Question-01: Investigating memory hierarchy in mobile and gaming devices

Many smartphones and gaming consoles achieve impressive performance despite tight power and thermal constraints. A key reason is their carefully designed memory hierarchies, which balance speed, capacity, energy efficiency and responsiveness.

Choose a real world platform – such as the Apple M1 SoC, Snapdragon 8 Gen 2, or PlayStation 5 — and analyze how its memory system is structured to support high-performance workloads like gaming, multitasking or video processing.

For example, Apple M1 integrates 128KB L1 caches, a 12-24 MB L2 chache, and up to 16 GB of LPDDR4X DRAM delivering ~68 GB/s bandwidth – all within a unified memory architecture shared by CPU and GPU. The PlayStation 5 pairs 16 GB of GDDR6 RAM (448 GB/s with a custom SSD capable of 5.5 GB/s throughput, streaming game assets in real time without preloading large datasets into DRAM).

In your report-

1. Describe the memory hierarchy of your selected device: registers, cache levels (with sizes and associvity if available), DRAM and storage.
2. Explain how the architecture exploits locality principles – for example, through unified memory, fast access, or streaming from SSD.
3. Analyze tradeoffs in speed, latency, energy use, and die area among SRAM, DRAM, and NAND.
4. Consider whether the CPU and GPU share memory, and how cache coherence or bandwidth sharing is handled.
5. Support your points with real specifications and performance data from valid sources.
6. Finaly reflect on how memory design decisions affect the device's responsiveness, energy efficiency, longevity and potential environmental impact.

Expected outcome:

1. Breakdown the memory hierarchy of a real system
2. Apply concepts such as temporal/spatial locality, latency hierarchy and cache trade-offs.
3. Use hardware specs to justify why certain memory types are used at different levels.
4. Think critically about how design choices affect both performance and long-term sustainability.