RX 7900 XTX勉强干掉RTX 4080！还浪费了16%性能

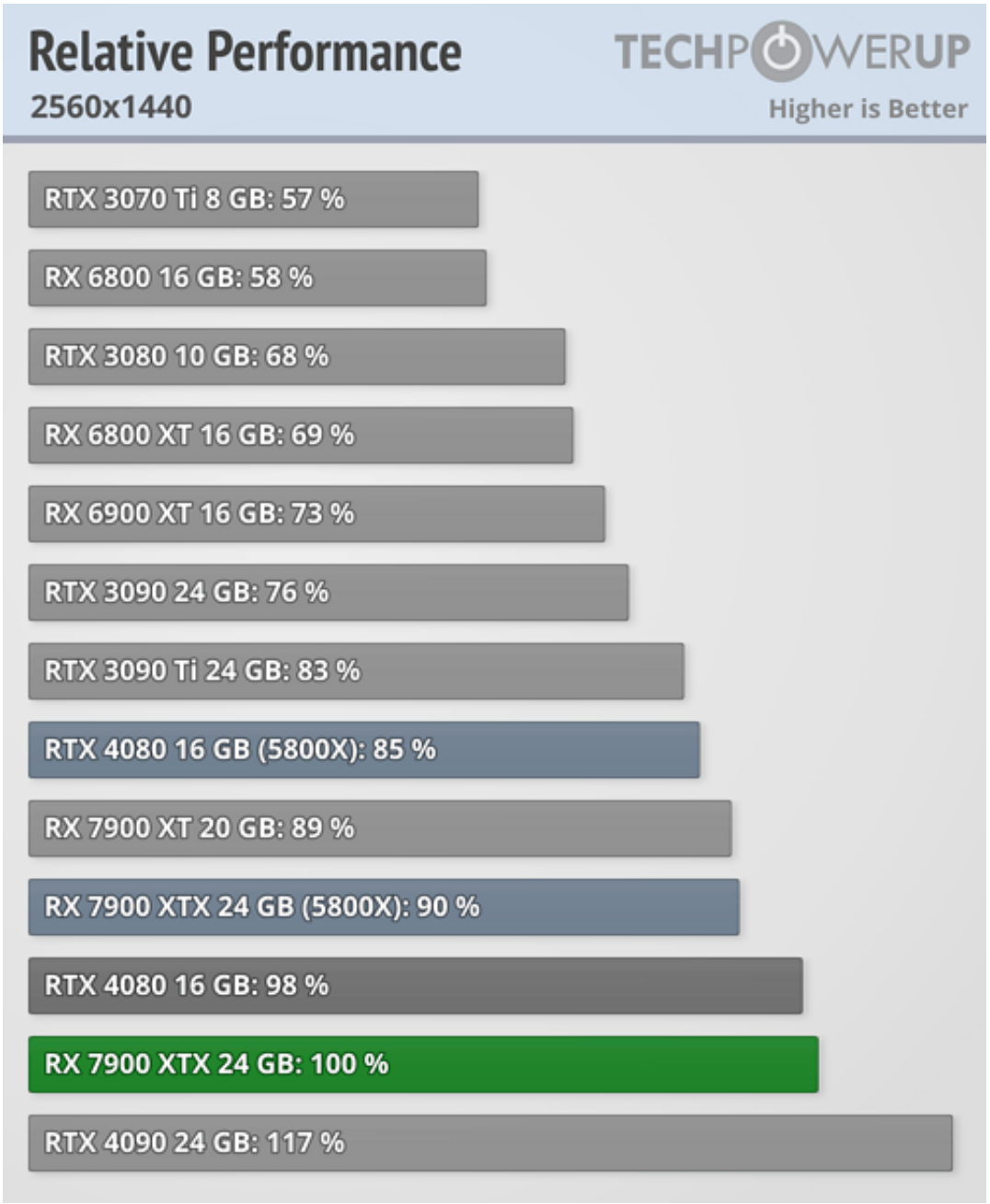
12月13日RX 7900
XTX显卡最新测试数据结果出来了，在这次测试结果中，RX 7900
XTX勉强干掉RTX
4080！综合性能勉强过关，平均幅度不到5%，非常吃功耗。

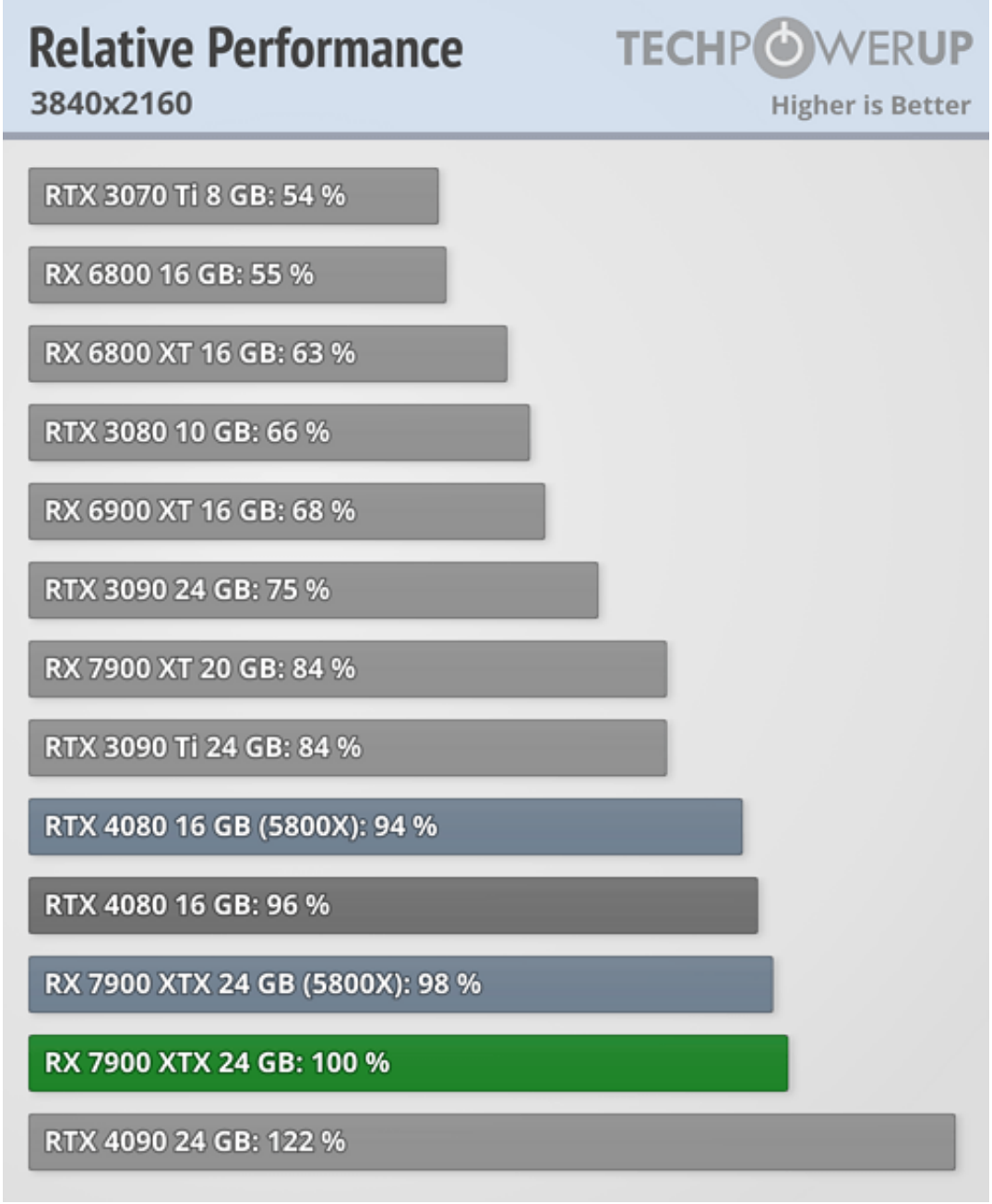
唯一的好消息就是，价格确实相对不高，RX 7900 XT、RX 7900
XTX分别做到了7399元、7999元原价起步。

其实，无论是从架构设计、规格参数还是官方宣传看，RX 7900 XTX的表现都不应该如此平平，那么问题出现在了哪里？是BIOS和驱动优化不到位吗？

游戏性能对比2560*1440									
测试项目	RTX 4090	RTX 4080	RTX 3090 Ti	RTX 3080 Ti	RTX 3080	RX 6950 XT	RX 7900 XTX	RX 7900 XTX vs RTX 4080	RTX 4080 vs RX 6950 XT
3DMark 图形分数	24744	17112	13901	12370	10918	14578	20104	17.5%	37.9%
3DMark SPY图形分数	19394	13851	11350	9896	8718	10224	14624	5.6%	43.0%
刺客信条：英灵殿	177	153	127	117	109	136	169	10.5%	24.3%
德军总部：新血脉	541	441	365	328	289	310	413	-6.3%	33.2%
地平线：零之曙光	202	197	179	165	145	178	193	-2.0%	8.4%
孤岛惊魂5	237	229	210	195	178	207	216	-5.7%	4.3%
孤岛惊魂6	181	167	150	136	129	156	186	11.4%	19.2%
古墓丽影：暗影	294	247	192	174	153	191	232	-6.1%	21.5%
极限竞速：地平线5	178	159	127	116	107	144	149	-6.3%	3.5%
赛博朋克2077	145	116	106	96	83	101	139	19.8%	37.6%
巫师3	264	213	168	154	136	179	237	11.3%	32.4%
无主之地3	225	166	140	122	104	146	189	13.9%	29.5%
战争机器5	206	192	156	137	121	153	204	6.3%	33.3%
性能百分比	112%	97%	82%	74%	66%	82%	100%	3%	20.3%
性能对比	-12%	+3%	+22%	+35%	+51%	+21%	0%	3%	20.3%
快科技 KKJ.CN									

游戏性能对比3840*2160									
测试项目	RTX 4090	RTX 4080	RTX 3090 Ti	RTX 3080 Ti	RTX 3080	RX 6950 XT	RX 7900 XTX	RX 7900 XTX vs RTX 4080	RTX 4080 vs RX 6950 XT
刺客信条：英灵殿	115	100	85	77	69	76	109	9.0%	43.4%
德军总部：新血脉	338	244	208	182	160	169	232	-4.9%	37.3%
地平线：零之曙光	159	118	108	96	84	94	120	1.7%	27.7%
孤岛惊魂5	183	141	124	111	95	127	161	14.2%	26.8%
孤岛惊魂6	139	106	95	84	78	96	124	17.0%	29.2%
古墓丽影：暗影	186	133	112	99	87	101	133	0.0%	31.7%
极限竞速：地平线5	151	121	96	89	79	105	117	-3.3%	11.4%
赛博朋克2077	75	58	54	48	40	45	68	17.2%	51.1%
巫师3	182	138	118	104	93	107	140	1.4%	30.8%
无主之地3	127	88	81	70	60	78	104	18.2%	33.3%
战争机器5	142	108	93	80	72	84	113	4.6%	34.5%
性能百分比	124%	94%	82%	73%	64%	76%	100%	6%	32.0%
性能对比	-24%	+6%	+22%	+38%	+56%	+32%	0%	6%	32.0%
快科技 KKJ.CN									



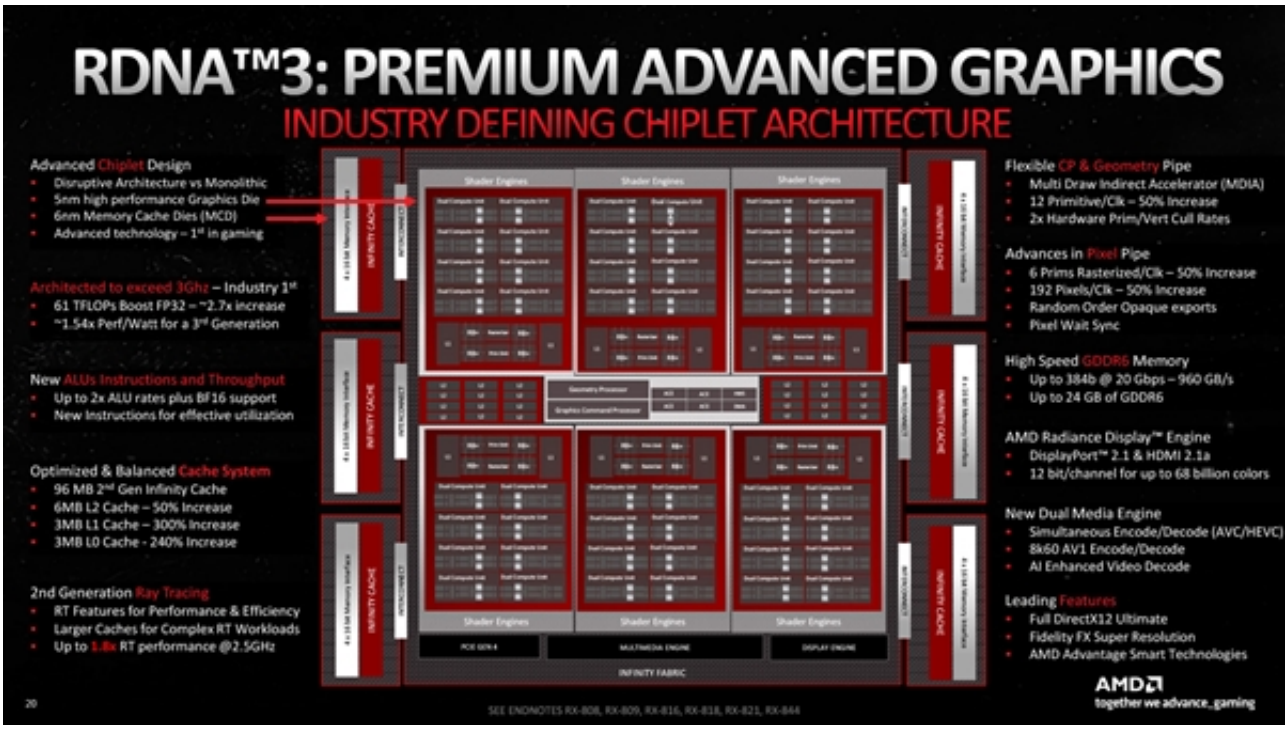


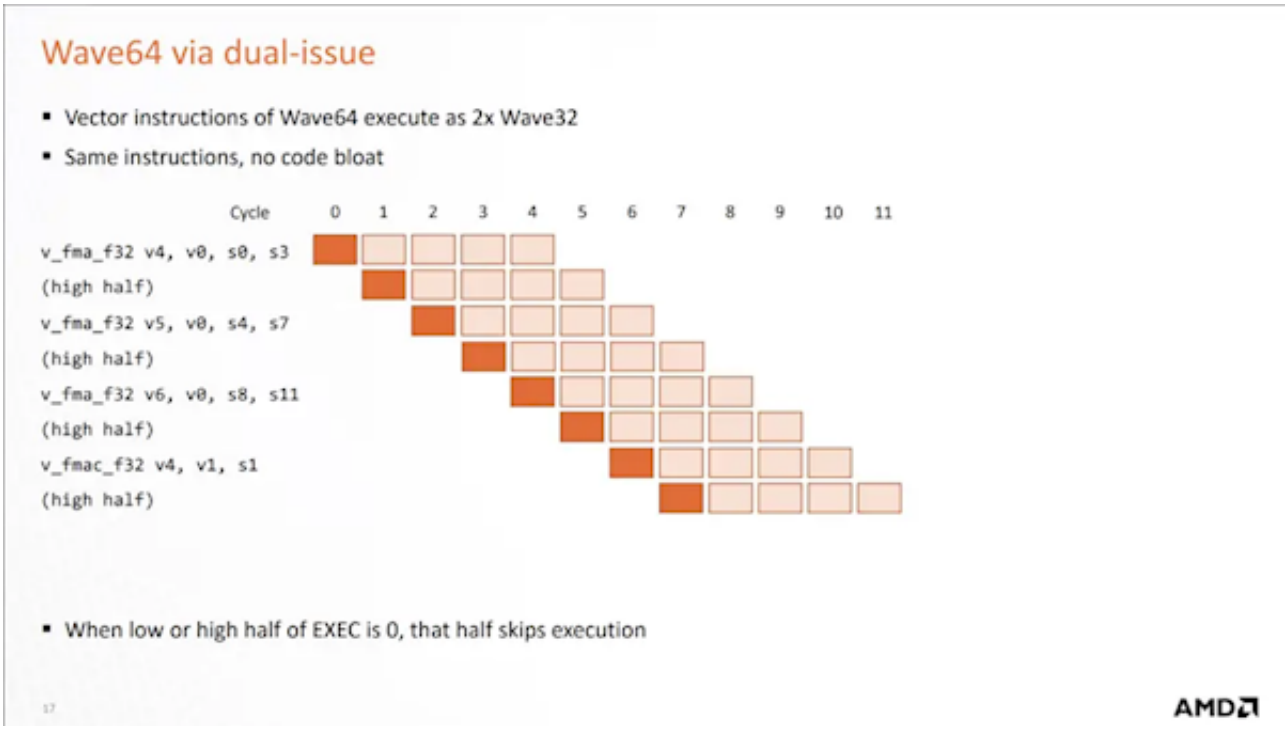
资深硬件玩家0x22h提出了一个观点，认为RDNA3架构设计存在重大缺陷，那就是增加了SIMD32单元之后，却没有相应地提升寄存器规格，导致Wave32单元内的VOPD格式下的两个FMA指令只能同时使用一个源操作数寄存器(operand register)、一个共享的立即数(immediate value)，Wave64下的情况也是类似。

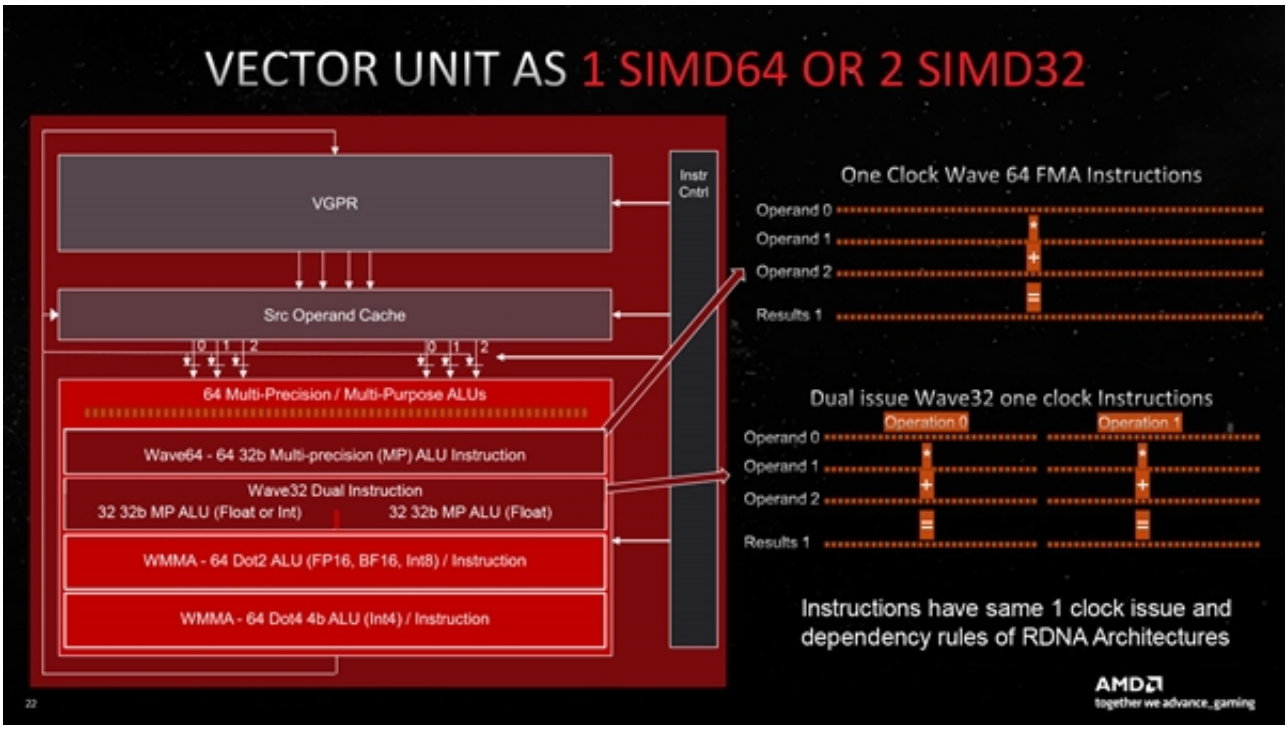
这就导致RDNA3 GPU的内部计算资源紧张，实际发挥出来的峰值性能，只有理论值的5/6，换言之就是浪费了超过16%的理论性能。

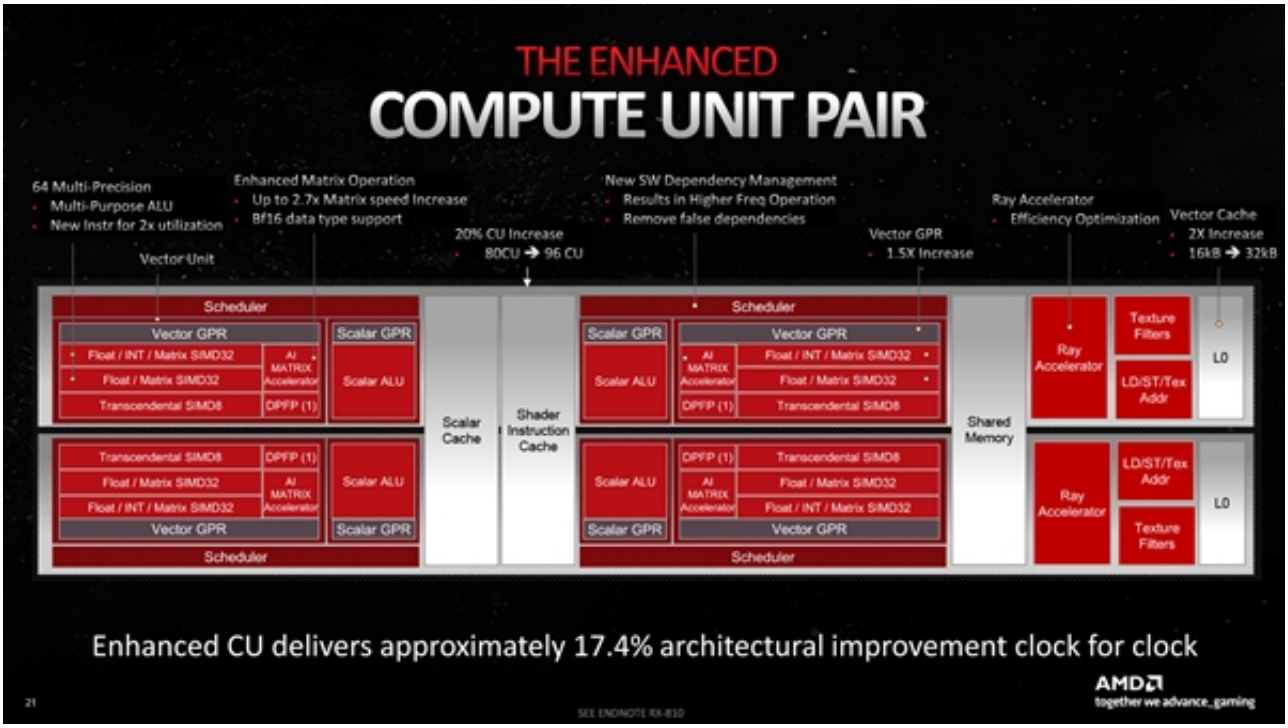
另外，RDNA3架构引入了双发射概念，流处理器等效数量翻番，但看样子并没有发挥出应有的实力。

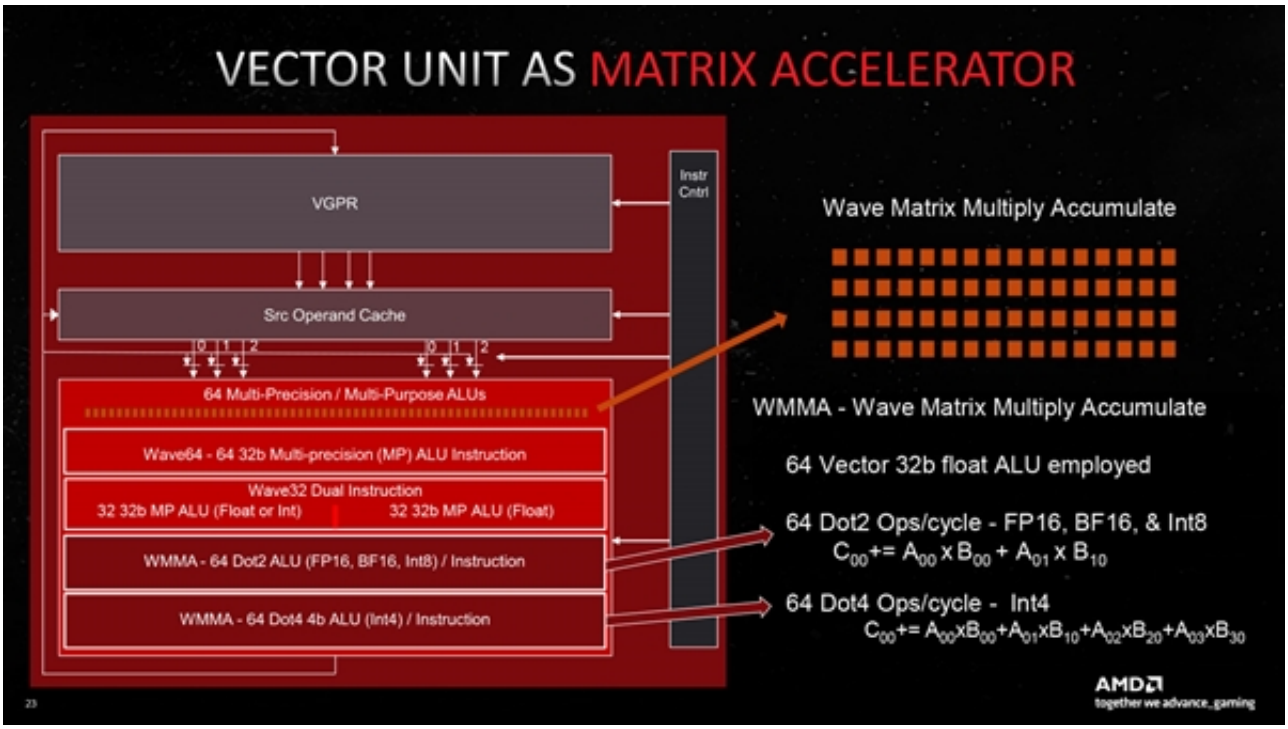
当然，以上观点都是猜测，尚无法得到AMD方面的证实。


RDNA™3: PREMIUM ADVANCED GRAPHICS
INDUSTRY DEFINING CHIPLET ARCHITECTURE
Advanced Chiplet Design
Disruptive Architecture vs Monolithic
5nm high performance Graphics Die
6nm Memory Cache Dies (MCD)
Advanced technology – 1st in gaming
Architected to exceed 3Ghz – Industry 1st
61 TFLOPs Boost FP32 – ~2.7x increase
~1.54x Perf/Watt for a 3rd Generation
New ALUs Instructions and Throughput
Up to 2x ALU rates plus BF16 support
New Instructions for effective utilization
Optimized & Balanced Cache System
96 MB 2nd Gen Infinity Cache
6MB L2 Cache – 50% Increase
3MB L1 Cache – 300% Increase
3MB L0 Cache - 240% Increase
2nd Generation Ray Tracing
RT Features for Performance & Efficiency
Larger Caches for Complex RT Workloads
Up to 1.8x RT performance @2.5GHz
Flexible CP & Geometry Pipe
Multi Draw Indirect Accelerator (MDIA)
12 Primitive/Clk – 50% Increase
2x Hardware Prim/Vert Cull Rates
Advances in Pixel Pipe
6 Prims Rasterized/Clk – 50% Increase
192 Pixels/Clk – 50% Increase
Random Order Opaque exports
Pixel Wait Sync
High Speed GDDR6 Memory
Up to 384b @ 20 Gbps – 960 GB/s
Up to 24 GB of GDDR6
AMD Radiance Display™ Engine
DisplayPort™ 2.1 & HDMI 2.1a
12 bit/channel for up to 68 billion colors
New Dual Media Engine
Simultaneous Encode/Decode (AVC/HEVC)
8k60 AV1 Encode/Decode
AI Enhanced Video Decode
Leading Features
Full DirectX12 Ultimate
Fidelity FX Super Resolution
AMD Advantage Smart Technologies
Shader Engines
INFINITY FABRIC
AMD
together we advance_gaming
20
SEE ENDNOTES RX-808, RX-809, RX-816, RX-818, RX-821, RX-844


Wave64 via dual-issue
Vector instructions of Wave64 execute as 2x Wave32
Same instructions, no code bloat
Cycle 0 1 2 3 4 5 6 7 8 9 10 11
v_fma_f32 v4, v0, s0, s3
(high half)
v_fma_f32 v5, v0, s4, s7
(high half)
v_fma_f32 v6, v0, s8, s11
(high half)
v_fmac_f32 v4, v1, s1
(high half)
When low or high half of EXEC is 0, that half skips execution
17
AMD


VECTOR UNIT AS 1 SIMD64 OR 2 SIMD32
VGPR
Instr Cntrl
Src Operand Cache
64 Multi-Precision / Multi-Purpose ALUs
Wave64 - 64 32b Multi-precision (MP) ALU Instruction
Wave32 Dual Instruction
32 32b MP ALU (Float or Int)
32 32b MP ALU (Float)
WMMA - 64 Dot2 ALU (FP16, BF16, Int8) / Instruction
WMMA - 64 Dot4 4b ALU (Int4) / Instruction
One Clock Wave 64 FMA Instructions
Operand 0
Operand 1
Operand 2
Results 1
Dual issue Wave32 one clock Instructions
Operation 0
Operation 1
Instructions have same 1 clock issue and dependency rules of RDNA Architectures
AMD
together we advance_gaming
22


THE ENHANCED
COMPUTE UNIT PAIR
64 Multi-Precision
Multi-Purpose ALU
New Instr for 2x utilization
Vector Unit
Enhanced Matrix Operation
Up to 2.7x Matrix speed Increase
Bf16 data type support
20% CU Increase
80CU → 96 CU
New SW Dependency Management
Results in Higher Freq Operation
Remove false dependencies
Vector GPR
1.5X Increase
Ray Accelerator
Efficiency Optimization
Vector Cache
2X Increase
16kB → 32kB
Scheduler
Vector GPR
Scalar GPR
Float / INT / Matrix SIMD32
Float / Matrix SIMD32
Transcendental SIMD8
AI MATRIX Accelerator
DPFP (1)
Scalar ALU
Scalar Cache
Shader Instruction Cache
Shared Memory
Ray Accelerator
Texture Filters
LD/ST/Tex Addr
L0
Enhanced CU delivers approximately 17.4% architectural improvement clock for clock
AMD
together we advance_gaming
21



本文链接：https://dqcm.net/zixun/16709122772788.html