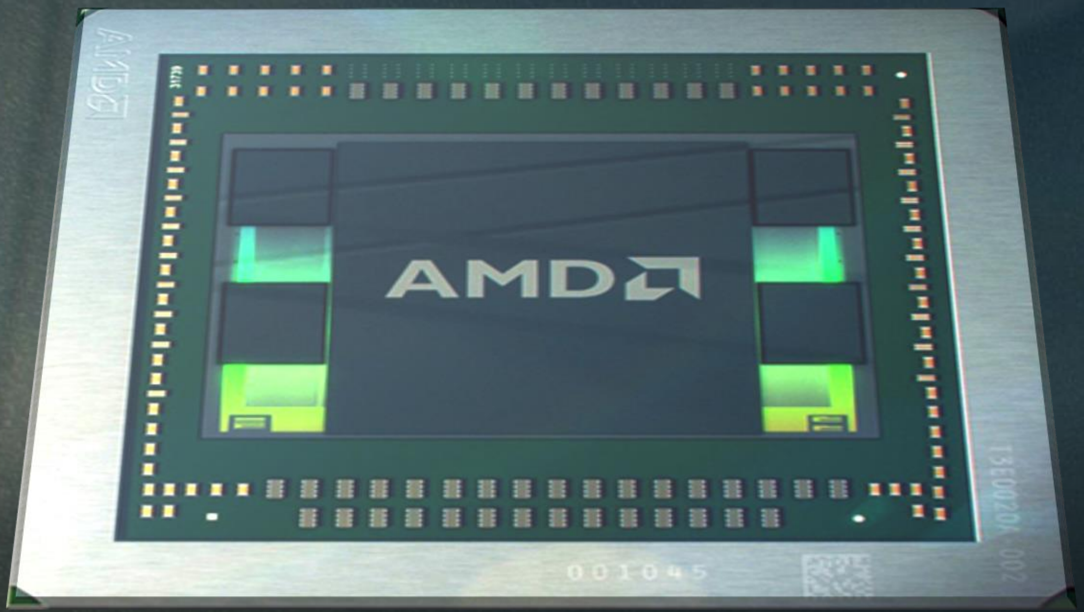
The Road to the AMD

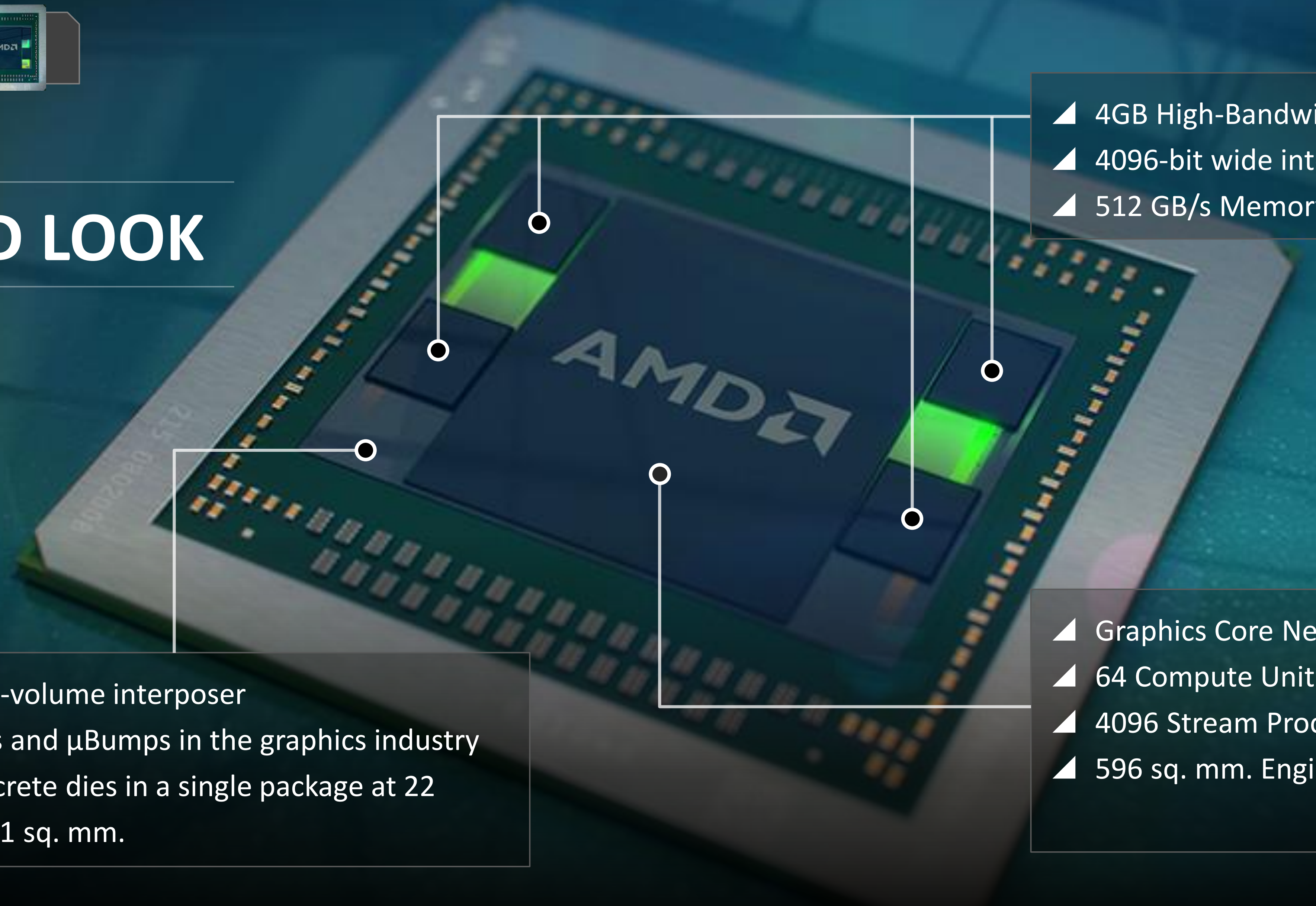
“Fiji” GPU

Featuring Die Stacking and HBM Technology



“Fiji” Chip

DETAILED LOOK



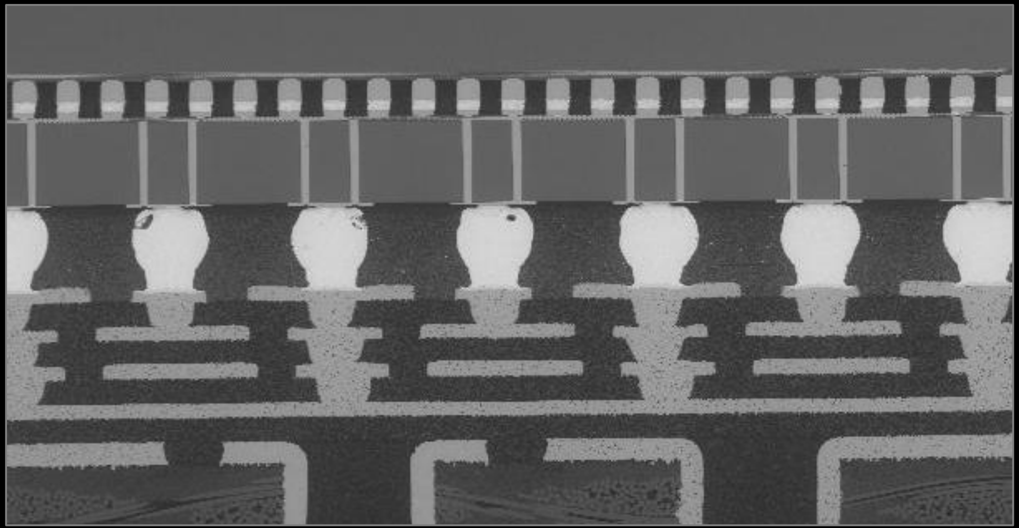
- 4GB High-Bandwidth Memory
- 4096-bit wide interface
- 512 GB/s Memory Bandwidth

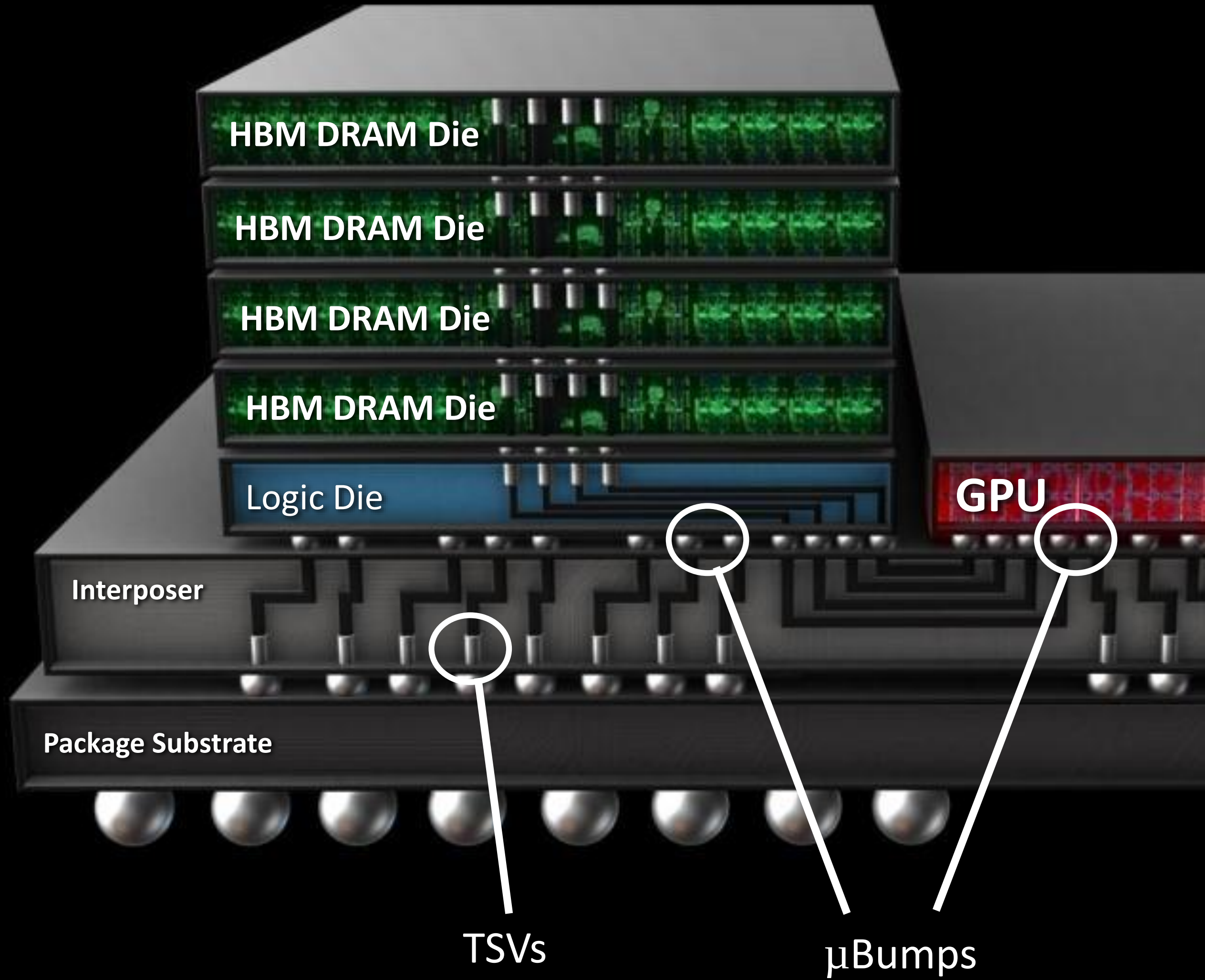
- First high-volume interposer
- First TSVs and µBumps in the graphics industry
- Most discrete dies in a single package at 22
- Total 1011 sq. mm.

- Graphics Core Next Architecture
- 64 Compute Units[8]
- 4096 Stream Processors
- 596 sq. mm. Engine

DIE STACKING TECHNOLOGY

- Die stacking facilitates the integration of discrete dies
- 8.5 years of development by AMD and its technology partners





WHY DID WE BUILD FIJI AND HBM?

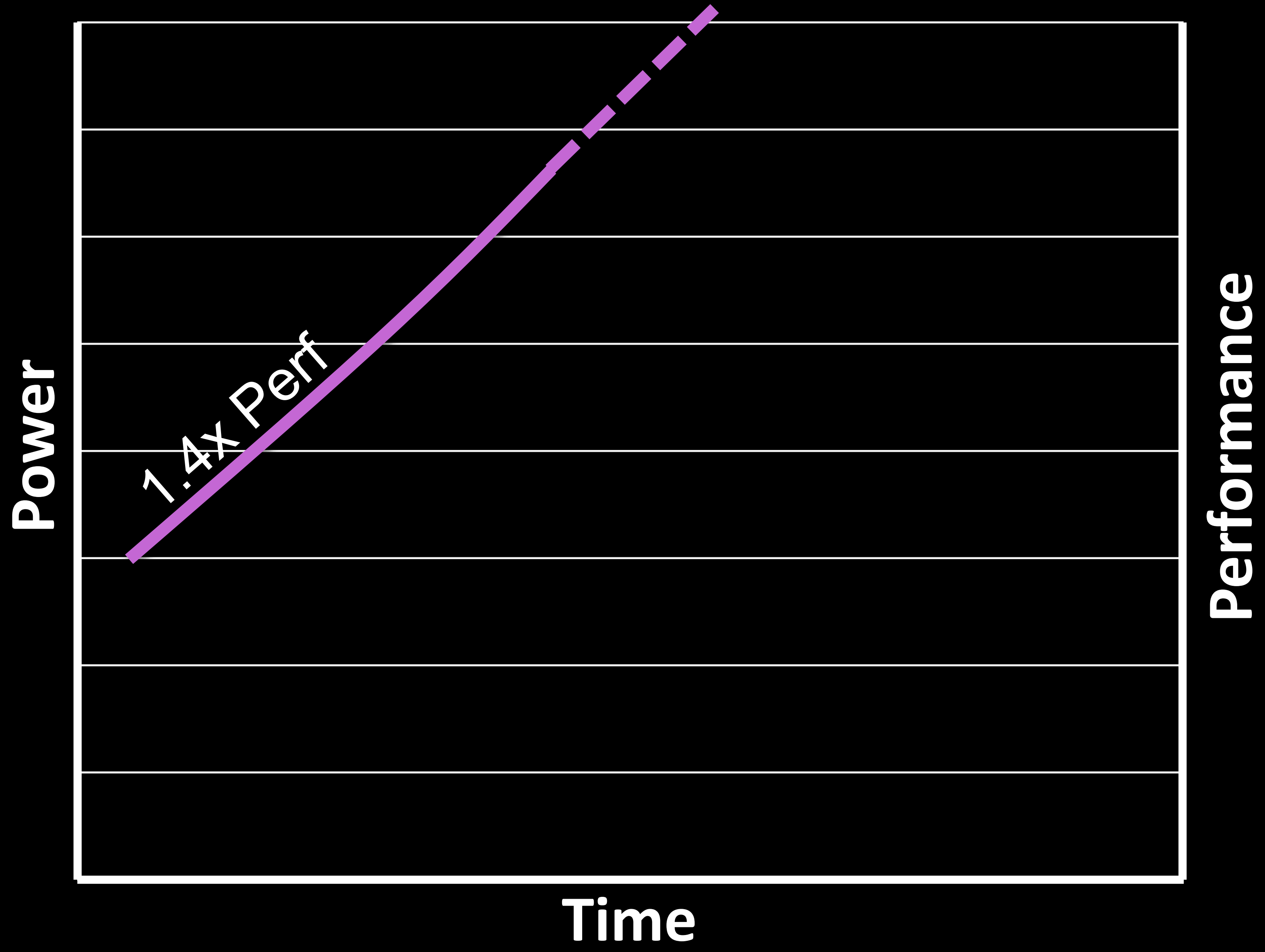
AN ANALYSIS FROM 2009

- Built a model to predict performance and power over time

WHY DID WE BUILD FIJI AND HBM?

AN ANALYSIS FROM 2009

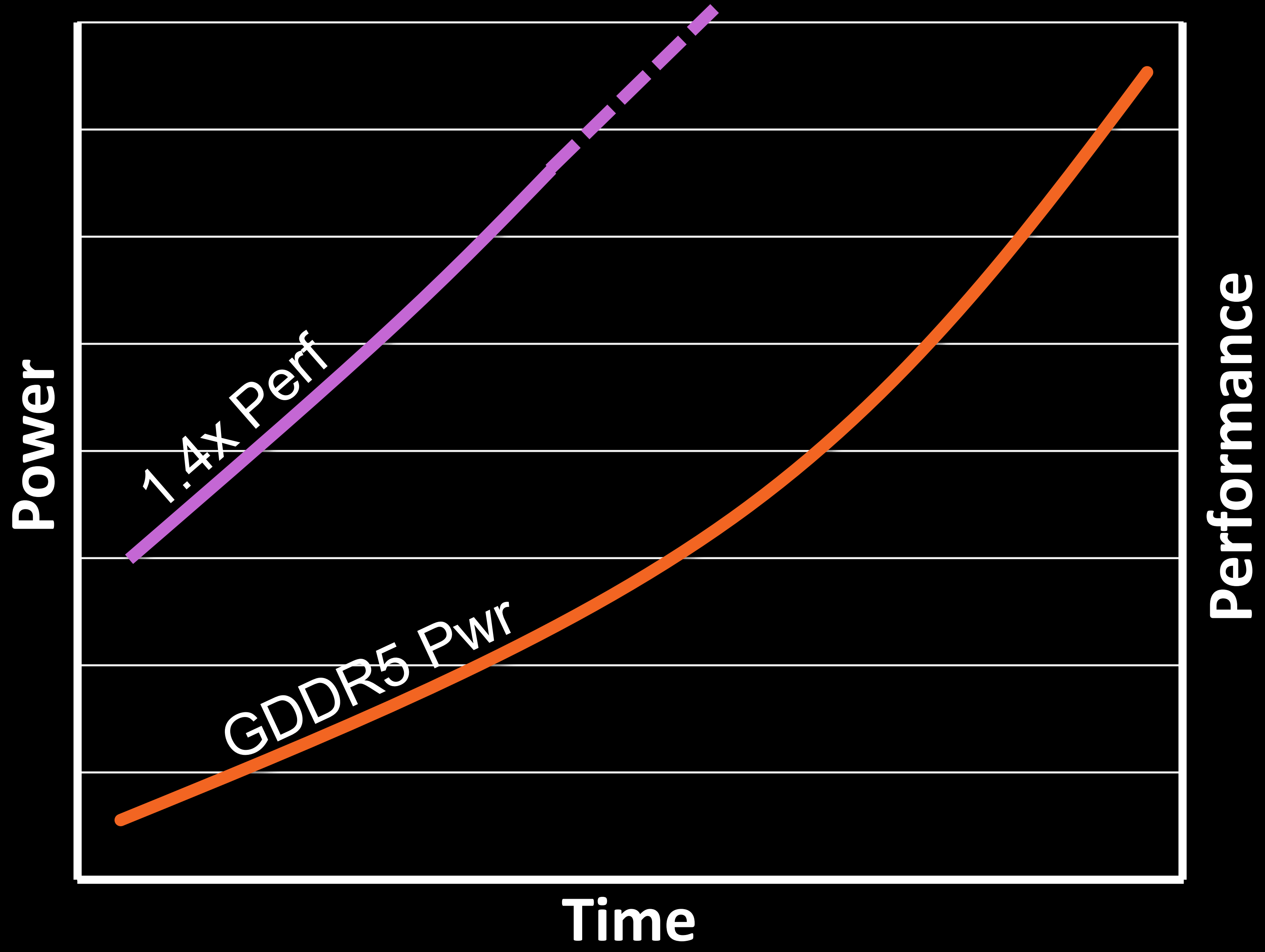
- Market performance demand requires 1.4x improvement per year



WHY DID WE BUILD FIJI AND HBM?

AN ANALYSIS FROM 2009

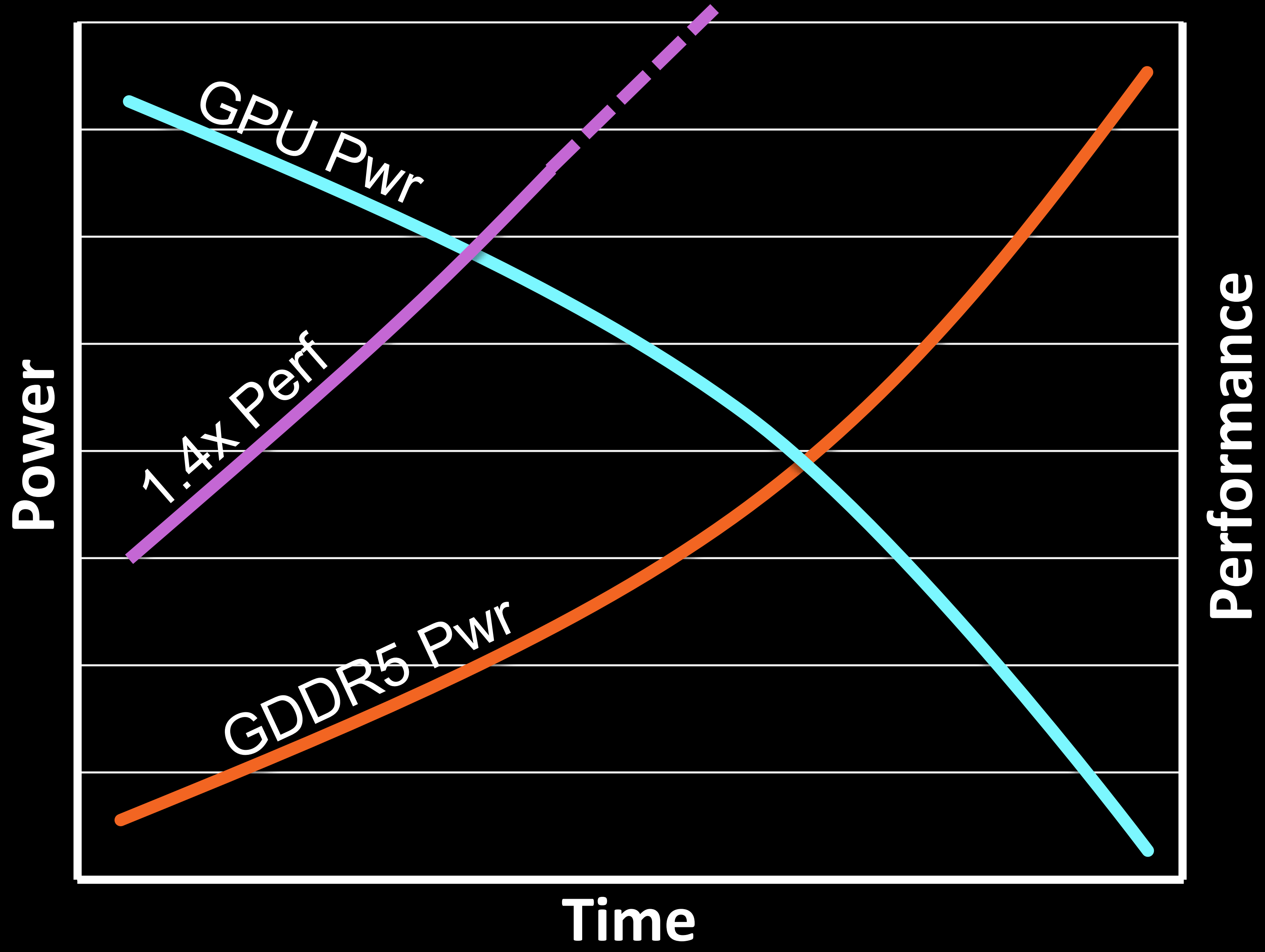
- GPU performance is proportional to memory BW
- Memory power increases with BW demand



WHY DID WE BUILD FIJI AND HBM?

AN ANALYSIS FROM 2009

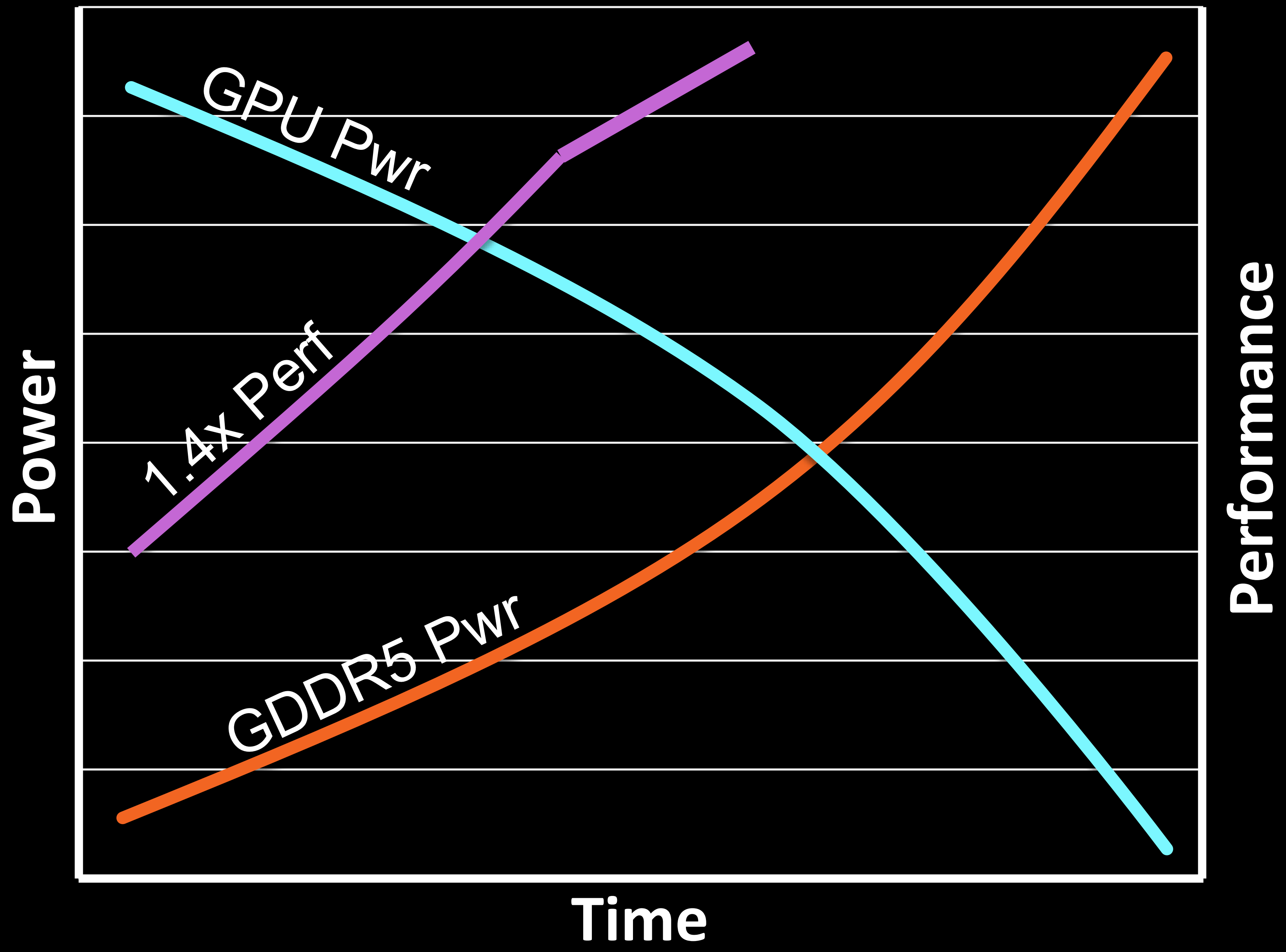
- System power is fixed in all platforms



WHY DID WE BUILD FIJI AND HBM?

AN ANALYSIS FROM 2009

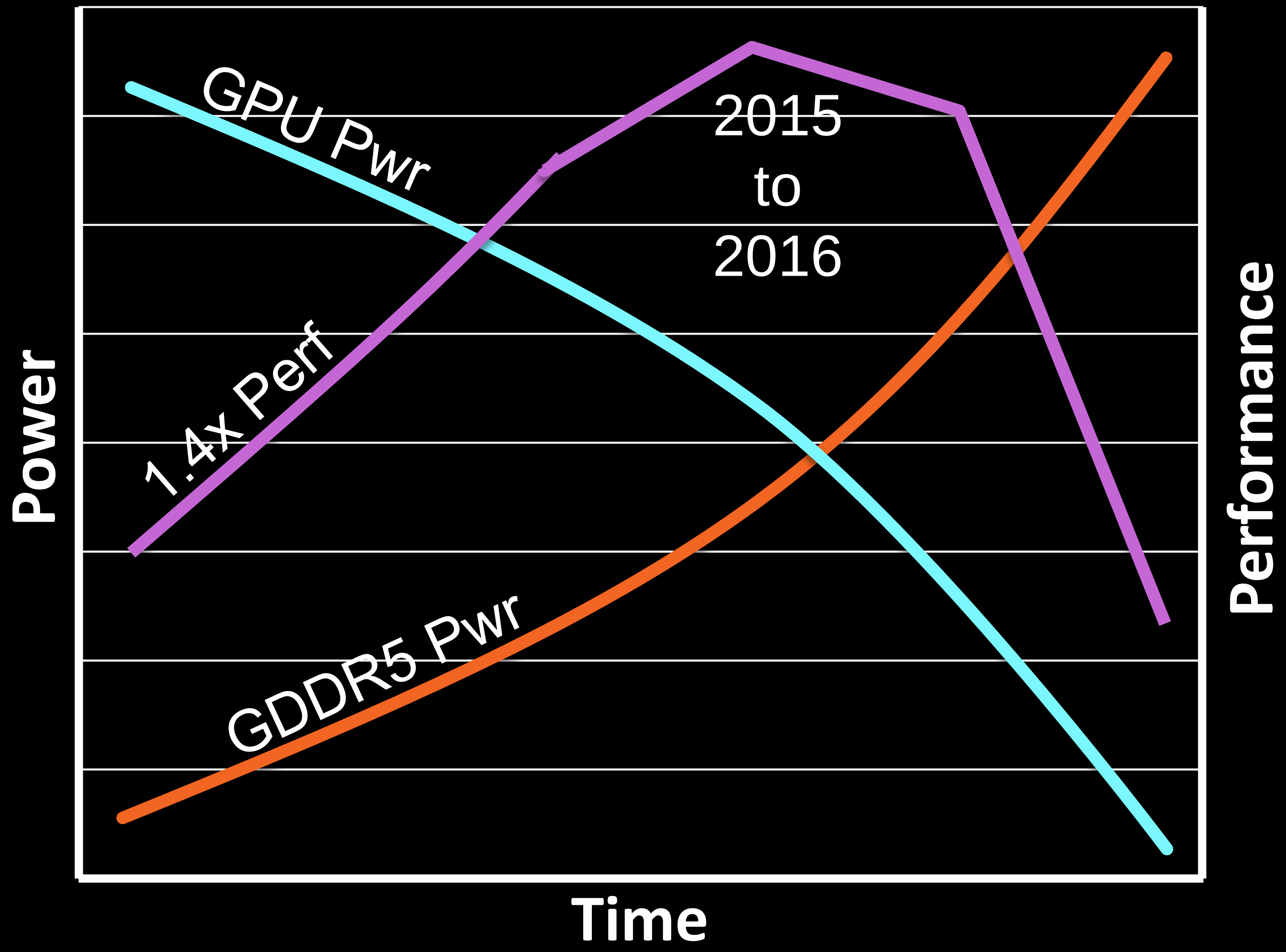
- As power is increasingly allocated to the memory system and taken away from compute performance growth slows



WHY DID WE BUILD FIJI AND HBM?

AN ANALYSIS FROM 2009

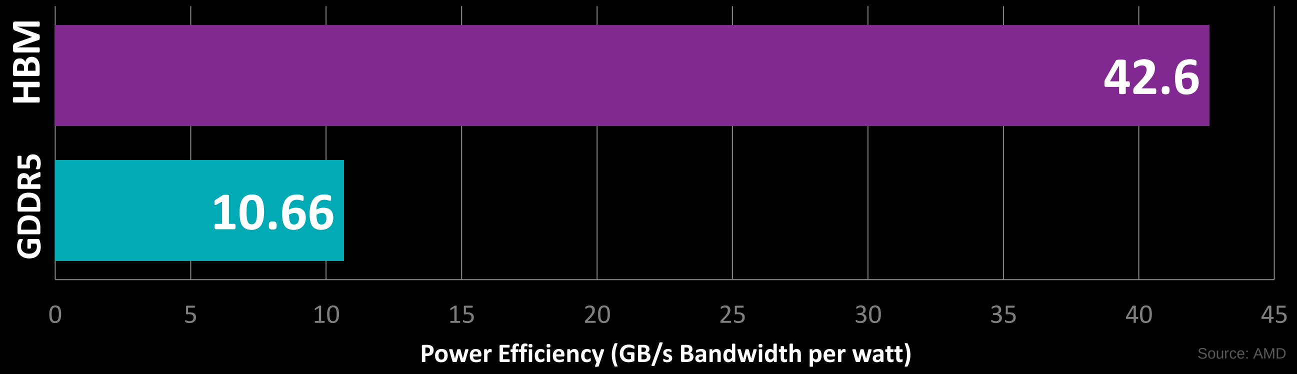
- At some point performance growth is not sustainable
- A new memory system with significantly better BW/W is required



“Fiji” Chip

HIGH-BANDWIDTH MEMORY

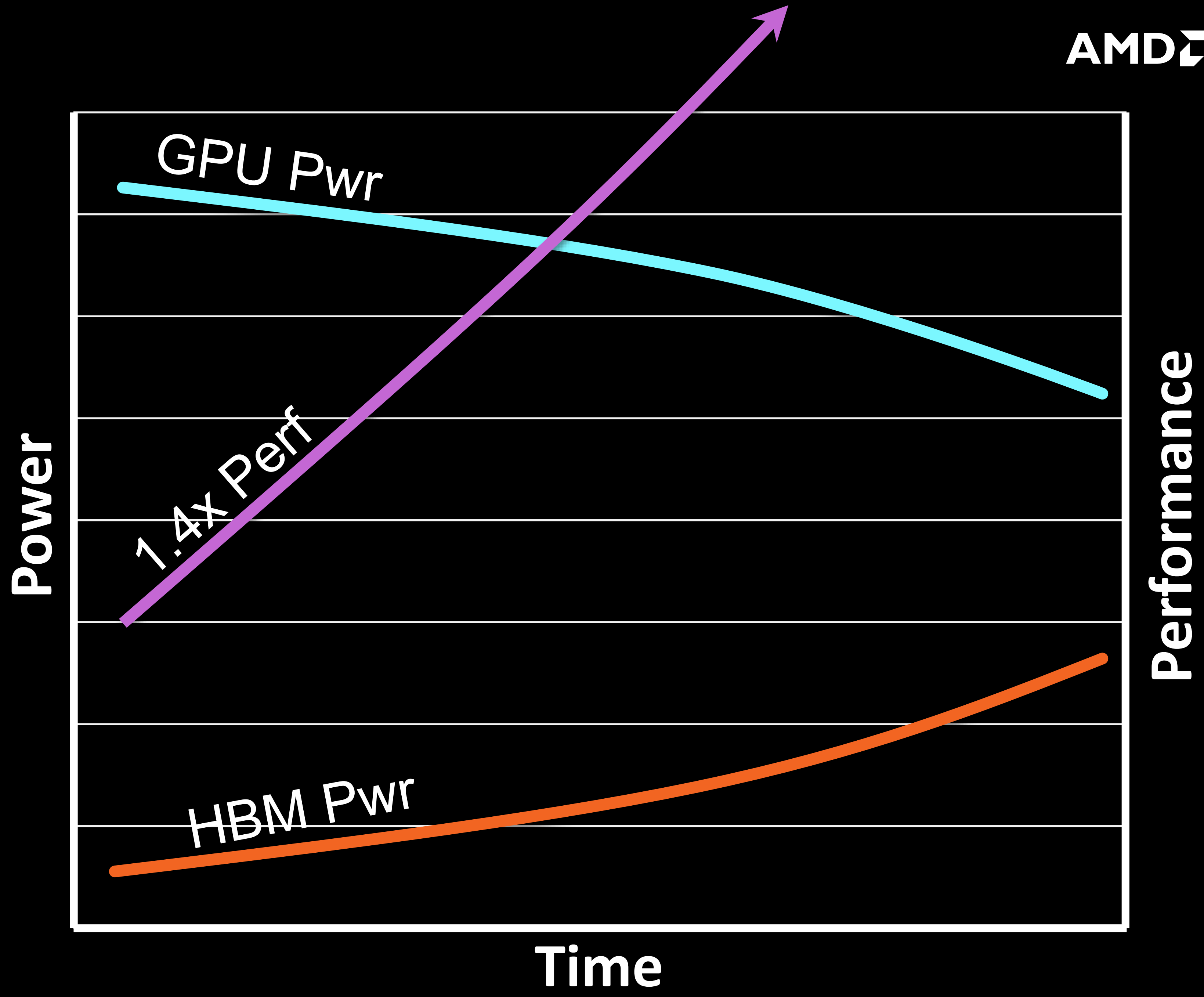
- Initiated with several DRAM partners 7 years ago
- SKhynix is in production supporting “Fiji”
- Benefits
  - 4096-bit memory interface with four stacks creating 512GB/s of bandwidth
  - 60% higher memory bandwidth[6] for 60% less power[7] than GDDR5
  - 4X Bandwidth per watt improvement from Radeon™ R9 290X
- Also required functional prototyping



NOW IN 2016

AMD

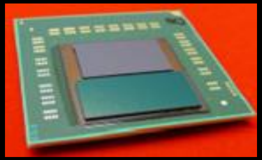
- HBM rolled the clock back and we have many years of performance scaling in front of us



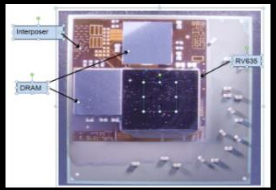
IT TOOK >15 PROTOTYPES OVER 8.5 YEARS

AMD

First Time Out



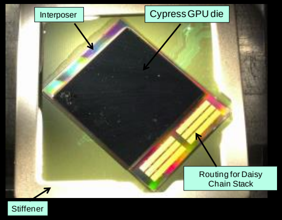
CPU + D3 Mech.



dGPU + G3

Primary Learning

345mm^2 ASIC
500mm^2 IP



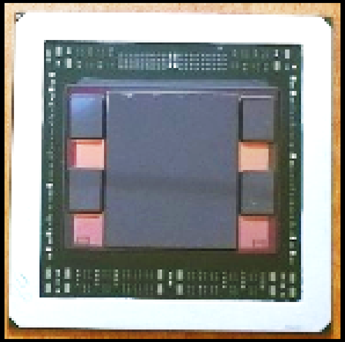
dGPU

502mm^2 ASIC
818mm^2 IP

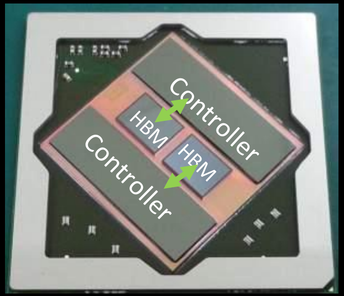
dGPU

Product Readiness

592mm^2 ASIC
1011mm^2 IP



"Fiji" Replica



Mission mode
HBM bringup

ESD | BLRT | Sort

Cost Down

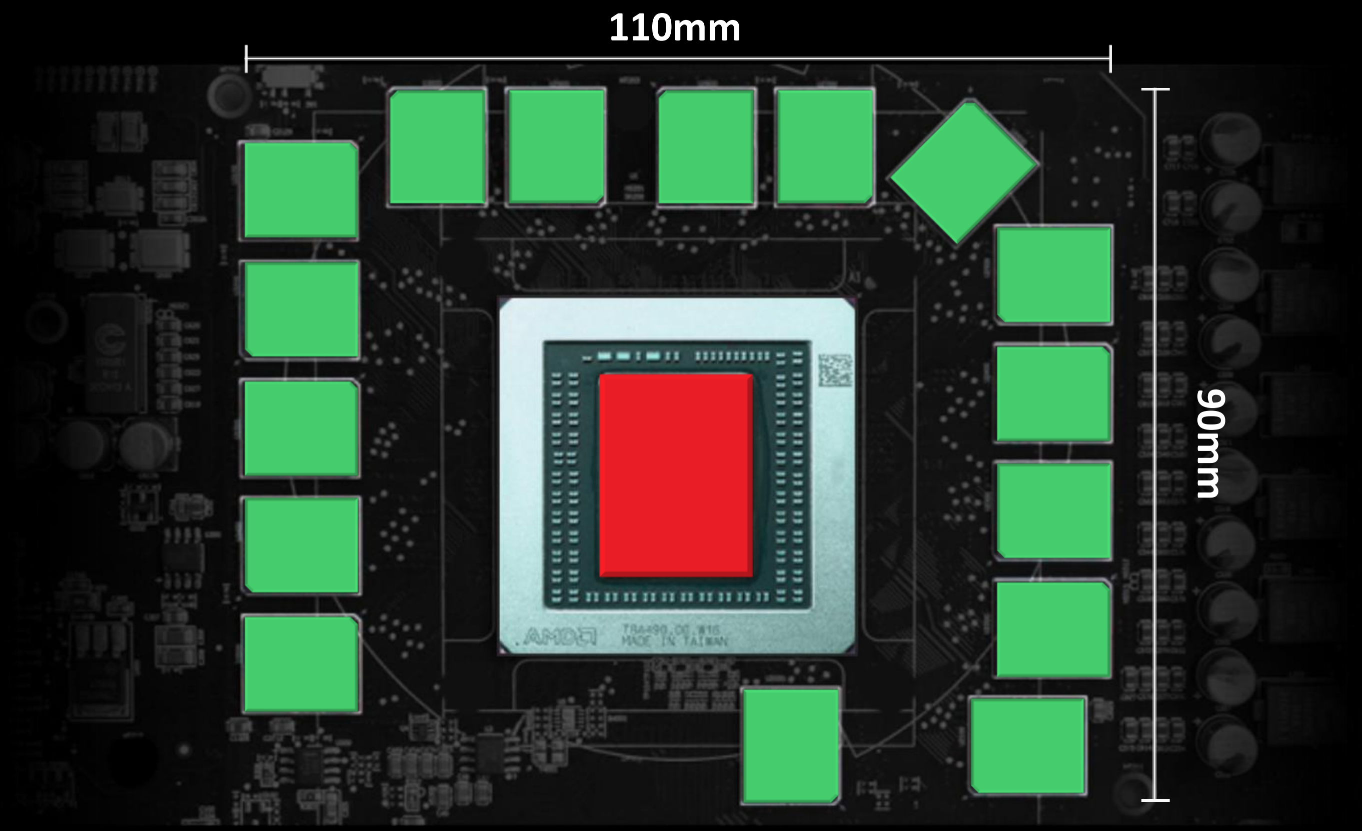
PwrCyc | uBump EM | TSV EM/SM

Component reliability: TC | uHAST | HTS

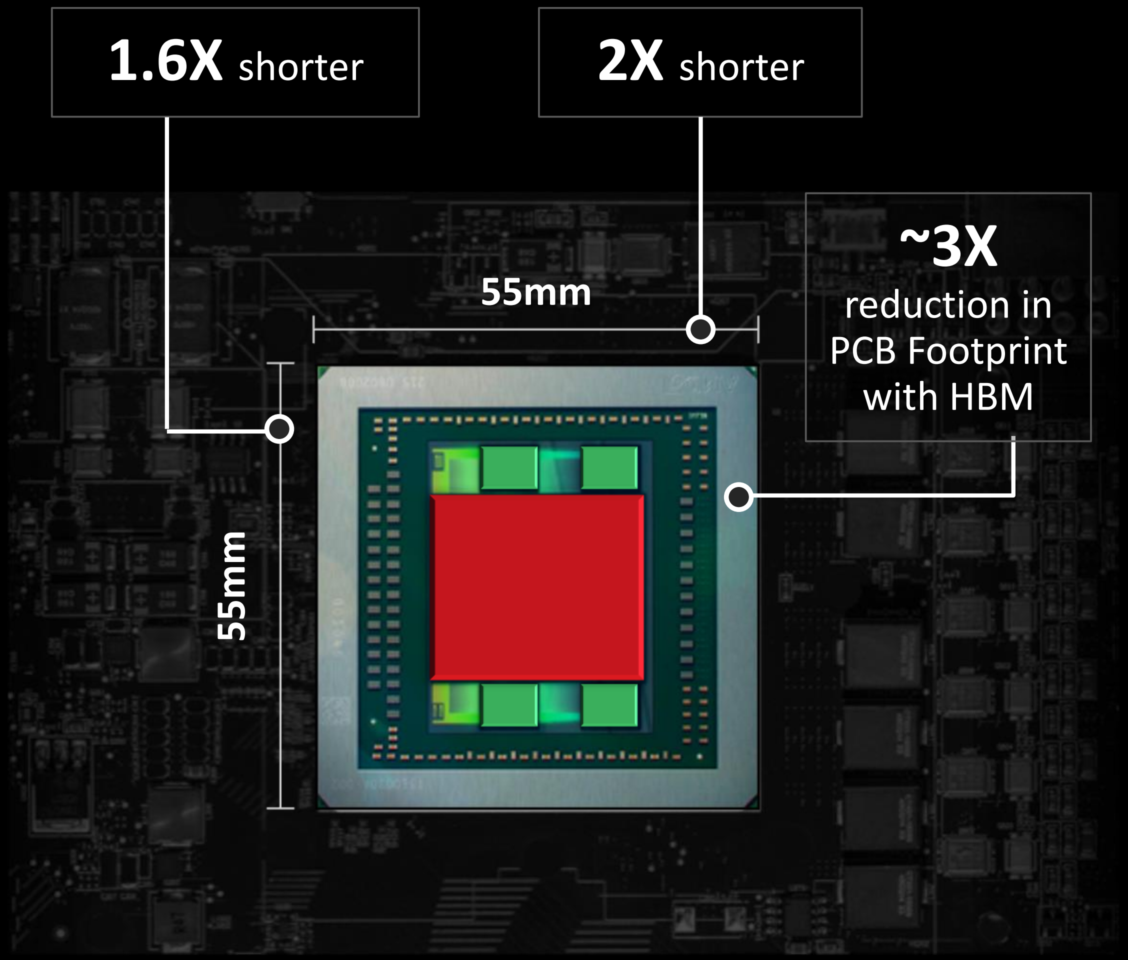
2007 **(100's of samples)** **2011** **(<5000 of samples)** **Jul'14** **(>5000 samples)** **2014**

“Fiji” Chip

EFFICIENT DESIGN



PCB area occupied by ASIC + Memory (Radeon™ R9 290X)



PCB area occupied by ASIC with HBM

WHAT IS NEXT FOR DIE STACKED MEMORY?

AMD

HBM1
&
DDR4

Expect Scaling

Traditional: BW with faster interfaces

Traditional: X-Y capacity with process node

New: BW with wider interfaces

New: Capacity with more die in the stack

A New
Stack

Focus

Cost reduction

3D usage model

DISCLAIMER & ATTRIBUTION

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. 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.

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 MERCHANTABILITY OR FITNESS FOR ANY PARTICULAR PURPOSE. IN NO EVENT WILL AMD BE LIABLE TO ANY PERSON FOR ANY 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.

ATTRIBUTIONS

FOOTNOTES

Required for all AMD Radeon™ graphics messaging: Additional hardware (e.g. Blu-ray drive, 4K monitor, TV tuner, wirelessly enabled HDTV) and/or software (e.g. multimedia applications) are required for the full enablement of some features. HD Video display requires an HD video source. Not all features may be supported on all components or systems - check with your component or system manufacturer for specific model capabilities and supported technologies.

1. Testing conducted by AMD engineering on the AMD Radeon™ R9 290X GPU vs. an HBM-based device. Data obtained through isolated direct measurement of GDDR5 and HBM power delivery rails at full memory utilization. Power efficiency calculated as GB/s of bandwidth delivered per watt of power consumed. AMD Radeon™ R9 290X (10.66 GB/s bandwidth per watt) and HBM-based device (42.66GB/s bandwidth per watt), AMD FX-8350, Gigabyte GA-990FX-UD5, 8GB DDR3-1866, Windows 8.1 x64 Professional, AMD Catalyst™ 15.20 Beta. HBM-1
2. Based on the product design, the Radeon™ R9 Nano is defined with an operating temperature target of 75°C while the Radeon™ R9 290X is defined with an operating temperature target of 95°C GRDT-75
3. Based on the product design, the Radeon™ R9 Nano is defined with a fan acoustic target of 42dBA while the Radeon™ R9 290X is defined with a fan acoustic target of 58dBA GRDT-77
4. Testing conducted by AMD Engineering on optimized AMD reference systems. PC manufacturers may vary configurations yielding different results. Far Cry 4 at 3840x2180, Ultra High preset, SMAA, 0XAF is used to simulate GPU performance; the Radeon™ R9 Nano on the system using the Intel® Core™ i7-5960X 3.0GHz processor, 16GB (4x4GB) DDR4 2666 MHz memory, Windows 10 64-bit, and AMD Catalyst Driver 15.201 scored 0.2169 fps/watt while the Radeon™ R9 290X on the same system and AMD Catalyst Driver 15.20 scored 0.1088 fps/watt GRDT-72
5. Testing conducted by AMD Engineering on optimized AMD reference systems. PC manufacturers may vary configurations yielding different results. Far Cry 4 at 3840x2180, Ultra High preset, SMAA, 0XAF is used to simulate GPU performance; the Radeon™ R9 Nano on the system using the Intel® Core™ i7-5960X 3.0GHz processor, 16GB (4x4GB) DDR4 2666 MHz memory, Windows 10 64-bit, and AMD Catalyst Driver 15.201 scored 0.2498 fps/mm while the Radeon™ R9 290X on the same system and AMD Catalyst Driver 15.20 scored 0.0989 fps/mm GRDT-71
6. Based on the memory bandwidth of the AMD Radeon™ R9 290X with a 1250MHz 512-bit GDDR5 interface (320GB/s) vs. AMD Radeon™ R9 Fury and R9 Fury X featuring HBM with a 500MHz 4096-bit interface (512GB/s). HBM-4
7. Testing conducted by AMD engineering on the AMD Radeon™ R9 290X GPU vs. the AMD Radeon™ R9 Fury X GPU. Data obtained through isolated direct measurement of GDDR5 and HBM power delivery rails at full memory utilization. AMD Radeon™ R9 290X and R9 Fury X GPU, AMD FX-8350, Gigabyte GA-990FX-UD5, 8GB DDR3-1866, Windows 8.1 x64 Professional, AMD Catalyst™ 15.20 Beta. HBM-3
8. Discrete AMD Radeon™ GPUs and AMD FirePro™ GPUs based on the Graphics Core Next architecture consist of multiple discrete execution engines known as a Compute Unit (“CU”). Each CU contains 64 shaders (“Stream Processors”) working in unison. GRT-5