

Game Developers Versus Cloud Architects: What the PS5 Can Teach Us About Cloud Infrastructure

Nigel Dsouza

The majority of my time as a cloud architect is spent considering deployment automation, scalability, and resilience. However, my PS5 was the source of the most astounding systems I've recently seen—not AWS or Kubernetes.

With more stringent resource limitations and higher performance requirements than the majority of enterprise systems will ever encounter, the most recent generation of gaming consoles—and the developers pushing them to their limits—are producing real-time graphics and physics.

It made me wonder: could cloud architects learn from the AAA game development industry?

1. Latency Is Critical

You lose the game if there is a 50 ms delay. In the cloud? The majority of systems ignore 300 ms without blinking. However, latency is now becoming a significant limitation in domains like fintech and IoT.

Cloud engineers are just starting to use the immediacy-focused techniques that game developers design for, such as memory streaming, state prefetching, and predictive loading.

Consider thinking like a game engine if the cold start of your Lambda function is degrading user experience.

2. Limitations of Hardware Breed Innovation

Cloud-native engineers are pampered. Without blinking, we scale horizontally.

Developers of video games? They manage to fit 8-core CPUs and 16GB of RAM into a cinematic experience.

Asynchronous loading, dynamic LOD (level of detail), and asset compression all require extreme creativity. This discipline is frequently lacking in cloud computing.

Perhaps it's time to limit our clusters and push ourselves to become more intelligent rather than merely larger.

3. Retry Logic Is Outperformed by Real-Time Resilience

Games don't try again. They don't "gracefully degrade." They instantly and imperceptibly recover.

The global state of multiplayer game servers is synchronized in milliseconds—real-time coordination of voice communication, state persistence, and collision handling.

Dead-letter queues and exponential backoff are still issues in cloud systems. What if we implemented authoritative clients and state machines modeled after video games?

4. Delight Should Be a KPI, But It Isn't

The player experience is the lifeblood of game developers. Each controller vibration, particle effect, and animation frame is meticulously designed.

We rarely discuss delight in cloud architecture; instead, we discuss throughput.

However, what if observability, developer experience, or API ergonomics were viewed similarly to gameplay? Would we create more effective systems?

5. A Boss Fight Mentality Is Needed in Both Worlds

One thing is true whether you're creating a multi-region disaster recovery plan or a battle royale: stress exposes design flaws.

Game developers are always simulating chaos—input lag, network jitter, frame drops. Cloud architects also require fault injection, simulated brownouts, and chaos engineering.

In Conclusion

Cloud architecture and gaming may seem like two different worlds. However, they both address the same issue: human intervention, performance at scale, and under pressure.

The unsung infrastructure experts are game developers. We could create more immersive, robust, and human-centered systems—not just faster ones—if cloud engineers took the time to study them.

About the Author

Nigel Dsouza works for Fidelity Investments as a Technical Lead and Principal Software Engineer.

He reverse-engineers the reason why *Elden Ring* never crashes when he's not creating fault-tolerant cloud infrastructure. Nigel infuses enterprise systems with cloud discipline and innovative ideas from the interactive entertainment industry.