CIS 371
Computer Organization and Design

Unit 15: Putting It All Together:
Anatomy of the XBox 360 Game Console

This Unit: Putting It All Together

- Anatomy of a game console
- Microsoft XBox 360
- Focus mostly on CPU chip
- Briefly talk about system
  - Graphics processing unit (GPU)
  - I/O and other devices

Sources
- Application-customized CPU design: The Microsoft Xbox 360 CPU story, Brown, IBM, Dec 2005
- Microprocessor Report
  - IBM Speeds XBox 360 to Market, Krewell, Oct 31, 2005
  - Powering Next-Gen Game Consoles, Krewell, July 18, 2005

What is Computer Architecture?
The role of a **computer** architect:

"Technology"
Logic Gates
SRAM
DRAM
Circuit Techniques
Packaging
Magnetic Storage
Flash Memory

Goals
Function
Performance
Reliability

Plans
Design

Cost/Manufacturability
Energy Efficiency
Time to Market

Computer
PCs
Servers
PDAs
Mobile Phones
Supercomputers
Game Consoles
Embedded
Microsoft XBox Game Console History

- **XBox**
  - First game console by Microsoft, released in 2001, $299
  - Glorified PC
    - 733 Mhz x86 Intel CPU, 64MB DRAM, NVIDIA GPU (graphics)
    - Ran modified version of Windows OS
  - ~25 million sold

- **XBox 360**
  - Second generation, released in 2005, $299-$399
  - All-new custom hardware
    - 3.2 Ghz PowerPC IBM processor (custom design for XBox 360)
    - ATI graphics chip (custom design for XBox 360)
  - 19+ million sold

Microsoft Turns to IBM for XBox 360

- Microsoft is mostly a software company
  - Turned to IBM & ATI for XBox 360 design
  - Sony & Nintendo also turned to IBM (for PS3 & Wii, respectively)

- Design principles of XBox 360 [Andrews & Baker]
  - Value for 5-7 years
    - \( \rightarrow \) big performance increase over last generation
  - Support anti-aliased high-definition video (720*1280*4 @ 30+ fps)
    - \( \rightarrow \) extremely high pixel fill rate (goal: 100+ million pixels/s)
  - Flexible to suit dynamic range of games
    - \( \rightarrow \) balance hardware, homogenous resources
  - Programmability (easy to program)
    - \( \rightarrow \) listened to software developers (quote)

More on Games Workload

- Graphics, graphics, graphics
  - Special highly-parallel graphics processing unit (GPU)
  - Much like on PCs today

- But general-purpose, too
  - "The high-level game code is generally a database management problem, with plenty of object-oriented code and pointer manipulation. Such a workload needs a large L2 and high integer performance." [Andrews & Baker]

- Wanted only a modest number of modest, fast cores
  - Not one big core
  - Not dozens of small cores (leave that to the GPU)
  - Quote from Seymour Cray

XBox 360 System

![XBox 360 System Diagram](image)
XBox 360 “Xenon” Processor

- ISA: 64-bit PowerPC chip
  - RISC ISA
    - Like MIPS, but with condition codes
    - Fixed-length 32-bit instructions
    - 32 64-bit general purpose registers (GPRs)
  - ISA++: Extended with VMX-128 operations
    - 128 registers, 128-bits each
    - Packed “vector” operations
    - Example: four 32-bit floating point numbers
      - One instruction: VR1 * VR2 → VR3
      - Four single-precision operations
    - Also supports conversion to MS DirectX data formats
    - Similar to Altivec (and Intel’s MMX, SSE, SSE2, etc.)
    - Works great for 3D graphics kernels and compression

- Peak performance: ~75 gigaflops
  - Gigaflop = 1 billion floating points operations per second

- Pipelined Superscalar processor
  - 3.2 Ghz operation
  - Superscalar: two-way issue
  - VMX-128 instructions (four single-precision operations at a time)
  - Hardware multithreading: two threads per processor
  - Three processor cores per chip

- Result:
  - 3.2 * 2 * 4 * 3 = ~77 gigaflops

XBox 360 “Xenon” Chip (IBM)

- 165 million transistors
  - IBM’s 90nm process
- Three cores
  - 3.2 Ghz
  - Two-way superscalar
  - Two-way multithreaded

“Xenon” Processor Pipeline
XBox 360 Memory Hierarchy

- 128B cache blocks throughout
- 32KB 2-way set-associative instruction cache (per core)
- 32KB 4-way set-associative data cache (per core)
  - Write-through, lots of store buffering
- 1MB 8-way set-associative second-level cache (per chip)
  - Special "skip L2" prefetch instruction
  - MESI cache coherence
- 512MB GDDR3 DRAM, dual memory controllers
  - Total of 22.4 GB/s of memory bandwidth
- Direct path to GPU (not supported in current PCs)
Graphics “Parent” Die (ATI)

- 232 million transistors
- 500 MHz
- 48 unified shader ALUs

GPU “daughter” die (NEC)

- 100 million transistors
- 10MB eDRAM
  - “Embedded”
- NEC Electronics
- Anti-aliasing
  - Render at 4x resolution, then sample
- Z-buffering
  - Track the “depth” of pixels
- 256GB/s internal bandwidth

Putting It All Together

- Unit 0: Abstraction and design goals
- Unit 1: ISAs
- Unit 2: Single cycle
- Unit 3: Performance
- Unit 4/5: Integer/floating point ALU
- Unit 6/7: Pipelining & Superscalar
- Unit 8: Static & Dynamic Scheduling
- Unit 9: Caches
- Unit 10: Shared memory & multicore
- Unit 11: Virtual memory and I/O
- Unit 12/13/14: Reliability/power/vectors