

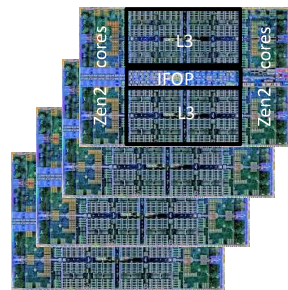
AMD Chiplet Architecture for High-Performance Server and Desktop Products

Samuel Naffziger

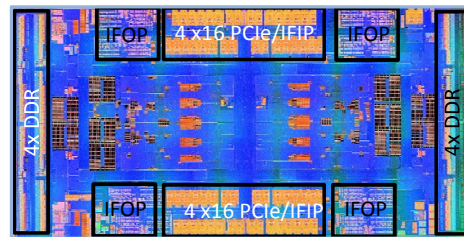


Outline

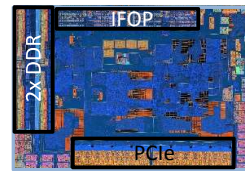
- **Motivation and architectural goals**
- **Engineering challenges and solutions**
 - Silicon-package co-design
 - Die-to-die interconnects
 - Shared IO die architecture
 - Power distribution and management
- **Results**



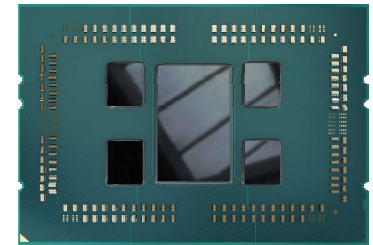
7nm Core Complex Die:
3.8 Billion FETs, 74 mm²



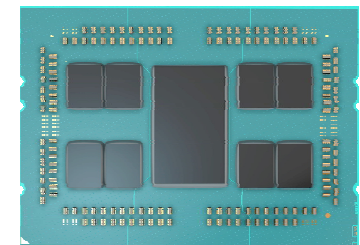
Server IO Die
8.34 Billion FETs, 416 mm²



Client IO Die
2.09 Billion FETs, 125 mm²



3rd Gen AMD Ryzen™
Threadripper™ Processor



2nd Gen AMD EPYC™



3rd Gen
AMD
Ryzen™
Processor



AMD X570 Chipset

Motivation and Architectural Goals

Primary goal:

Achieve leadership performance, performance/Watt and performance/\$ in server and desktop markets

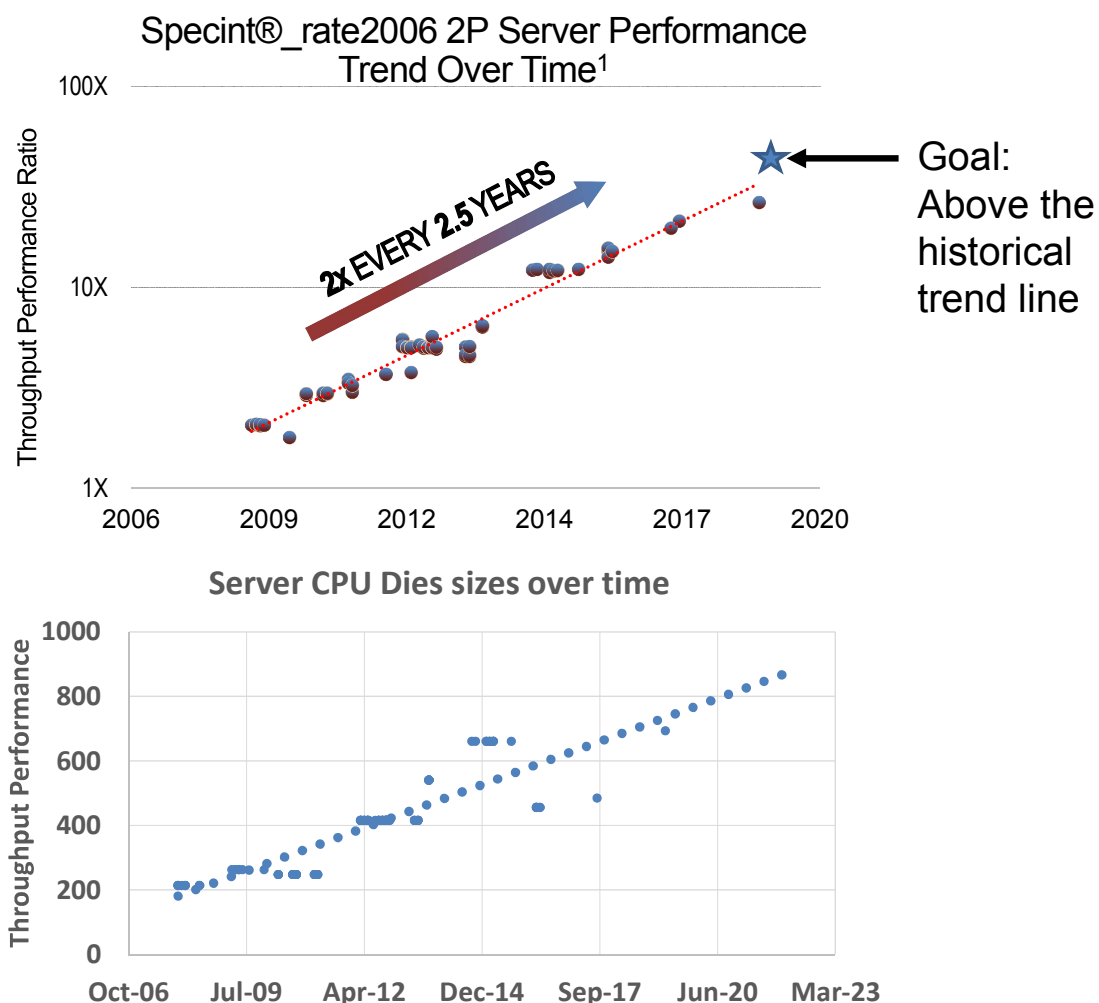
- **This required**
 - Exploiting advanced 7nm technology for better performance and performance/Watt
 - Packing more silicon into the package than traditional approaches enable
- **While also**
 - Enabling scalable performance/\$ up to performance levels otherwise not achievable
 - Improving memory and IO latency
 - Supporting leverage across markets by re-using IP and SOCs

Background: Performance and Die Size Trend

- Generational performance improvements are an exponential trend
- Holding to this trend has required increasing core counts and die sizes
- Bumping up against the reticle limit and becoming too costly

1. Su, Lisa "Delivering the Future of High-Performance Computing", Hot Chips 31 (2019)

© 2020 IEEE
International Solid-State Circuits Conference



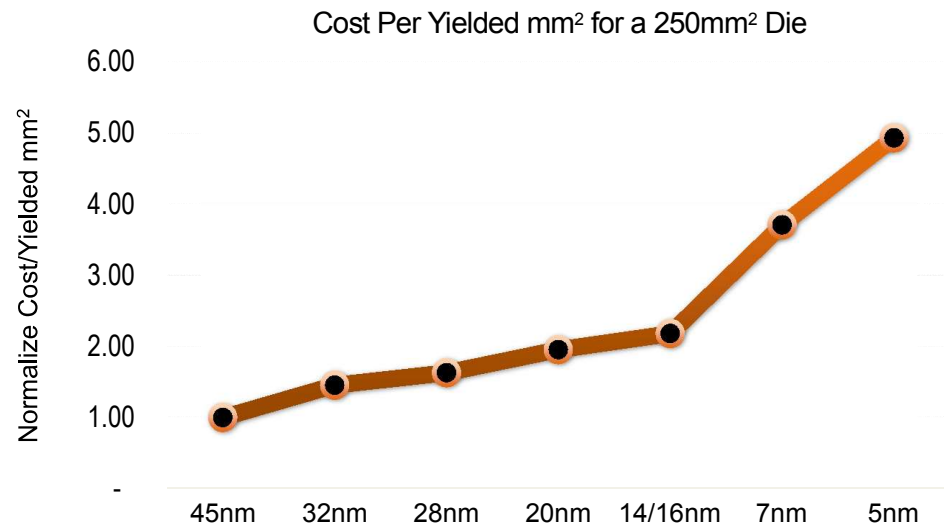
2.2: AMD Chiplet Architecture for High-Performance Server and Desktop Products

Exploiting 7nm Technology

- Leadership performance requires 7nm benefits
- Yet the cost of advanced technologies are increasing
- Traditional approaches of large die sizes not viable
- Innovation required

2X DENSITY¹ **>1.25X FREQUENCY¹** **0.5X POWER¹**
(same power) (same performance)

7nm Compute Efficiency Gains

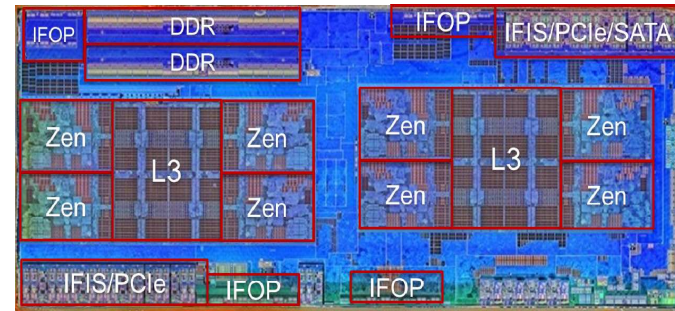


1. Based on June 8, 2018 AMD internal testing of same-architecture product ported from 14 to 7 nm technology with similar implementation flow/methodology, using performance from SGEMM.

7nm Scaling

- High-performance server and desktop processors are IO-heavy
- Analog devices and bump pitches for IO benefit very little from leading edge technology, and that technology is very costly
- **Solution:** Partition the SOC, reserving the expensive leading-edge silicon for CPU cores while leaving the IO and memory interfaces in N-1 generation silicon

Prior Generation RYZEN™ Processor Die



- CPU core + L3 on this die comprises 56% of the area
- These circuits see huge 7nm gains
- Remaining 44% sees very little performance and density improvement from 7nm



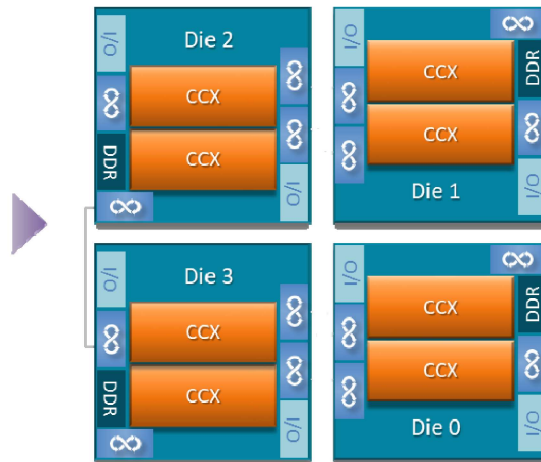
7nm CCD is
86% CPU + L3

Chiplets Evolved – Hybrid Multi-die Architecture

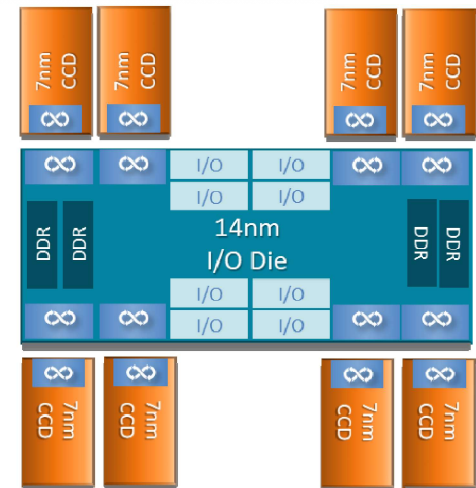
Traditional Monolithic



1st Gen EPYC



2nd Gen EPYC



Use the Most
Advanced Technology
Where it is Needed
Most

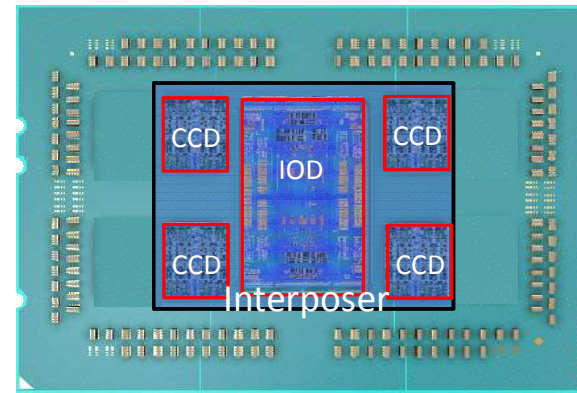
Each IP in its Optimal
Technology, 2nd Gen
Infinity Fabric™
Connected

Centralized I/O Die
Improves NUMA

Superior
Technology for
CPU Performance
and Power

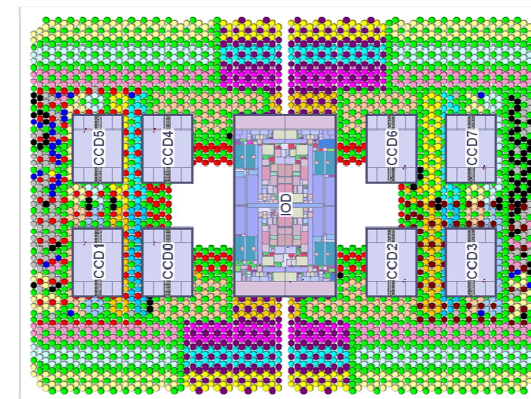
Connecting the Chiplets

Theoretical Interposer-based



- Silicon interposers and bridges provide high wire density, but have limited reach
- Only supports die edge connectivity which limits number of chiplets and cores that can be supported
- Performance goals required more Core Complex Die (CCDs) than can be tiled adjacent to the IOD
- **Solution** is to retain the on-package SerDes links for die-die connections

Selected MCM Approach



CPU Compute Die (CCD) Floorplan

2 CCX core complexes

- 4 core and 16MB L3 each
- Comprise 86% of CCD area

System Management Unit (SMU)

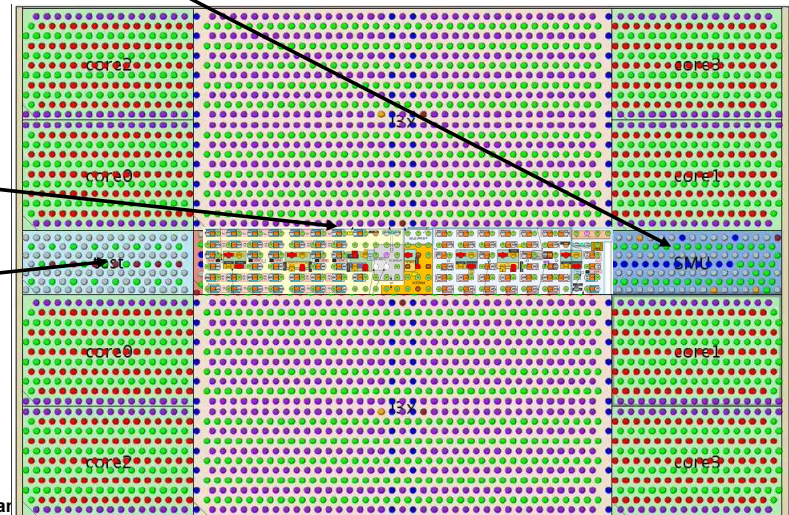
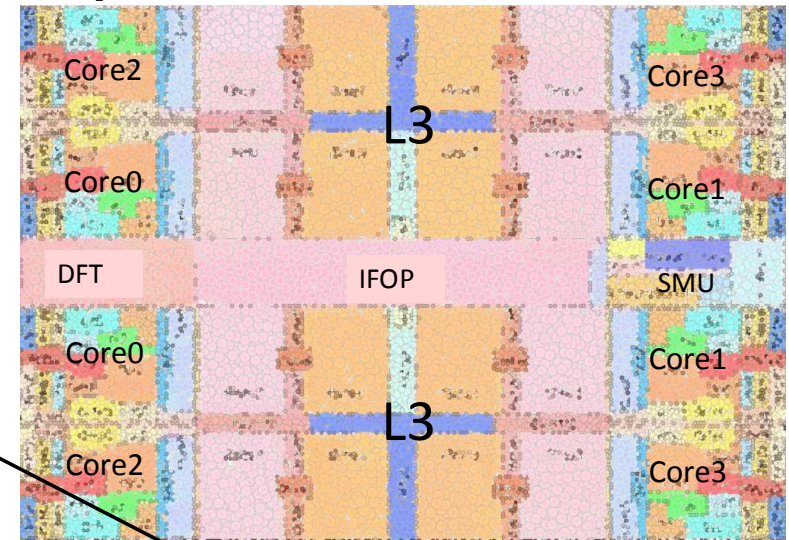
- Microcontroller
- Power management
- Clocks and reset
- Fuses
- Thermal monitor and control

Infinity Fabric On-Package (IFOP) Links

- 14.6 GT/s (packing 10 bits at 1.46Ghz)
- 39 RX lanes – 2 clock lanes – 1 clock gating lane
- 31 TX lanes – 1 clock gating lane
- 4 lanes for control traffic – 2 clock lanes

DFT and Debug

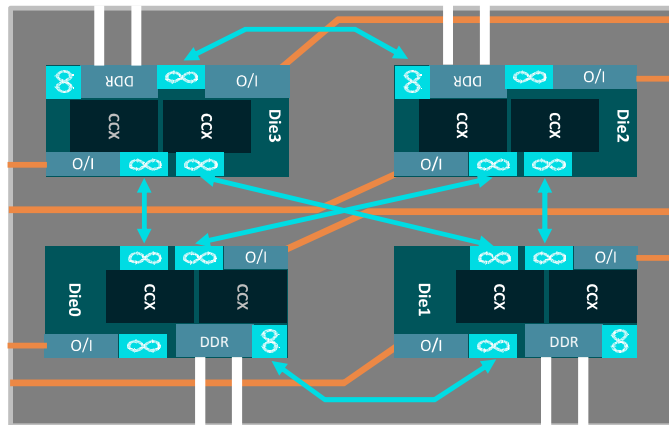
Wafer test bumps



IFOP GEN2 KEY FEATURE SUMMARY AND COMPARISON

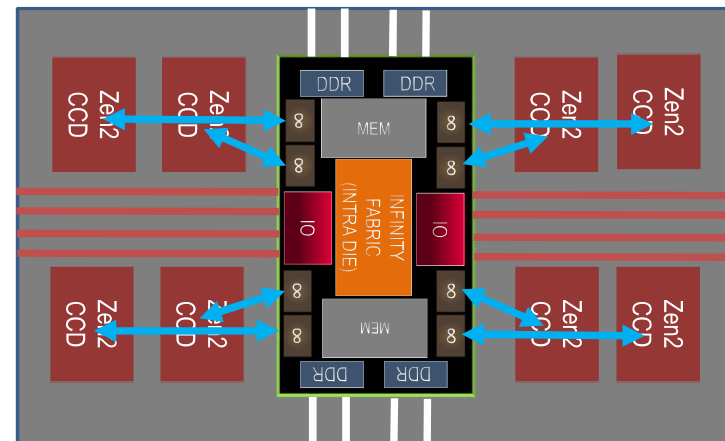
Gen1 14nm

Max per lane datarate 6.4Gbp/s	Local clock alignment and global tracking
4:1 Serialization/Deserialization	50/100/200 Ohm drive strength and termination
Forwarded clocks	PHY Regulated through package



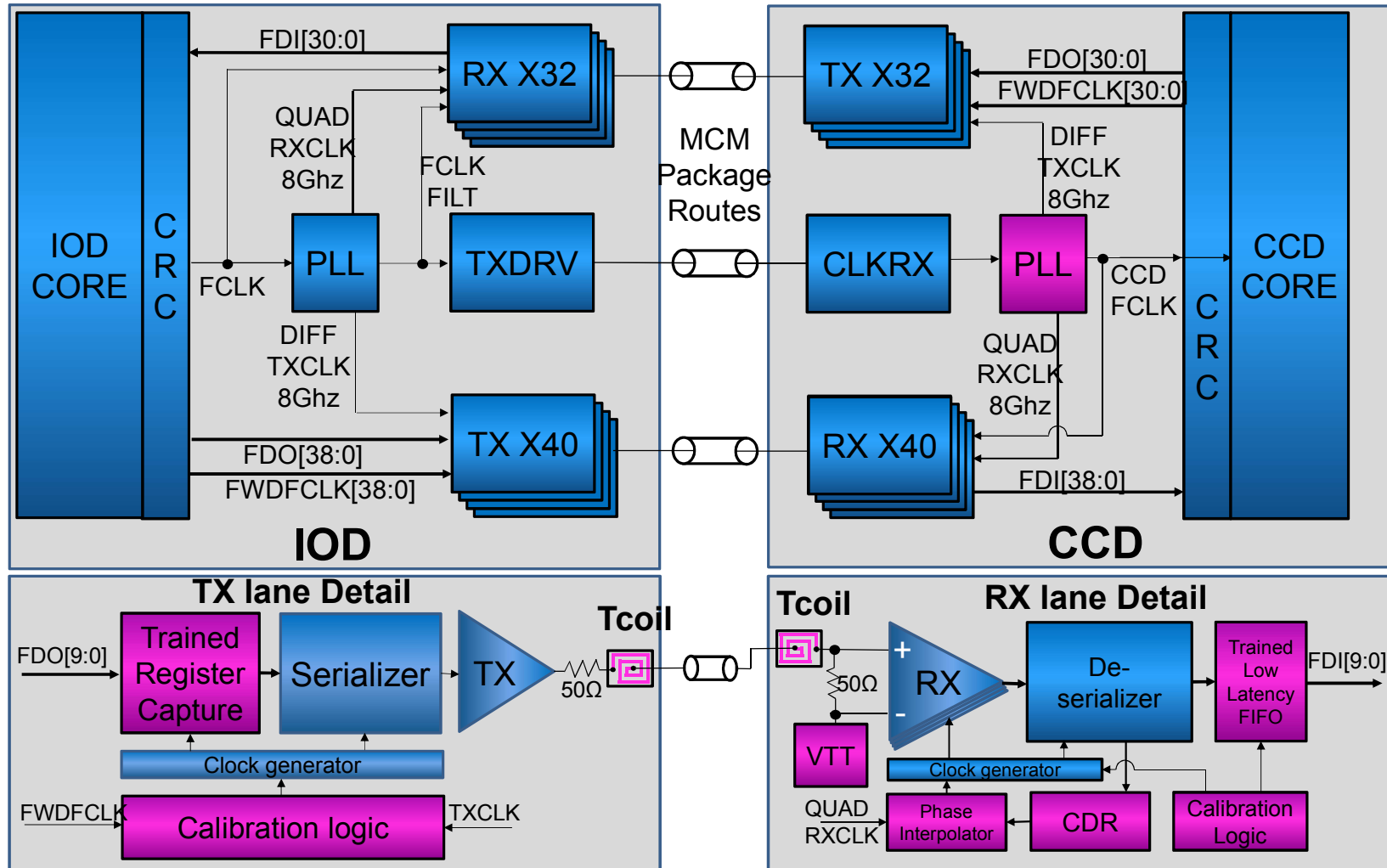
Gen2 14nm IOD, 7nm CCD

Max Per lane Datarate 14.6Gbps	Synchronous clock crossing Local CDR
10:1 Serialization/Deserialization	50 Ohm fixed drive strength and termination
TX and RX T-Coil	Local PHY Regulators
VTT Termination	Pseudo-Diff Single Ended Receiver



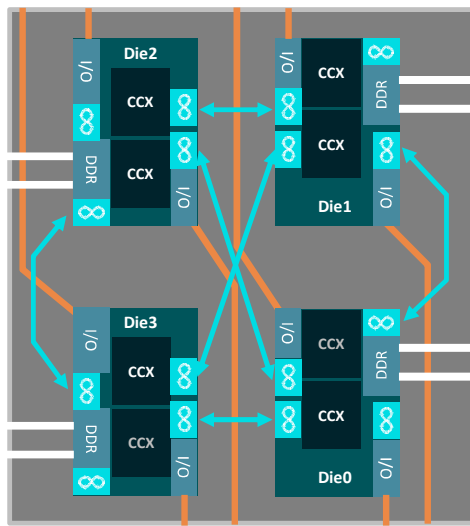
IFOP SerDes Architecture

 New Gen2 Features

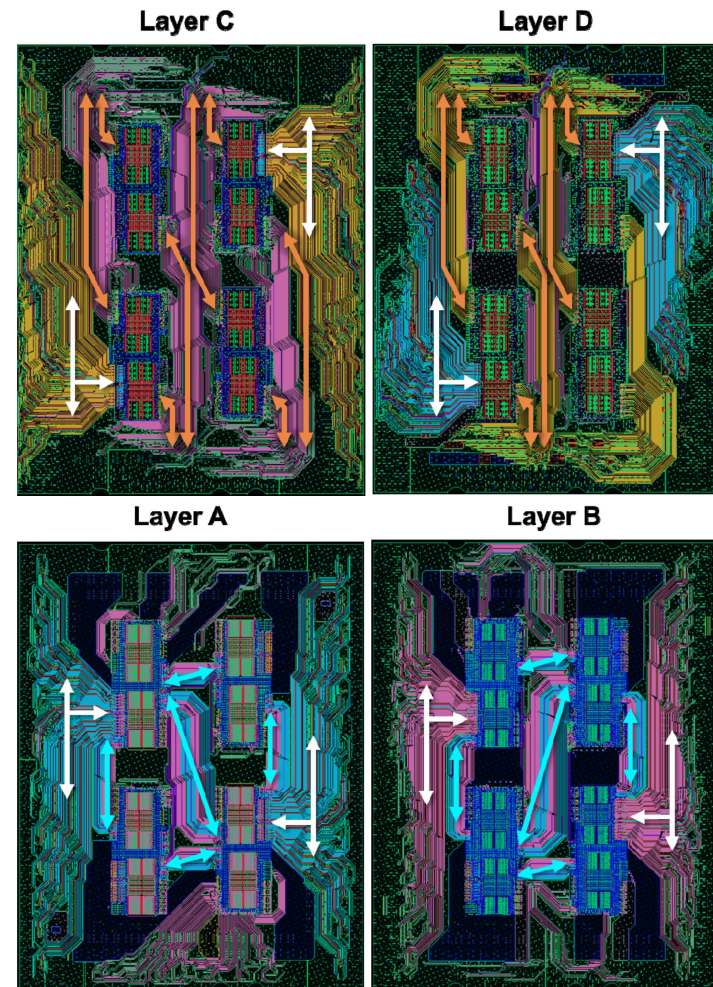


Package Routing Challenges

- Prior generation already consumed all package routing resources for memory and IO
- Connecting 9 chiplets in the same package requires innovation



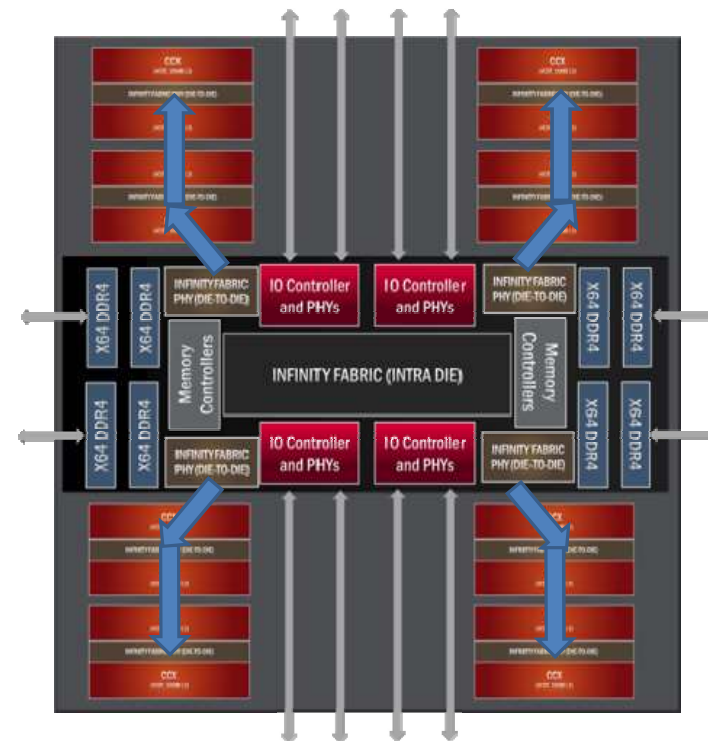
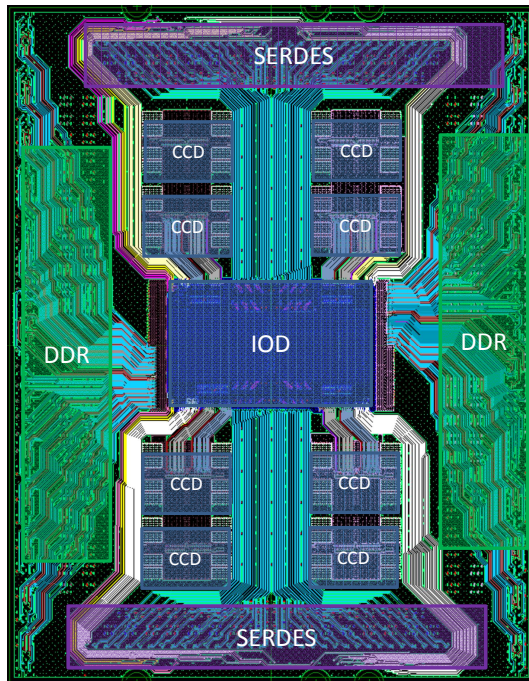
1st Gen AMD EPYC™



2.2: AMD Chiplet Architecture for High-Performance Server and Desktop Products

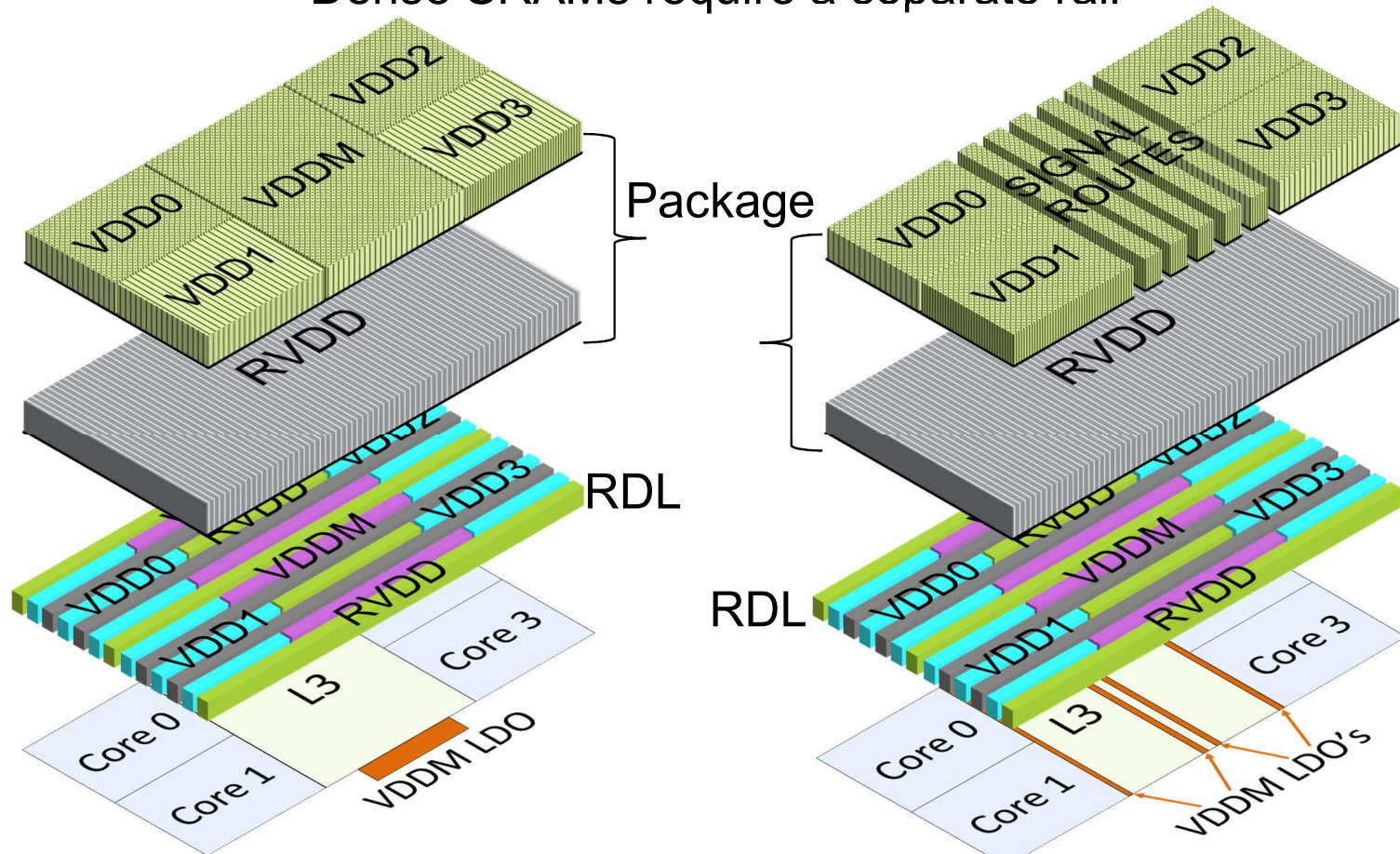
Under-CCD Routing

Routing Infinity Fabric on Package (IFOP) SerDes links from IOD to the 2-deep chiplets required sharing routing layers with off-package SerDes and competing with power delivery requirements



Zen vs. Zen 2 VDDM Distribution

Dense SRAMs require a separate rail



Zen VDDM distribution via package plane

Zen 2 VDDM distribution via RDL only

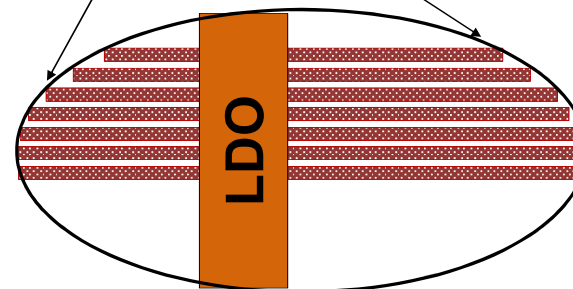
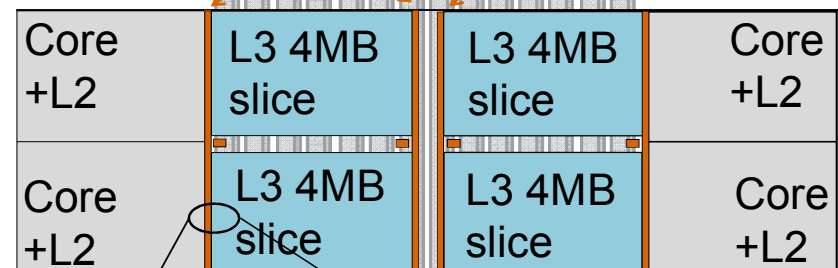
2.2: AMD Chiplet Architecture for High-Performance Server and Desktop Products

Zen 2 VDDM Design Challenges

- RDL is more resistive than a dedicated package layer
- Therefore we reduced overall VDDM current draw by 80% compared to Zen ([Singh ISSCC 2020])
- New, smaller, and distributed LDO design
- Ensured sufficient routing porosity through the integrated LDO's to enable critical routing
- These improvements kept the IR drop to $\approx 10\text{mV}$ impact

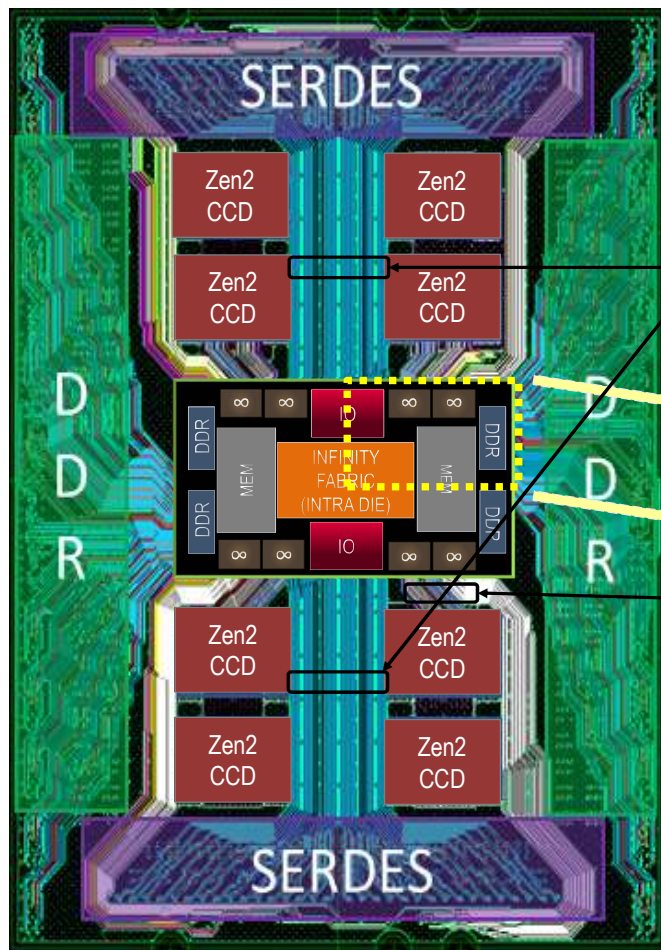
Enables 80 IFOP package routed signals under the CCD

4 VDDM LDO's inside the L3

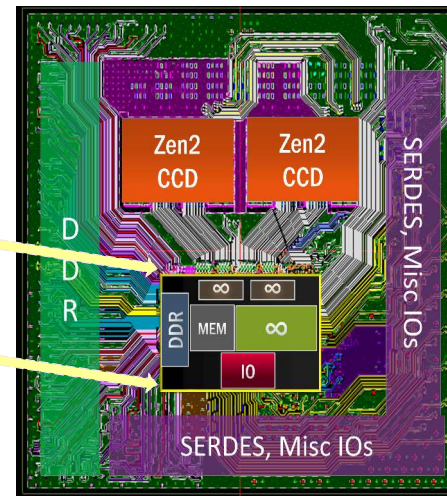


VDDM
RDL
spanning
L2 and L3




Package Integration, Server, and Desktop

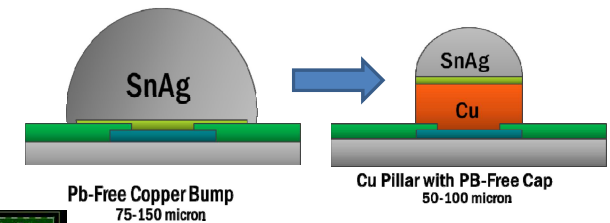


2nd Gen AMD EPYC™ Server Processor



3rd Gen AMD Ryzen™ Processor

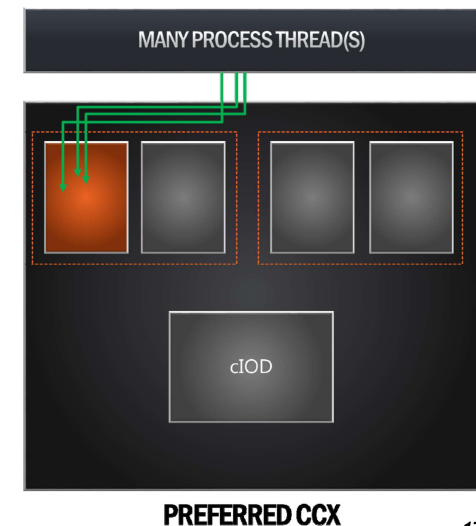
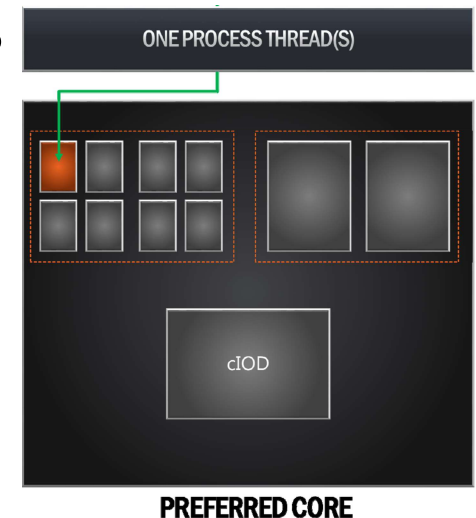
-  Infinity Fabric (die-to-die)
-  IO Controllers and PHYs
-  2 x DDR4 PHYs



- Bump pitch for 14nm and 7nm is 150um and 130um respectively
- Transitioned IOD from solder bumps to copper pillars, enabling a common interface for IOD+CCD
 - Conducive to tighter bump pitches (compact)
 - Enabled common die height after assembly
 - Higher max current (electromigration) limits

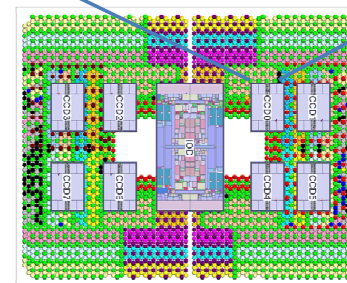
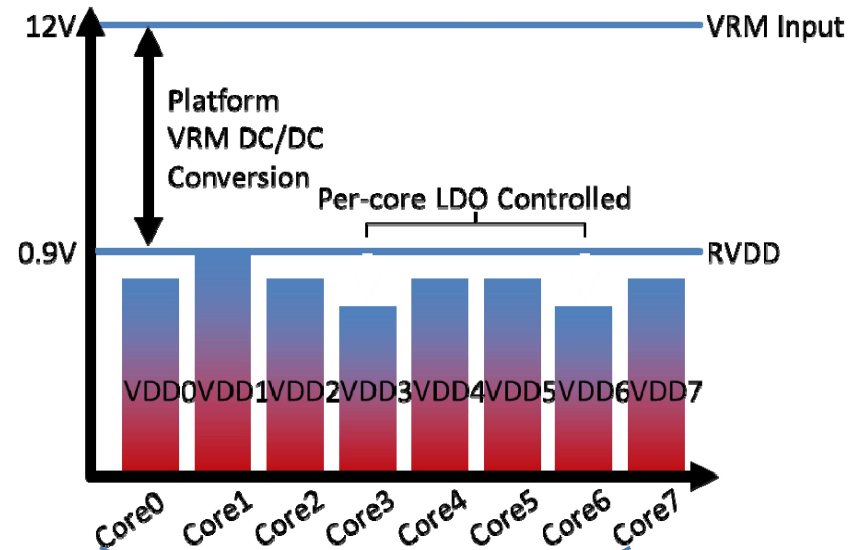
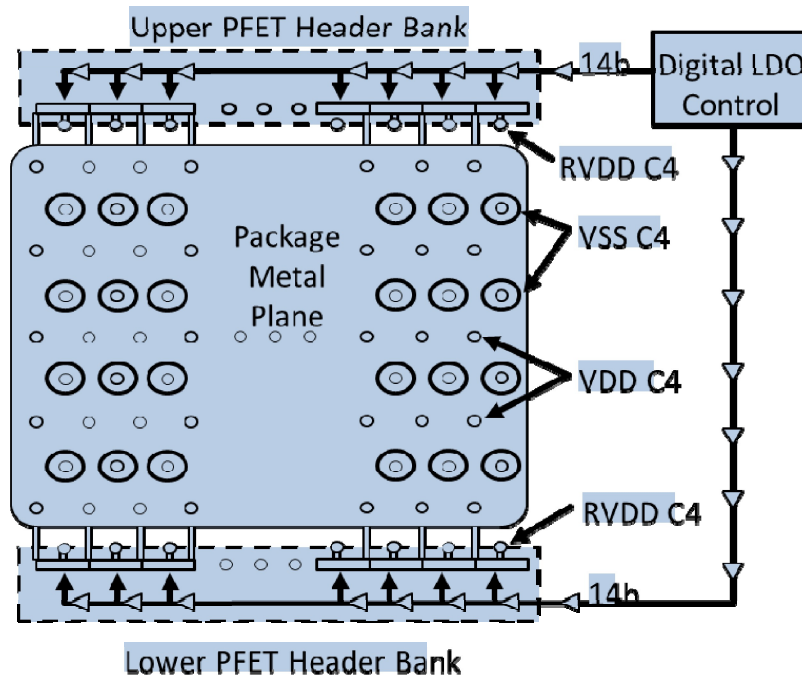
Operating System Scheduler Optimizations

- **Growing number of cores and the advent of chiplets resulted in a wider range of frequency responses to process, voltage and temperature variations**
 - Up to 200MHz core-to-core Fmax upside within a CCD
 - Legacy boost approaches don't take advantage of the faster cores
- **Preferred Core Ordering maximizes performance**
 - New algorithm characterizes the capabilities of the cores at boot time under various system parameters and generates a list of cores in an order of frequency capability
 - The core ordering is modified according to the usage policy detected
 - Single threaded applications scheduled to the fastest cores
 - Multi-threaded applications scheduled toward the fastest core cluster (CCX), maximizing L3 cache sharing
 - This core ordering is expressed to the OS allowing for an efficient, dynamic, HW-directed selection of processors for a given workload



Per-Core Linear Regulation

Regulating the voltage per-core enables power savings by adapting the voltage to each core's capability and compensating for power delivery gradients across-package

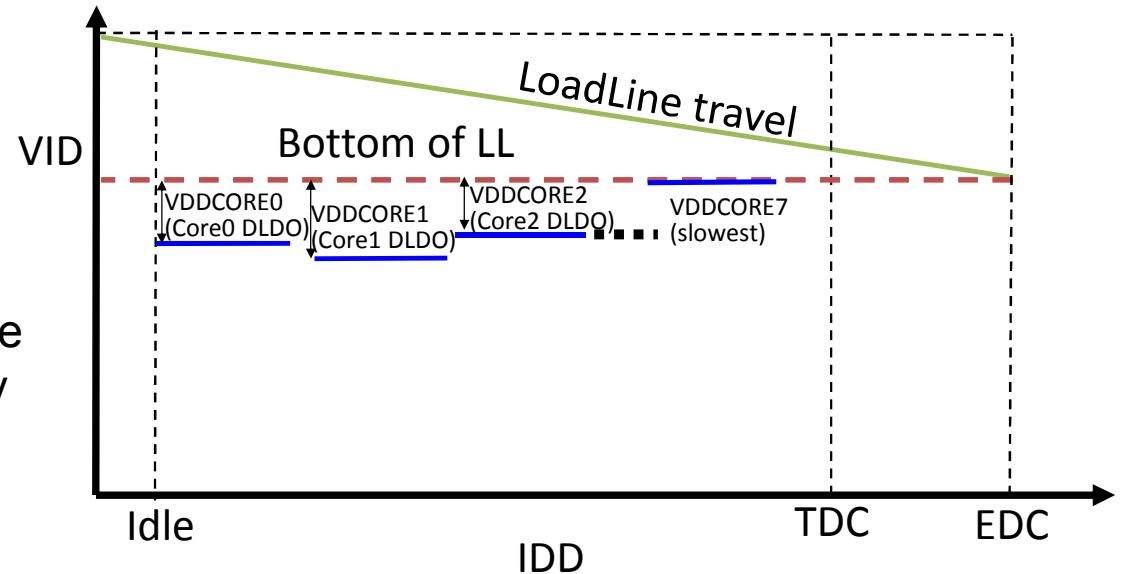


8 cores per chiplet, each with separate VDD
64 total core-specific voltages

- Digitally controlled LDO enables setting voltage based on per-core speed capability for a given frequency
- Droops mitigated with fast-response charge injection from RVDD for cores with a drop-out

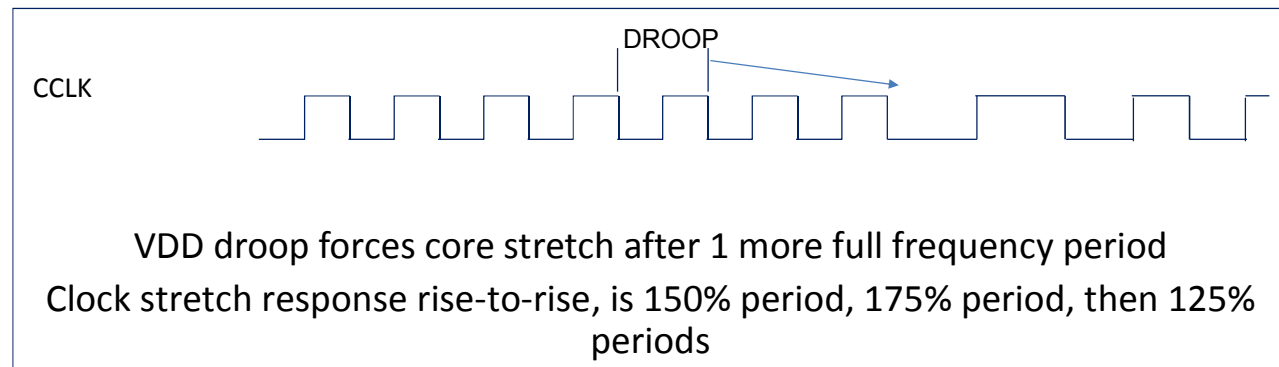
Clock Stretching and Per-Core Voltage

- Droop detection with a fast analog comparator
- Separate threshold for LDO Charge Injection (CI level) and for clock stretching (CKS)
- These work synergistically to lower the required voltage for a given frequency



	Same-frequency power savings through voltage reduction
No LDO, no CKS	0%
LDO only	19%
CKS only	19%
LDO and CKS	25%

Based on AMD internal testing of 64C AMD EPYC "Rome" processor operating at 2.5GHz, synthetic di/dt pattern



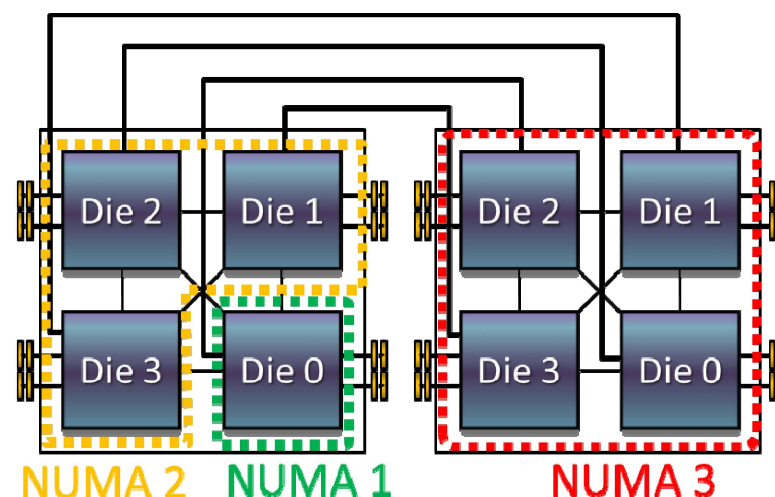
Improving Memory Performance

- Server memory latency is a key factor in performance
- A goal for 2nd Gen AMD EPYC™ was to improve on the 2017 1st Gen EPYC™ design
- Non-Uniform-Memory-Access (NUMA) behaviors are a result of memory interfaces being distributed across die
- Significant deltas from NUMA1 to NUMA2 impact performance for some applications

1: AMD internal testing with DRAM page miss

2: 75% NUMA 2 + 25% NUMA 1 traffic mix

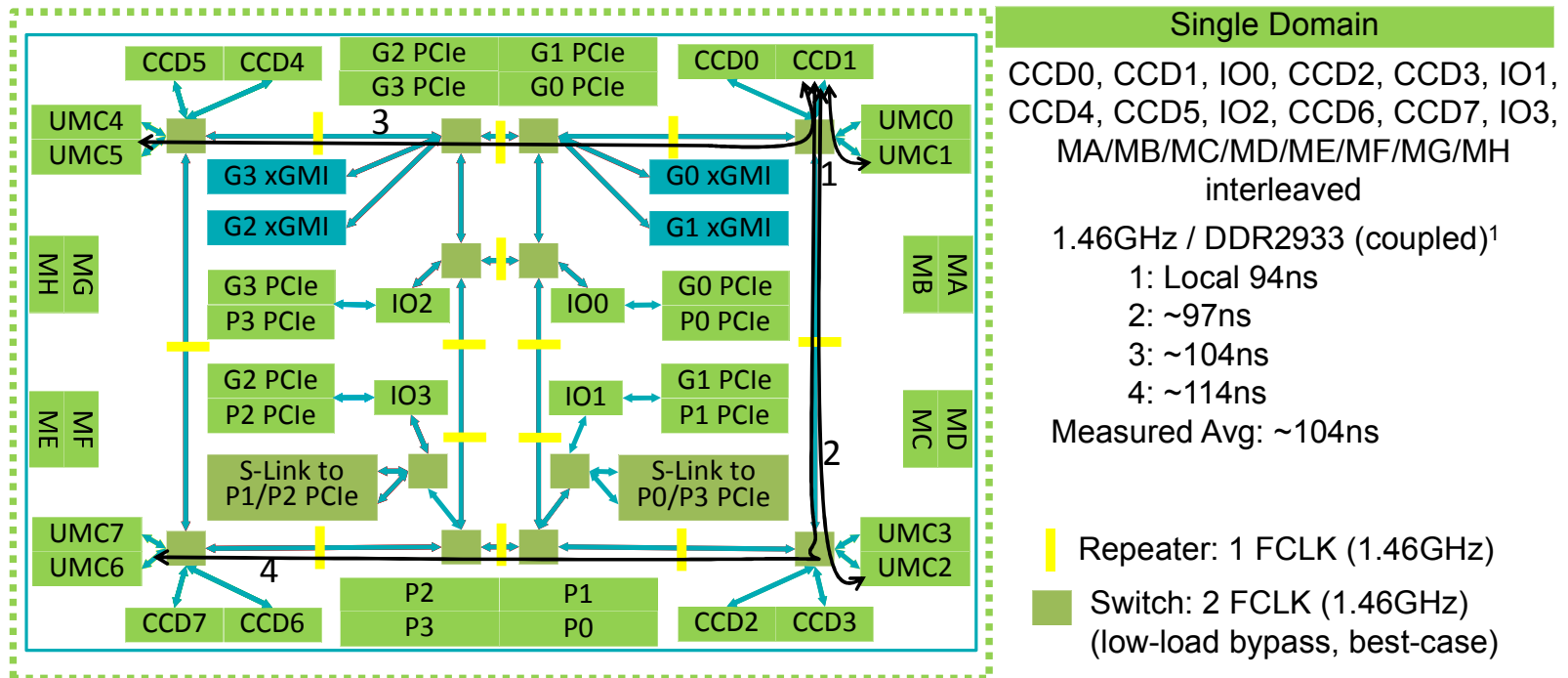
Prior Generation
(EPYC™ 7001 Series Processors)



Domain	Latency ¹ (ns)
NUMA1	90
NUMA2	141
NUMA3	234
Avg. Local ²	128

2nd Gen AMD EPYC™ Improved Memory Latency

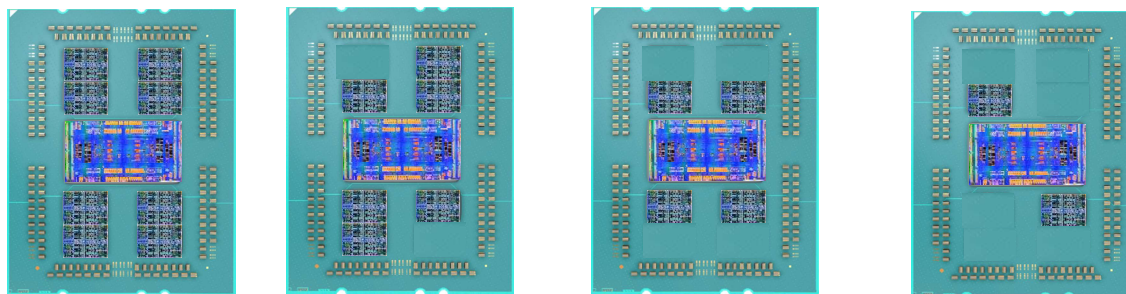
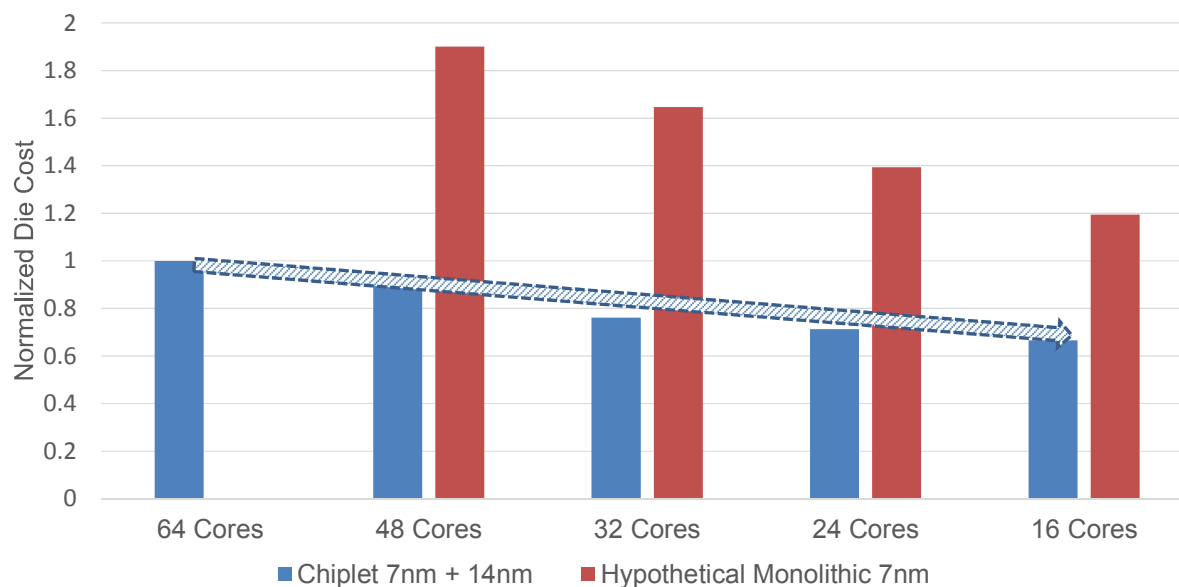
- Central IOD enables a single NUMA domain per socket
- Improved average memory latency¹ by 24ns (19%)²
- Minimum (local) latency only increases 4ns with chiplet architecture



1: AMD internal testing with DRAM page miss
2: EPYC 7002 Series NUMA 1 vs EPYC 7001 Series Avg. Local; EPYC 7002 Series NUMA2 vs EPYC 7001 Series NUMA 3

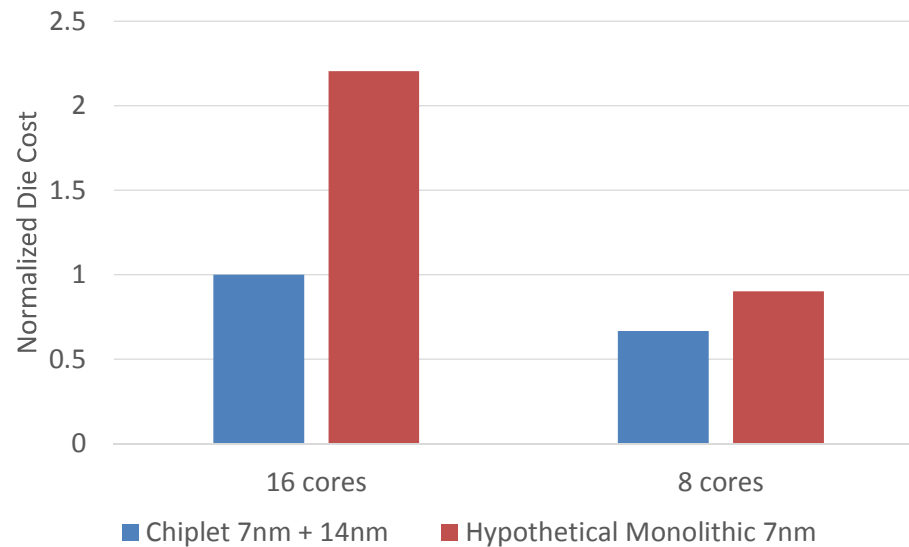
2nd Gen AMD EPYC™ Chiplet Performance vs. Cost

- Higher core counts and performance than possible with a monolithic design
- Lower costs at all core count / performance points in the product line
- Cost scales down with performance by depopulating chiplets
- 14nm technology for IOD reduces the fixed cost



2.2: AMD Chiplet Architecture for High-Performance Server and Desktop Products

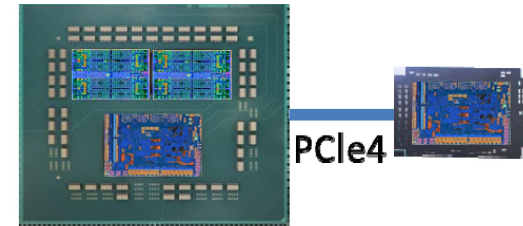
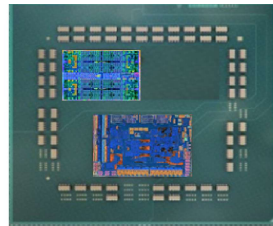
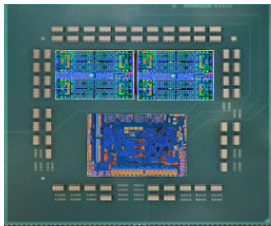
3rd Gen AMD Ryzen™ Processor Chiplet Performance vs. Cost



Similar cost savings and scalability for desktop

Re-using the client IO die for the X570 Chipset expander enables optional additional connectivity for higher end systems

- PCIe, SATA, USB



Performance Results

Chiplet architecture enables leadership performance and performance/Watt in server and desktop markets

1. Testing as of 12/13/2019 by AMD Performance Labs using a Ryzen 9 3950X with 16 cores vs. a Ryzen 7 2700X with 8 cores in the Cinebench R20 1T benchmark test. Results may vary. RZ3-102

Metric at 105W TDP ¹	Ryzen 2700X (8C)	Ryzen 3950X (16C)	Improvement (%)
Cinebench r15 1T	177	216	22%
Cinebench r20 1T	434	527	21%
Cinebench r15 NT	1802	3928	118%
Cinebench r20 NT	4020	8862	120%
1T Fmax (Max Boost)	4.3	4.7	9%
NT Base Freq (All-core) ¹	3.9	3.95	1%

Metric	EPYC 7601 (32C 2P 180W TDP)	EPYC 7742 (64C 2P 225W TDP)	Improvement (%)
SPECrate®2017_int_base ²	272	663	144%
SPECrate®2017_fp_base ²	259	511	97%
NT Base Freq	2.2	2.5	14%

2: Results obtained from the SPEC® website as of Jan 3, 2020.

EPYC 7601 SPECrate®2017_int_base: <https://www.spec.org/cpu2017/results/res2017q4/cpu2017-20171114-00833.html>

EPYC 7601 SPECrate®2017_fp_base: <https://www.spec.org/cpu2017/results/res2017q4/cpu2017-20171114-00845.html>

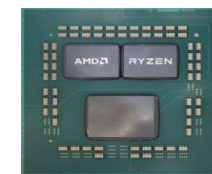
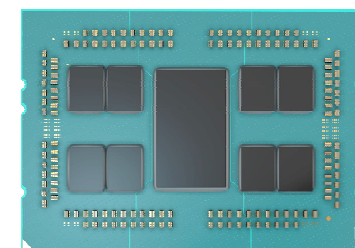
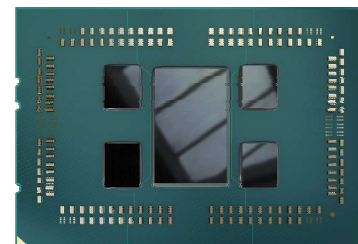
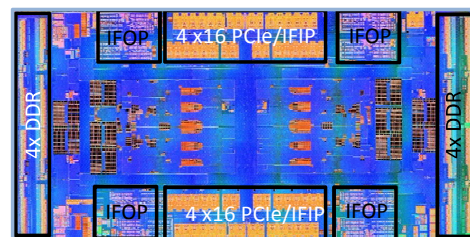
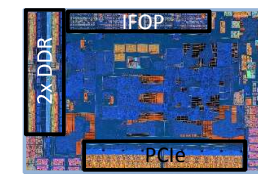
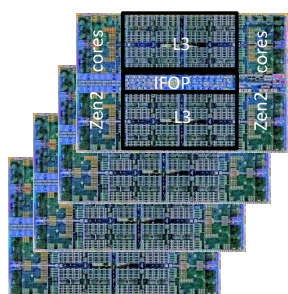
EPYC 7742 SPECrate®2017_int_base: <https://www.spec.org/cpu2017/results/res2019q4/cpu2017-20191028-19261.html>

EPYC 7742 SPECrate®2017_fp_base: <https://www.spec.org/cpu2017/results/res2019q4/cpu2017-20191028-19237.html>

More information about SPEC CPU® 2017 can be obtained from <https://www.spec.org/cpu2017>. SPEC®, SPEC CPU® and SPECrate® are registered trademarks of the Standard Performance Evaluation Corporation.

Summary

- Chiplet architecture has proven key to achieving leadership performance, performance/\$ and performance/Watt across multiple market segments
- Many significant innovations were required:
 - Package + Silicon co-design for optimizing complex routes and heterogeneous technology chiplet die
 - Package level fabric and interconnect architecture
 - Power delivery and voltage adaptation



Acknowledgment

We would like to thank our talented AMD design teams across Austin, Bangalore, Boston, Fort Collins, Hyderabad, Markham, Santa Clara, and Shanghai.

Disclaimer and Endnotes

DISCLAIMER

The information contained herein is for informational purposes only, and is subject to change without notice. While every precaution has been taken in the preparation of this document, it may contain technical inaccuracies, omissions and typographical errors, and AMD is under no obligation to update or otherwise correct this information. Advanced Micro Devices, Inc. makes no representations or warranties with respect to the accuracy or completeness of the contents of this document, and assumes no liability of any kind, including the implied warranties of noninfringement, merchantability or fitness for particular purposes, with respect to the operation or use of AMD hardware, software or other products described herein. No license, including implied or arising by estoppel, to any intellectual property rights is granted by this document. Terms and limitations applicable to the purchase or use of AMD's products are as set forth in a signed agreement between the parties or in AMD's Standard Terms and Conditions of Sale. GD-18

©2020 Advanced Micro Devices, Inc. All rights reserved. AMD, the AMD Arrow logo, EPYC, RYZEN, Threadripper, 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.