINE LEARNING



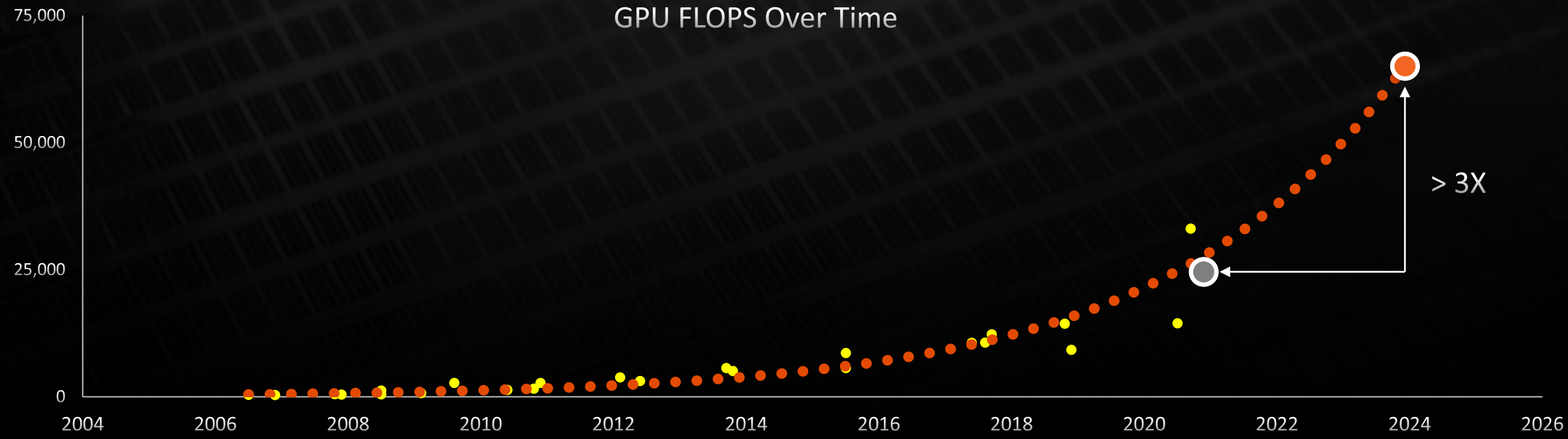
REAL TIME SIMULATION

Naffziger, DAC Keynote, 2021

AMD

INSATIABLE DEMAND FOR GRAPHICS COMPUTE

… CONTINUES TO INCREASE EXPONENTIALLY

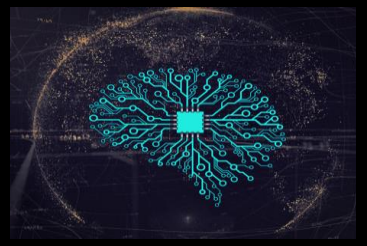


PERSONAL GAMING

ENTERTAINMENT

PROFESSIONAL GAMING

CREATORS



MACHINE LEARNING

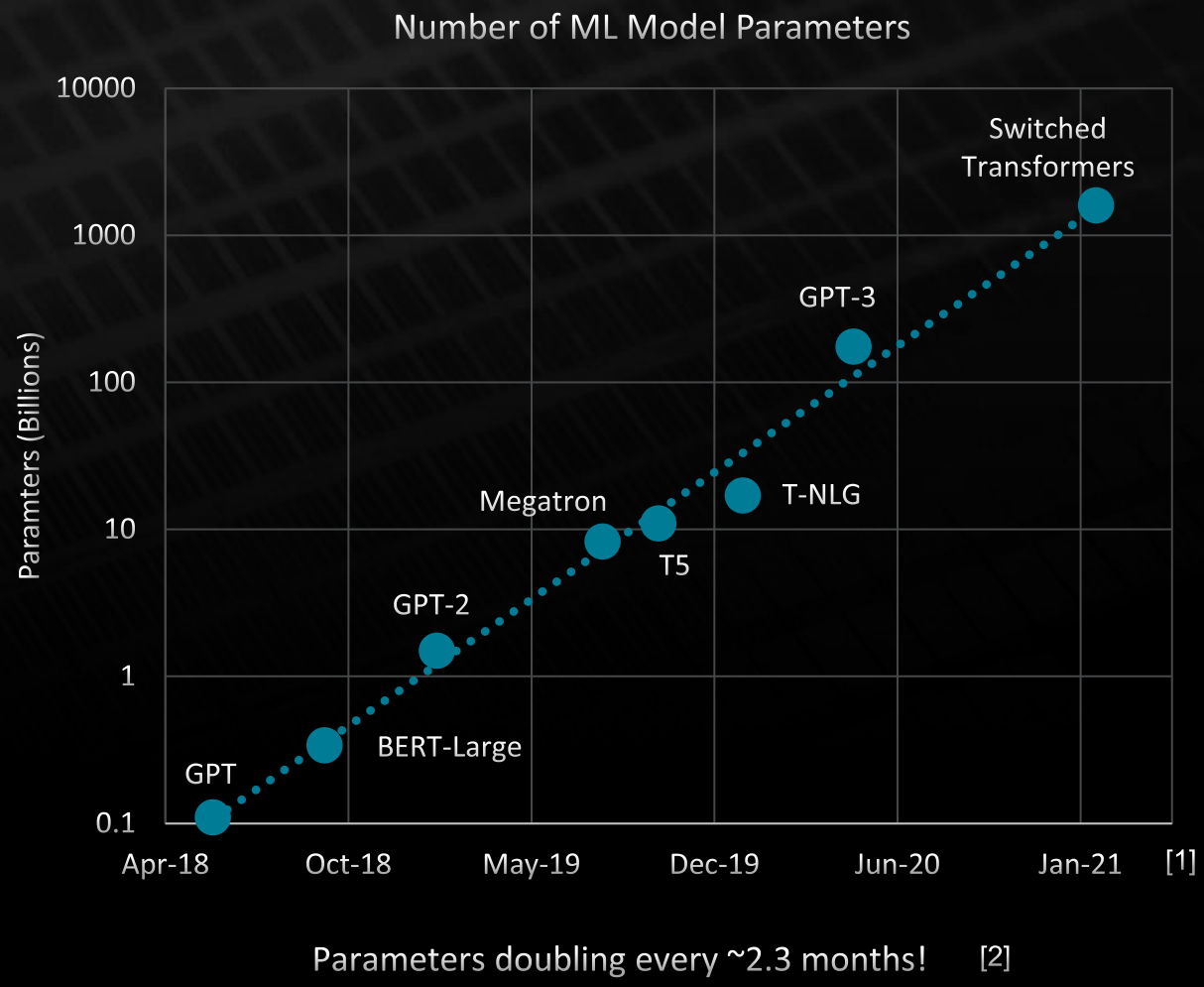
PHOTOREALISTIC GAMING

Naffziger, DAC Keynote, 2021

AMD

GROWTH OF MACHINE LEARNING

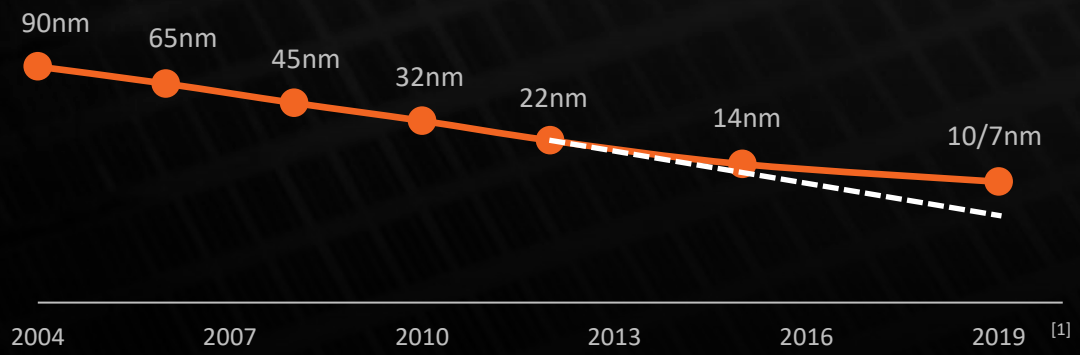
- Exponentially growing model sizes driving massive growth in compute and memory
- Underlying algorithms are evolving rapidly
- Required compute doubling every ~3.4 months[2]
- Ever-growing capabilities that deliver major quality of life improvements such as multimodal learning



[1] Based on published parameter counts of leading training models, [2] https://openai.com/blog/ai-and-compute Naffziger, DAC Keynote, 2021

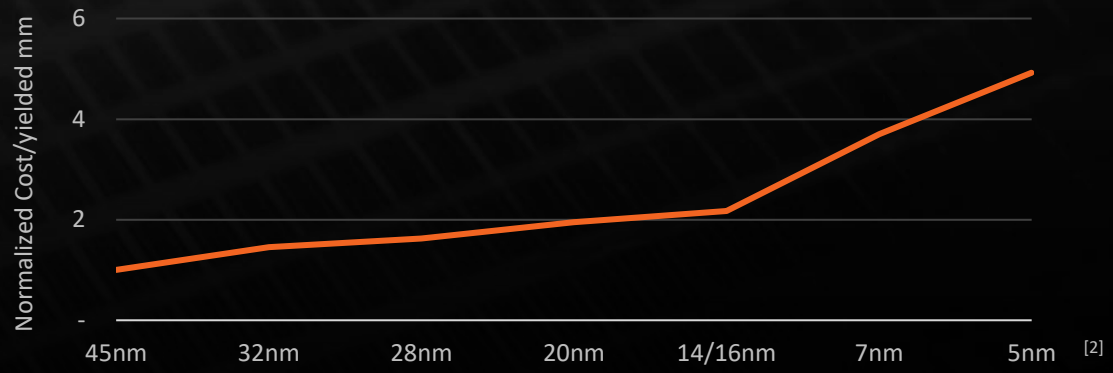
TECHNOLOGY HEADWINDS TO MEETING THE DEMAND

Slowing of Moore's Law



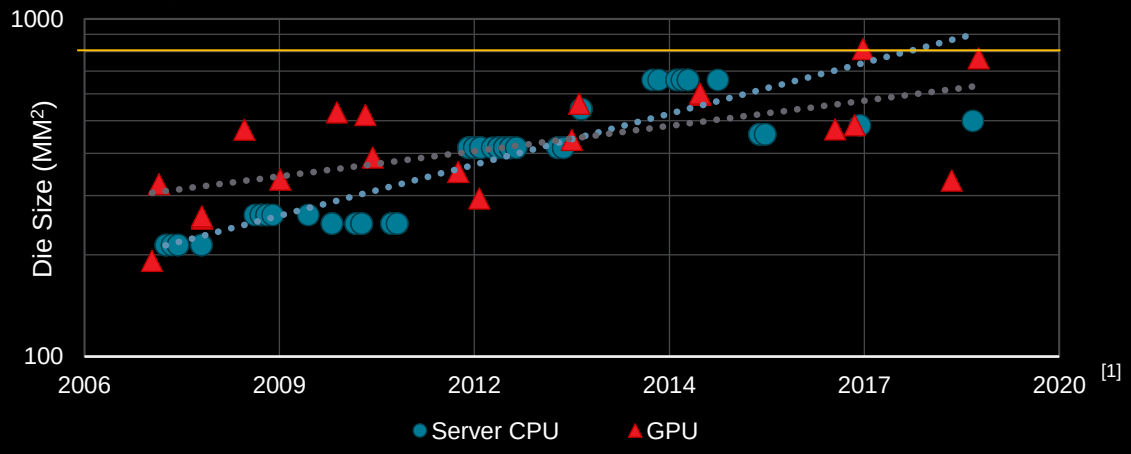
[1]

Increasing Cost

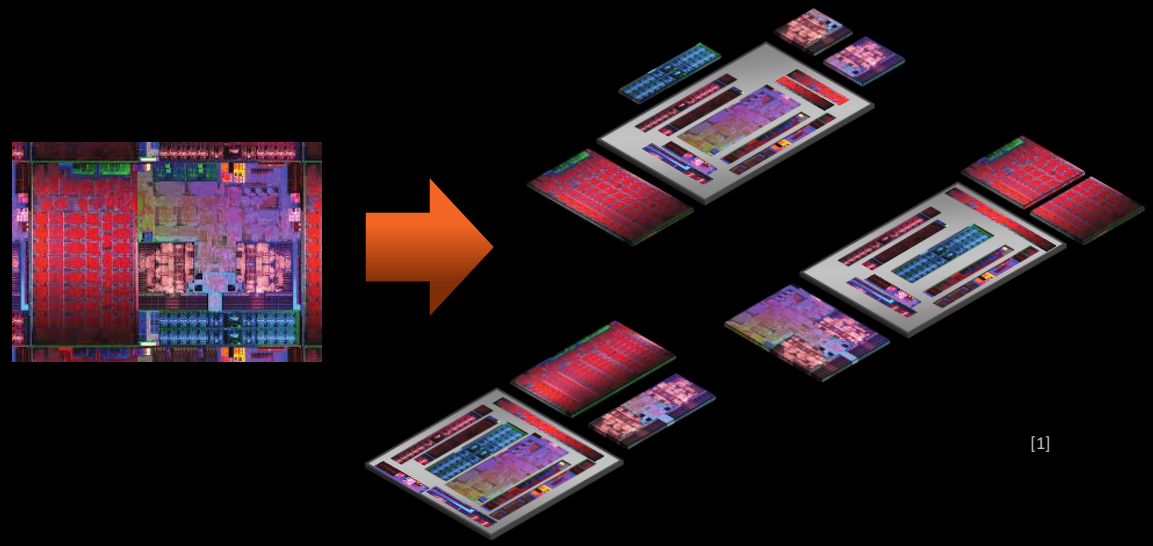


[2]

Reticle Limit



[1]

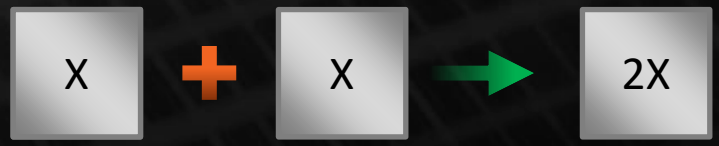

[1]

[1] Naffziger, VLSI Short Course, 2020, [2] Cost per yielded mm² for a 250mm² die

AMD

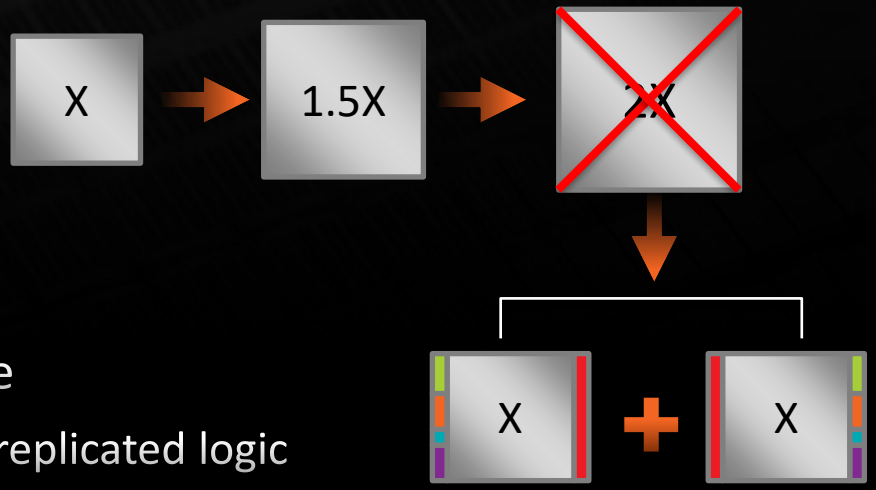
CHIPLETS BACKGROUND

Historically, except for the largest systems,
Moore's Law was sufficient to meet compute needs



Current trends require a new approach

One Generation Later



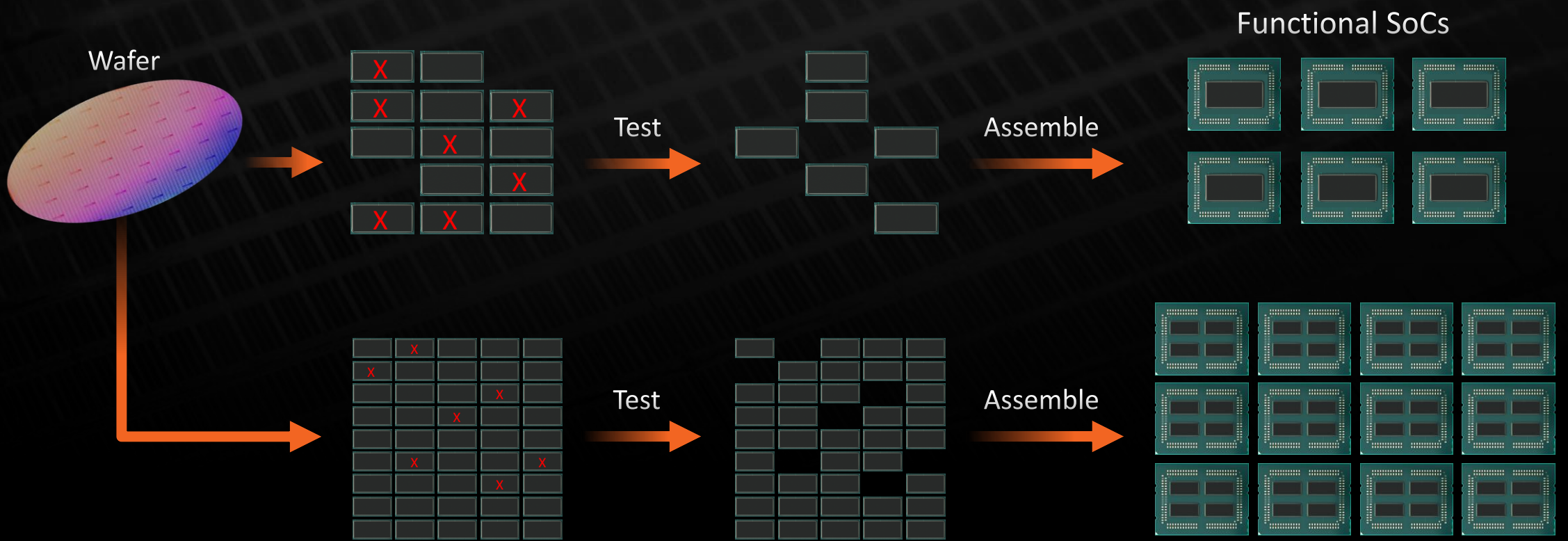
2X Device Functionality
Costs > 2X Silicon Area

However, chiplets are not free

- Additional area for interfaces, replicated logic
- Additional design effort, complexity
- Past methodologies less suited for chiplets

Naffziger, DAC Keynote, 2021

HIGH-LEVEL APPROACH TO CHIPLETS



Many More Functional SoCs

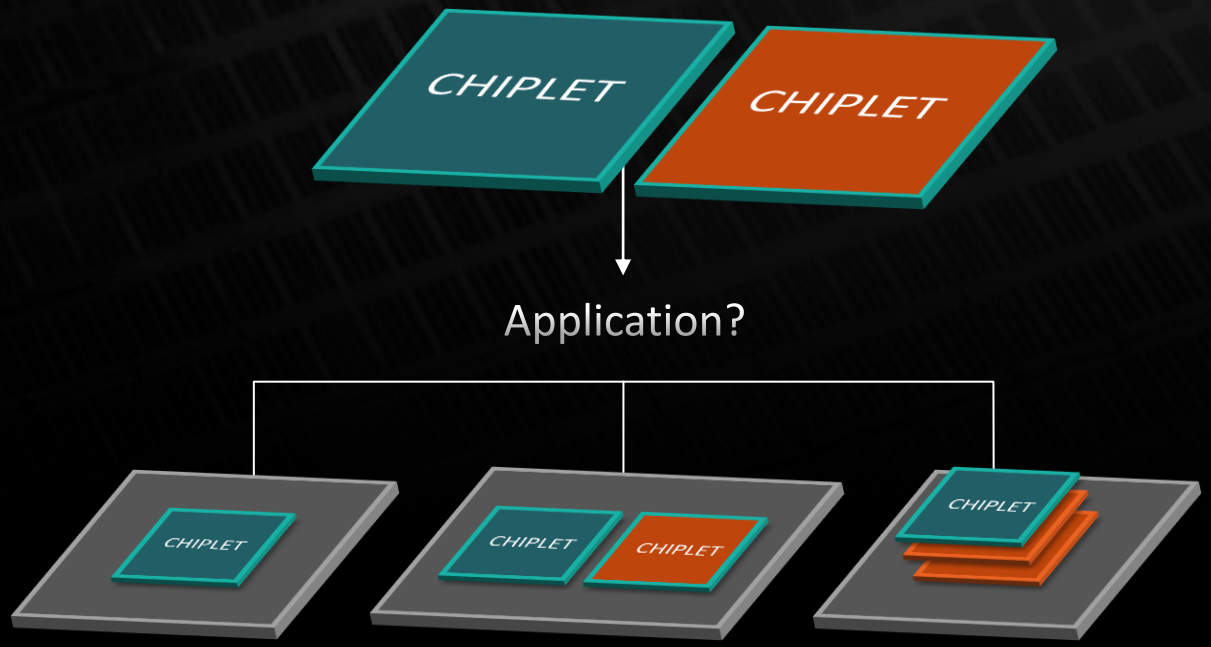
Ability to mix and match at a finer grained level

Naffziger, DAC Keynote, 2021

AMD

MODULAR ARCHITECTURE GOALS

ENABLING A MORE FLEXIBLE APPROACH

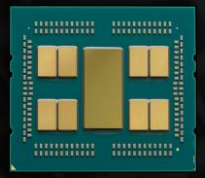


- We want to build tailored products for specific markets by mixing and matching chiplet types
- We can now specialize a domain specific chiplet and include more or fewer of them for a given product
- More domain-specific products at higher yields … provided we can build low-overhead chiplets

Naffziger, DAC Keynote, 2021

PACKAGE ARCHITECTURES FOR CHIPLETS

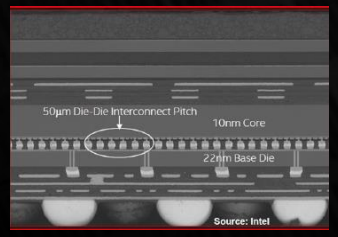
MCM



INFO-R

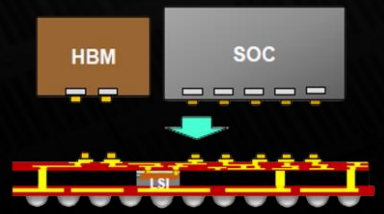
INFO-L

Foveros



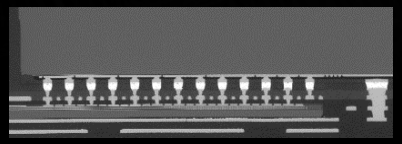
Intel: Foveros

CoWoS-L

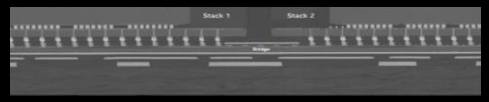


TSMC: INFO-R/-L, CoWoS-L

EMIB

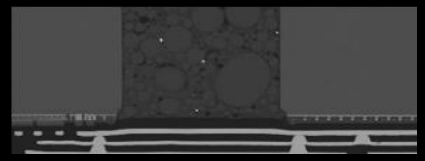


Co-EMIB



Intel: EMIB and Co-EMIB

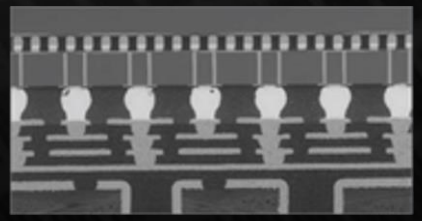
FoCoS



ASE: FoCoS

Micro Bump 3D

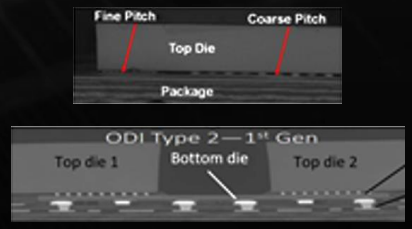
Si Interposer + TSV



AMD Fiji GPU

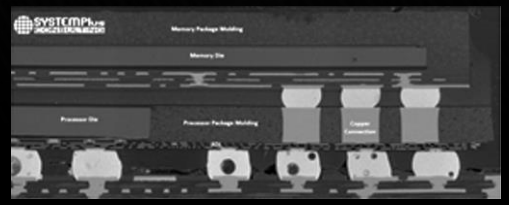
Courtesy: TechSearch

Foveros-ODI



Intel: Omni-directional interconnect

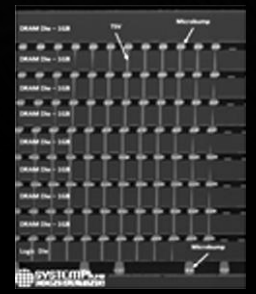
INFO-POP



Apple A10 on FO+POP

Courtesy: SystemPlus Consulting

W2W stacking+ TSV+uBump



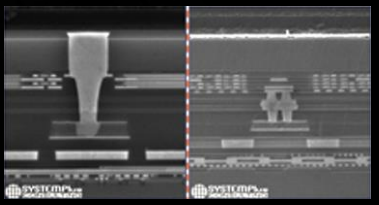
Samsung: HBM2 memory

Courtesy: SystemPlus Consulting

WoW

W2W F2F di-electric bonding+ TSV

W2W F2F hybrid bonding w/o TSV



Sony: CMOS Image Sensors

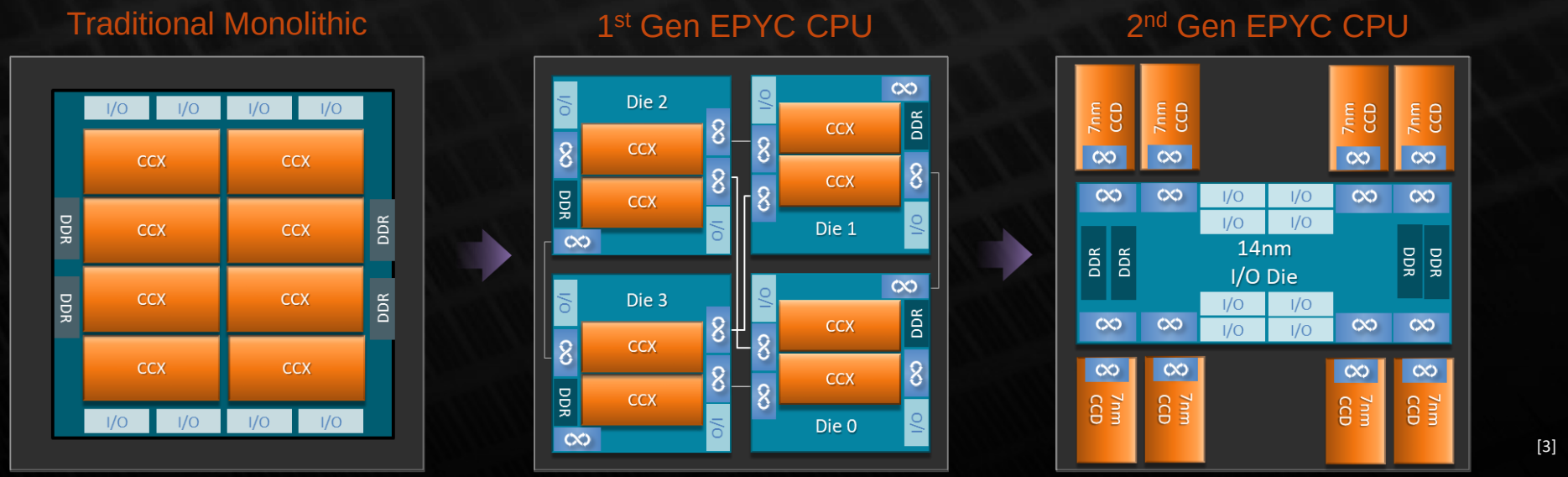
Courtesy: SystemPlus Consulting

No single package architecture works for all products - choice based on product PPAC

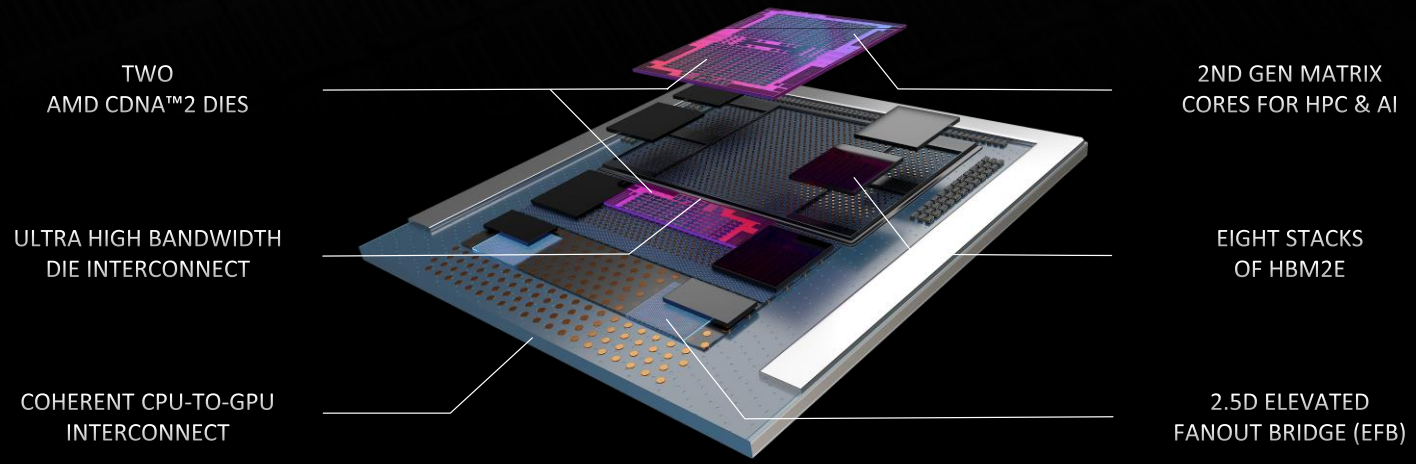
Images courtesy: Intel, TSMC, Amkor, ASE from public material
Naffziger, DAC Keynote, 2021

AMD

AMD'S CHIPLETS EVOLUTION



[3]



[4]

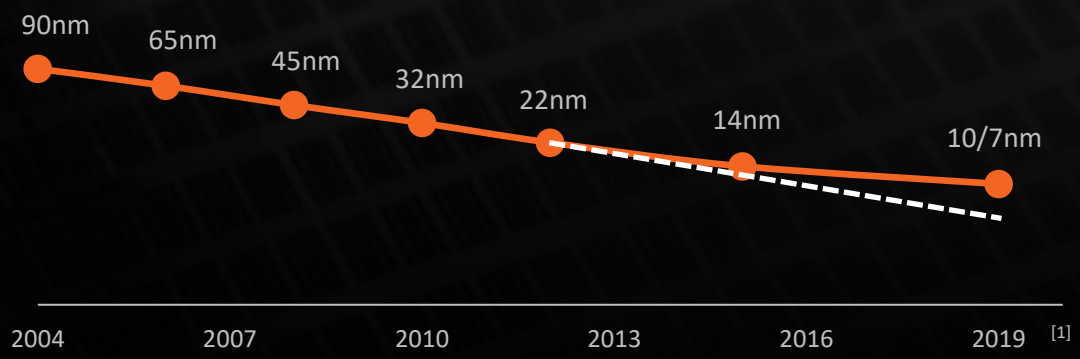
AMD INSTINCT™ MI200 OAM SERIES

[3] Naffziger, ECTC, 2021, [4] Naffziger, DAC Keynote, 2021

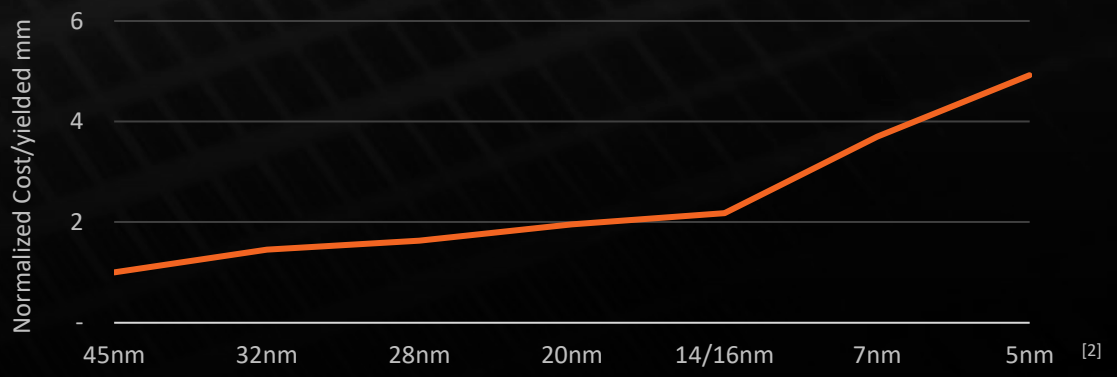
AMD

TECHNOLOGY HEADWINDS MEET OPPORTUNITY

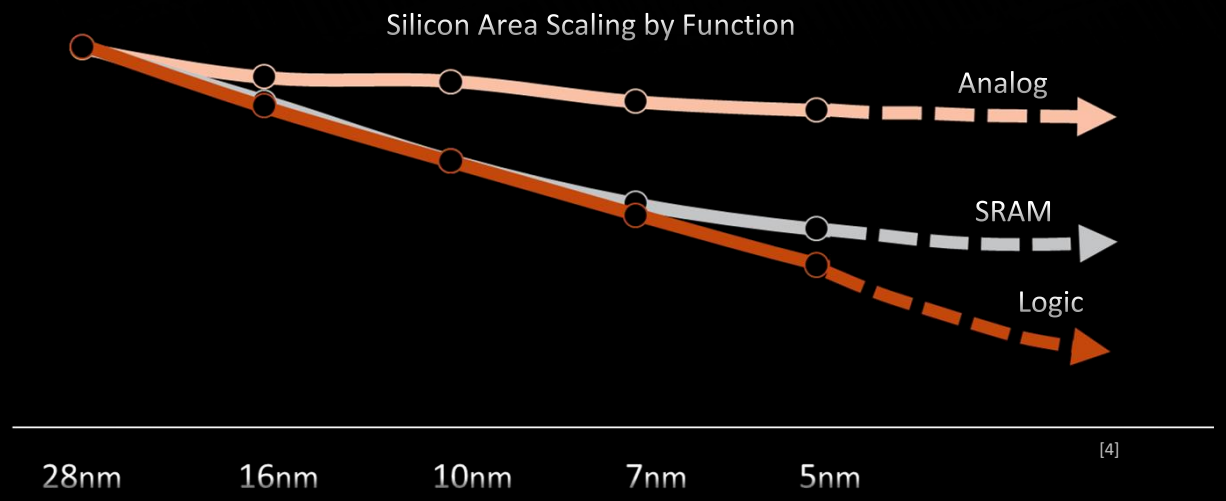
Slowing of Moore's Law



Increasing Cost



Limited and Divergent Scale Factors

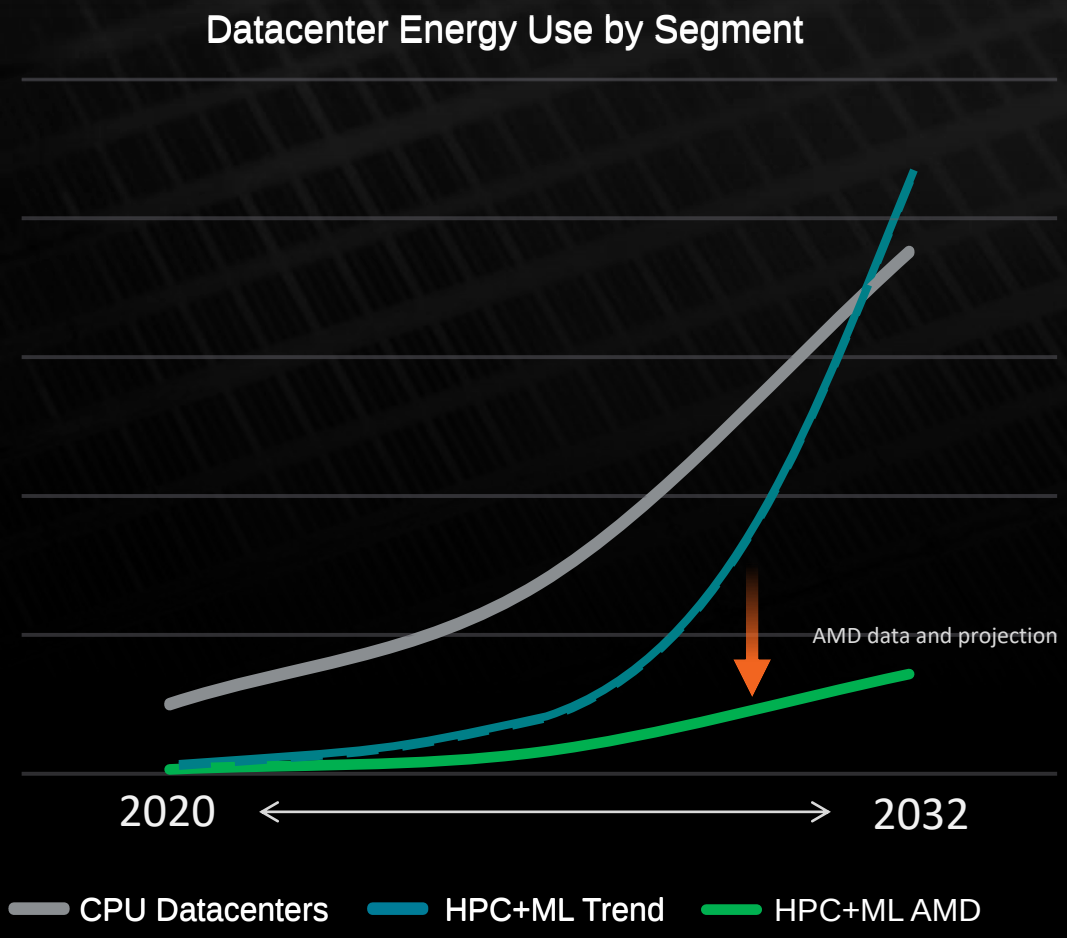


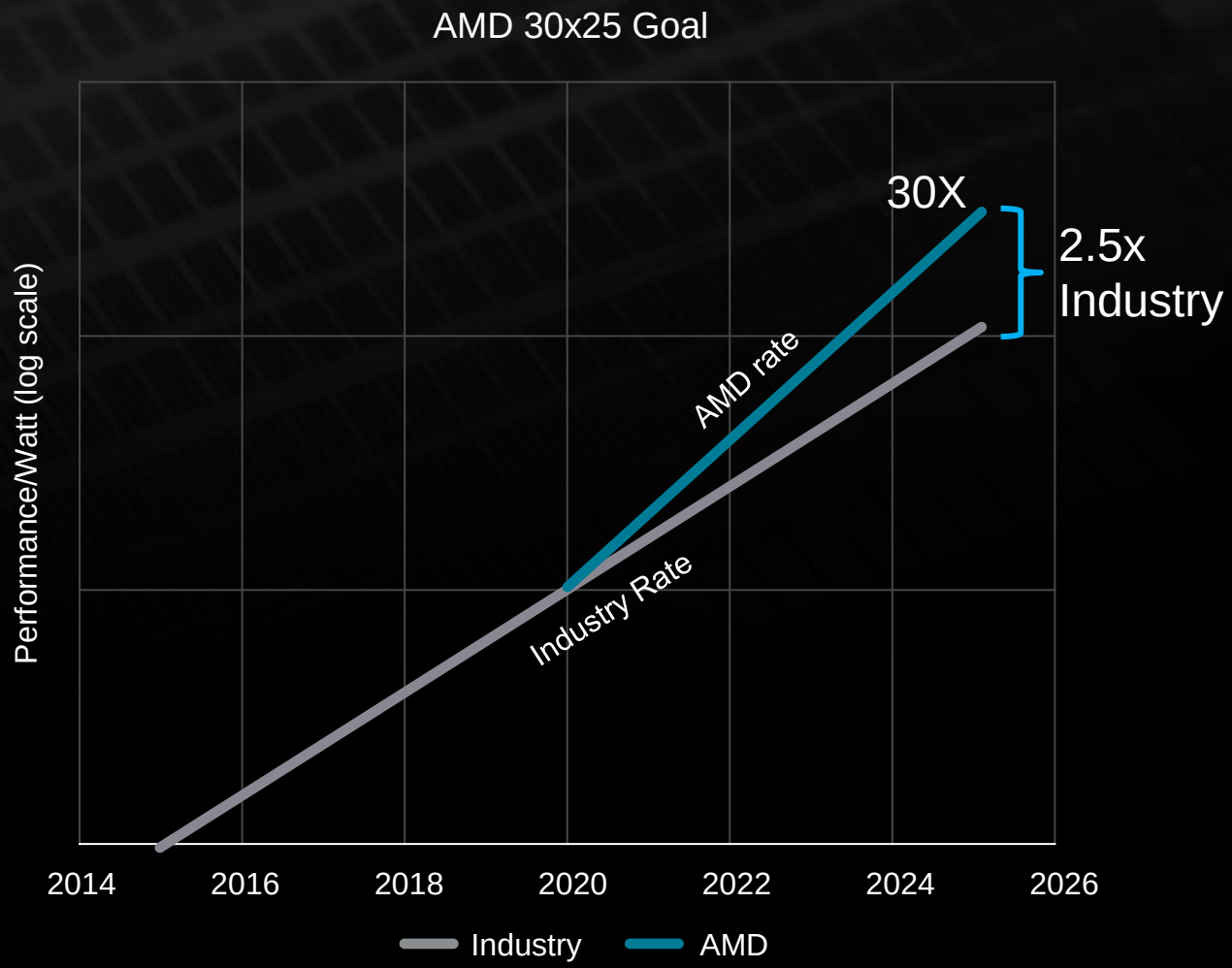
Large L3 Caches Provide IPC Uplift

[1] Naffziger, VLSI Short Course, 2020, [2] Cost per yielded mm² for a 250mm² die, [4] Naffziger, DAC Keynote, 2021, [5] Wuu et al, ISSCC 26.4, 2022

AMD

ACCELERATING DATA CENTER SUSTAINABILITY



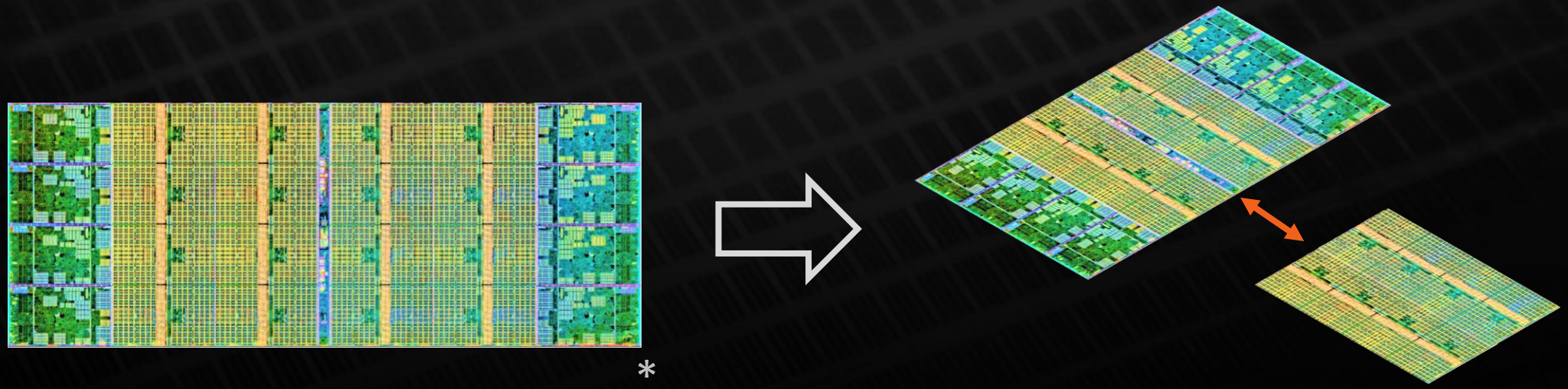


Server growth is from independent research data further extrapolated to 2030. This sets a compute demand. This is then combined with Performance per Watt trends to calculate the power consumed for that compute demand, both for the "industry trend" line and the "AMD projection" line. Performance per Watt is calculated by the CPU socket and GPU node power consumptions are based on segment-specific utilization (active vs. idle) percentages then multiplied by PUE to determine actual energy use for calculation of the performance per Watt.

Based on 2015-2020 industry trends in energy efficiency gains and data center energy consumption in 2025.
* Includes AMD high performance CPU and GPU accelerators used for AI training and High-Performance Computing in a 4-Accelerator, CPU hosted configuration. Goal calculations are based on performance scores as measured by standard performance metrics (HPC: Linpack DGEMM kernel FLOPS with 4k matrix size. AI training: lower precision training-focused floating point math GEMM kernels such as FP16 or BF16 FLOPS operating on 4k matrices) divided by the rated power consumption of a representative accelerated compute node including the CPU host + memory, and 4 GPU accelerators.

Naffziger, MIT Symposium on Climate and Computing, 2022

MORE MORE-THAN-MOORE



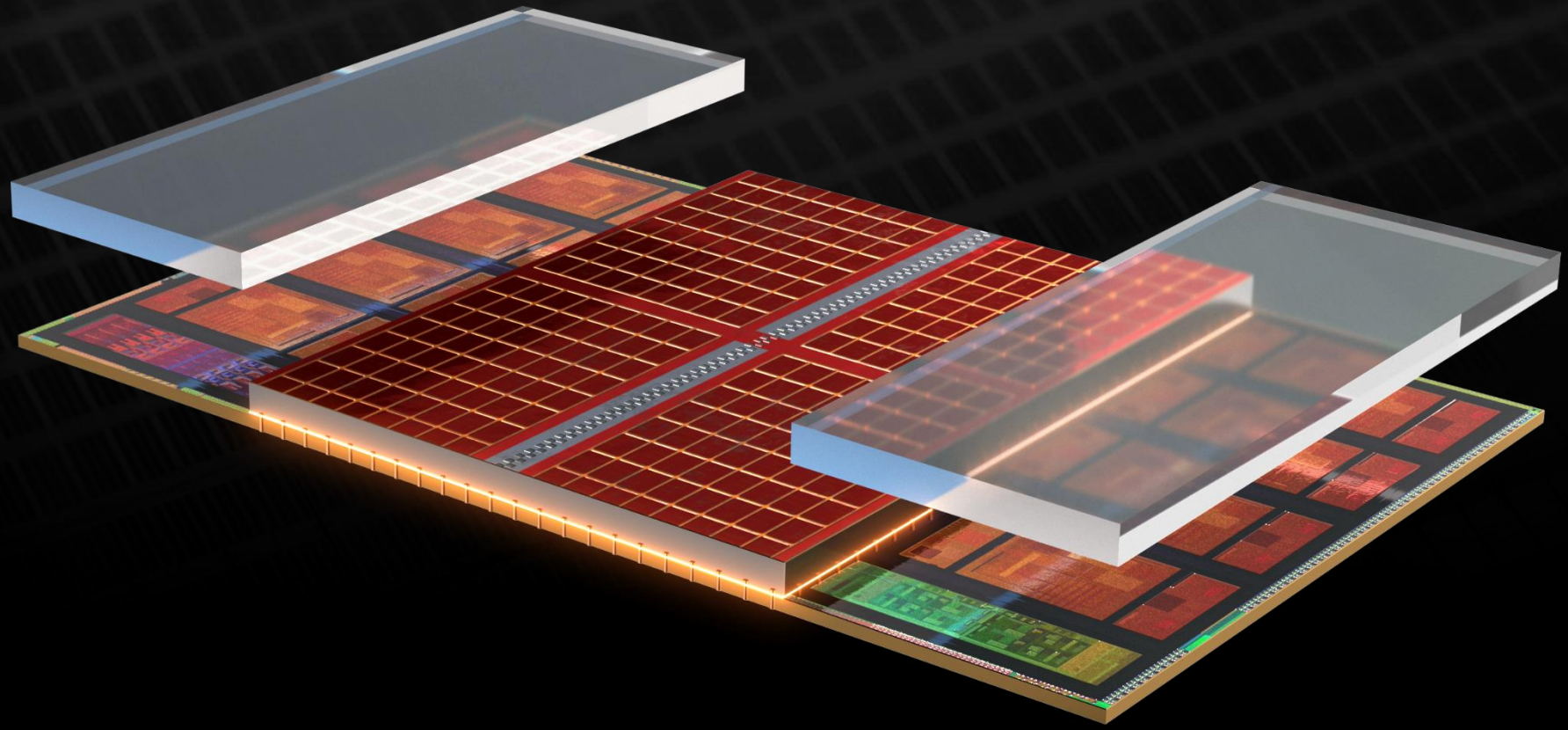
*

- 2.5D chiplets can provide product flexibility and reduce cost
- However, 3D can be even better!
  - Improves effective memory latency
  - Reduces long datapath and I/O's dynamic powers
  - Fits more transistors within a given package cavity size

*Hypothetical processor with large cache

Wuu et al, ISSCC 26.4, 2022

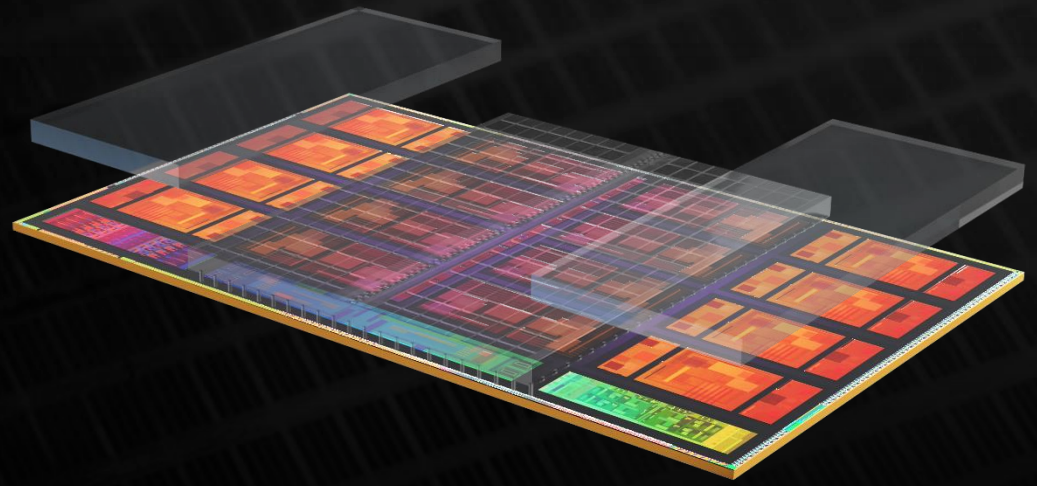
AMD 3D V-CACHE™

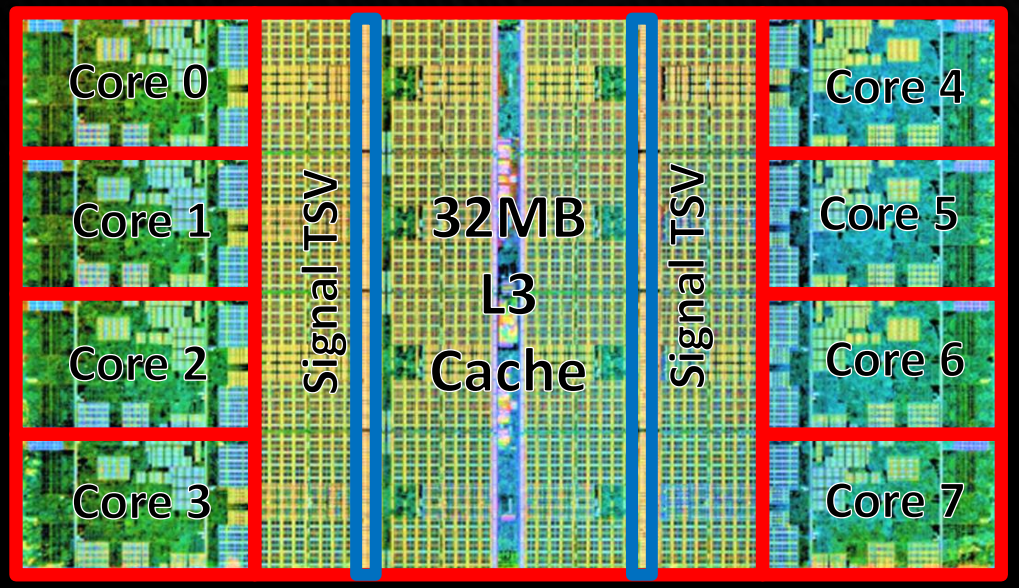


- Industry’s first high-performance processor product with Hybrid Bonded 3D cache die

Wuu et al, ISSCC 26.4, 2022

AMD 3D V-CACHE™ COMPONENTS: CCD



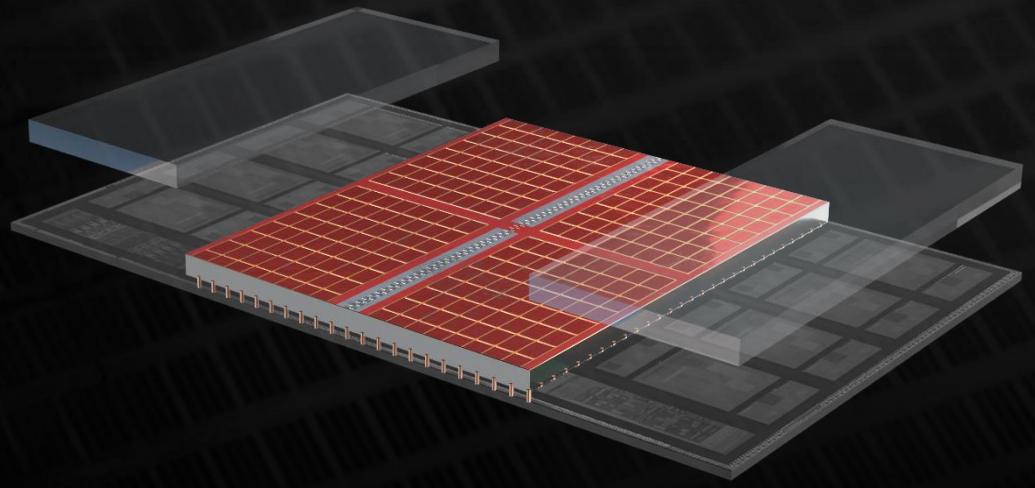


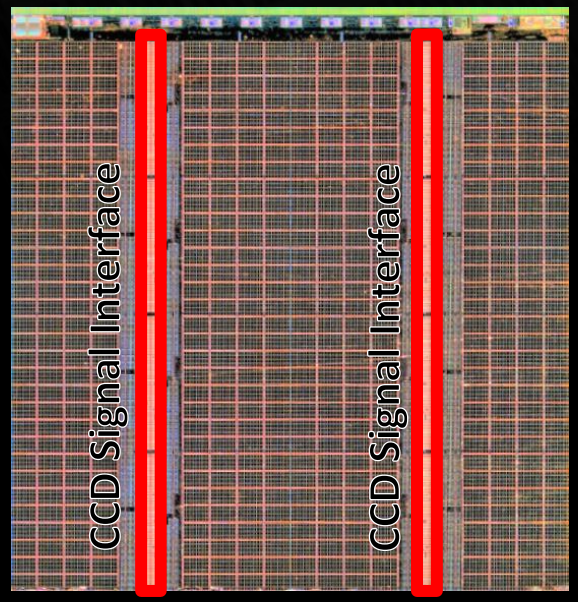
- "Zen 3" x86-64 CPU Core Complex Die (CCD)
- TSMC 7nm technology
- 8 cores per Core Complex (CCX)
- 32MB shared L3 Cache
- +19%[1] IPC (Ave) vs. "Zen 2"
- 81mm^2
- AMD 3D V-Cache™ support integrated from Day 1

[1]SEE ENDNOTES: R5K-003

Wuu et al, ISSCC 26.4, 2022

AMD 3D V-CACHE™ COMPONENTS: L3D

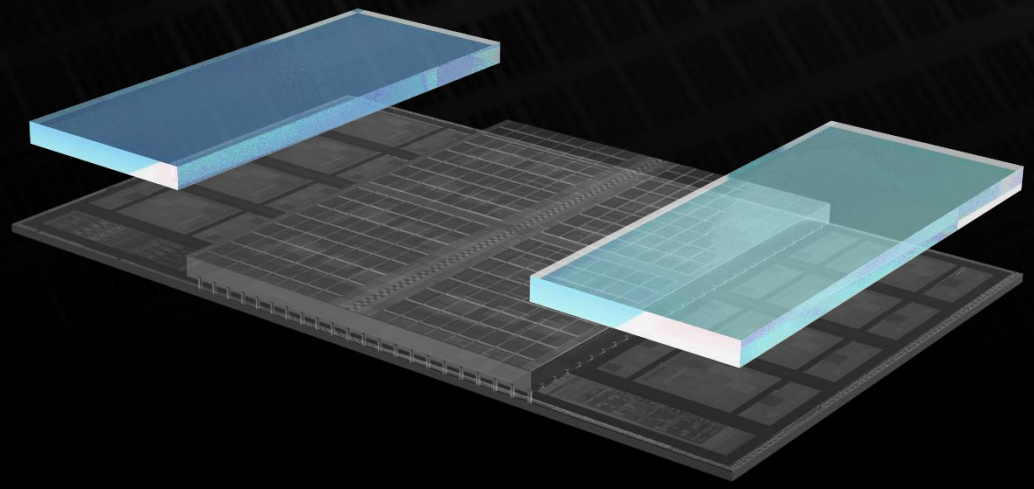




- AMD 3D V-Cache™ extended L3 Die (L3D)
- TSMC 7nm FinFET Technology
- 13 layers Cu + 1 layer Al metal stack
- 64MB L3 Cache Extension
- 41mm^2

Wuu et al, ISSCC 26.4, 2022

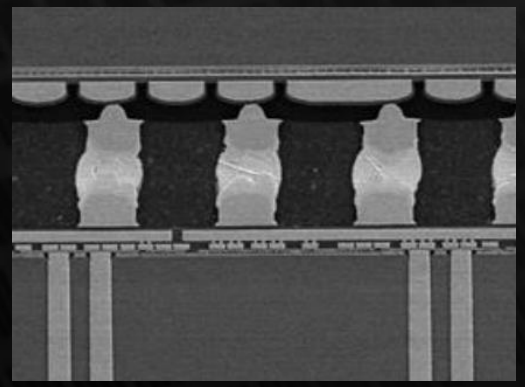
AMD 3D V-CACHE™ COMPONENTS: STRUCTURAL DIE



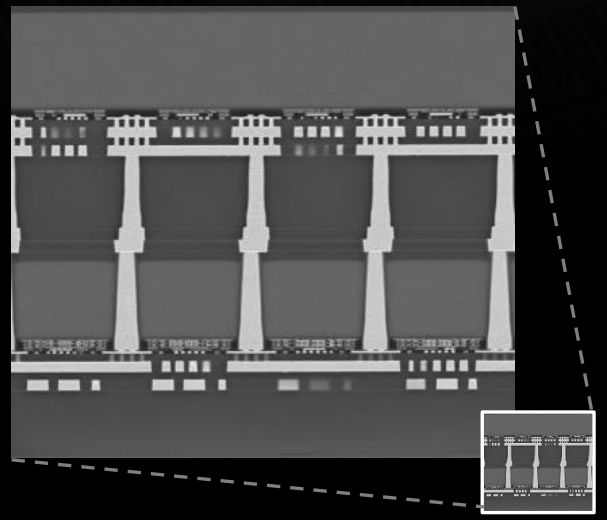
- AMD 3D V-Cache™ Structural Dies
- Structural support for thinned CCD
- Thermal dissipation for CPU cores

Wuu et al, ISSCC 26.4, 2022

MICRO BUMP VS. HYBRID BOND



Micro Bump 3D

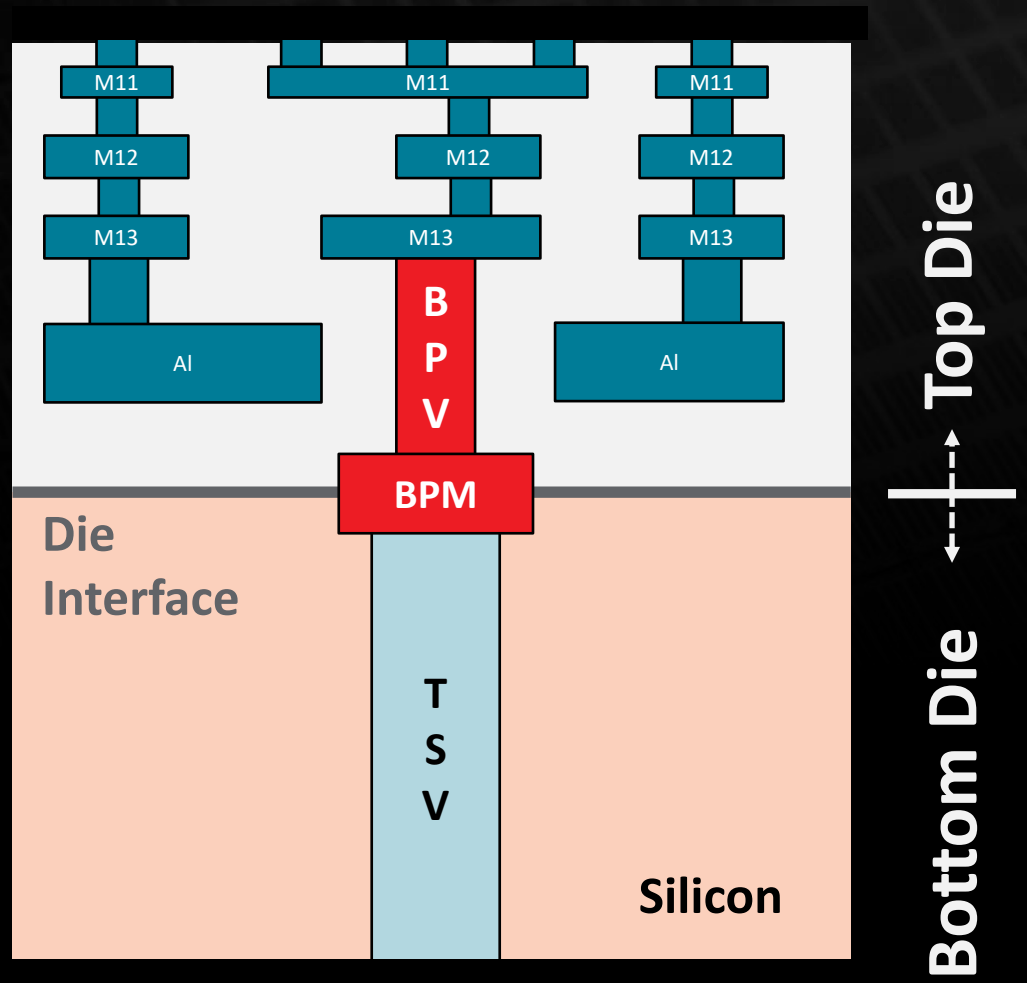


- Compared to Micro Bump 3D solutions, Hybrid Bond offers
  - >15x interconnect density
  - >3x interconnect energy efficiency
  - Superior thermal conductance

SEE ENDNOTES: EPYC-026, EPYC-027

Wuu et al, ISSCC 26.4, 2022

3D V-CACHE™ BONDING TECHNOLOGY

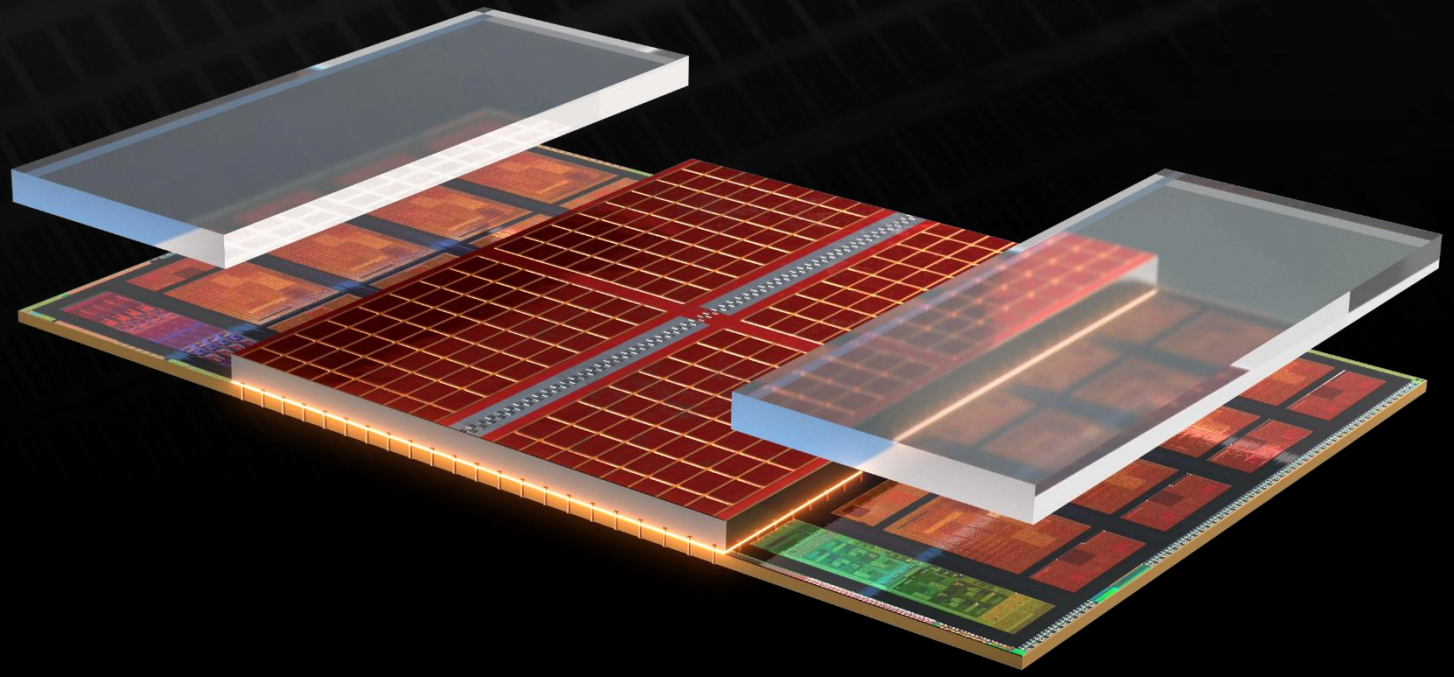


- TSMC SoIC process
- Cu bonded using Bond Pad Metal (BPM) pads
- BPM interfaces with TSV
- Bond Pad Via (BPV) connects BPM to M13
- 9um minimum TSV pitch

Wuu et al, ISSCC 26.4, 2022

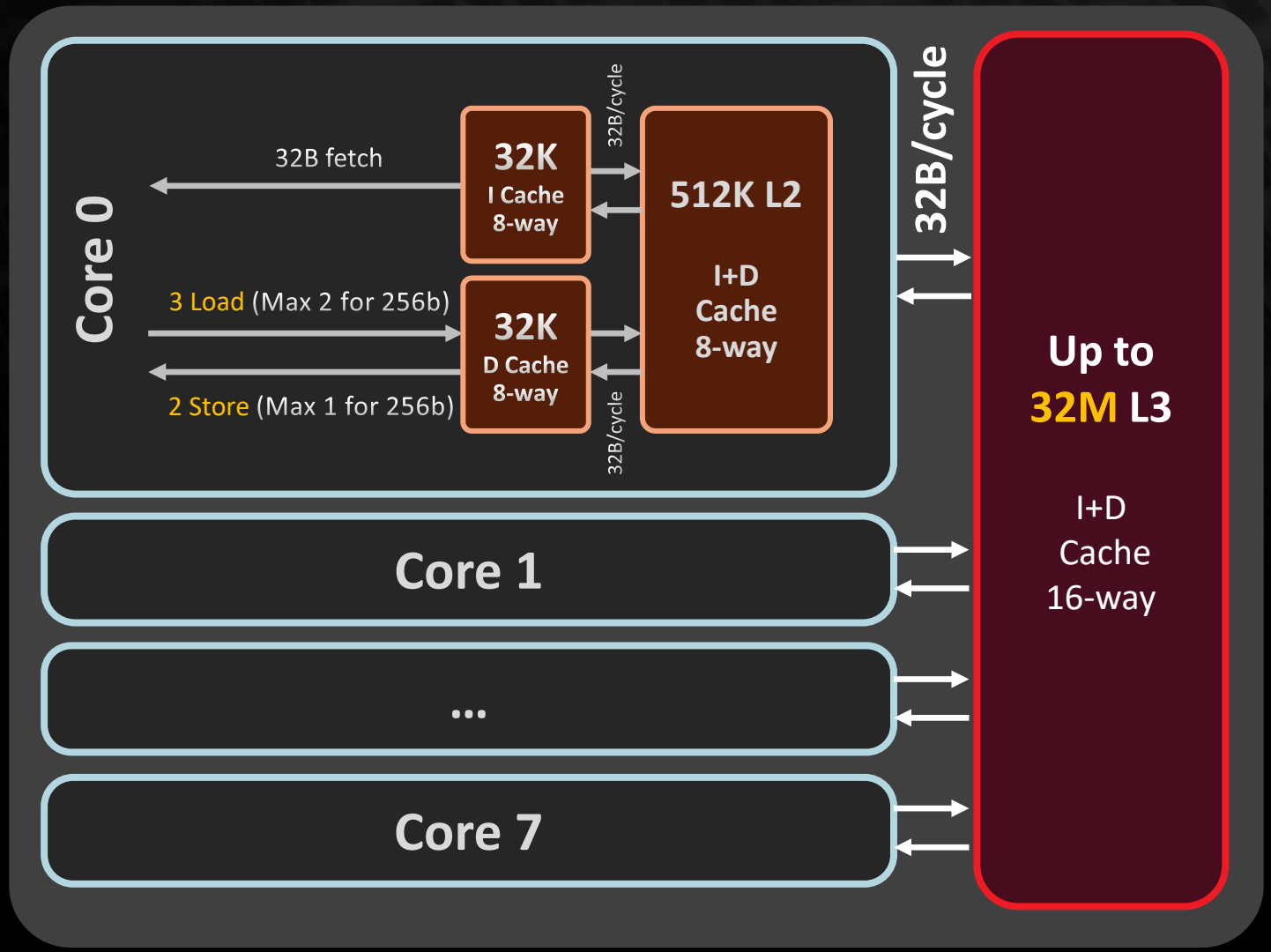
3D V-CACHE™ PHYSICAL ORGANIZATION

- CCD face-down
  - C4 interface to substrate
  - TSV interface to L3D
- L3D face-down
  - Hybrid Bonded (HB) to CCD
- Structural Dies
  - Oxide bonded to CCD



Wuu et al, ISSCC 26.4, 2022

"ZEN 3" CACHE HIERARCHY

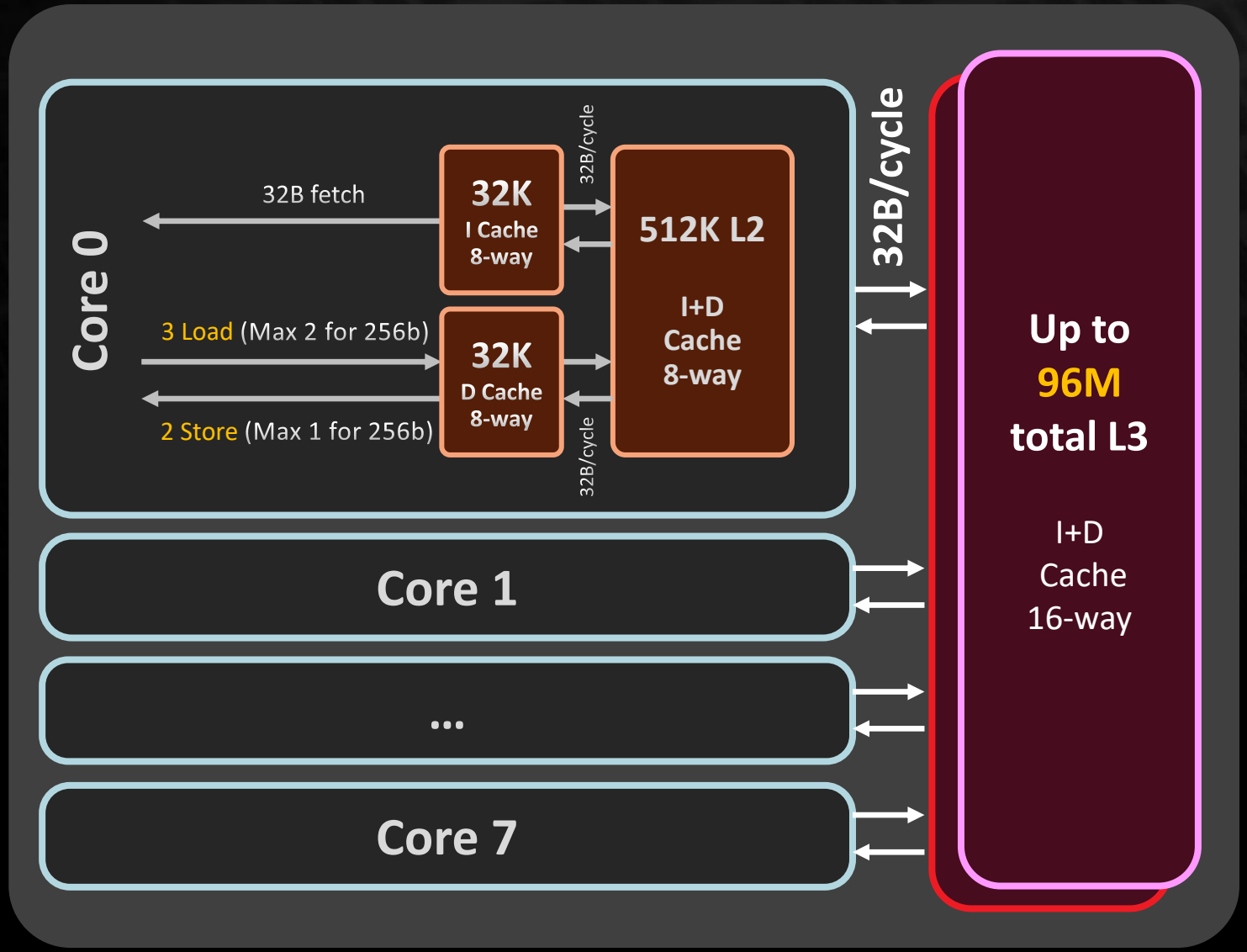


- 8 Cores per CCD
  - 32K I-Cache + 32K D-Cache
  - Private 512K L2 per core
- Shared 32MB L3 between 8 cores
  - 16-way set associative
  - 32B/cycle interface to each core
- DECTED ECC for enhanced data reliability

Wuu et al, ISSCC 26.4, 2022

AMD

"ZEN 3" + AMD 3D V-CACHE™

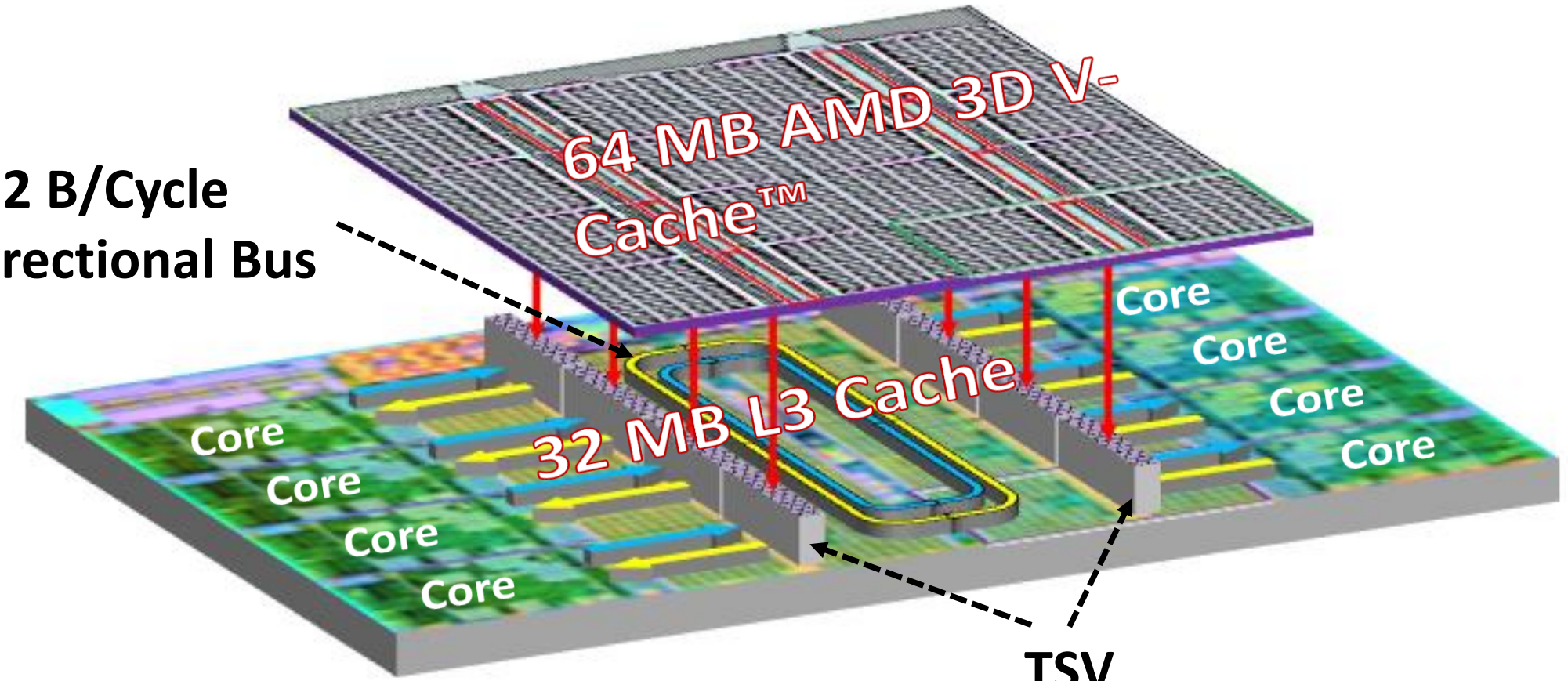


- 96MB shared L3 Cache between 8 cores
  - 16-way set associative
  - 32B/cycle interface to each core
- >2 TB/s L3 bandwidth
- +4 cycles latency
- Each die's L3 includes its own
  - Data arrays
  - Tag arrays
  - LRU arrays

Wuu et al, ISSCC 26.4, 2022

AMD

3D V-CACHE™ INTERFACE ILLUSTRATION



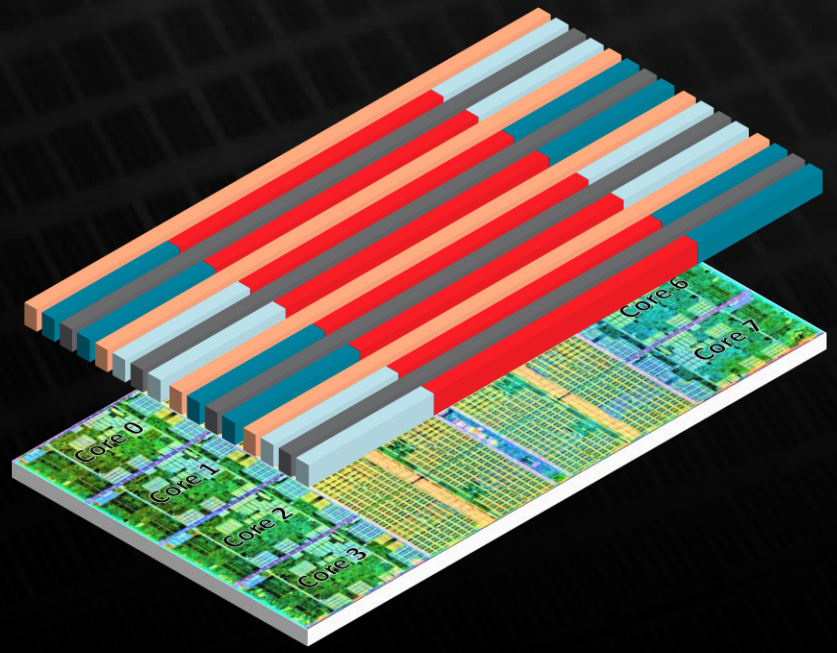
[6]

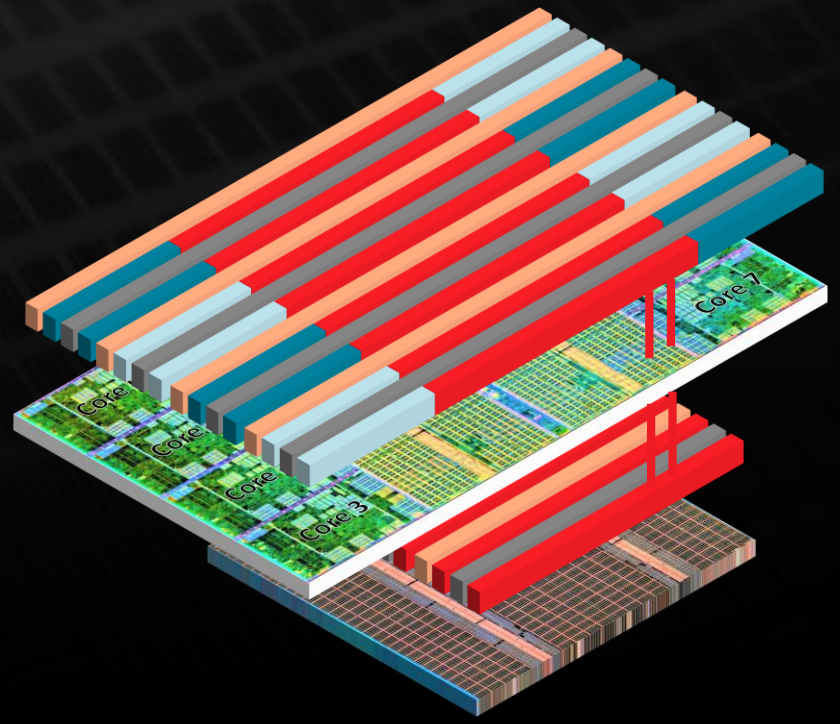
[6] Burd, ISSCC 2.6, 2022

Wuu et al, ISSCC 26.4, 2022

AMD

CCD & L3D POWER DELIVERY





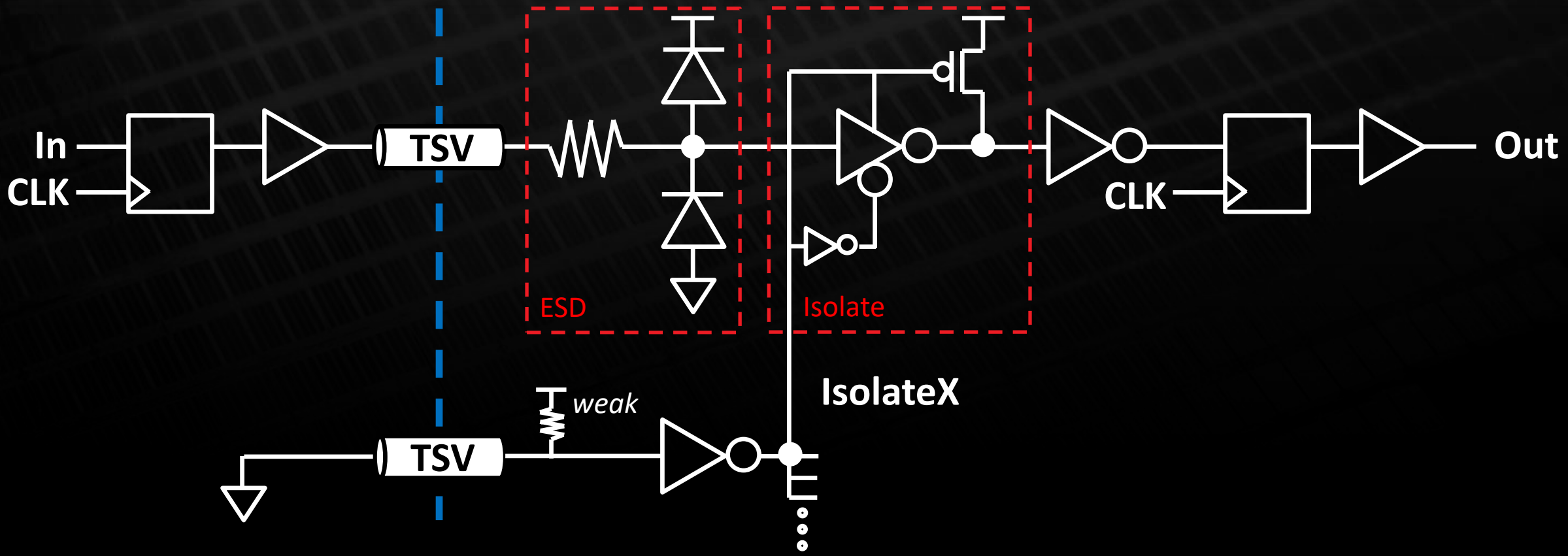
- CCD primary supplies
  - RVDD: Ungated supply
  - VDD: Per-core gated supply
  - VDDM: Gated L2/L3 SRAM supply

- RVDD and VDDM delivered to L3D through power TSVs
- Power TSVs in channels between CCD array macros

Wuu et al, ISSCC 26.4, 2022

AMD

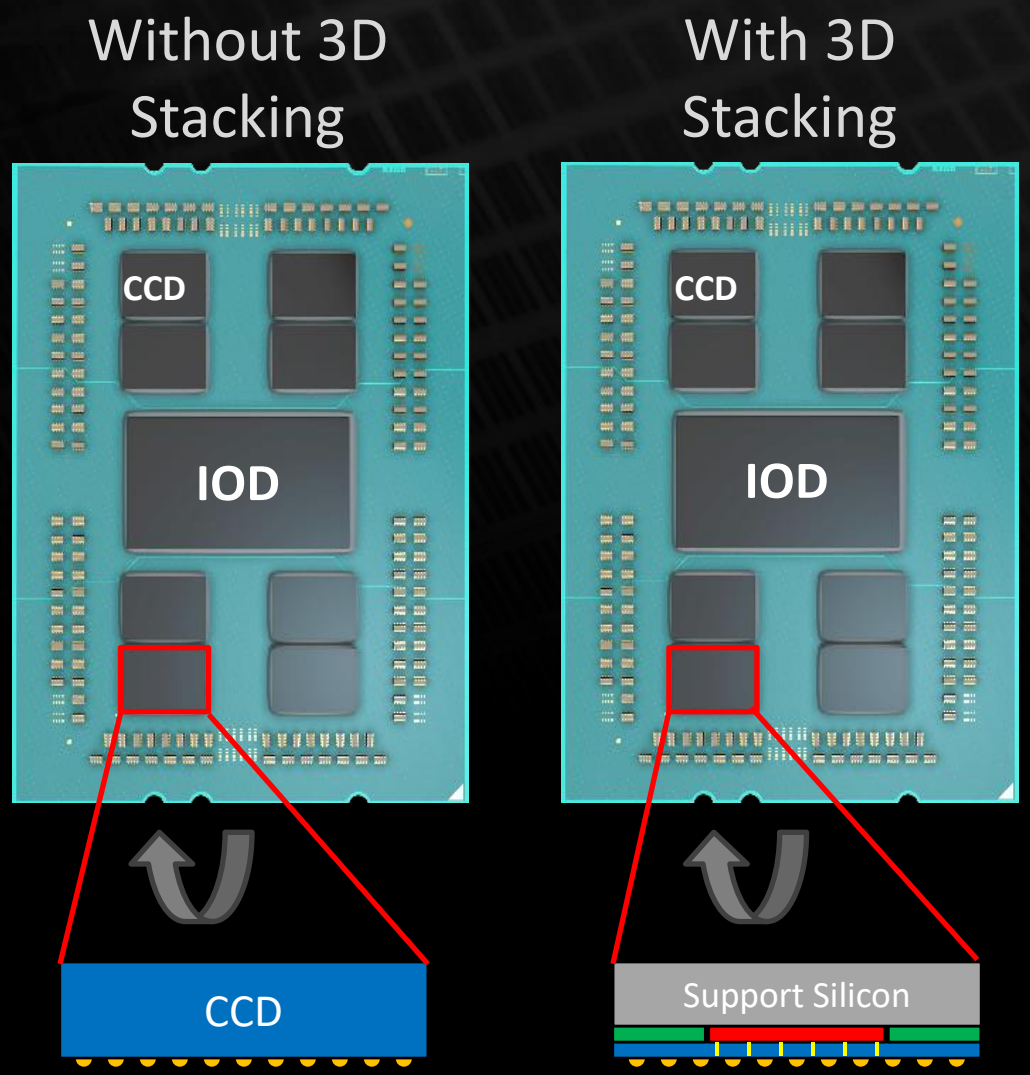
3D SIGNAL INTERFACE



- Simple digital signal interface between dies
  - Enabled by HB technology's low parasitics

Wuu et al, ISSCC 26.4, 2022

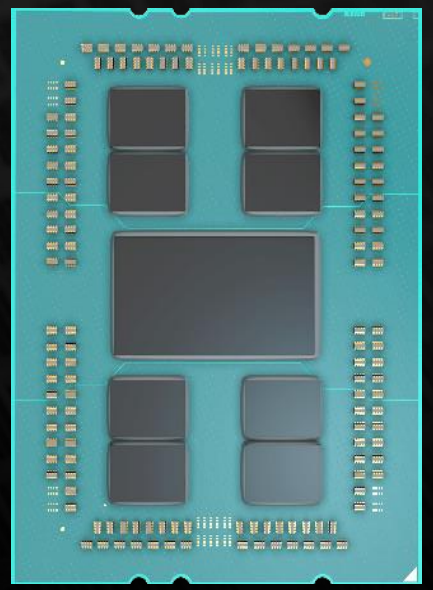
3D V-CACHE™ SERVER CONFIGURATION



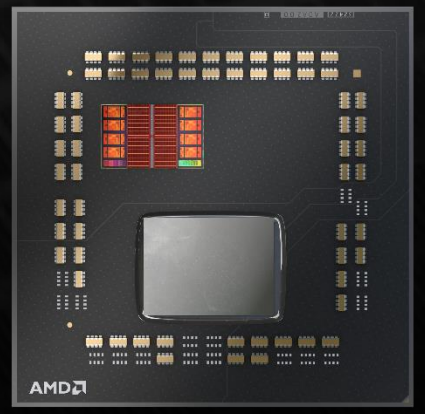
- AMD 3rd Gen EPYC™ Server CPU
  - Up to 8 "Zen 3" CCDs
  - 1 I/O Die (IOD)
- AMD 3rd Gen EPYC™ Server CPU with AMD 3D V-Cache™
  - Up to 8 thinned CCDs + L3Ds
  - Support silicon added to match 2D CCD Z-height
- Both designs compatible with the same package

Wuu et al, ISSCC 26.4, 2022

AMD 3D V-CACHE™ PRODUCT PORTFOLIO



AMD 3rd Gen
EPYC™ Server CPU



AMD RYZEN™ 7 5800X3D
Gaming CPU

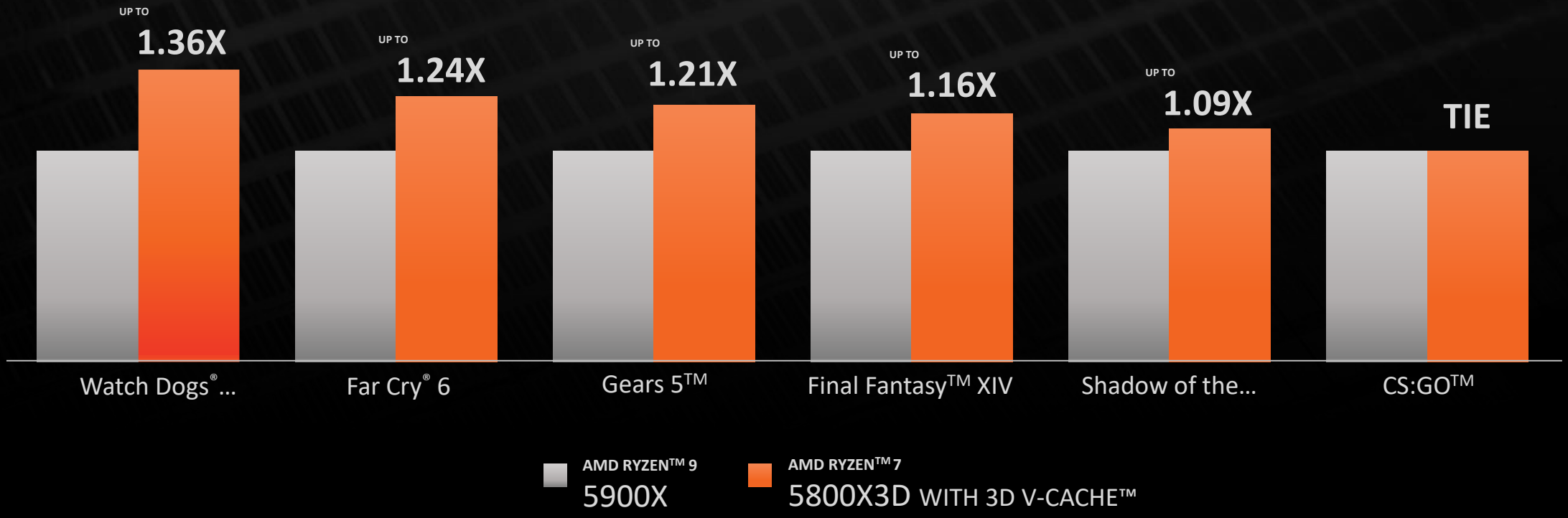
- AMD 3D V-Cache™ supports L3 Cache extension for both server and desktop product families

Wuu et al, ISSCC 26.4, 2022

AMD

DESKTOP PERFORMANCE

AMD RYZEN™ 7 5800X3D WITH AMD 3D V-CACHE™



~15% faster gaming at 1080p high

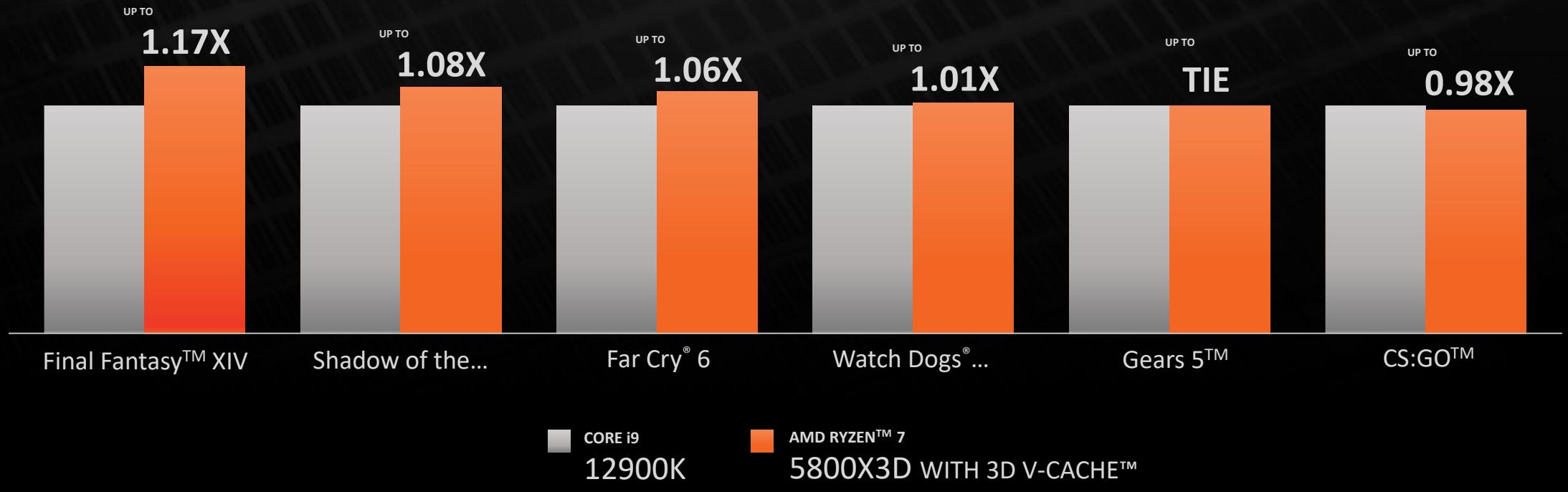
SEE ENDNOTES: R5K-106

Wuu et al, ISSCC 26.4, 2022

AMD

STATE OF THE ART COMPARISON

AMD RYZEN™ 7 5800X3D WITH AMD 3D V-CACHE™



World’s fastest gaming processor

SEE ENDNOTES: R5K-107

Wuu et al, ISSCC 26.4, 2022

AMD

SERVER PERFORMANCE

24.4
JOBS/HOUR

3RD GEN AMD EPYC™ 16-CORE
WITHOUT AMD 3D V-CACHE™

~66%
FASTER RTL VERIFICATION

SYNOPSYS® VCS®

40.6
JOBS/HOUR

3RD GEN AMD EPYC™ 16-CORE
WITH AMD 3D V-CACHE™

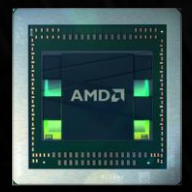
RESULTS MAY VARY. SEE ENDNOTES: MLNX-001R

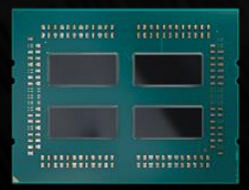
Wuu et al, ISSCC 26.4, 2022

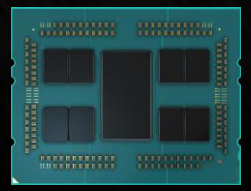
AMD

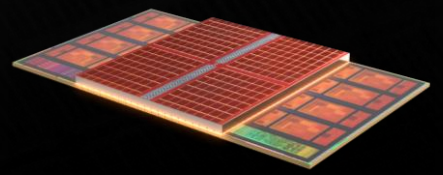
AMD PACKAGING SOLUTIONS

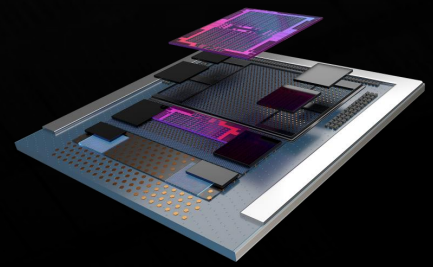
2015	2017	2019	2021	
2.5D HBM	MULTICHIP MODULE	CHIPLETS	3D CHIPLETS	2.5D CHIPLETS
			(Chiplet + Advanced 3D Stacking)	











Led Industry in HBM, 2.5D & Chiplet Architecture

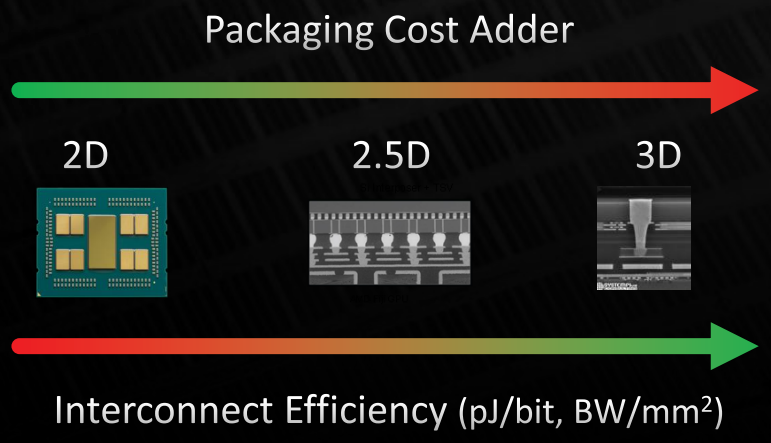
Aggressive Roadmap for Chiplet & 3D Integration

Manufacturing Friendly Elevated Fanout Bridge (EFB)

Naffziger, DAC Keynote, 2021

FINDING THE OPTIMAL SOLUTION

Chiplet package architecture selection requires balancing a complex equation…



[Value of modular product]
+
[Chiplet yield and technology cost benefit]

>

[Chiplet die overhead adder]
+
[Packaging cost adder]
+
[Power cost of interconnect]
+
[Engineering complexity of solution]

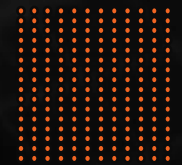
Architectural need for bandwidth, die partition options and package technology create a multi-disciplinary optimization equation

Naffziger, DAC Keynote, 2021

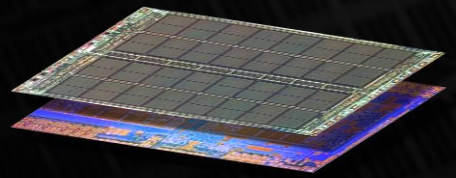
AMD

FUTURE OF 3D STACKING

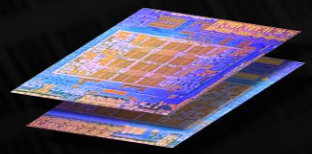
TSV Pitch



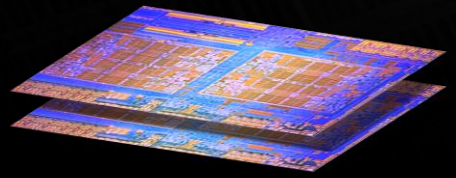
SRAM on CPU



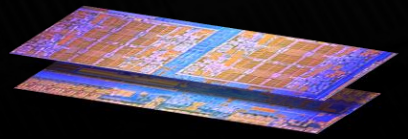
Cores on Cores



CPU on CPU



Cores on Uncore

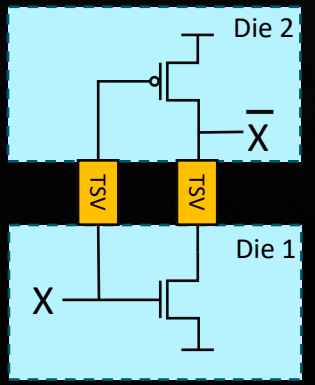


Full die-to-die

IP on IP

Macro on Macro

IP Folding/Splitting



Circuit Slicing

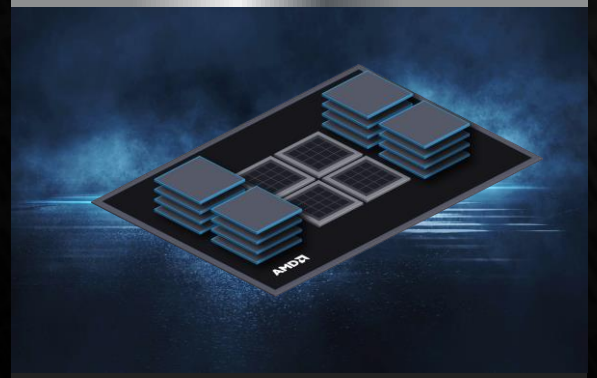
Advanced packaging can enable integration schemes not possible with monolithic designs

Based on AMD engineering internal analysis, May 2021

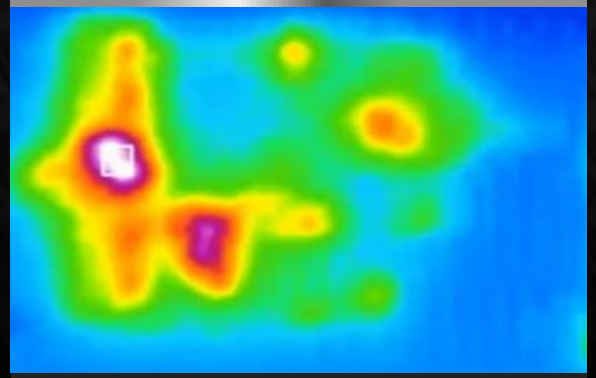
Naffziger, DAC Keynote, 2021

AMD

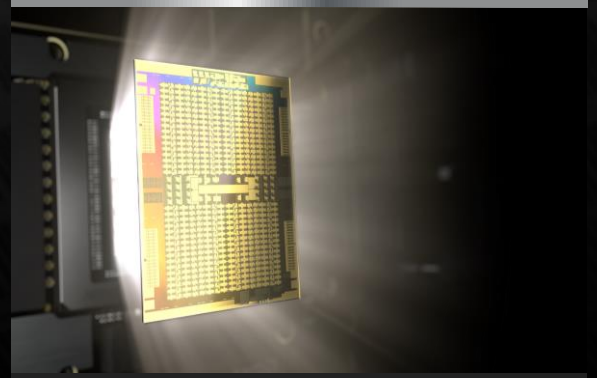
MANY AREAS NEED FURTHER INNOVATION



TEST AND KGD



THERMALS



POWER DELIVERY

SYSTEM-LEVEL INTEGRATION

Naffziger, DAC Keynote, 2021

ENDNOTES

R5K-003: Testing by AMD performance labs as of 09/01/2020. IPC evaluated with a selection of 25 workloads running at a locked 4GHz frequency on 8-core "Zen 2" Ryzen 7 3800XT and "Zen 3" Ryzen 7 5800X desktop processors configured with Windows® 10, NVIDIA GeForce RTX 2080 Ti (451.77), Samsung 860 Pro SSD, and 2x8GB DDR4-3600. Results may vary.

EPYC-026: Based on calculated areal density and based on bump pitch between AMD hybrid bond AMD 3D V-Cache stacked technology compared to AMD 2D chiplet technology and Intel 3D stacked micro-bump technology.

EPYC-027: Based on AMD internal simulations and published Intel data on “Foveros” technology specifications.

R5K-106: Based on testing by AMD as of 12/14/2021. Performance evaluated with Watch Dogs Legion, Far Cry 6, Gears 5, Final Fantasy XIV, Shadow of the Tomb Raider and CS:GO. All games test at 1920x1080 resolution with the HIGH in-game quality preset (or equivalent). System configuration: Ryzen 7 5800X3D and AMD Reference Motherboard, Ryzen 9 5900X and ASUS Crosshair VIII Hero with BIOS 3801. Both systems configured with 2x8GB DDR4-3600, GeForce RTX 3080 with 472.12 driver, Samsung 980 Pro 1TB, NZXT Kraken X62, and Windows 11 28000.282.

R5K-107: Based on testing by AMD as of 12/14/2021. Performance evaluated with Watch Dogs Legion, Far Cry 6, Gears 5, Final Fantasy XIV, Shadow of the Tomb Raider and CS:GO. All games test at 1920x1080p resolution with the HIGH in-game quality preset (or equivalent). System configuration: Ryzen 7 5800X3D and AMD Reference Motherboard with 2x8GB DDR4-3600. Core i9-12900K and ROG Maximus Z690 Hero motherboard with BIOS 0702 and 2x16GB DDR5-5200. Both systems configured with GeForce RTX 3080 on driver 472.12, Samsung 980 Pro 1TB, NZXT Kraken X62, Windows 11 28000.282.

MLNX-001R: EDA RTL Simulation comparison based on AMD internal testing completed on 9/20/2021 measuring the average time to complete a test case simulation. comparing: 1x 16C 3rd Gen EPYC CPU with AMD 3D V-Cache Technology versus 1x 16C AMD EPYC™ 73F3 on the same AMD “Daytona” reference platform. Results may vary based on factors including silicon version, hardware and software configuration and driver versions

DISCLAIMER AND COPYRIGHT

The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions, and typographical errors. The information contained herein is subject to change and may be rendered inaccurate for many reasons, including but not limited to product and roadmap changes, component and motherboard version changes, new model and/or product releases, product differences between differing manufacturers, software changes, BIOS flashes, firmware upgrades, or the like. Any computer system has risks of security vulnerabilities that cannot be completely prevented or mitigated. AMD assumes no obligation to update or otherwise correct or revise this information. However, AMD reserves the right to revise this information and to make changes from time to time to the content hereof without obligation of AMD to notify any person of such revisions or changes.

This information is provided "as is." AMD makes no representations or warranties with respect to the contents hereof and assumes no responsibility for any inaccuracies, errors, or omissions that may appear in this information. AMD specifically disclaims any implied warranties of non-infringement, merchantability, or fitness for any particular purpose. In no event will AMD be liable to any person for any reliance, direct, indirect, special, or other consequential damages arising from the use of any information contained herein, even if AMD is expressly advised of the possibility of such damages.

AMD


AMD