# The Road to Data Center Power Efficiency

Tawfik Rahal-Arabi & Anshuman Mittal
Advanced Micro Devices, Inc.
Radeon Power and Performance Group
tawfik.arabi@amd.com & anshuman.mittal@amd.com

Acknowledgments: Raja Swaminathan, Tony Vivenzio, Yue Cao



AMD

# 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 and technology as well as 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.

# Outline

- Fundamentals and trends of power and performance (PnP)
- Data center power and performance trends
- The Mobile World
- Power, power delivery & power management opportunities with chiplets
- Summary



AMD

# Outline

- Fundamentals and trends of power and performance (PnP)
- Data center power and performance trends
- The Mobile World
- Power, power delivery & power management opportunities with chiplets
- Summary



AMD

# Moore's Law

Linear dimensions scale by 0.7X

Area scales by 0.7x0.7=0.5 or number of transistors per unit area doubles

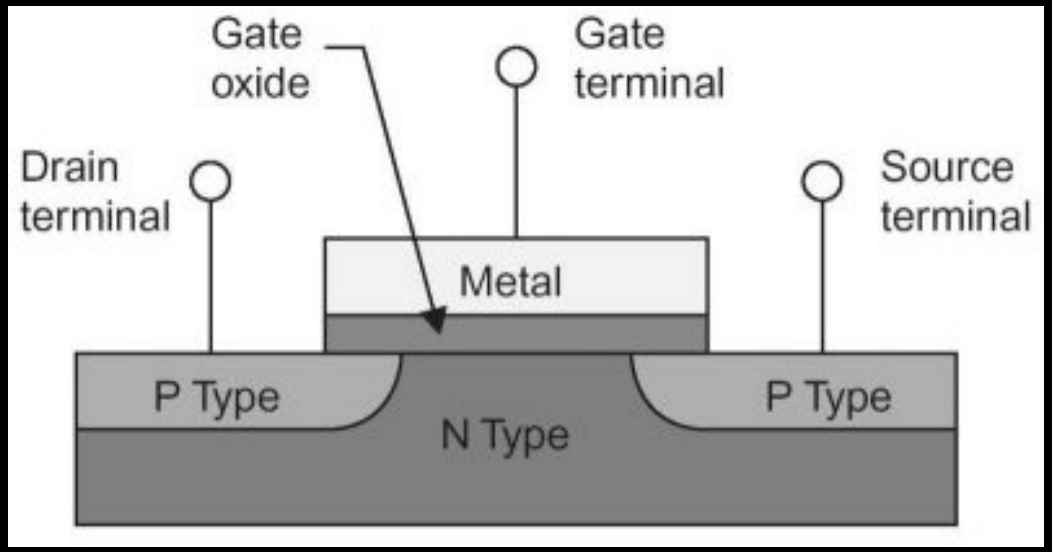
Voltage $(\mathrm{E} = \frac{v}{d})$ scales by 0.7x to maintain a constant electric field in the oxide

Capacitance $\varepsilon\frac{A}{d}$ scales by 0.7

Delay $\frac{1}{f}$ scales by 0.7X

Power ($cv^2f$) scales by *0.5*

***Number of transistors doubles, frequency improves by 40%, and power reduces by ½***

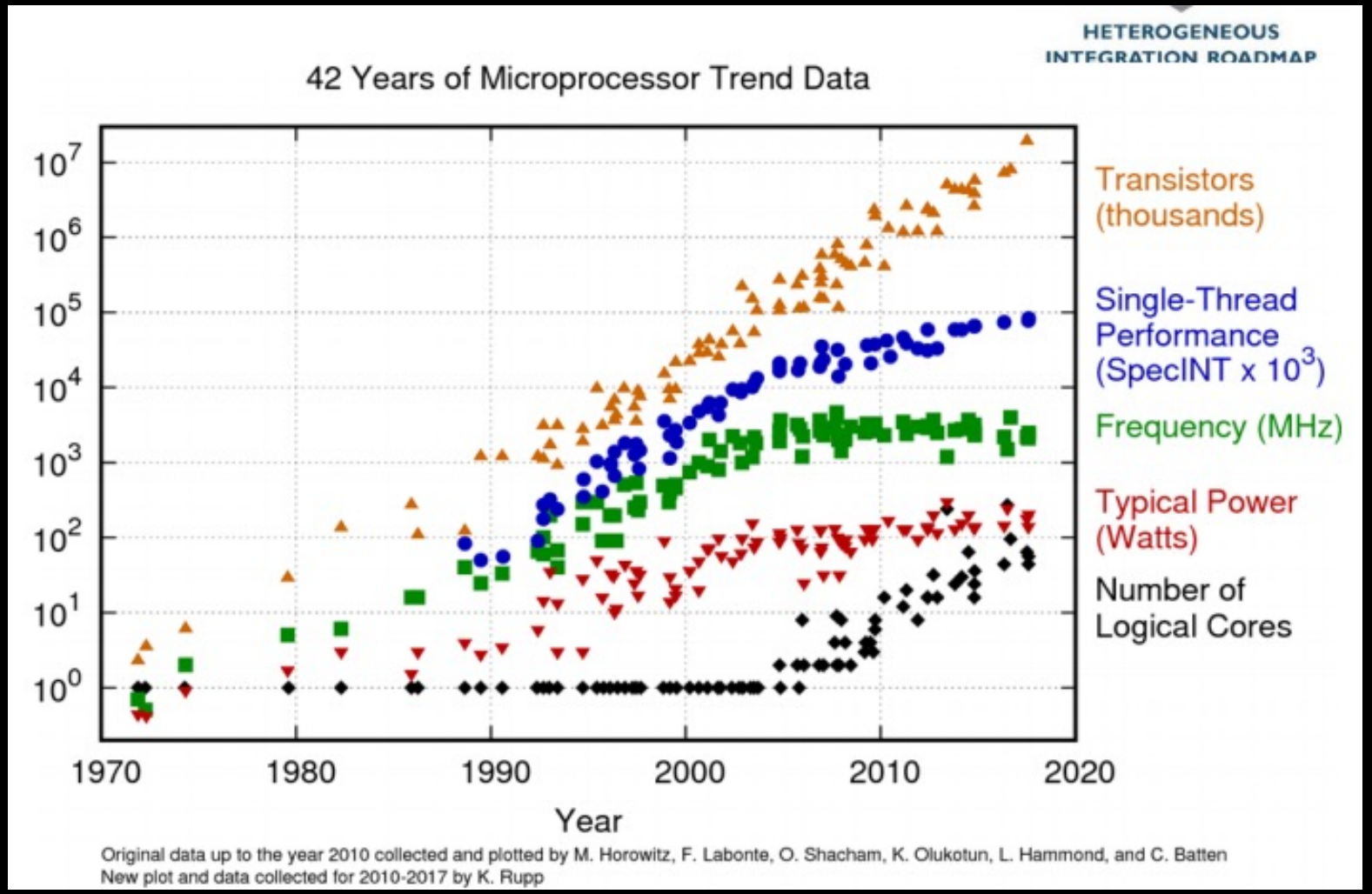



$$E = \frac{v}{d}$$



AMD

# The slowing down of Moore's law




Dale Becker and Kanad Ghose: Heterogenous integration for HPC and data centers



AMD

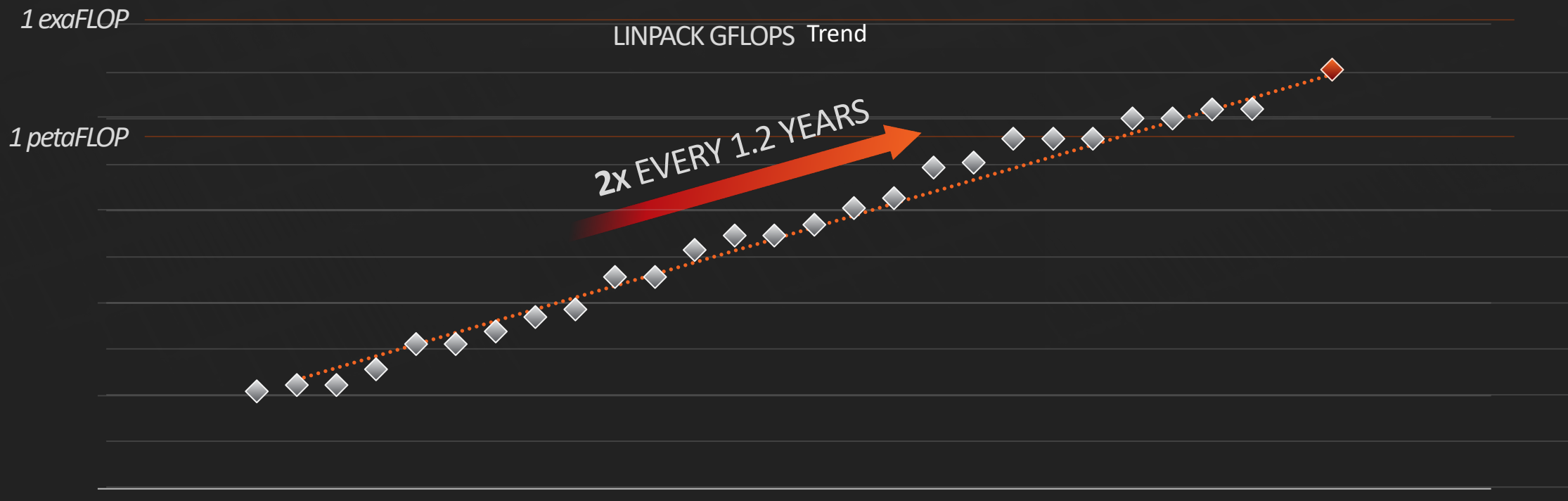
# Outline

- Fundamentals and trends of power and performance (PnP)
- Data center power and performance trends
- The Mobile World
- Power, power delivery & power management opportunities with chiplets
- Summary

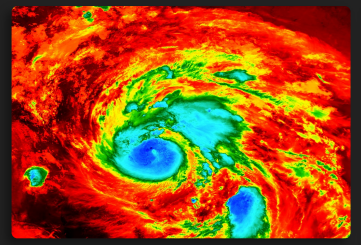


AMD

# THE RELENTLESS DEMAND FOR MORE COMPUTE..



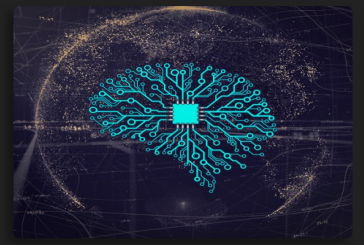

SPACE EXPLORATION



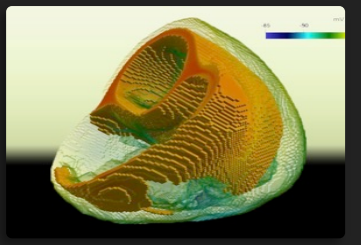
CLIMATE CHANGE

CHEMICAL SCIENCES

ENERGY SOLUTIONS



MACHINE LEARNING



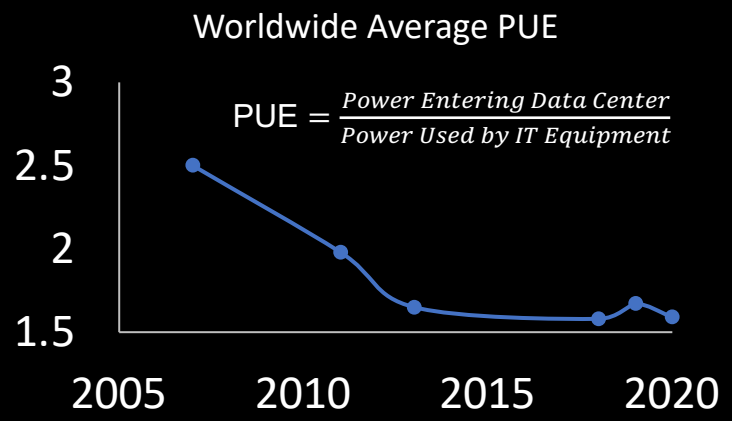
REAL TIME SIMULATION



Source: “Enabling Moore’s Law’s next frontier through 3DIC heterogeneous integration” by Raja Swaminathan, International Symposium on 3DIC, April,26-28, 2022

AMD

# Data center power efficiency trends




Today

- Data centers consume ~200TWH of total energy
- ~ 1 billion cores in use
- Best in class PUE ~ 1.1 but average industry @1.59
- Long way from compute energy proportionality
- Peak power is the number one limiter to performance of data centers, driver of e-waste, and provisional power

2025-2030

- 20% (~2% today) of total electricity in data centers by 2030 (IT Renew, OCP San Jose, Nov.2021)
- 5 billion cores by 2025 (Intel-Zane Ball, OCP San Jose, Nov. 2021)



AMD

# Outline

- Fundamentals and trends of power and performance (PnP)
- Data center power and performance trends
- The Mobile World
- Power, power delivery & power management opportunities with chiplets
- Summary



AMD

# Mobile devices from the 90's to today

## Early 90's – first mobile 'laptops'

Very primitive power and thermal management schemes

- DOS-based 286 through Pentium® processor designs
- BIOS and then SMM-based power management

Hard to figure out basic things

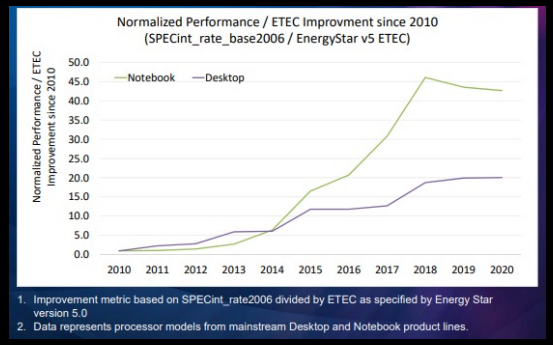
- Is the user present? (timer/trap on KBD/mouse I/O)
- Is the HDD in use? (timer trap on disk I/O)
- Is the system compute busy? (timers based on all relevant I/O and processing)
- Is the CPU idle? (I don't know, OS is always doing something!)

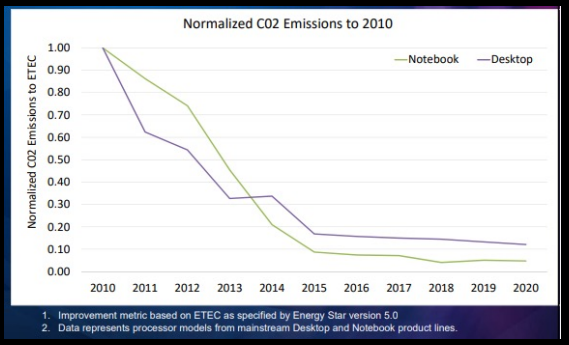
2-3 hours of battery life was great…


CLIENT COMPUTING GROUP

intel | 3

- > 10 Hrs. of active use battery life
- > Weeks of airplane mode/ Modern Standby
- Sophisticated platform power management




Barnes Cooper: Bits 2018 distinguished speaker



AMD

# Client vs. data center power management

- Order of magnitude power reduction with single client power management
- Idle power is a very small percentage of total power
  - Leakage reduction and power management
  - Battery life is an end user experience
- Energy Star like benchmarks for environmentally conscious end users

- Multi-client power management is harder
- Idle power in data centers considerable percentage of total
  - Leakage reduction
  - No end user experience equivalent of battery life
- No end user sustainability metric equivalent to Energy Star
  - Energy Star VM?



AMD

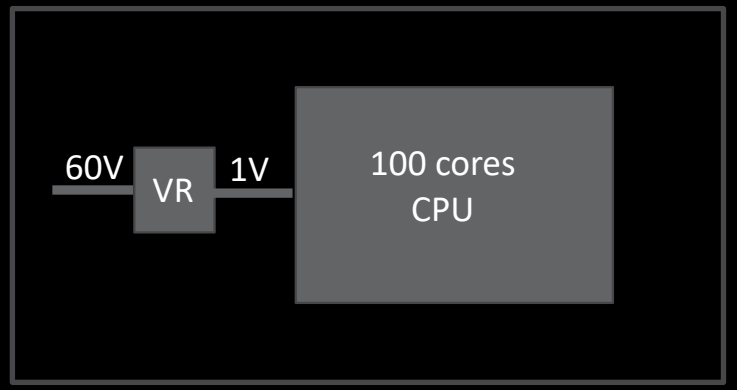
# Outline

- Fundamentals and trends of power and performance (PnP)
- Data center power and performance trends
- The Mobile World
- Power, power delivery & power management opportunities with chiplets
- Summary

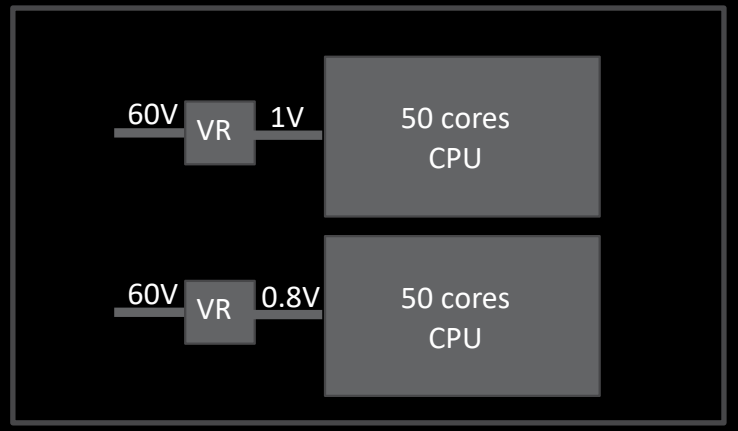


AMD

# Power savings with chiplets




Server Motherboard




Server Motherboard

25% Power savings with 2 CPUs and 2 VRs instead of 1CPU and 1 VR

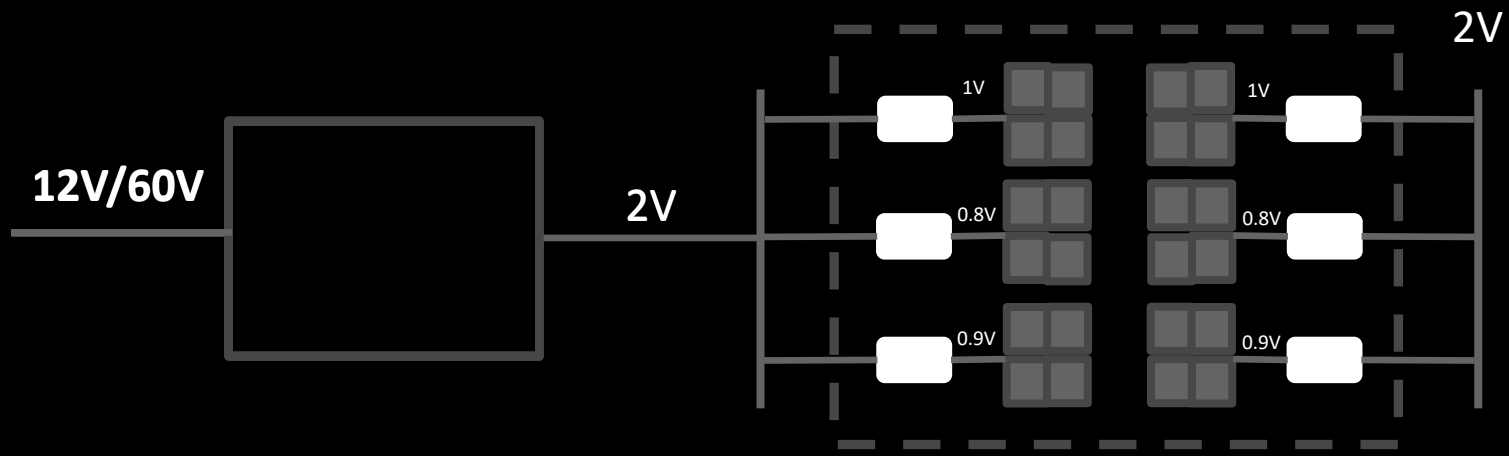
Challenges

- 2 VRs and 2 CPUs, higher cost and bigger real estate
- Cores allocation

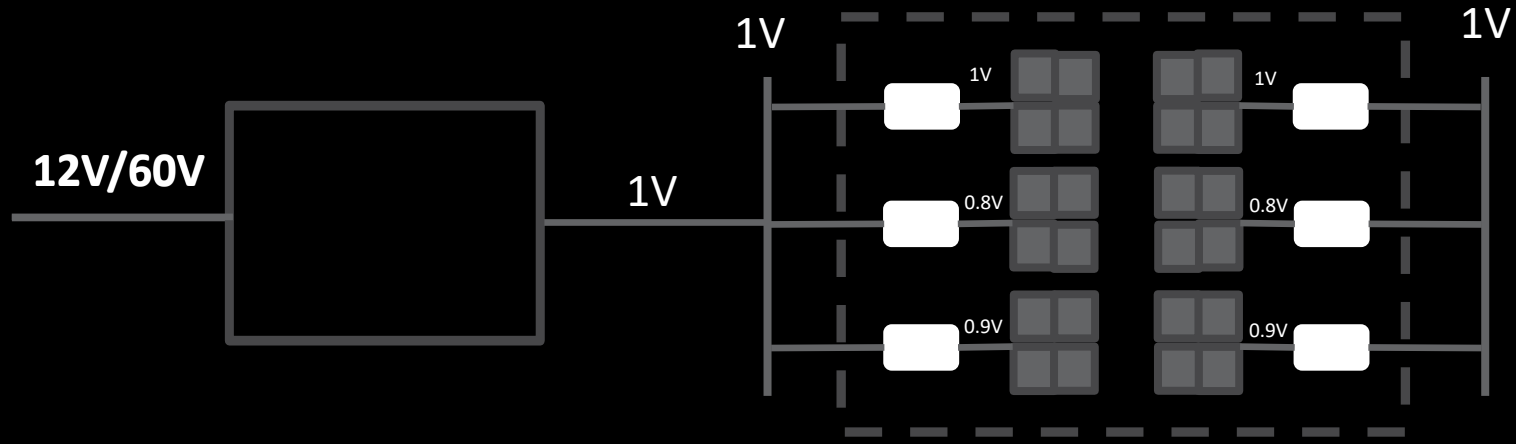


AMD

# Finer grain power distribution with more chiplets and/or integrated voltage regulators




FIVR
Higher complexity
Lower current into the package
Higher efficiency over larger dynamic range




LDOs
Lower complexity
Lower efficiency over large dynamic range
More efficient when dropout % is less than FIVR loss %



AMD

# Finer grain allocation software

| VM | Cluster | # of Cores | Voltage | Performance |
|---|---|---|---|---|
| A | Low Power | 6 | LV | 2.5 GHZ |
| B | HPC | 7 | HP | 3.5 GHZ |
| C | GP | 7 | MV | 3 GHZ |

Cluster

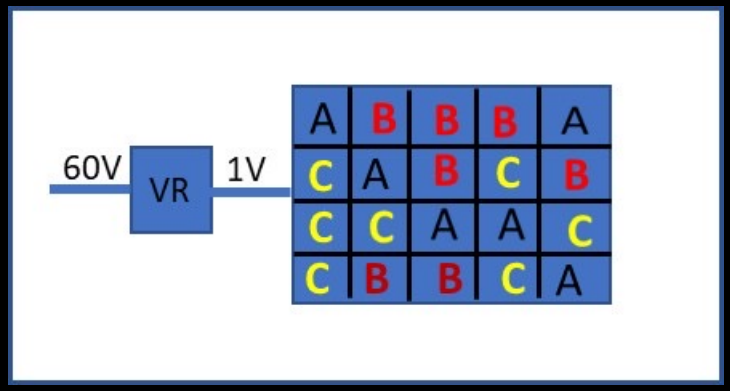
Rack

Blade

Board


Processor

Better MI based strategies for improved power management considering both the s/w & h/w aspects of the system






AMD

# Outline

- Fundamentals and trends of power and performance (PnP)
- Data center power and performance trends
- The Mobile World
- Power, power delivery & power management opportunities with chiplets
- Summary



AMD

# Summary

- Moore’s law slowed but performance did not
- Heterogenous integration (chiplets) helps performance scaling, but can add power and complexity
- Chiplet technology is inherently  better suited for finer grain power delivery and management
- Fine grain (HW and SW) power management can reduce overall data center power significantly



AMD

# COPYRIGHT AND DISCLAIMER

©2022 Advanced Micro Devices, Inc.  All rights reserved.

AMD, the AMD Arrow logo, EPYC, Ryzen, Infinity fabric, and combinations thereof are trademarks of Advanced Micro Devices, Inc.
Other product names used in this publication are for identification purposes only and may be trademarks of their respective companies.

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