In this module we learn about thread-level parallelism (TLP), a method processors use to increase the usage of pipeline hardware when a single program cannot fully use it due to data dependencies or because it is waiting for memory. The idea is that the same processor hardware can operate on multiple programs or sub-programs (threads). When one thread cannot use all of the hardware, it can be used by other threads. In this module we'll focus on a technique called simultaneous multithreading (SMT).

By the end of this module you will be able to:

- Describe how processor hardware can be shared among threads.
- Demonstrate how SMT can increase the use of pipeline resources in the presence of data dependences.
- Discuss how instructions from multiple threads can share a single register file and issue queue.

INTRODUCING TLP AND SMT

5:23 / 5:23 1.0x

Download transcript .txt

Show Discussion

New Post

Help