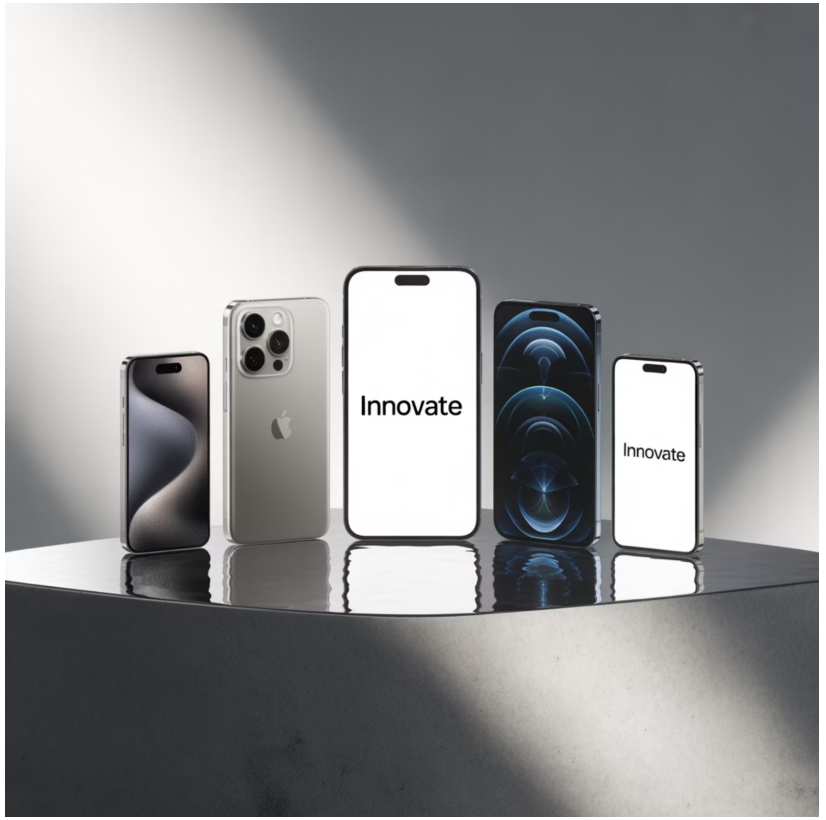
Five Mainstream Computing Classes

Class	System Price	Processor Price	Critical Design Issues
Personal Mobile Device (PMD)	$100-$1000	$10-$100	Cost, energy, media performance, responsiveness
Desktop	$300-$2500	$50-$500	Price-performance, energy, graphics performance
Server	$5K-$10M	$200-$2000	Throughput, availability, scalability, energy
Clusters/Warehouse Scale	$100M-$200M	$50-$250	Price-performance, throughput, energy proportionality
Embedded	$10-$100K	$0.01-$100	Price, energy, application-specific performance

Personal Mobile Devices (PMDs)

- PMDs include cell phones, tablet computers, and other wireless devices with multimedia interfaces. With consumer prices of a few hundred dollars, cost is a prime concern.
  - Energy efficiency driven by battery use and cooling limitations
  - Flash memory instead of magnetic disks due to energy and size requirements
  - Responsiveness and predictability critical for media applications
  - Real-time performance requirements for video and audio processing
  - Emphasis on code size optimization to minimize memory costs



Desktop and Server

Desktop Computing

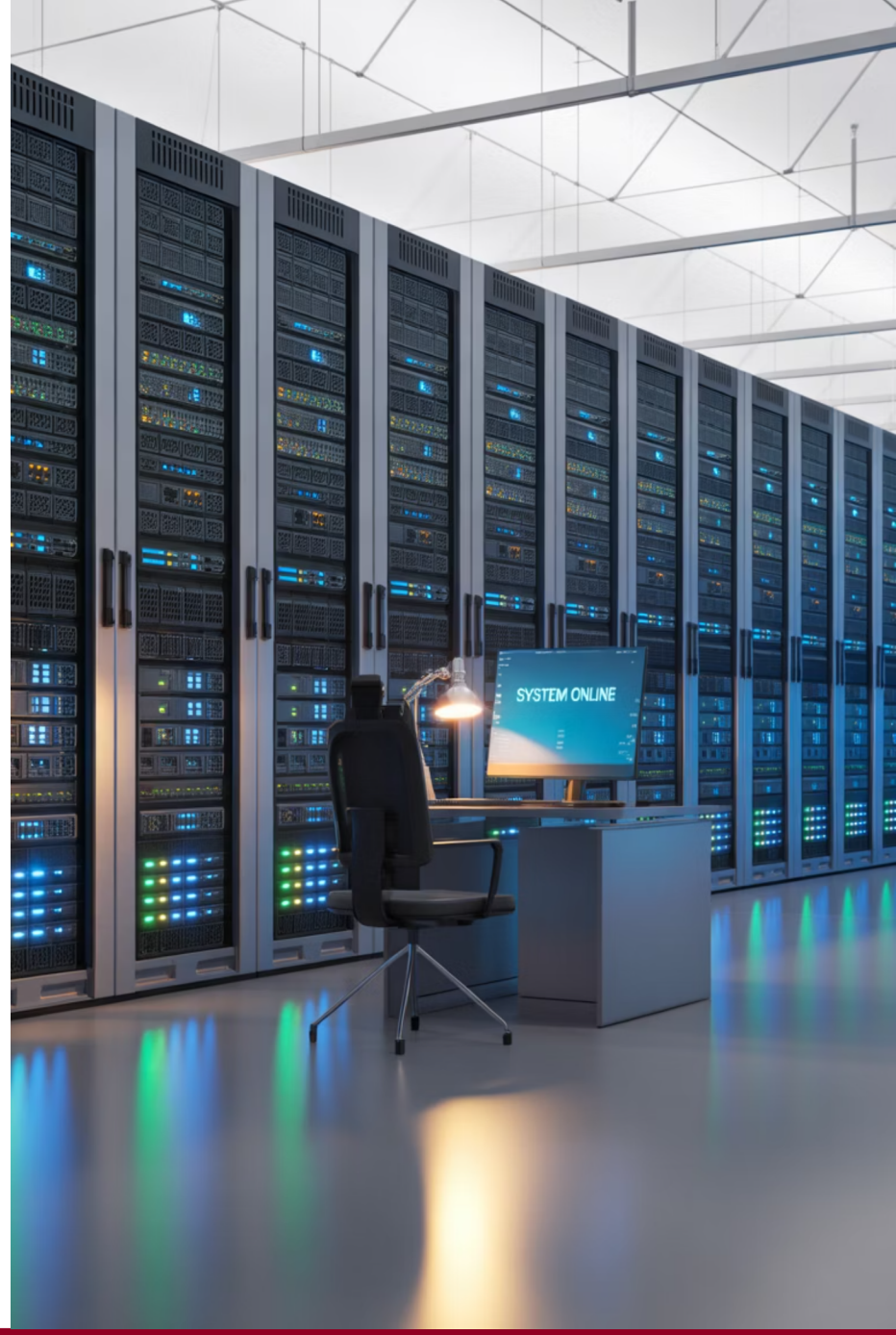
Desktop computing spans from low-end netbooks to high-end workstations, with price-performance as the primary driver.

- Newest microprocessors often appear first in desktops
- Increasingly focused on web-centric, interactive applications
- 350 million desktop PCs sold in 2010

Server Computing

Servers provide large-scale, reliable file and computing services as the backbone of enterprise computing.

- Availability is critical - downtime can cost millions
- Scalability to grow with increasing demand
- Efficiency measured by throughput (transactions/minute)
- 20 million servers sold in 2010



Clusters and Warehouse-Scale Computers (Cloud)

- The growth of Software as a Service (SaaS) for applications like search, social networking, and online shopping has led to the development of clusters and warehouse-scale computers (WSCs).
  - Collections of computers connected by local area networks
  - Each node runs its own operating system
  - WSCs designed for tens of thousands of servers acting as one
  - Price-performance and power efficiency are critical
  - 80% of a $90M warehouse cost is for power and cooling
  - Uses redundant inexpensive components with software handling failures

Embedded Computing

Diverse Processing Power

From 8-bit processors costing less than a dime to network switches at $100 executing billions of instructions per second.

Price Sensitivity

Meeting performance needs at minimum price is the primary goal, rather than achieving higher performance at higher cost.

Ubiquitous Presence

Found in microwaves, washing machines, printers, networking switches, and all cars. Nearly 19 billion embedded processors sold in 2010.

Unlike PMDs, embedded computers typically cannot run third-party software and have more limited hardware and software sophistication. This distinction separates embedded from non-embedded computers.

Classes of Parallelism

Instruction-Level Parallelism

Exploits data-level parallelism at modest levels with compiler help using pipelining and at medium levels using speculative execution.

Vector Architectures and GPUs

Exploit data-level parallelism by applying a single instruction to a collection of data in parallel (SIMD).

Thread-Level Parallelism

Exploits either data-level or task-level parallelism in a tightly coupled hardware model allowing interaction among parallel threads (MIMD).

Request-Level Parallelism

Exploits parallelism among largely decoupled tasks specified by the programmer or operating system (loosely coupled MIMD).

The challenge for the 21st century is the restructuring of applications to exploit explicit parallelism. In some instances, this is easy; in many, it is a major new burden for programmers.