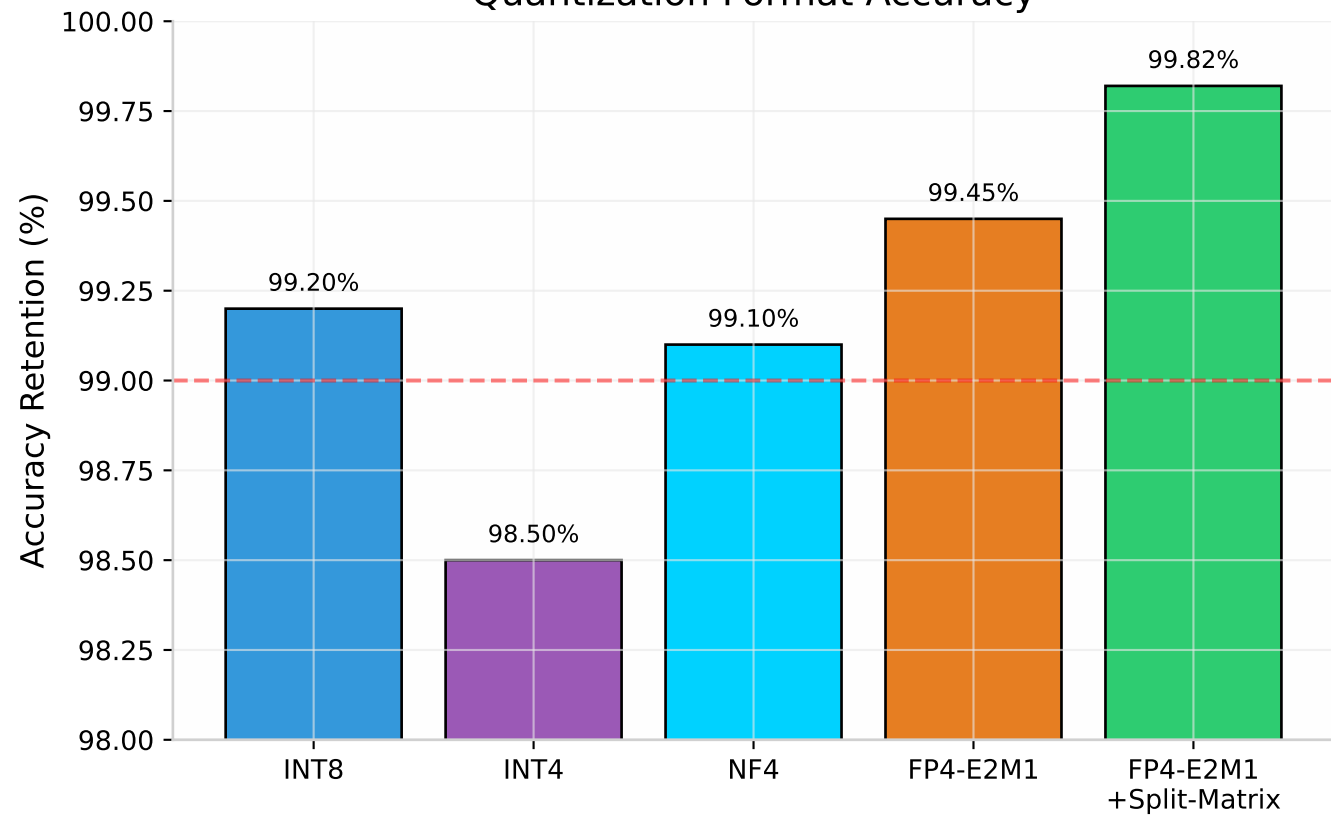
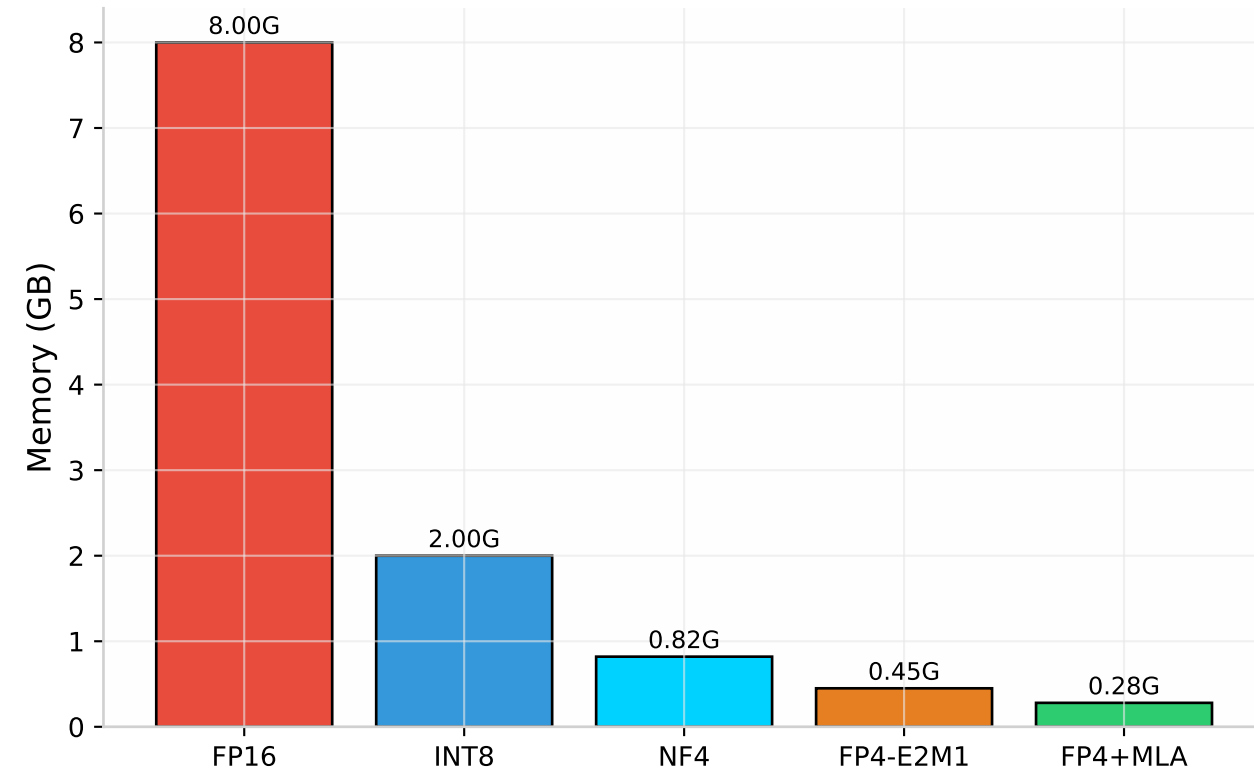


AGi v5.1 Advanced — FP4-E2M1 Benchmarks

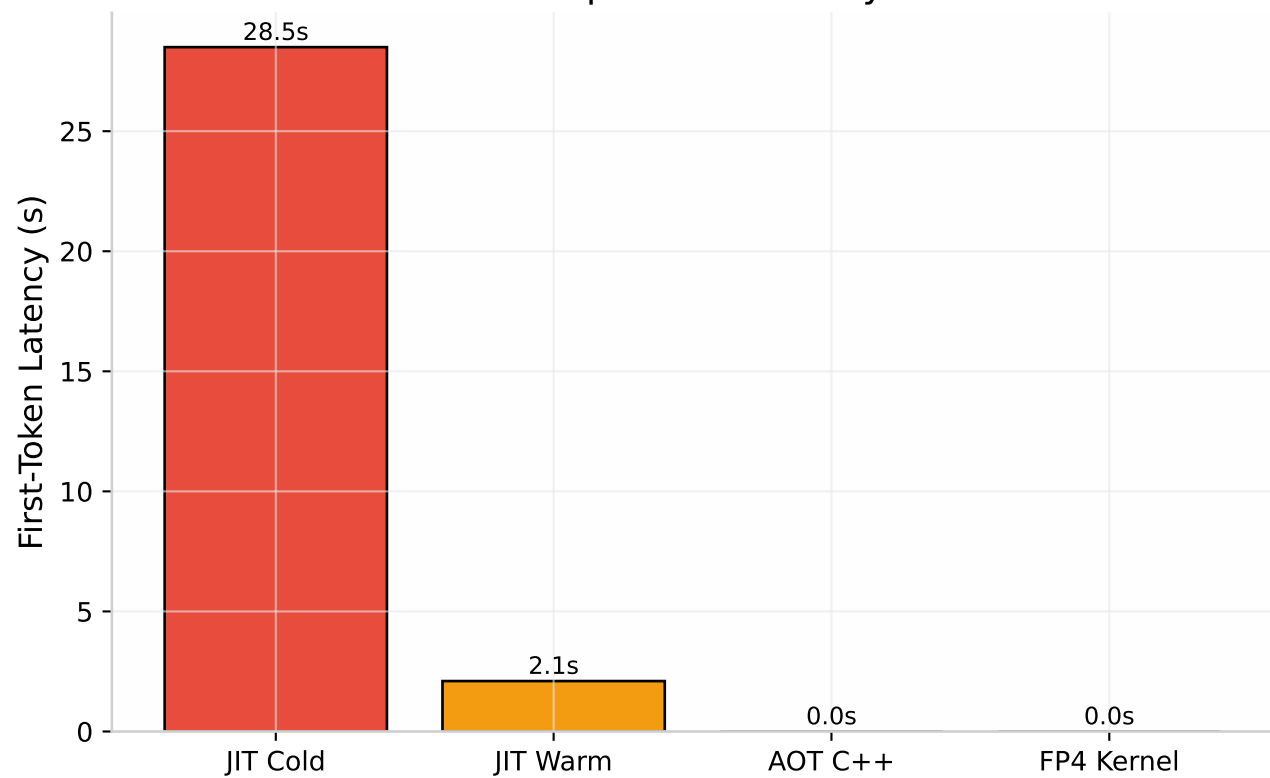
Quantization Format Accuracy



7B Model Memory Footprint



Compilation Latency



Inference Throughput

