Why Multiprocessors?

Moore’s Law predicted a doubling of processor performance every couple of years
• true until about 2000

Limits on the performance of a single processor: what are they?

Why Multiprocessors

Utilizes coarser granularities than ILP
Lots of workload opportunity
• Scientific computing/supercomputing
  • Examples: weather simulation, aerodynamics, protein folding
  • Each processor computes for a part of the grid
• Server workloads
  • Example: airline reservation database
  • Many concurrent updates, searches, lookups, queries
  • Processors handle different requests
• Media workloads
  • Processors compress/decompress different parts of image/frames
• Desktop workloads …
• Gaming workloads …

What would you do with a billion transistors? Or more?
Multiprocessors

Low-end

- bus-based
  - simple, but a bottleneck
  - broadcast-based cache coherency protocol
- physically centralized memory
- uniform memory access (UMA machine)
- most of today’s small CMPs (Intel Core 2 Quad, AMD Quad-Core Operon “Barcelona”, SunFire (16))
**Multiprocessors**

**High-end**
- multiple-path interconnect
  - higher bandwidth
  - longer memory latencies
  - more scalable
  - point-to-point cache coherency protocol
- physically distributed memory
- non-uniform memory access (NUMA machine)
- could have processor clusters
- SGI Origin, AMD HyperTransport, Cray T3D, IBM SP-2, Intel Paragon

![High-end MP](image-url)
Comparison of Issue Capabilities

Superscalar horizontal waste

Single-chip Multiprocessor

vertical waste

Thread 1
Thread 2
Thread 3
Thread 4
Thread 5

Spring 2011
CSE 471 - Multiprocessors