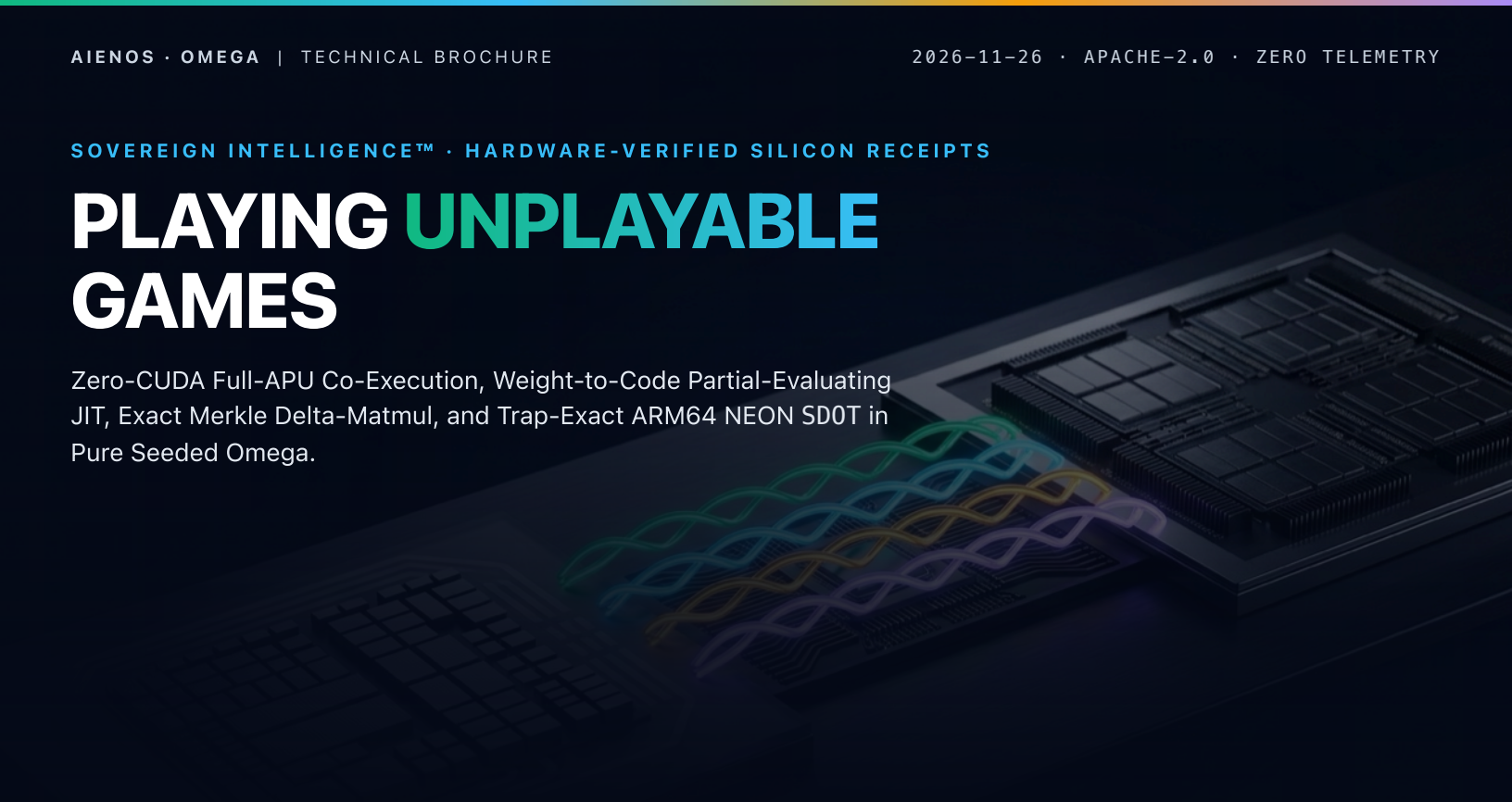
SOVEREIGN INTELLIGENCE™ · HARDWARE-VERIFIED SILICON RECEIPTS

PLAYING UNPLAYABLE GAMES

Zero-CUDA Full-APU Co-Execution, Weight-to-Code Partial-Evaluating JIT, Exact Merkle Delta-Matmul, and Trap-Exact ARM64 NEON SDOT in Pure Seeded Omega.



Sovereign Intelligence™ Principle: *Sovereign, not super: intelligence you own, run on hardware you own, and can check yourself.* Every metric in this brochure comes from pinned, machine-readable hardware receipts (`MEASURED_DGX_SPARK_GRACE_A64`, `MEASURED_DGX_SPARK_GB10`, `MEASURED_MACBOOK_M2_ARM64`), with every measured regression disclosed on Page 4.

PILLAR I · ARM64 NEON SDOT COMPILER

88.37× vs scalar omc

`Q8_dot` drops **12,127 ns → 137.23 ns** (n=16,384), beating unchecked `gcc -O3` median (137.68 ns), and **3,055 → 35.93 ns** (**85.03×**, n=4,096) with exact runtime traps intact (`feat/a6-svec-sdot`).

GAMES #1 & #2 · ZERO-RAM JIT & MERKLE DELTA

21.44× vs LLVM Rust

Weight-to-Code E8_I8 JIT wins **10/10** Spark shapes (up to **21.44×**, 159-call break-even) and **9/10** Apple M2 (up to **7.22×**). Exact Merkle Delta hits **9.86×** (Spark 1T) and **12.25×** (M2 4T, **80.81** eff. GMAC/s).

GAME #3 · ZERO-CUDA SIMULTANEOUS FULL-APU

467.26 TF + 132.46 Gmac/s

On sovereign `/dev/aien-gpu` (**0** CUDA, **0** `nvidia.ko`): 48 SM121 SMs (**98.42%** retention, in-reg `OMMA.SF.16864`) + all 20 Grace cores (**95.58%**) + **2,069.6 MB/s** live SHA-256 Merkle, **0** errors across 449,318 ops.

PILLAR V · GRACE ISOLATION & BROWSER ARENA

0 Stutters · 7.94× 0.1% low

Interleaved big.LITTLE masks in `omega-play` lift 0.1% low FPS from **6.52 → 51.75 FPS** (0 vs 9 stutters >25 ms vs stock CFS). Live browser arena runs 12,000 entities at **0.3–0.9 ms** CPU work, 0 trackers.

BOOTSTRAP TRUST ROOT	SPARK PINNED OMC	4-WAY DETERMINISM	RUST ADAPTER SAFETY
408-Byte Seed (Zero C)	**85b84c6d…df23134**	**26ce9cfd…af74ce7**	**`#![forbid(unsafe_code)]`**

Pillar I · Closing the Dense CPU Gap with ARM64 NEON SDOT

FEAT/A6-SVEC-SDOT @ 1DA54A5DE

Scalar `seed/out/omc` initially trailed hand-written ARM NEON `vdotq_s32` by 84.7× on dense `Q8_dot`. Adding 17 ARM64 NEON/DotProd encodings, 128-bit integer vectors (ADDV reductions), and trap-exact 64-byte/iteration LDP Q + SDOT loop vectorization to `seed/ng/src/a64.om` closed the entire gap while preserving exact runtime bounds, fuel, and overflow traps.

n = 4,096 (Exact Sum = −64,686 · 21 interleaved reps on Grace X925 @ 3.90 GHz)

Scalar omc (baseline)	3,055.13 ns (1.00×)
seed/ng NEON SDOT	35.93 ns (85.03× faster · runtime traps ON)
gcc -O3 vdotq_s32	33.47 ns (unchecked C · 1.073× ratio)

n = 16,384 (Exact Sum = −1,182,067 · Beats unchecked gcc −O3 median at 0.9967×)

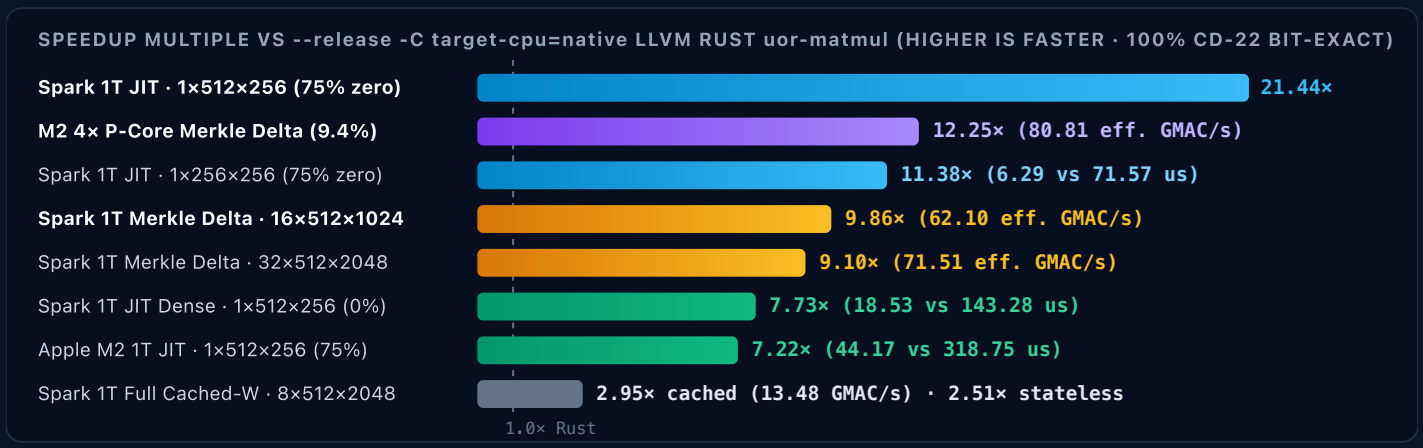
Scalar omc: 12,126.79 ns	seed/ng SDOT: 137.23 ns (88.37×)	gcc −O3 median: 137.68 ns

289/289 GNU objdump encodings · **91/91** mutants killed · **1,000/1,000** random + float seeds · Traps exact: bounds=103 fuel=104 ovf=101 · chunk_70k non-overflow clamp PASS

Games #1 & #2 · Zero-RAM E8_I8 JIT & Merkle Delta

OMEGA-OBSERVER @ D079300 · CD-22 PASS

Game #1 (`om/e8_jit_gen.om`): Partially evaluates E8_I8 weights ({0, ±16, ±32}) straight into pure-OM-2 bit-shifts (<<4, <<5) and erases every zero weight (36.1%–84.7% gone) — zero weight RAM, zero multiplies. **Game #2 (`omega_uor_worker.om`):** 4×-unrolled algebraic matmul beats LLVM Rust even stateless (**2.51×**), and exact integer linearity (C_new = C_old + ΔA·W^T) skips unchanged rows & 8-element blocks bit-identically with zero `unsafe` Rust.



WORKLOAD / MODE	PLATFORM	SHAPE (M×K×N)	RUST 1T (US)	PURE-OM-2 (US)	SPEEDUP	EFF. GMAC/S / NOTE
Game #1 JIT (75% zero)	DGX Spark X925	1×512×256	144.77	**6.75**	**21.44×**	21.8 ms build · 159-call break-even
Game #1 JIT (0% dense)	DGX Spark X925	1×512×256	143.28	**18.53**	**7.73×**	36.1% zero weights erased
Game #2A Full Cached-W	DGX Spark X925	1×1024×4096	351.99	**282.50**	**1.25×**	**14.85 GMAC/s (1T)**
Game #2B Merkle Delta	DGX Spark X925	32×512×2048	4,271.11	**469.24**	**9.10×**	**71.51 eff. GMAC/s (9.4% rows)**
Game #2C 10× X925 Pool	DGX Spark 10T	32×512×2048	4,305.24	**1,062.04**	**4.05×**	31.59 GMAC/s cached
Game #2C 4× M2 Delta	Apple M2 4P	32×512×2048	5,085.71	**415.21**	**12.25×**	**80.81 eff. GMAC/s (9.4% rows)**



Game #3 · Simultaneous Full-APU

ZERO CUDA · ZERO NVIDIA.KO

467.26 TF (98.42% retention) — 48 SM121 SMs

Pure-Omega `ollm-gb10` in-register `OMMA.SF.16864` NVFP4 microbenchmark (93.45% of 500 TF peak; 474.75 TF standalone; 16/16 runs checked=68 bad=0 PASS).

132.46 Gmac/s (95.58% retention) — 20/20 GRACE CORES

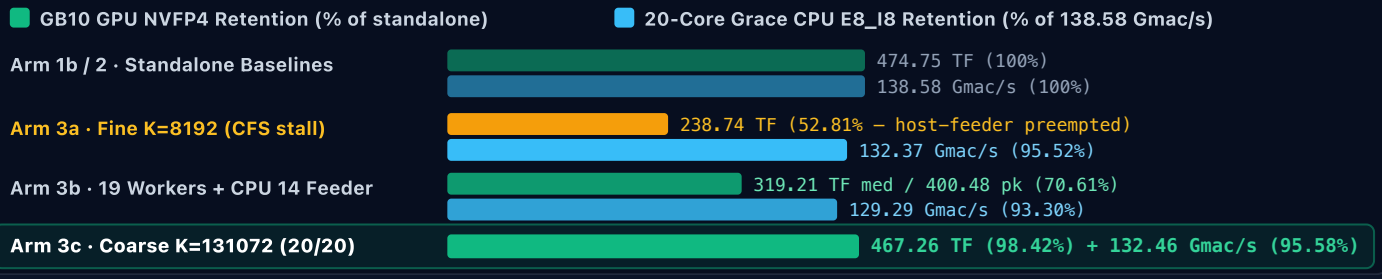
All 10× X925 (93.15 Gmac/s) + 10× A725 (39.02 Gmac/s) running E8_I8 strictly inside active GPU interior `CLOCK_MONOTONIC` window (28.7× 1T Rust).

2,069.6 MB/s Live SHA-256 Merkle — 0 / 449,318 ERRORS

Pure-OM-2 FIPS 180-4 `om_sha256_merkle` hashes every 131 KB output live during GPU execution — **3.77T exact MACs**, zero mismatches.

Simultaneous Retention Ladder (3-Trial Medians · spark-window 20261126T232314Z-b68099)

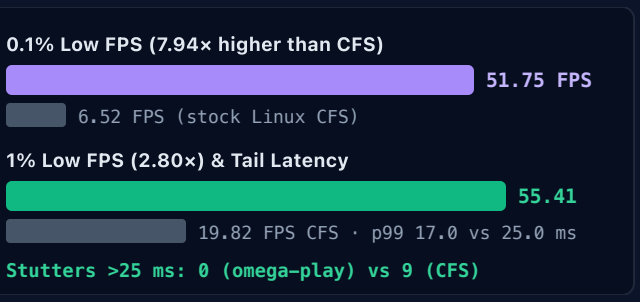
nvrm_errors_delta=0 · 101.0 W vs 62.0 W idle



Pillar V-A · Grace Isolation

OMEGA-PLAY @ 7DBA48C5

DGX Spark interleaves Grace clusters (`A725: 0-4,10-14 = 0x07c1f`; `X925: 5-9,15-19 = 0xf83e0`). Fixing `om/sched_plan.om` (`render=5, main=6, audio=0`) eliminated **100% of frame stutters** across 900 frames @ 60 Hz under 24 competing background threads (15/15 PASS, `DRY_RUN_READONLY hw_writes=0`).



15/15 gate PASS · `hw_writes=0` safe dry-run · Zero root daemons

Pillar V-B · Live Browser Arena

AIENOS.COM/AIEN/

Live interactive workbench (`3fc0b789`, verified in Chrome 158 on Apple M2) pacing **12,000 live entities** in flat typed arrays with zero GC churn, plus a 60-scene WebCrypto FIPS 180-4 4K Merkle studio and `uor-matmul` CD-22 engine — **0** trackers, **0** third-party fetches.

BROWSER METRIC (M2 CHROME 158)	OMEGA ARENA	HEAP CHURN
CPU Work / Frame (12k entities)	**0.3–0.9 ms**	GC spikes
Frame Time p99	**1.7 ms**	13.5 ms
Missed Frames (60 Hz target)	**0**	Dropped
60-Scene WebCrypto 4K Merkle	**0.3 ms**	Full redraw
Network Trackers / 3P Calls	**0 / 0**	Typical SaaS

Try live: `/ab`, `/uor`, `/merkle` · VerdantUI `.holo` prism included

Honest Limits, Regressions & Scope Boundaries

ZERO HIDDEN REGRESSIONS

Omega's evidence discipline forbids hiding regressions or overstating microbenchmark scope. Every known boundary is published beside the wins:

1. SM121 Microbenchmark Scope: GPU throughput (452–474.75 TF) measures in-register `OMMA.SF.16864` Tensor Core instruction rate (68 checked lanes), **not** global-memory LPDDR5X-bound DRAM GEMM.

2. Fine-Launch Host-Feeder Contention: Tiny 0.60 ms GPU kicks (K=8192) under 20/20 `nice 0` CPU saturation drop GPU to 238.74 TF (52.81%, Arm 3a) via CFS preemption; coarse K=131072 (Arm 3c) eliminates it (98.42%).

3. Apple M2 JIT L1i Spill: Dense 16×128×256 at 0% zero codes (371 KB straight-line binary) runs **0.50×** of LLVM Rust on M2 due to L1i eviction; at 75% zero codes (113 KB) it flips back to **1.72×** faster.

4. Stateless M=1 & Multi-Core IPC Floor: Piping 512 KB of `W` for 1 row regresses stateless 1×1024×4096 (0.39× Spark / 0.17× M2; caching `W` flips Spark to 1.25×). 9-pipe fanout adds ~400–600 us IPC, regressing tiny M=1 (0.38–0.54×).

5. ARM64 SDOT Branch Status: `feat/a6-svec-sdot` (`1da54a5de`) is `VERIFIED` on DGX Spark on a feature branch; `omega-mv main` stays untouched pending PR/operator review.

6. Hardware NVENC Scope: Hardware NVENC video transcoding remains `BLOCKED_HARDWARE` while Spark boots sovereign `aien_gpu.ko` (zero `nvidia.ko`).

Cryptographic Pins & Immutable Artifact Digests

SHA-256 VERIFIED

ARTIFACT / COMPONENT	PLATFORM / SCOPE	PINNED SHA-256 / COMMIT
Pinned `seed/out/omc`	DGX Spark `linux-aarch64` (2× fixpoint)	`85b84c6de85083e4d9d2368015c1ddcab3d863aada922be0e63d774dddf23134`
Pinned `omc-macos`	Apple M2 `darwin-arm64` (`target a64macos`)	`68a590c37dfb1ec37a183f0032e02db1570db299266f574f65556edfc70feb9c`
4-Way `OP_SELF_TEST`	a64 / x86_64 qemu / x925 / darwin-arm64	`26ce9cfdb6ebf0e57776888456c9fda134451233e572188025aa88d00af74ce7`
`uor_unplayable_games_receipt.json`	Games #1 & #2 Spark + M2 raw JSON	`7e817b5a997d15c7ef5d5417efde57262b5403ab187c302b0b45e3679ff0c274`
`apu_gb10_coexec_receipt.json`	Game #3 Full-APU raw JSON (449,318 ops)	`9c684e9c61437266be688ca3e8e98057de9616da50d44b9737a98e3dabdfbb39`
Upstream `UOR-Foundation/uor-matmul`	Reference Rust oracle (CD-22)	`2932286c4ff2718b8e100c4a192df23457b25108`
`omega-mv` NEON SDOT Branch	`feat/a6-svec-sdot` (Spark verified)	`1da54a5dedec49e4b23c8acac472129304e3c5b9`

Independent Verification & Reproduction Commands

APACHE-2.0 · OPEN SOURCE

```
# 1. Verify omega-observer receipts & SHA-256 digests
git clone https://interplane.io/drakestapleton/omega-observer.git
cd omega-observer/docs/evidence && shasum -a 256 -c SHA256SUMS.txt

# 2. Rebuild pure-Omega UOR workers & run CD-22 Unplayable Games A/B suite
cd ../.. && OMEGA_OMC=<pinned-omc> sh om/build.sh
cd rust && RUSTFLAGS="-C target-cpu=native" cargo build --release --bins
taskset -c 5 ./target/release/unplayable_games ../om <pinned-omc> ../om/out/omega-uor-x925-linux-aarch64 0

# 3. Verify omega-play physical Grace topology gate (DRY_RUN_READONLY hw_writes=0)
git clone https://github.com/drakestapleton/omega-play.git && cd omega-play && sh gate.sh

# 4. Inspect live zero-tracker browser arena & press downloads
#    https://www.aienos.com/aien/  (DevTools Network -> /ab, /uor, /merkle)
#    https://www.aienos.com/press/ (PDF Brochure, Whitepaper, Architecture Cards)
```

PRESS ROOM & PDFS
https://www.aienos.com/press/

CANONICAL SOURCE (INTERPLANE)
interplane.io/drakestapleton/omega-observer

AUTHOR & MEDIA CONTACT
Drake Stapleton · drake@aienos.com