摘要

许多现代计算机系统,包括同构架构(homogeneous architecture)和异构架构(heterogeneous architecture),在硬件上都支持共享内存(shared memory)。在共享内存系统中,每个处理器核心都能对一个共同的地址空间进行读写。而 consistency model 就定义了共享内存机器中的,内存系统中可见的体系结构行为。

consistency 定义了有关加载(load)和存储(store)(或者内存读写(memory read/write))操作的规则,以及它们是如何作用到内存中的。作为 consistency 模型的一部分,许多机器还提供了 cache coherence protocol,以确保数据的多个缓存副本(cached copies)能够同时保持最新状态。

本导论的目的就在于,让读者对 consistency 和 coherence 获得一个基本的了解。其中既包括必须解决的问题,也包括各种问题的解决方案。在提供高层次抽象概念的同时,我们也会提供来自现实世界系统中的具体例子。

本书的第二版反映了自第一版以来,十年间发生的进展。除较小的改动外,第二版还增加了两个新章节:一个关于非CPU加速器的 consistency 和 coherence(重点在于GPU),另一个则关于和 consistency 和 coherence 相关的形式化(formalization)工作和工具。

关键词

computer architecture(计算机体系结构), memory consistency, cache coherence, shared memory (共享内存), memory systems(内存系统), multicore processor(多核处理器), heterogeneous architecture(异构架构), GPU(图形处理单元), accelerators(加速器), semantics(语义学), verification(验证)