AMD Next-Generation FPGA Built From Chiplets
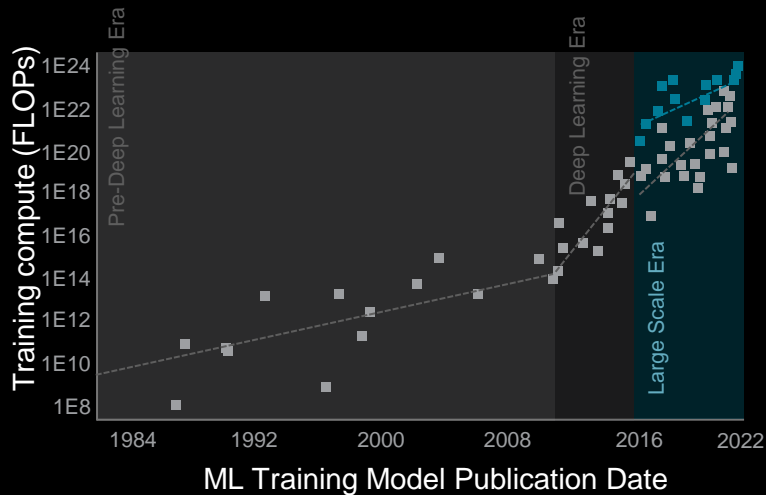
Dinesh Gaitonde

Hot Chips 2023

AI Chips & Advanced Integration Require Next Generation Solutions for Emulation & Prototyping
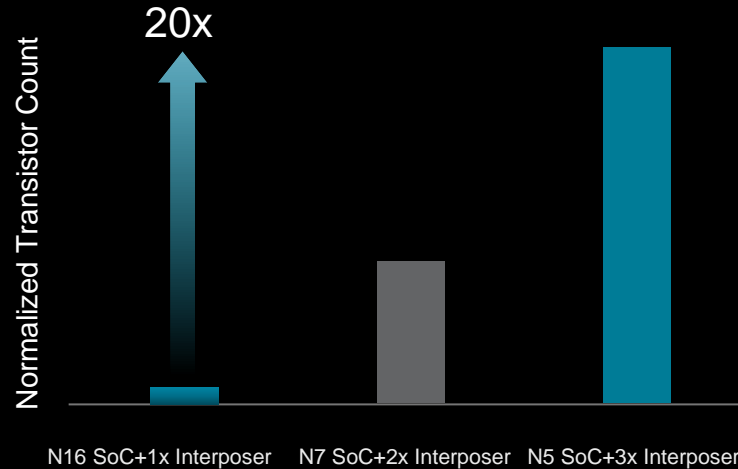
AI Demands Exponential Growth in Compute Capacity

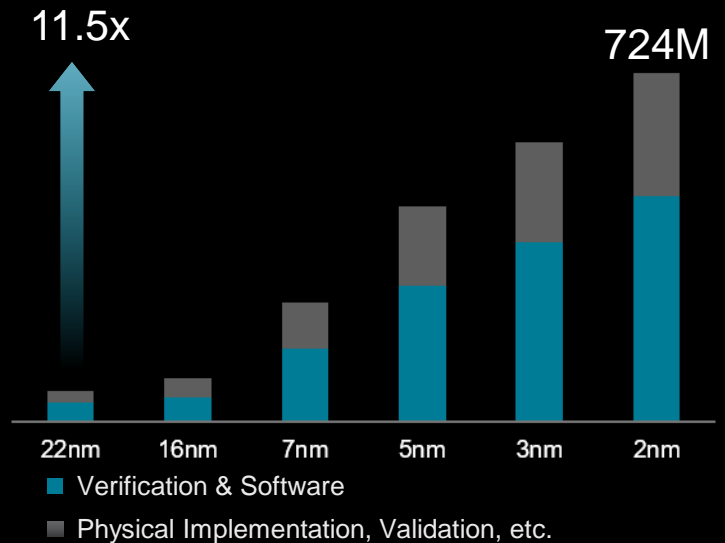### Driving Introduction of Massive Soc Designs

Advanced Integration Promises Scaling of Performance, Cost

Heterogeneous Integration, Chiplets Increase Verification Scope

20x

N16 SoC+1x Interposer   N7 SoC+2x Interposer   N5 SoC+3x Interposer

Majority of Design Cost in Software & Verification

Requires Earlier Software Development to Maintain Time-to-market

11.5x   724M

22nm   16nm   7nm   5nm   3nm   2nm

■ Verification & Software
■ Physical Implementation, Validation, etc.

[1] Compute Trends Across Three Eras of Machine Learning by Jaime Sevilla, Tama Besiroglu, Lennart Heim, Marius Hobbhahn, Anson Ho, Pablo Villalobos; 2022. (Source)
[2] P. K. Huang et al., "Wafer Level System Integration of the Fifth Generation CoWoS®-S with High Performance Si Interposer at 2500 mm2," 2021 IEEE 71st Electronic Components and Technology Conference (ECTC), San Diego, CA, USA, 2021, pp. 101-104, doi: 10.1109/ECTC32696.2021.00028. (Source)
[3] Cost of Advanced Designs (Source: IBS)

AMD

Design Verification Market

- FPGAs are key for verification acceleration

- Larger FPGAs provide better verification performance

- Therefore, build large FPGAs

AMD ◢

Key Requirements for Verification Hardware

Large Devices → Larger Canvas

- Keep up with growing design sizes
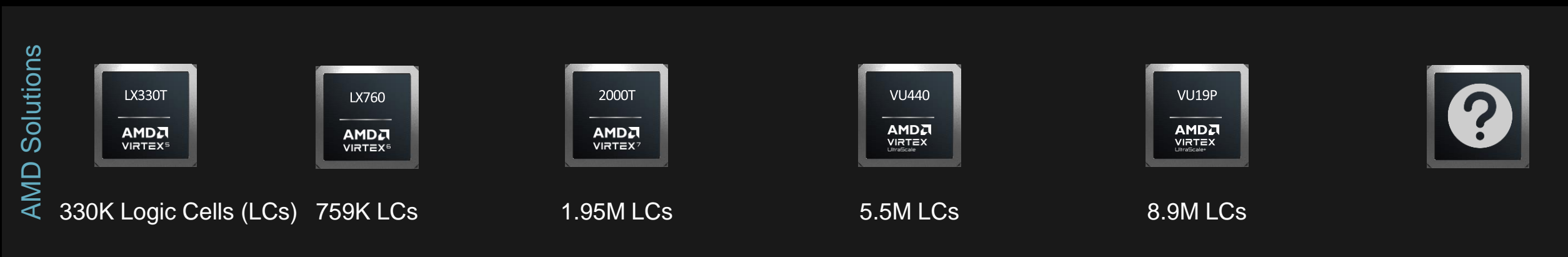
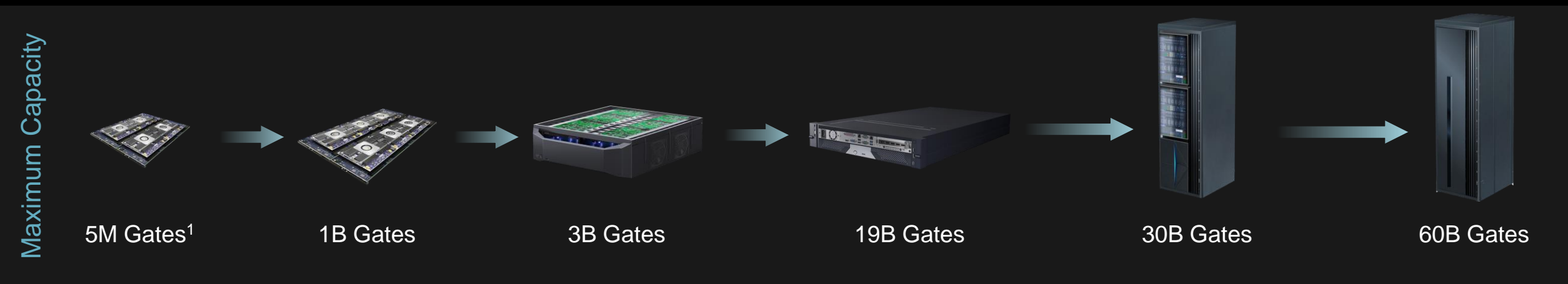## Abundant Low Latency Interfaces → Large Design Performance

- Large design system performance depends critically on latency of interfaces

Improved Debug Performance

- Save and restore design state

AMD

Driving Performance and Capacity Leadership

AMD Foundational Compute Technology Has Enabled High Performance Emulation and Prototyping Solutions for 17+ Years

Maximum Capacity

| 5M Gates[1] | 1B Gates | 3B Gates | 19B Gates | 30B Gates | 60B Gates |

AMD Solutions

| LX330T AMD VIRTEX5 | LX760 AMD VIRTEX6 | 2000T AMD VIRTEX7 | VU440 AMD VIRTEX UltraScale | VU19P AMD VIRTEX UltraScale+ | ? |

| 330K Logic Cells (LCs) | 759K LCs | 1.95M LCs | 5.5M LCs | 8.9M LCs | |

2006 ➝ 2023

AMD Next-Generation FPGA Built From Chiplets | August, 2023   1 Shows the maximum possible numbers

AMD

AMD Versal™ Premium VP1902 Adaptive SoC

The World's Largest Adaptive SoC

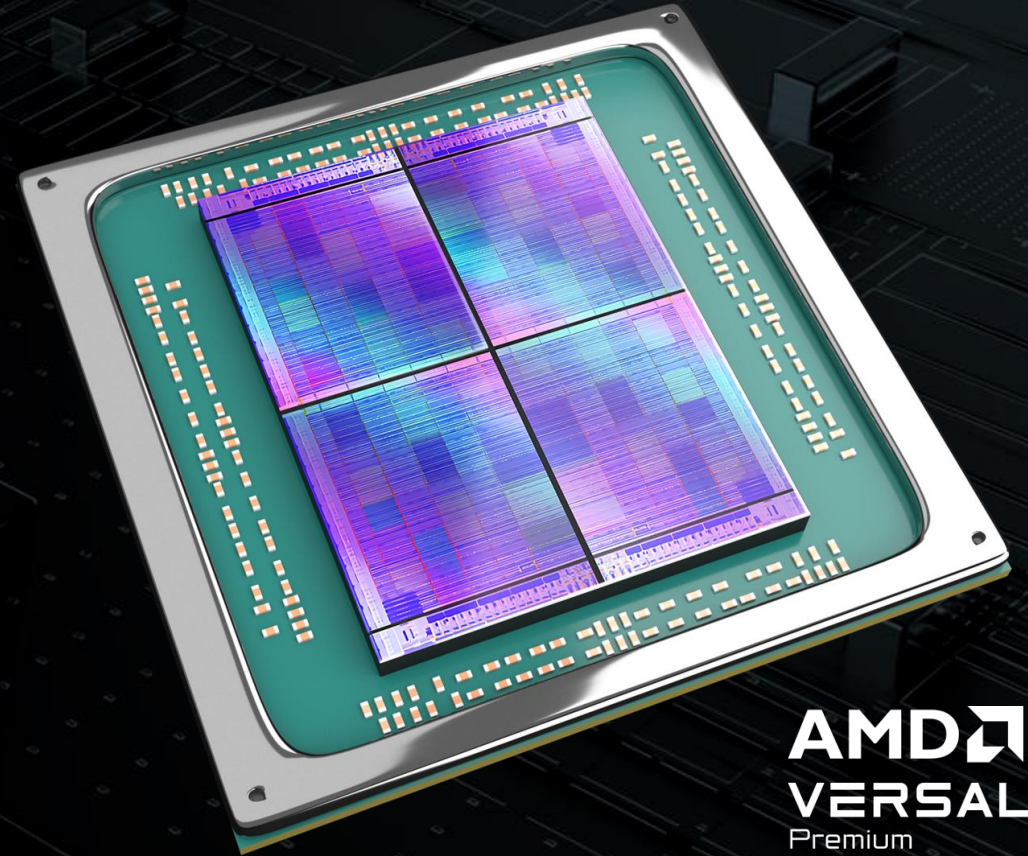### ~2x Higher Logic Density and Connectivity

- Industry's Highest Capacity for Emulation and Prototyping of Next-Generation SoCs

8x Faster Debug Performance with Versal™ Architecture

- Outstanding Debug Capabilities for Rapid Iteration

Novel two-by-two SLR (Chiplet) Arrangement for Enhanced Routability and Lower Latency

- Architectural Innovation Based on 4th gen SSI (*Stacked Silicon Interconnect*) Technology to Maximize Performance

AMD♦ VERSAL Premium

Available in Q3 '23

AMD Next-Generation FPGA Built From Chiplets | August, 2023   All comparisons in this slide are to previous generation AMD part (VU19P)   See Endnotes VER-001, VER-002, VER-003, VER-004   **AMD♦**

Introducing Versal Premium VP1902 Adaptive SoC – Largest Capacity Adaptive SoC

~2x **Logic Density**
Emulate Complex ASIC/SoC Designs

~2x **I/O Bandwidth**
Partition Designs Across Multiple Devices

2x **Transceiver Count**
Scale From Desktop to Enterprise

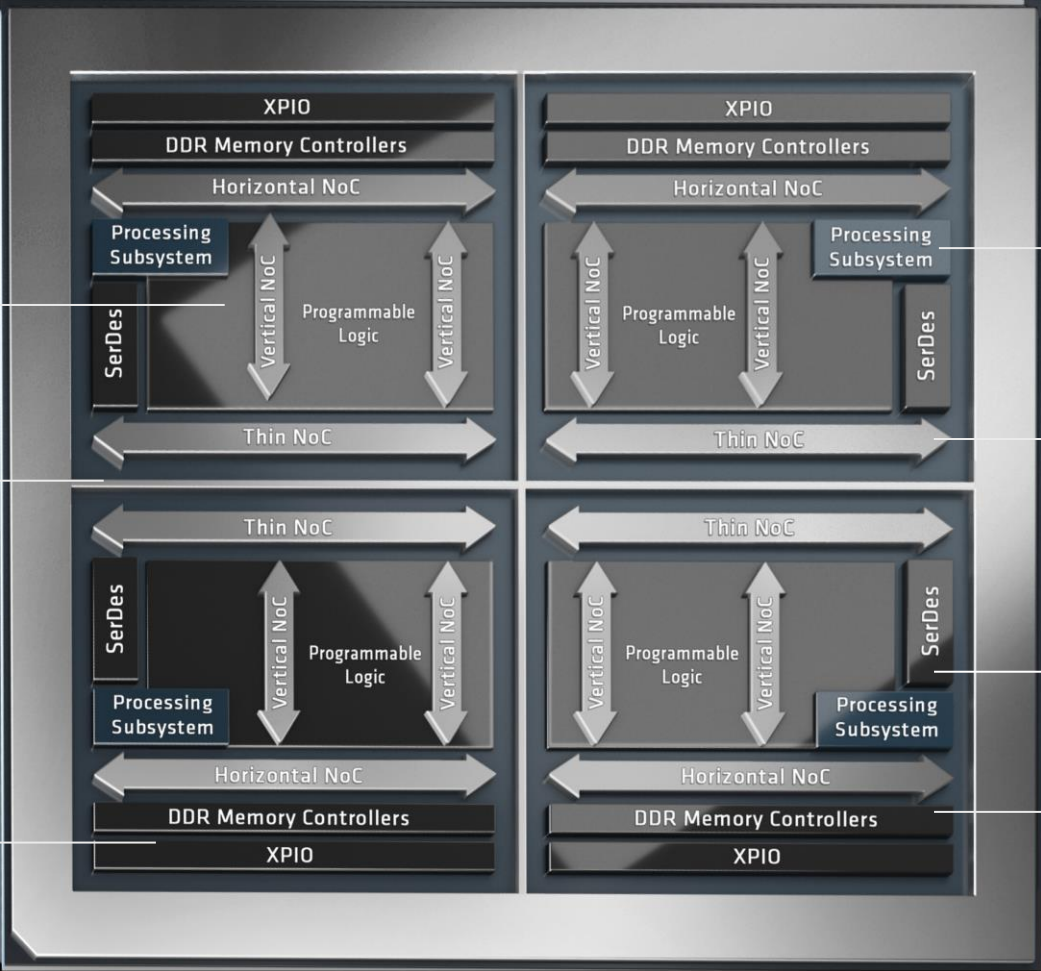**~2x** **Transceiver Bandwidth**
High Speed Signal Multiplexing

	AMD Virtex™ UltraScale+™ VU19P FPGA	AMD Versal™ Premium VP1902 Adaptive SoC
Logic Density	8.9M Logic Cells	18.5M Logic Cells
Aggregate I/O Bandwidth with XPIOs	2.7 Tbps	5.6 Tbps
Number of Transceivers	80	160
Aggregate Transceiver Bandwidth	5.4 Tbps	12.2 Tbps

AMD

VP1902 Adaptive SoC: Purpose Built for Emulation and Prototyping

Programmable Logic
Custom emulated designs in RTL

Quadrant-Based Architecture
Enhanced routability, low latency, and congestion

XPIOs
Partition design across multiple VP1902 devices with chip-to-chip interfacing at up to 36% lower latency*

Processing Subsystem
Flexible boot and control for SoC-based verification

Programmable NoC
High-bandwidth debug traffic over hardened infrastructure

Transceivers
Scale connectivity between boards for massive designs

Memory Controller
Easy external memory access with hard and soft controllers

* Versus HP I/O in Virtex Ultrascale+ FPGA. See Endnotes VER-008

AMD

Building Large FPGAs Requires Solving Difficult Challenges

Large FPGAs Desirable

- Large FPGAs → better system performance

Large FPGAs Not Economical Without Innovating

- Programmability tax → worse for FPGAs than custom silicon
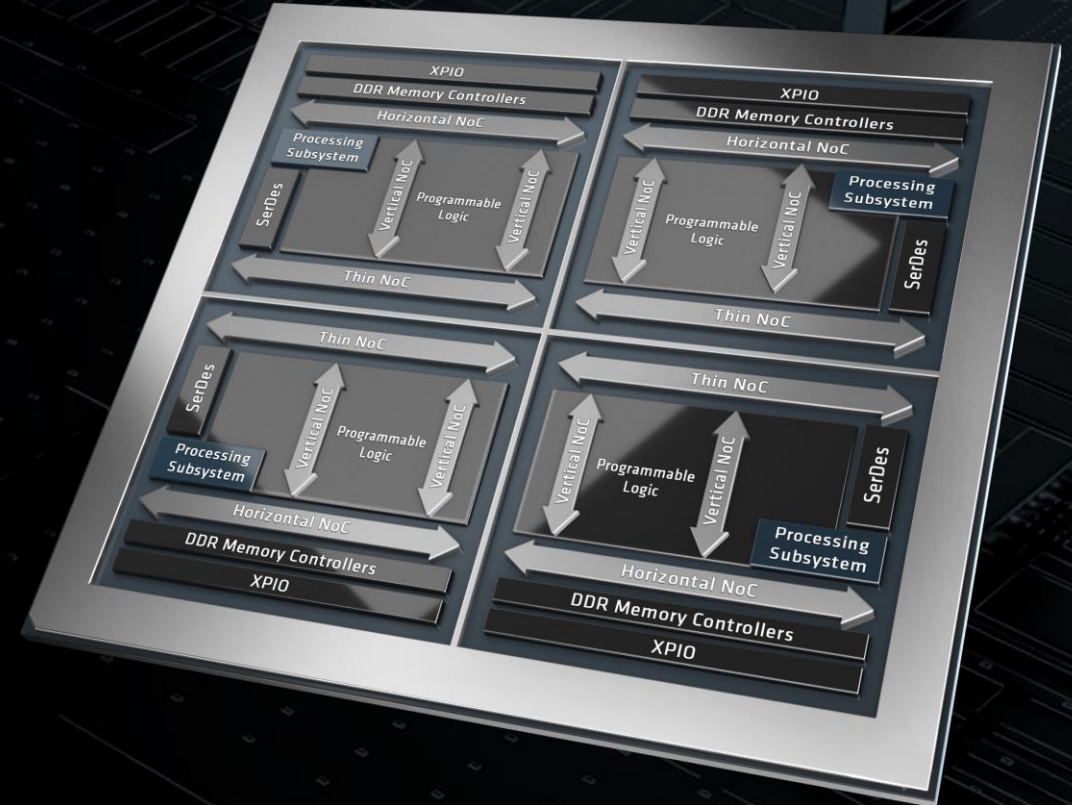
Chipletize → Scale Capacity & Manage Cost

AMD Next-Generation FPGA Built From Chiplets | August, 2023

AMD

Chiplets in the Context of FPGAs

- Inter die connectivity → bunch of wires – user view of uninterrupted large canvas

- Protocol, if any, is overlayed by end-user

- Inter-die interconnect simply extends intra-die interconnect

- Tools manage partitioning customer designs into various SLRs
  - Maintain user experience of uninterrupted canvas
  - Manage the cut and performance

High-Bandwidth, Low-Latency Connections

Microbumps

Through-Silicon Vias (TSV)

C4 Bumps

FPGA Die (SLR)

Silicon Interposer

Package RDL

Package Substrate

BGA Solder Balls

SLR3 SLR2 SLR1 SLR0

AMD
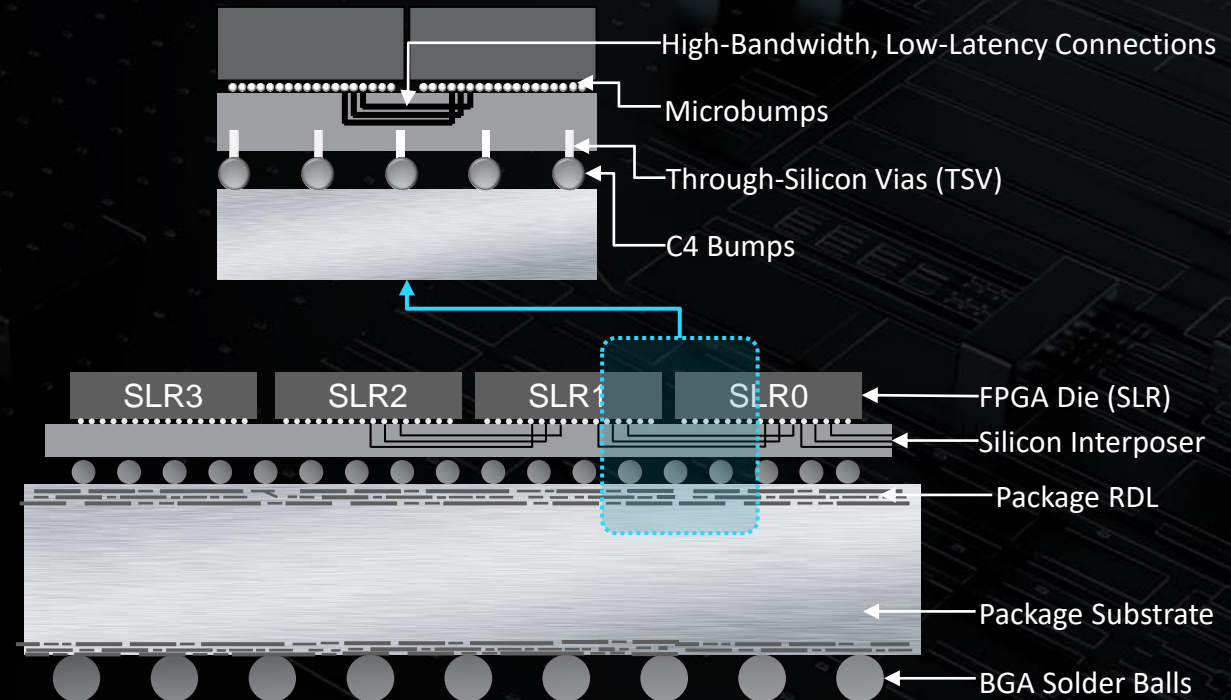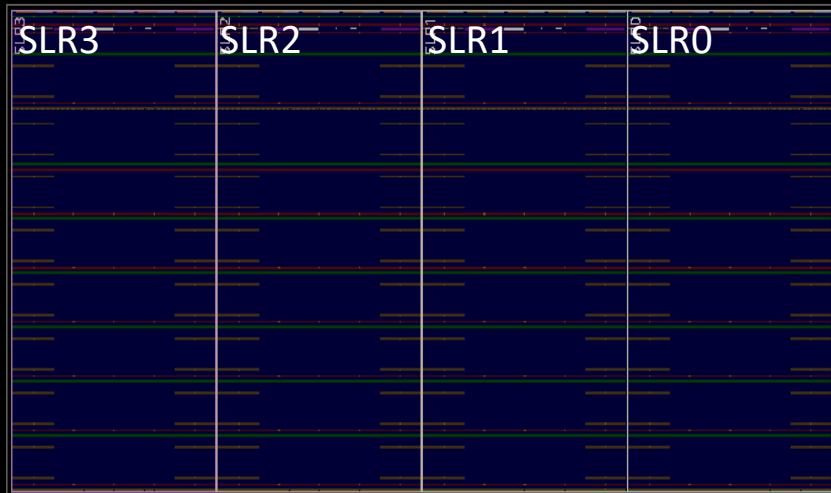
Traditional SLR Arrangement

1xN Arrangement

- Only 1 tapeout
- Can productize 1x1, 1x2, … variants

Extend interconnect and global signaling vertically

- Core Interconnect, clocking, configuration

I/O arrangement issues with 1xN

- I/O scaling trends competing with arrangement

Inter-SLR Bandwidth management

- Inter-SLR BW utilized less efficiently with 1xN

VU19P Device

SLR

SLR

SLR

SLR

AMD
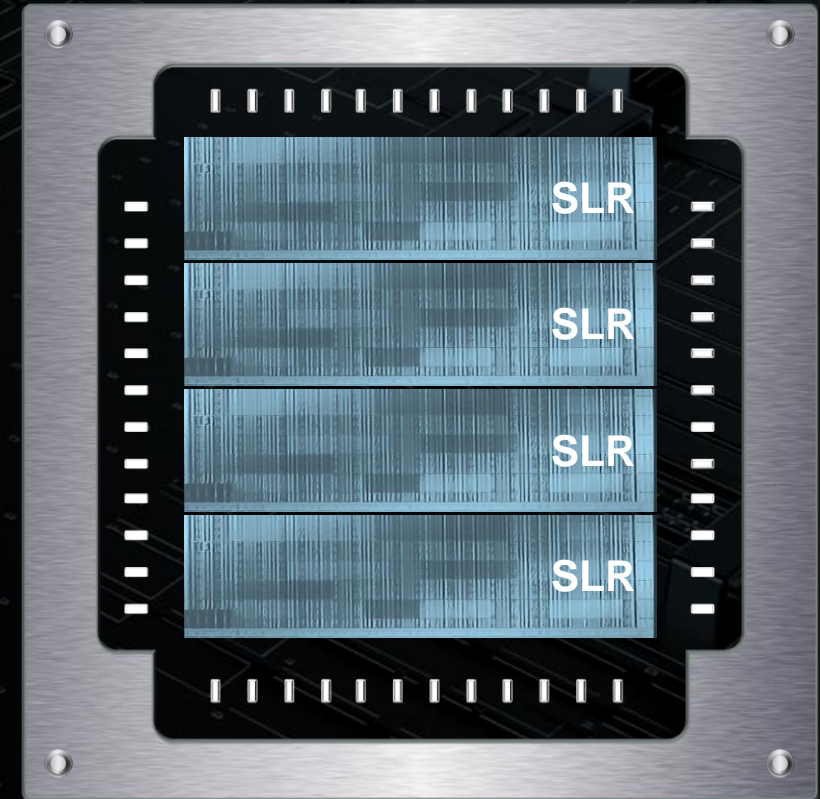
AMD FPGA Previous Generation I/O Arrangement

- Previous Generation → I/Os in columns
  - Each column of I/O in each SLR just replicated in other SLRs

- Need more I/Os → Add more columns

- But size of I/O column not scaling with technology
  - Logic scales, I/Os do not

- Delay Scaling
  - Delay from one logic tile to another improves with technology
  - Delay crossing I/O tile does not scale

- I/O crossing management becomes a software problem

- I/Os have to support higher speeds
  - Bonding out from the middle of the device is challenging

AMD
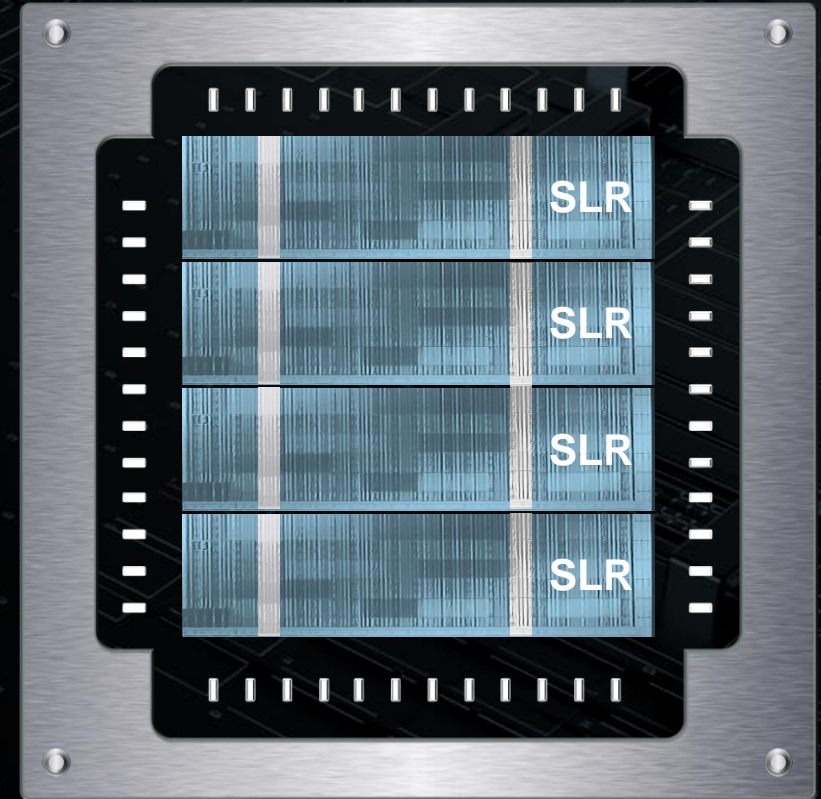
Reconfiguring Array Construction

- I/Os have to move to the periphery
- One possible construction could have been 1x4
- But delays for inter-die wires would be too large
- Difficult to bond out as many I/Os as we would like – middle rows
- Instead build a 2x2 structure

Possible Construction

New 2x2 Construction

AMD Next-Generation FPGA Built From Chiplets | August, 2023

AMD
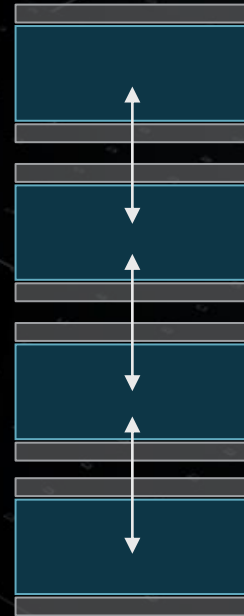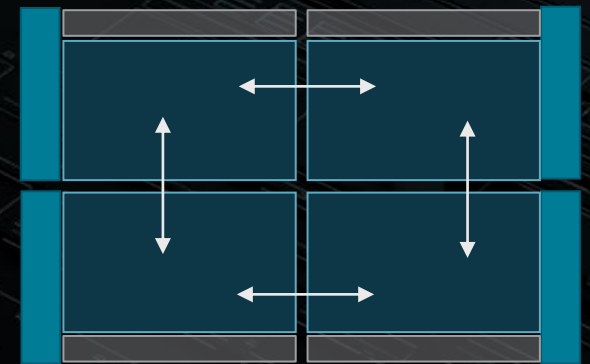
Cutting Down on Metal Demand

- Micro-Bump pitch not scaling, thus, 1x4 arrangement is more demanding than 2x2 arrangement
- 2x2 arrangement results in up to 40% reduction in track demand

1x4 Arrangement
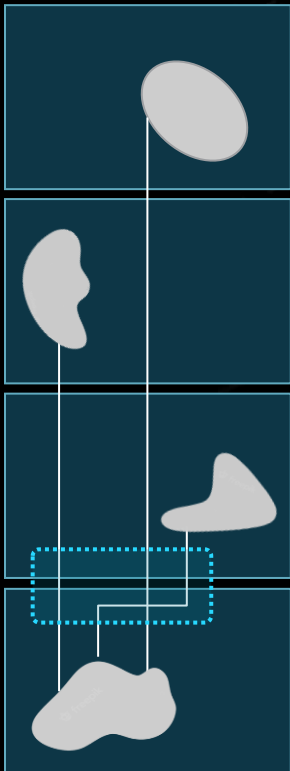
2x2 Arrangement
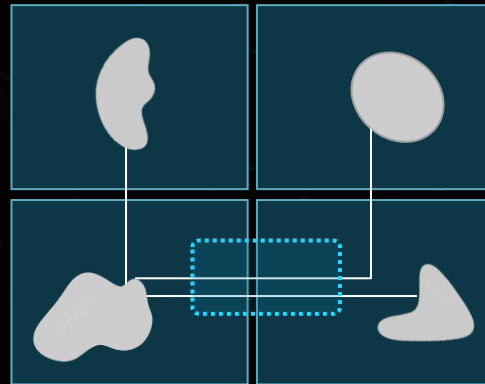
Cut Demand 2x2 over 1x4 (lower is better)

Designs

AMD Preliminary Engineering Data

AMD Next-Generation FPGA Built From Chiplets | August, 2023

AMD

Building Devices Beyond the Reticle Limit

54 mm

45 mm

Keep Out Overlap

Metal Traces
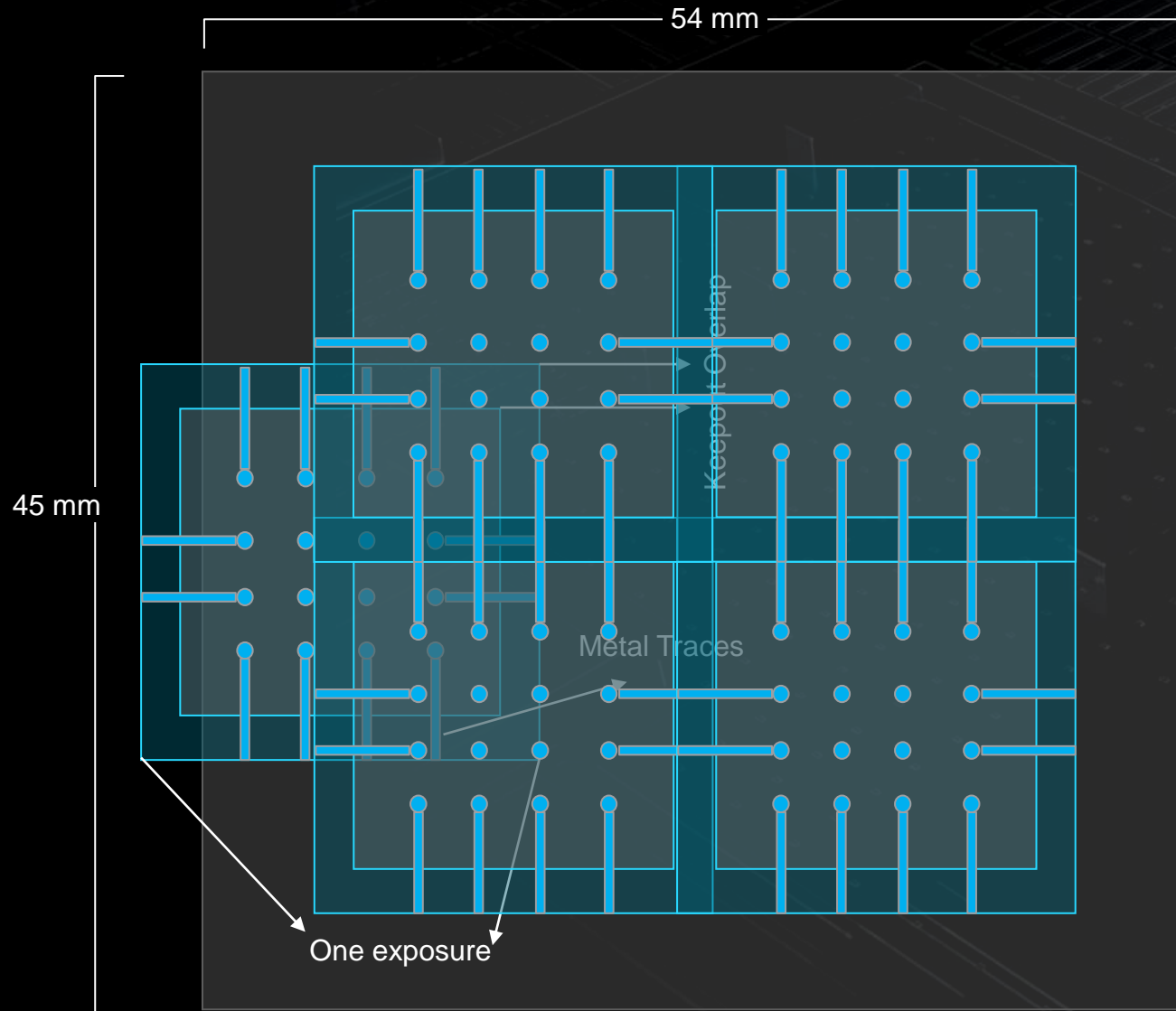
One exposure

Multiple overlapping exposures

Build up to 4x reticle size devices

Smaller in practice – Mechanical reasons

AMD

Performance Criteria for Verification

Implementation with a Single Device

- Mapping a small design on one VP1902 device
- Fabric Performance is a key
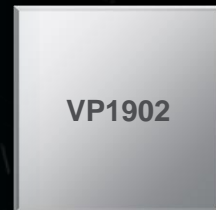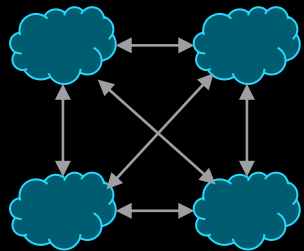
Implementation with Multiple Devices

- Mapping a large design on multiple VP1902 devices
- Interface Performance is a key
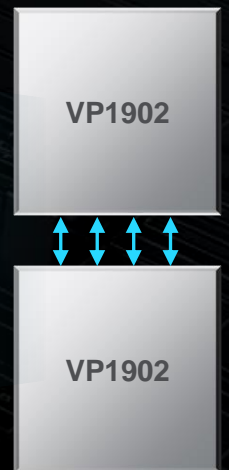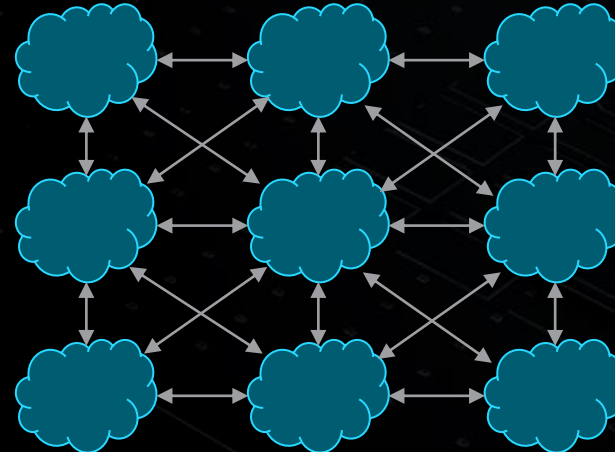
Core Performance

Performance scaling with technology from 16nm to 7nm devices

- Architecture modified to be more routable than 16nm devices to support larger device

2x Device size for better performance

- More designs fit in one device

2x2 Arrangement less demanding on metal

- Tools have easier time managing cut
- Do not need to sacrifice performance to manage cut

Use of hardened programmable network on chip (NoC)
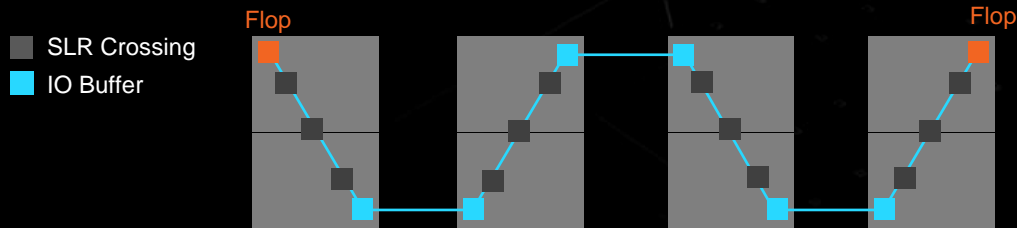
- Infrastructure can use NoC. Canvas available for user design. More routable

AMD

Capacity and SLR Architecture Drive System Performance
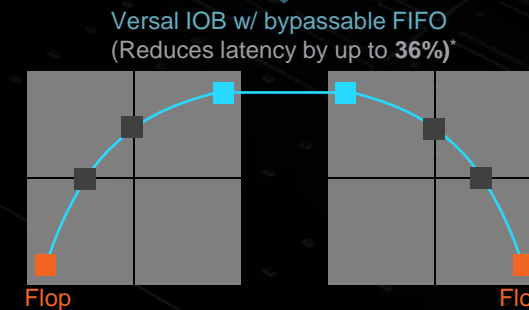
Virtex™ UltraScale+™ VU19P FPGA

Versal™ Premium VP1902 Adaptive SoC
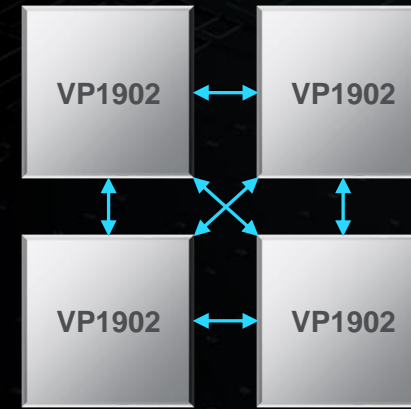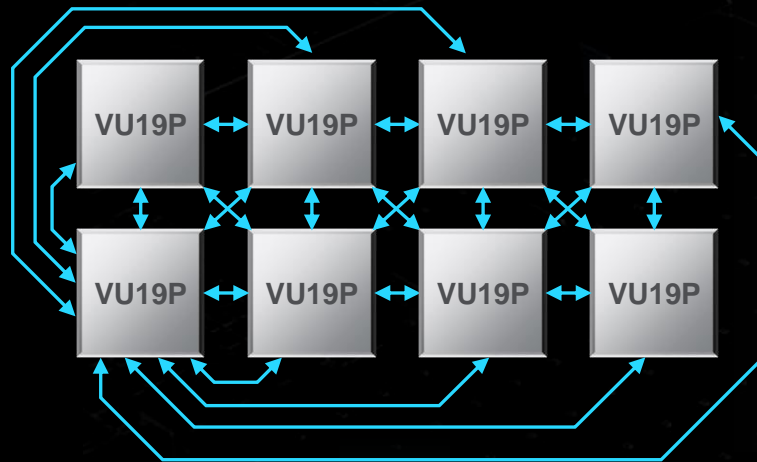
Higher Density, Simplified System Design

Versal IOB w/ bypassable FIFO
(Reduces latency by up to **36%**)*

Shorter critical path improves system performance

- SLR Crossing
- IO Buffer

Critical Data path with 4 - VU19P FPGA

3 – Inter FPGA Interfaces   12 - SLR Crossings

Critical Data path with 2 - VP1902 adaptive SoC

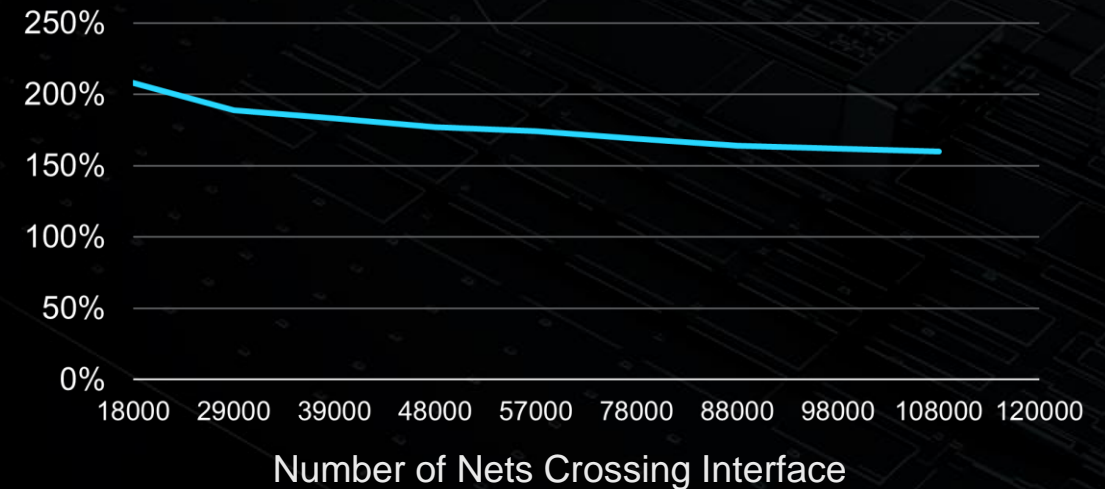1 – Inter FPGA Interface   4 - SLR Crossings

AMD Next-Generation FPGA Built From Chiplets | August, 2023

* Versus HP I/O in Virtex Ultrascale+ FPGA. See Endnotes VER-08

AMD

Interface Performance

- Bandwidth within an SLR >> bandwidth across SLRs
- Bandwidth across SLRs >> bandwidth across I/Os SerDes
- Signals across devices >> number of physical I/Os
  - Multiplex signals across I/Os – IO latency key performance metric
- Versal I/O latencies ~ 0.64 of UltraScale+™ I/O latencies
 - Programmability added in the Versal IO to optionally cut latencies even more

Mux User Signals | High Speed I/O | Board
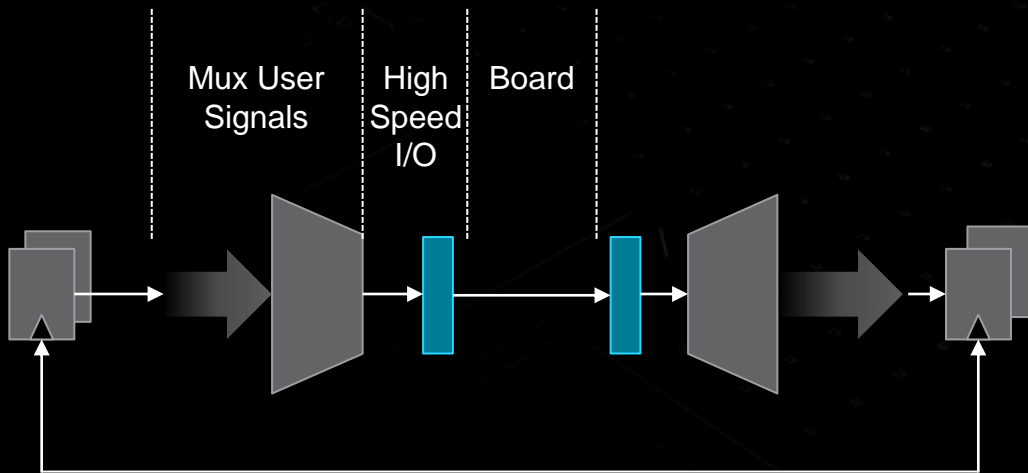
Interface Performance Improvement
(Versal Adaptive SoC over UltraScale+ FPGA)

Number of Nets Crossing Interface

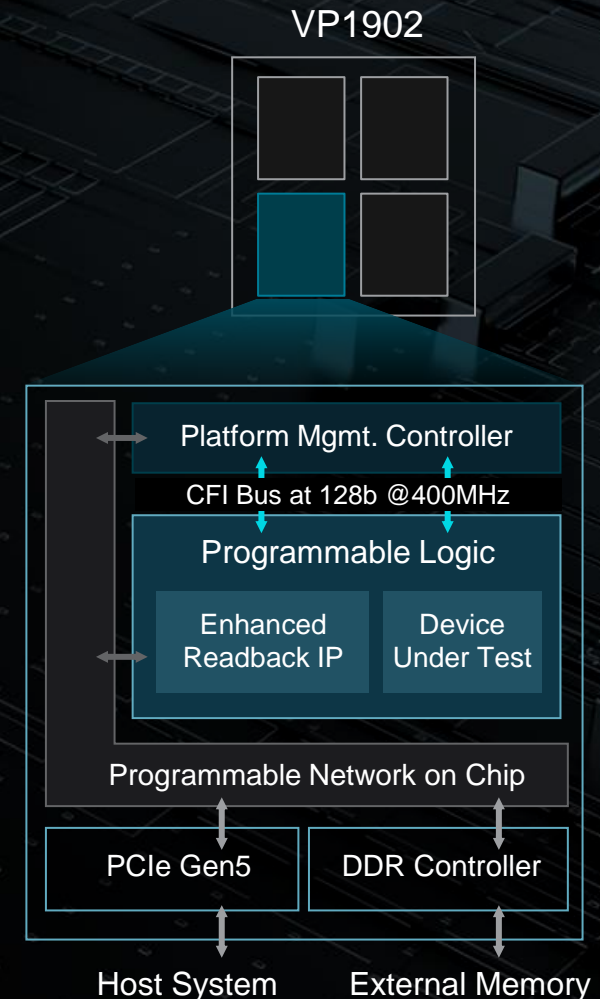AMD Preliminary Engineering Data

AMD

Readback Performance

- Readback a key requirement for FPGA fabric used in verification
- Users set breakpoints and expect to be able to read a large fraction of the state
- Users expect to be able to resume from a given state
- Faster readback → better system performance
- Readback functionality in the Versal™ devices
  - Improved traditional readback by 8x compared to VU19P
  - Using the programmable NoC, can further improve readback performance

AMD Next-Generation FPGA Built From Chiplets | August, 2023

See Endnote VER-04

AMD
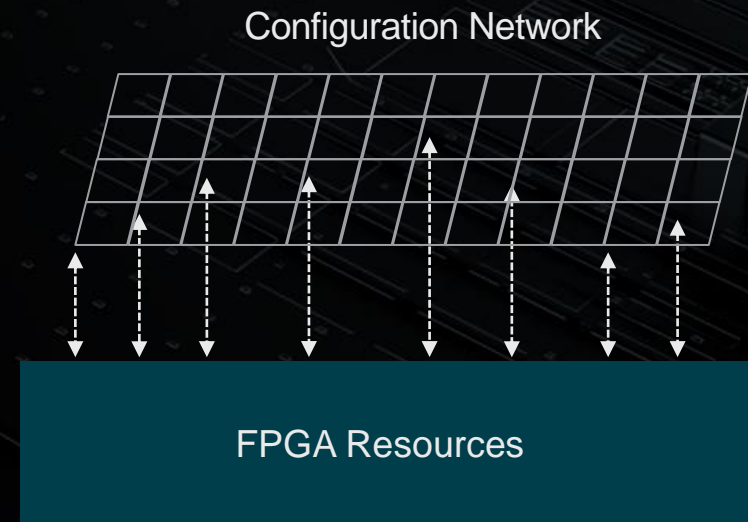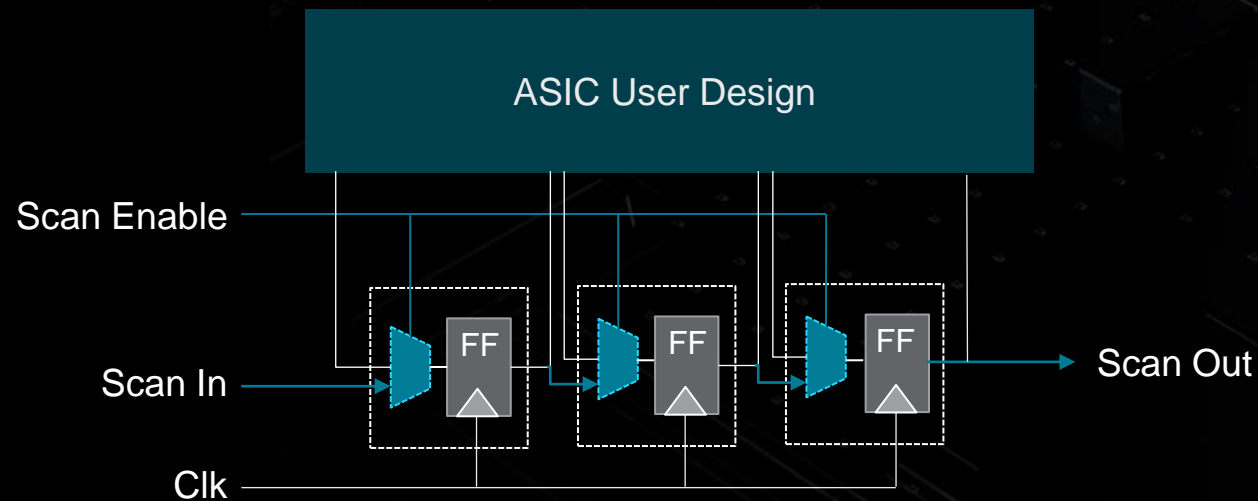
Providing 8x Faster Debug with Enhanced Readback IP

Leveraging AMD Versal™ architecture for improved debug

- New "Configuration Data Interface" (CDI) IP enables readback and writeback of device state

- Gives designers full visibility into every register of the design for offline analysis

- Enabled by Versal "Configuration Frame Interface" (CFI) running 2X faster at 400 MHz with 4x wider bus width compared to Ultrascale+ platform

- Debug data offloaded via programmable NoC and hardened DDR memory controllers, freeing up PL for emulated design

VP1902

Platform Mgmt. Controller

CFI Bus at 128b @400MHz

Programmable Logic

| Enhanced Readback IP | Device Under Test |

Programmable Network on Chip

| PCIe Gen5 | DDR Controller |

Host System   External Memory

Summary of Versal Architecture Readback Enhancements

CFI Bus runs @ 400 Mhz
- Previous generation ICAP was 200Mhz and SMAP 125Mhz
- Improvement of 2x to 3.2x over previous generation

CFI Bus width quadruples from 32 bit to 128 bit (4x)

~8x raw bandwidth improvement

CLE FF Frame Data Efficiency goes from 25% to 100% (4x)
- Only readback flip flop data and not other unrelated data that need to get filtered out by user

Designed for **8x** efficiency improvement

Single frame read efficiency increases from 50% to 100% (2x)
- No dummy frames. Useful when small number of frames are being read

AMD

Scan-Chain Methodology

- Build soft scan-chains of shadow registers
- Very high performance (Fmax)
- Funnel data from nearby chains into NoC ports
- Relatively non-intrusive, high performance readback

- ~12x readback performance using scan-chains[*]
- Leverage unused registers and NoC
- Much better performance than native hardware

User Design
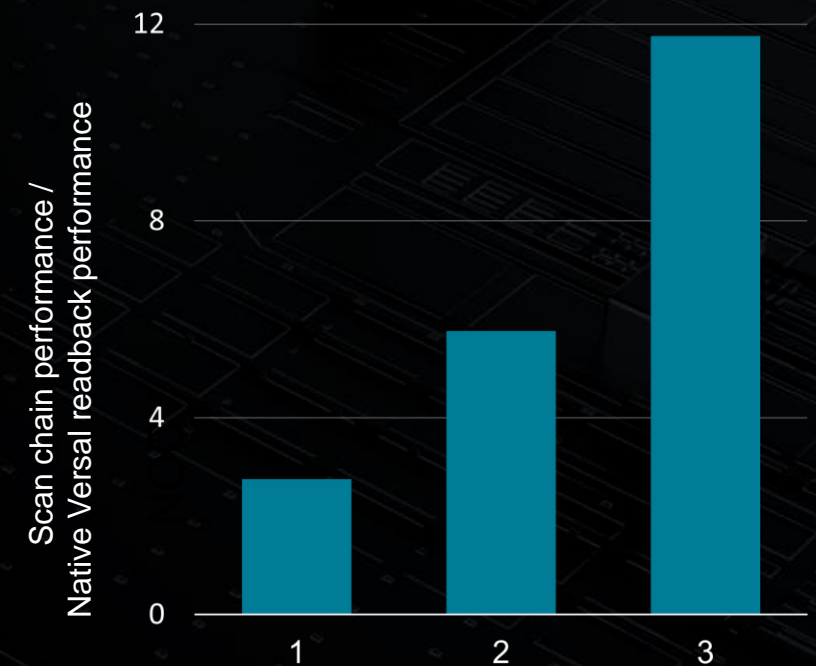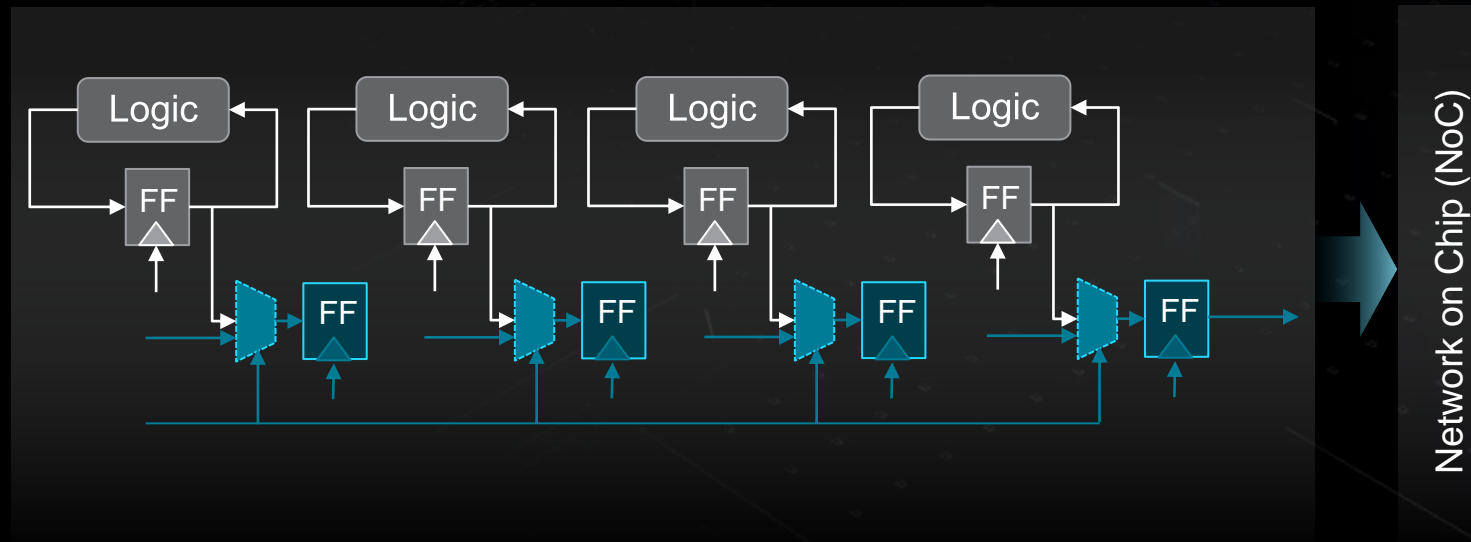
Network on Chip (NoC)

Scan chain performance / Native Versal readback performance

AMD Preliminary Engineering Data

Multi-Partition Flow with Enhanced Place and Route for Bigger Designs

Automatic Partition for Fast Place and Route

- Enabled by default, no required interaction
- Vivado™ Design Suite handles partition-level constraint budgeting
- Partition boundaries can be adjusted for cross-boundary optimization

Partition 1

Partition 2

Partition 3

AMD
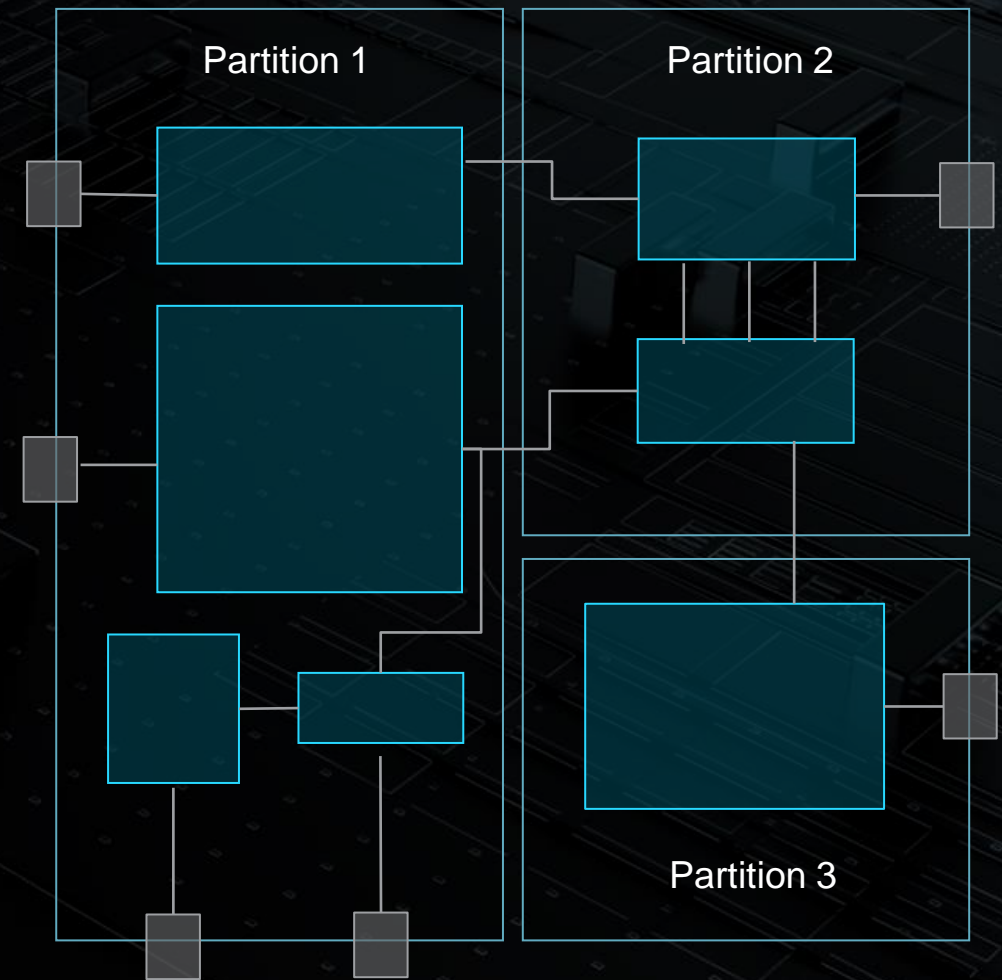
Versal™ Premium VP1902 Adaptive SoC

The World's Largest Adaptive SoC, Architected for Next-Gen Emulation and Prototyping

Industry-Leading Capacity and Connectivity

- ~2X higher logic density with 18.5M system logic cells
- ~2X aggregated I/O bandwidth and 2.3X transceiver bandwidth

Backend Design Tool Innovation to Meet Evolving Demands

- 8X faster debug performance with enhanced readback IP
- Continuous tool improvements that enable fast design iteration

17+ Years of Technology Leadership

- AMD's 6th generation emulation-class device
- Trusted partner for top EDA vendors

AMD
VERSAL
Premium

Available in Q3 '23

AMD
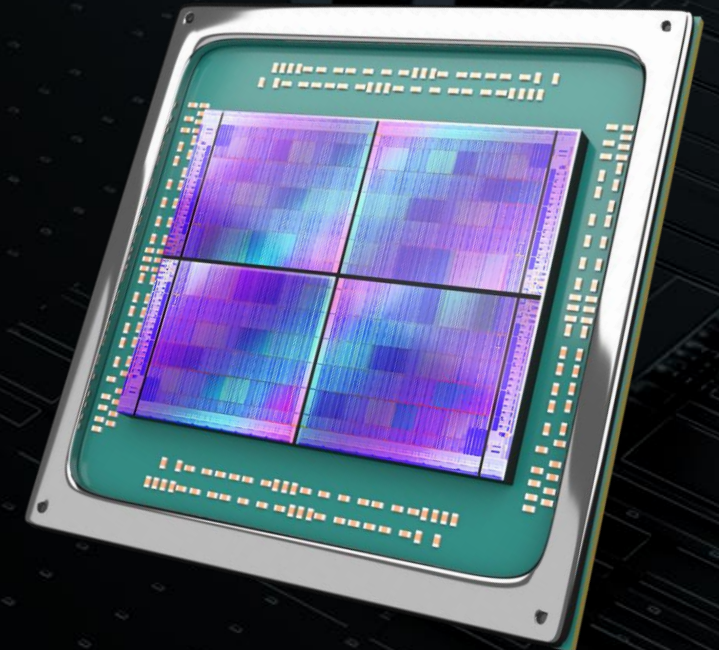
Endnotes

VER-001: Based on AMD internal analysis May 2023, comparing the number of system logic cells of the Versal Premium VP1902 device versus the Virtex UltraScale+ VU19P device.

VER-002: Based on AMD internal analysis in May 2023 with a 6-input LUT count to compare the Versal Premium VP1902 device versus the Intel Stratix 10 GX 10M FPGA.

VER-003: Based on AMD Labs testing using an A6865 package to simulate the XPIO data rate performance of an AMD Versal Premium VP1902 device versus the published data rate of an AMD Virtex UltraScale+ VU19P FPGA. Actual results will vary.

VER-004: Based on AMD internal analysis in May 2023, comparing the readback/writeback performance of an AMD Versal adaptive SoC CFI interface versus an AMD Virtex UltraScale+ FPGA ICAP interface. Actual performance will vary.

VER-005: Based on AMD Labs calculation in May 2023 of aggregate transceiver bandwidth of a Versal Premium VP1902 device B6865 package versus a Virtex UltraScale+ VU19P device B3824 package, assuming GTY/GTYPs running at 32G and GTMs running at 56G.

VER-008: Based on AMD Labs internal analysis in May 2023, comparing the latency in nanoseconds of an AMD Versal adaptive SoC XPIO in an 8:1 mux configuration with bypass FIFO mode enabled to a Virtex UltraScale+ FPGA HP I/O with no bypass FIFO option. Actual results will vary.

AMD

Disclaimer and Attributions

AMD