

June 2024



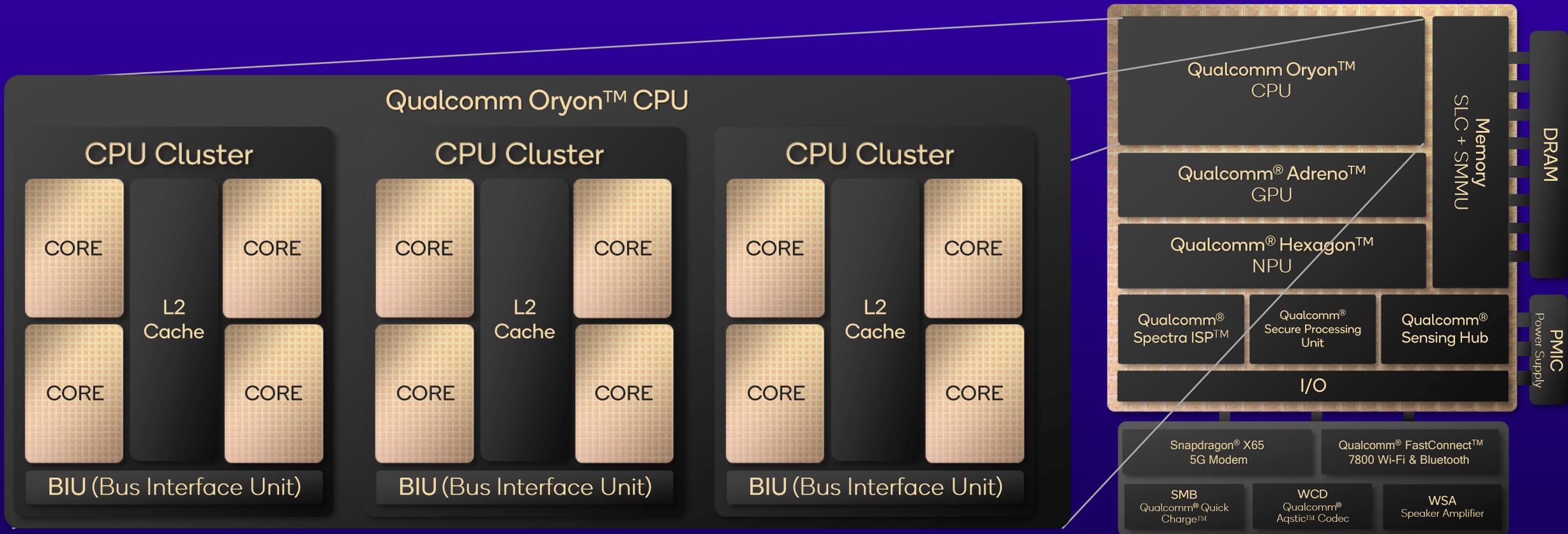
# QUALCOMM ORYON™ CPU

**Gerard Williams**

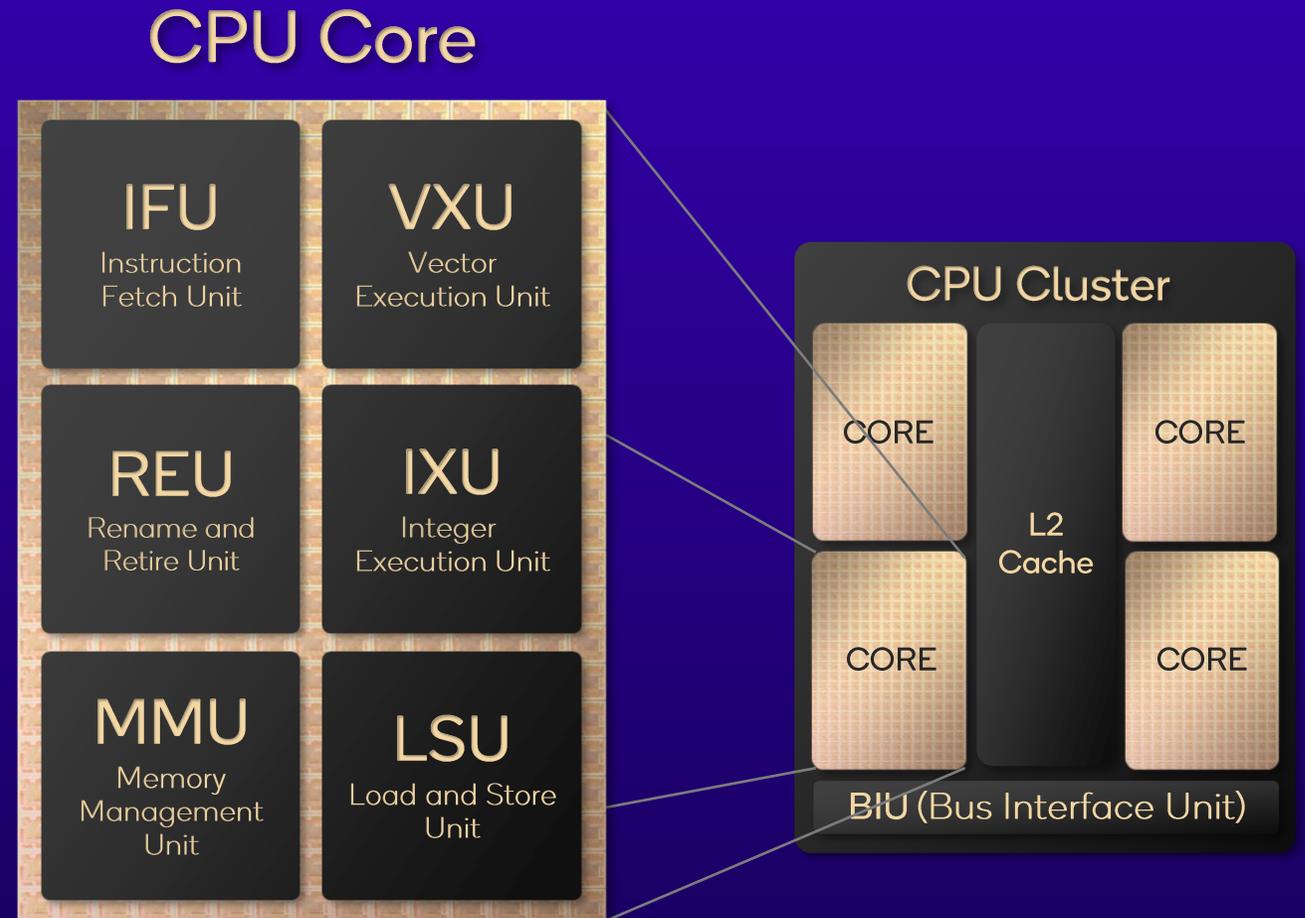
Senior Vice President of Engineering  
Qualcomm Technologies, Inc.

Custom CPU built on a 4nm process  
from the ground up for modern  
compute workloads

# Qualcomm Oryon CPUs Powering the SoC



Efficient  
performance starts  
at the CPU Core



# CPU Micro-Architecture

# Fetch & Decode

## Instruction Cache

- Fully coherent 192KB 6-way L1 ICache
- Fetches up to 16 instructions per cycle

## Instruction Translation, L1 iTLB

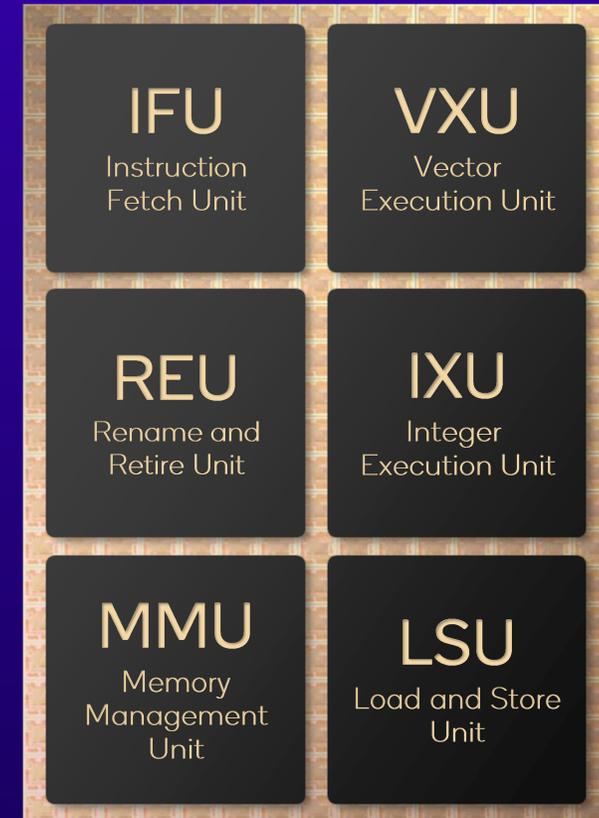
- Manages address translation for Instruction Fetch
- 256 entry 8-way buffer
- Supports both 4KB and 64KB translation granules

## Multiple Branch Prediction Tables

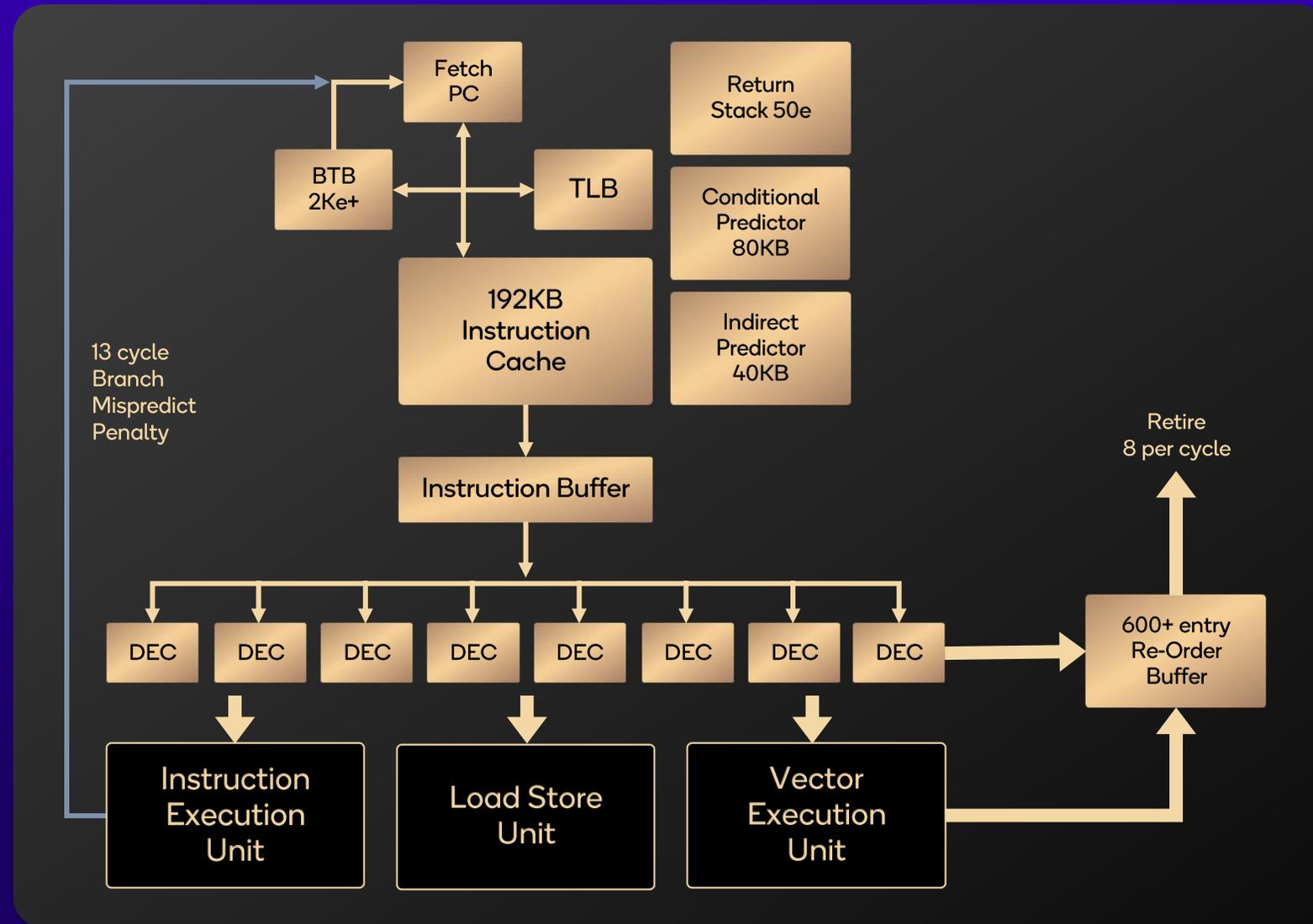
- Single Cycle Branch Target Buffer predicting the next fetch group PC
- Conditional Branch Predictor predicting the direction of the next branch
- Indirect Branch Target Predictor
  - Predicts the Branch Target Address
- Branch mispredict latency is designed for 13 cycles
- 50 entry Return Stack predictor

## Decode

- Up to 8 instructions / cycle decoded
- Map decoded instructions to uOps



# Fetch Pipeline, including Decode, Rename, and Retire



# Rename, Dispatch & Execution Pipes

## Register Rename

- Separate rename pools for IXU (integer) and VXU (vector) registers
- Integer pool is 400+ registers
- Vector pool is 400+ registers

## Reservation Stations

- IXU is 6-wide 64 bit pipeline, each with a 20e queue
- VXU is 4-wide 128 bit pipeline, each with a 48e queue
- LSU is 4-wide 128 bit pipeline, each with a 16e queue

## Integer Execution Pipes

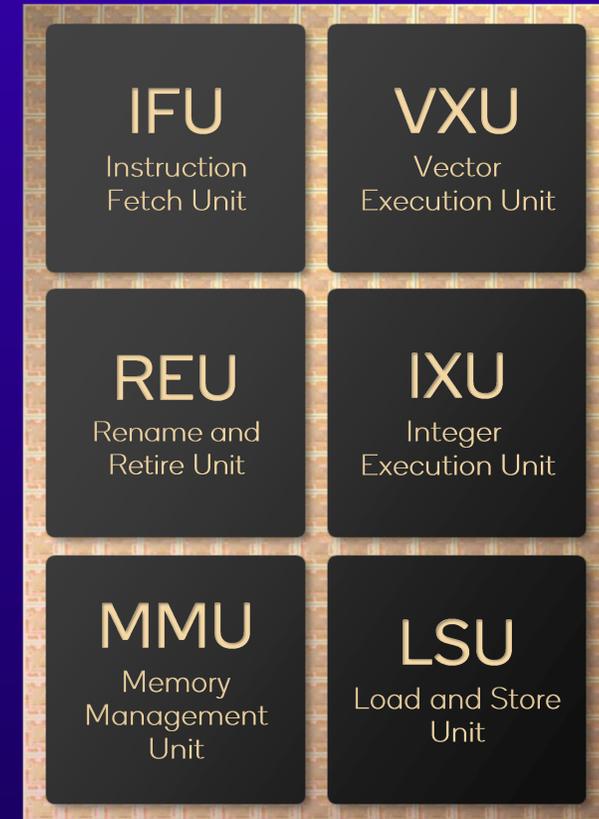
- Up to 6 ALU operations per cycle
- Up to 2 branches per cycle
- Up to 2 multiply per cycle

## Vector Execution Pipeline for FP and NEON SIMD

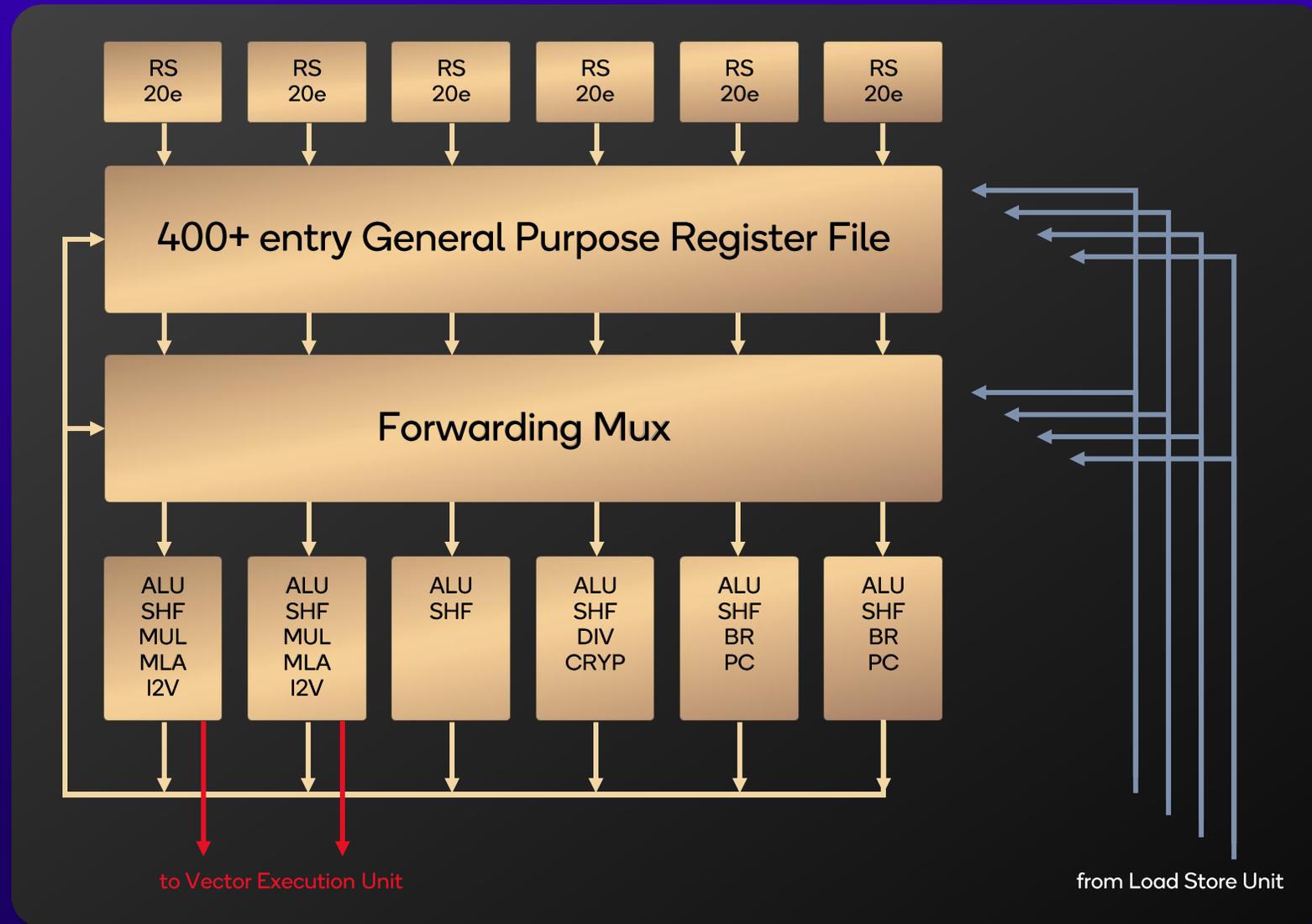
- Each pipe is up to 128-bit wide
- Up to 4 FP32-MLA/MUL, FP32-ADD, INT-MLA, and INT-ALU ops per cycle
- Supports all types : INT8, INT16, etc. and BF16, FP16, FP32, FP64

## Instruction Retirement

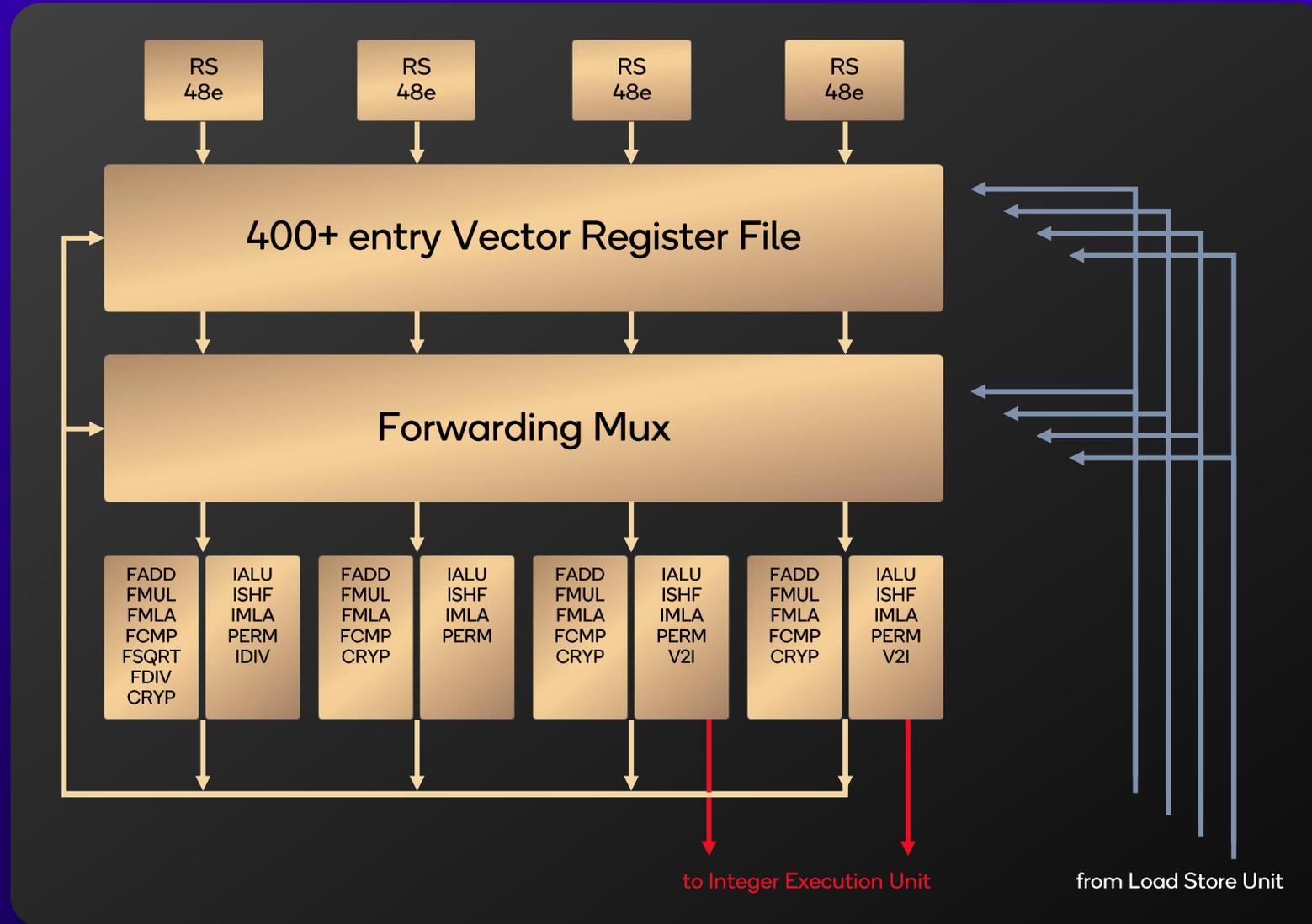
- Performed in-order, with up to 8 uOps per cycle
- Re-Order Buffer is 650+ uOps



# Instruction Execution Pipeline



# Vector Execution Pipeline



# Load-Store

## Data Cache for Load-Store

- Fully coherent 96KB 6-way L1 cache with 64B coherency granule
- Multi-ported and finely banked to support all access sizes

## Data Translation for Load-Store, L1 dTLB

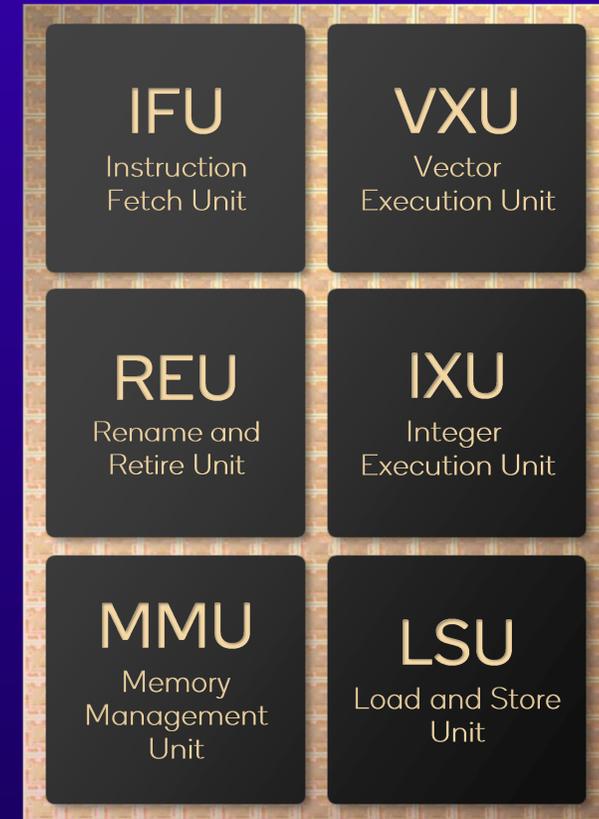
- Manages address translation for Load-Store
- 224 entry 7-way buffer
- Supports both 4KB and 64KB translation granules

## Load-Store Execution Pipes

- Up to 4 (of any combination of) Load-Store operations per cycle
- Full support for Store to Load forwarding
- 192 entry Load Queue and 56 entry Store Queue
- Tightly integrated with the large L2 cache, having 64B fills

## Prefetching

- Many advanced prefetching techniques applied
- Algorithms covering adjacent lines, stride, pointers, arrays, and patterns for both L1 data cache and data translation buffers





# Memory Management Unit

## L1 Instruction TLB and L1 Data TLB

- Supports VA-to-PA address translations for all traffic.
- 1 cycle access.

## L2-TLB Unified

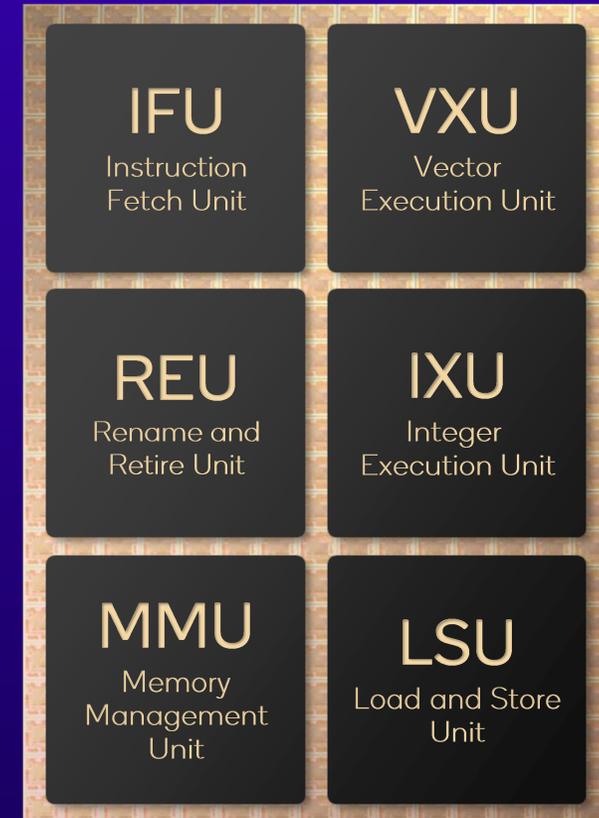
- Sized to handle Applications with large memory footprint
- >8K entry 8-way structure

## Table-Walk Caches

- Serve to cache descriptors from intermediate nodes of the page tables
- Supports 4KB and 64KB Translation Granules
  - And multiple-page sizes under each Granule
- Virtualization and 2-stage Translation
- Nested Virtualization: a Guest VM can host its own Guest-Hypervisor.

## Hardware Table Walker

- Multiple outstanding Table Walks in-flight per Core.



# Memory

## DRAM

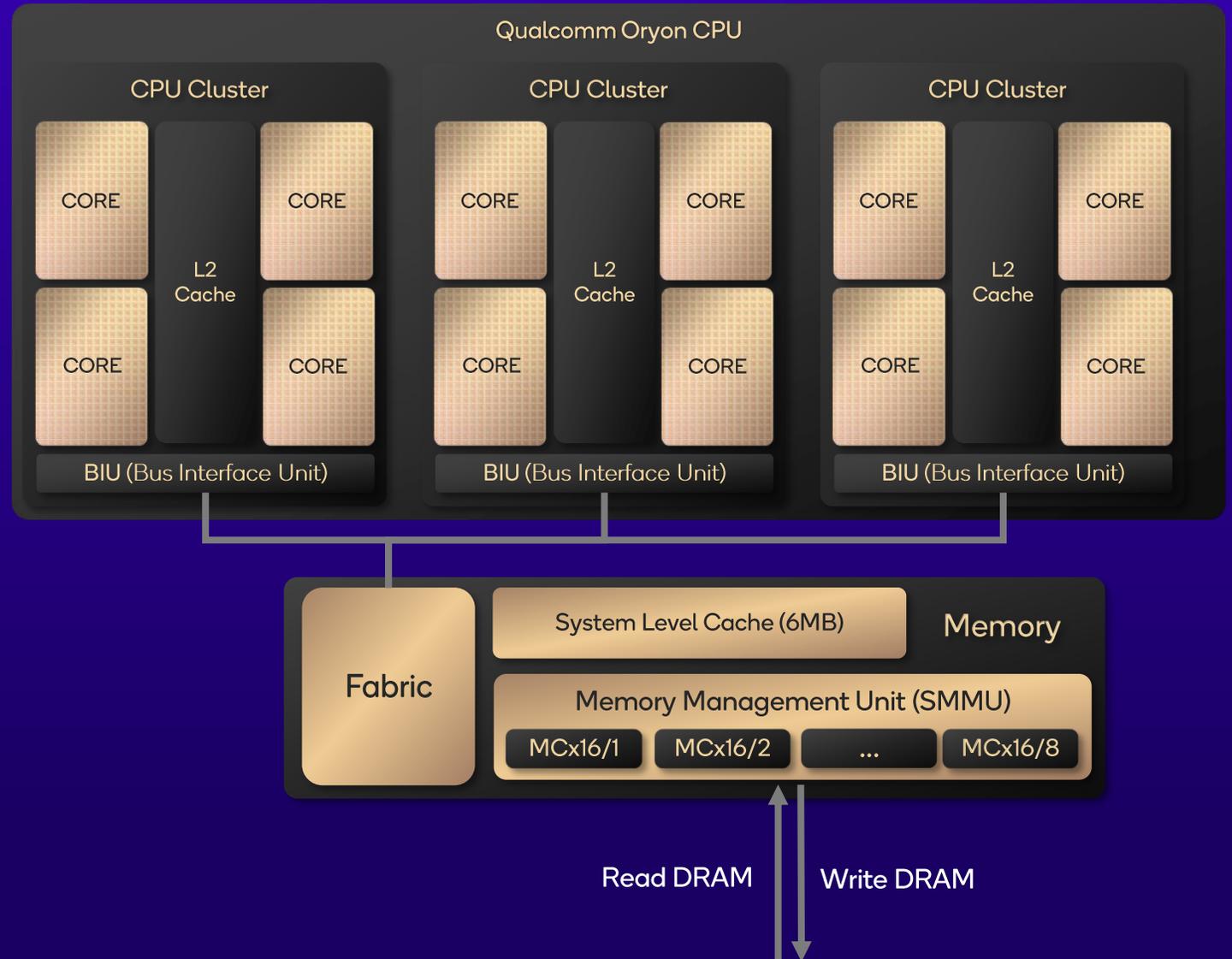
LPDDR5x

8448 MT/s, with 8-channel and 16-bit  
135 GB/s Bandwidth

Up to 64 GB

## System Level Cache (SLC)

6MB



# CPU Security

## Control Flow Integrity Measures

- Pointer Authentication
- Branch Target ID

## Side-Channel Mitigations built into Hardware

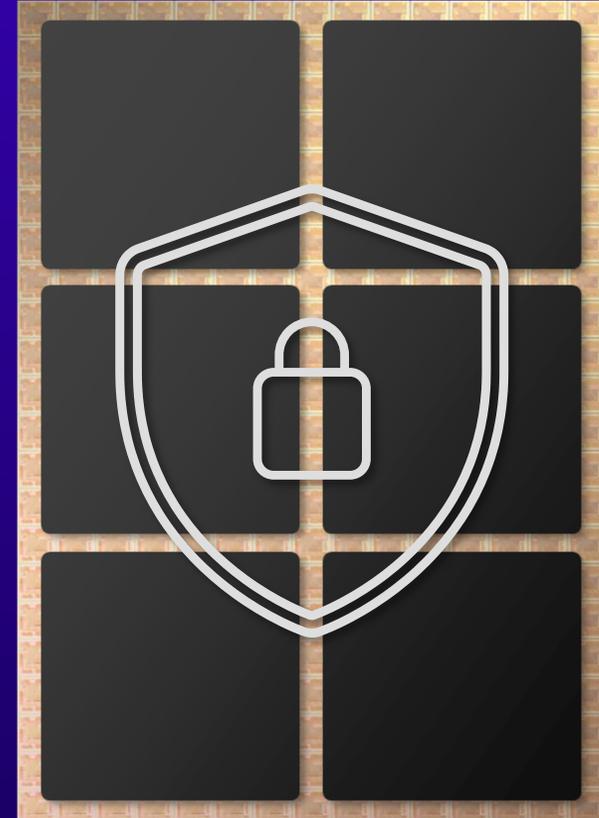
- Against all known attack vectors
- Meltdown, Spectre, and many more
- Use of internal block-ciphers for micro-architectural obfuscation

## Dedicated Random Number Generator per Cluster

## State of the Art Security Architectures

- Trustzone
- Cryptographic Extensions
- Pointer Authentication
- Branch Target ID
- Speculation Barriers & RCTX instructions
- Random Number Generation

## Not susceptible to GoFetch attacks



# CPU Benchmarks

# CPU Performance for GBv6 and Spec2017 (Windows, WSL, and Linux)

Geekbench 6.3.0 (Windows)	Score
File Compression	2313
Navigation	2362
HTML5 Browser	3331
PDF Renderer	2977
Photo Library	2885
Clang	3451
Text Processing	2688
Asset Compression	2782
Object Detection	2983
Background Blur	2760
Horizon Detection	2986
Object Remover	2633
HDR	3482
Photo Filter	3504
Ray Tracer	2417
Structure from Motion	2826
Benchmark Summary	
Single-Core Score	2868
Integer Score	2840
Floating Point Score	2920

Geekbench 6.3.0 (WSL)	Score
File Compression	2564
Navigation	2491
HTML5 Browser	3026
PDF Renderer	3063
Photo Library	2897
Clang	3566
Text Processing	2772
Asset Compression	2790
Object Detection	2991
Background Blur	2784
Horizon Detection	2962
Object Remover	3038
HDR	3446
Photo Filter	3860
Ray Tracer	2483
Structure from Motion	3108
Benchmark Summary	
Single-Core Score	2953
Integer Score	2892
Floating Point Score	3070

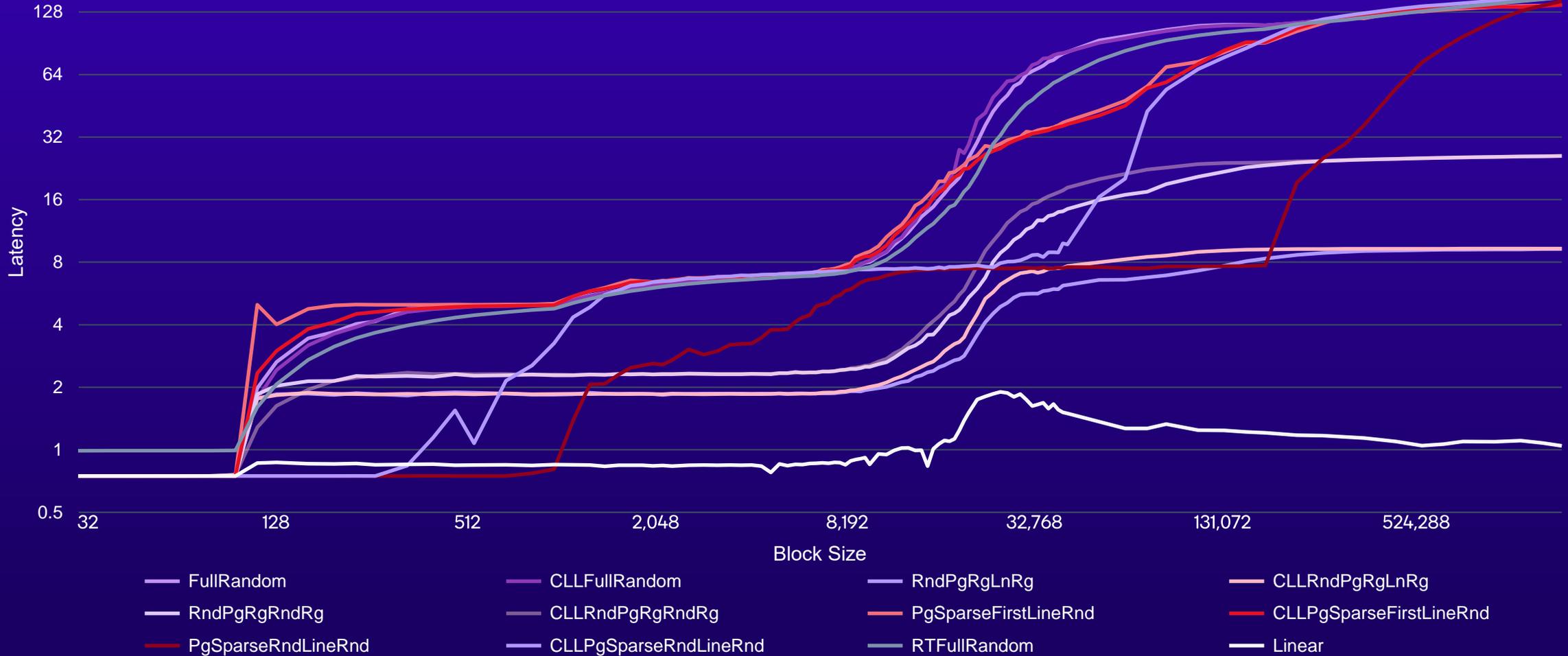
Geekbench 6.3.0 (Linux)	Score
File Compression	2800
Navigation	2591
HTML5 Browser	3287
PDF Renderer	3274
Photo Library	3030
Clang	3865
Text Processing	3015
Asset Compression	2944
Object Detection	3174
Background Blur	2921
Horizon Detection	3101
Object Remover	3431
HDR	3911
Photo Filter	4274
Ray Tracer	2657
Structure from Motion	3221
Benchmark Summary	
Single-Core Score	3169
Integer Score	3091
Floating Point Score	3320

CPU2017 Marketing (WSL)	Score
500.perlbench_r	9.1
502.gcc_r	12.7
505.mcf_r	10.04
520.omnetpp_r	6.27
523.xalanbmk_r	10.85
525.x264_r	24.9
531.deepsjeng_r	7.7
541.leela_r	7.48
548.exchange2_r	16.08
557.xz_r	4.04
IntRate	9.70
503.bwaves_r	136.75
507.cactuBSSN_r	9.97
508.namd_r	11.9
510.parest_r	9.94
511.povray_r	15
519.lbm_r	10.84
521.wrf_r	20.96
526.blender_r	11.98
527.cam4_r	17.63
538.imagick_r	14.62
544.nab_r	14.41
549.fotonik3d_r	27.83
554.roms_r	11.22
FPRate	16.66

CPU2017 Marketing (Linux)	Score
500.perlbench_r	10.08
502.gcc_r	13.98
505.mcf_r	10.63
520.omnetpp_r	7.20
523.xalanbmk_r	11.55
525.x264_r	26.43
531.deepsjeng_r	8.92
541.leela_r	8.10
548.exchange2_r	16.89
557.xz_r	4.61
IntRate	10.64
503.bwaves_r	154.24
507.cactuBSSN_r	10.21
508.namd_r	12.79
510.parest_r	10.57
511.povray_r	16
519.lbm_r	11.34
521.wrf_r	22.56
526.blender_r	12.95
527.cam4_r	18.73
538.imagick_r	15.35
544.nab_r	15.62
549.fotonik3d_r	29.15
554.roms_r	11.95
FPRate	17.77

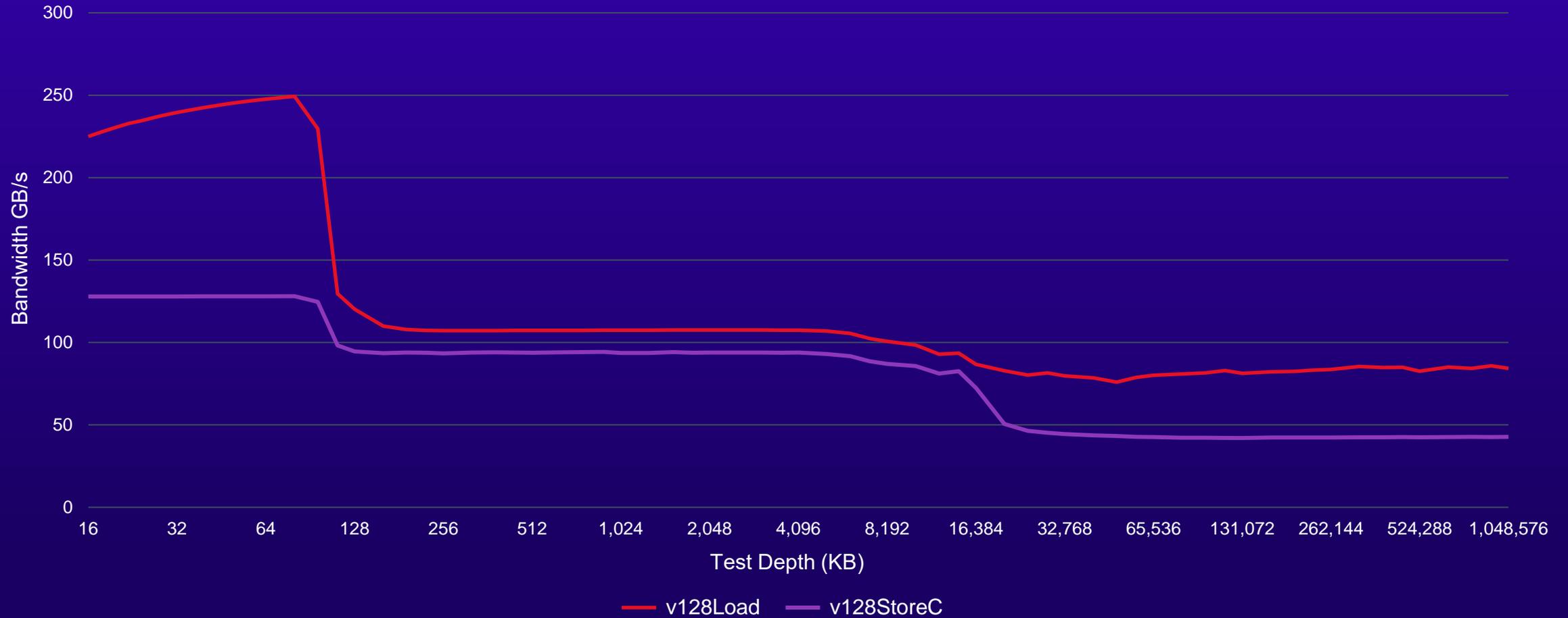
GeekBench6.3.0 and SPEC CPU2017 Singlethread (ST) run in June 2024. Snapdragon X Elite was tested using a Qualcomm reference design running Windows 11 OS. WSL numbers were run with Linux binaries under Windows Subsystem for Linux

# Memory System Latency



# Memory System Bandwidth

## Bandwidth 1T



# Scaling Qualcomm Oryon CPU technology to our business segments

Qualcomm Oryon  
CPU Technology



Segment-Specific  
Design Customizations

- Custom micro-architecture
- Unique implementations
- Modified Core/Cache configurations
- Memory/Cache hierarchy



# THANK YOU



Follow us on:       

For more information, visit us at:

[snapdragon.com](https://www.snapdragon.com) & [snapdragoninsiders.com](https://www.snapdragoninsiders.com)

All data and information contained in or disclosed by this document is confidential and proprietary information of Qualcomm Technologies, Inc. and/or its affiliated companies and all rights therein are expressly reserved. By accepting this material the recipient agrees that this material and the information contained therein will not be used, copied, reproduced in whole or in part, nor its contents revealed in any manner to others without the express written permission of Qualcomm Technologies, Inc. Nothing in these materials is an offer to sell any of the components or devices referenced herein.

©2018-2023 Qualcomm Technologies, Inc. and/or its affiliated companies. All Rights Reserved.

Qualcomm and Snapdragon are trademarks or registered trademarks of Qualcomm Incorporated. Other products and brand names may be trademarks or registered trademarks of their respective owners.

References in this presentation to "Qualcomm" may mean Qualcomm Incorporated, Qualcomm Technologies, Inc., and/or other subsidiaries or business units within the Qualcomm corporate structure, as applicable. Qualcomm Incorporated includes our licensing business, QTL, and the vast majority of our patent portfolio. Qualcomm Technologies, Inc., a subsidiary of Qualcomm Incorporated, operates, along with its subsidiaries, substantially all of our engineering, research and development functions, and substantially all of our products and services businesses, including our QCT semiconductor business.

Snapdragon and Qualcomm branded products are products of Qualcomm Technologies, Inc. and/or its subsidiaries. Qualcomm patented technologies are licensed by Qualcomm Incorporated.