Gabriel da Silva

Worldwide (Remote) - Time Zone: EST/CST Aligned tracerofphotons@gmail.com dasilvagf.github.io
gfelipess7 dasilvagf

Introduction

High-Performance Systems Architect focused on the intersection of Low-Level GPU Optimization, GPU Microarchitecture, and Neural Rendering. With over 14 years of experience, I specialize in bridging the gap between hardware constraints (Assembly/ISA) and application performance, optimizing critical pipelines from the driver stack up to the user interface.

Throughout my career, I have engineered core technology adopted by tier-1 organizations including NASA, Formula One, and the U.S. Government. Currently, I am architecting "Zero-Copy" infrastructure for next-generation GPUs, delivering next-generation visualization infrastructure that impacts billions of users and devices worldwide.

Experience

Embedded Engineering Expert I - R&D GPU Architecture *Texas, USA (Remote)*
Qubika *Jul 2025 – Present*

- Architect high-performance compute kernels using ISA-specific assembly to saturate next-generation hardware units, specifically optimizing data flow for Ray Tracing (RT) Cores and AI Acceleration engines.
- Engineer bare-metal firmware solutions that bypass standard driver abstraction, directly managing the memory hierarchy (LDS/L1/L2) and instruction scheduling to achieve theoretical peak throughput in compute-bound workloads.
- Solve critical register pressure (SGPR/VGPR) and latency bottlenecks, implementing manual wave-level optimizations that increase SIMD occupancy and parallelism across the GPU Compute Units.

Senior Embedded Engineer II - R&D Graphics/AI/Compute *Texas, USA (Remote)*
Qubika *Mar 2025 – Jul 2025*

- Designed System Level Tests (SLT) for high-end APU silicon in order to validate error rate, performance and heat dissipation before production.
- Implemented low-level compute kernels using ISA specific assembly, directly managing hardware resources including Ray Tracing (RT) Cores, AI Cores, and the memory hierarchy (LDS/Caches).
- Drove significant performance gains in compute-bound workloads, enhancing throughput for AI and Graphics pipelines. These optimizations were deployed to the global production fleet, delivering measurable efficiency improvements to millions of consumer and enterprise GPUs.

Senior Software Engineer - R&D Video Mixing *Texas, USA (Remote)*
Vizrt Group *Aug 2021 – Mar 2025*

- Architected a high-performance compute pipeline to bridge Chromium Embedded Framework (CEF) and the core rendering engine. Designed a custom IPC mechanism for direct writing to shared GPU memory surfaces, effectively eliminating CPU-side buffer copies and drastically reducing latency for real-time HTML5 graphics.
- Solved critical synchronization anomalies in the real-time video cropping subsystem, diagnosing complex race conditions in the frame scheduler to ensure deterministic behavior and artifact-free broadcasting across the flagship product line (TriCaster TC1/TC2).
- Re-engineered the 2D asset virtualization pipeline (Adobe PSD), moving from simple rasterization to a vector-based implementation. Enabled live, non-destructive editing of high-resolution assets directly within the mixing engine without performance penalties.
- Led the architectural modernization of legacy subsystems, refactoring monolithic C++ workflows into modular components. This reduced technical debt and improved build/iteration times for the engineering team while maintaining strict backward compatibility for the existing user base.

Lead Rendering Engineer *Belo Horizonte, Brazil*
P0P Render Team *Jan 2017 – Jul 2021*

- Designed and led the development of an internal real-time DirectX / OpenGL rendering engine for Windows, utilizing high-performance, clean, and modern C++ in a multithreaded context.
- Conducted CPU and GPU profiling, debugging, optimizations, and inspections of code and native binaries.
- Performed academic research on advanced rendering techniques and subsequently implemented a high-performance rendering pipeline.

Undergraduate Researcher *São João Del-Rei, Brazil*
UFSJ *Feb 2019 – Dec 2019*

- Lead a comparative research on Bidirectional Reflectance Distribution Functions (BRDFs) for real-time rendering, evaluating their theoretical accuracy and implementation complexity.
- Assessed the computational complexity, memory requirements, and optimization challenges of BRDF implementations in real-time scenarios.

Software Engineer *Brazil*
Freelancer *Oct 2010 – Jan 2017*

Worked in a diverse range of products and projects, including government, academia, on-demand, open-source, blogs, and other types of initiatives. During this period, I utilized a broad spectrum of technologies and tools, such as Visual Basic, DirectX, OpenGL, C/C++, SQL, Flash and HTML5.

Education

UFSJ *Jan 2017 – Dec 2020*
BS in Computer Science

Skills

Languages and APIs: C++, C, CUDA, Asm (x86/Arm/ISA), HLSL, GLSL, HIP, Vulkan, DX12, Win32

Technologies and Tools: NVIDIA Nsight, RenderDoc, ROCm, Intel VTune, Ghidra, x64dbg