第 36 卷 第 4 期 计 算 机 工 程 2010 年 2 月
Vol.36 No.4 Computer Engineering February 2010
·博士论文· 文章编号：1000—3428(2010)04—0004—03 文献标识码：A 中图分类号：TP333

多核处理器片上存储系统研究

黄安文，高 军，张民选
(国防科技大学计算机学院并行与分布处理国家重点实验室，长沙 410073)

摘 要：针对多核处理器计算能力和访存速度间差异不断增大对多核系统性能提升的制约问题，分析几款典型多核处理器存储系统的设计特点，探讨多核处理器片上存储系统发展的关键技术，包括延迟造成的非一致 cache 访问、核与 cache 互连形式对访存性能的束缚以及片上 cache 设计的复杂化等。

关键词：多核；存储系统；非一致 cache 访问

Research on On-chip Memory System of Multi-core Processor

HUANG An-wen, GAO Jun, ZHANG Min-xuan
(National Key Laboratory for Parallel and Distributed Processing, School of Computer Science,
National University of Defense Technology, Changsha 410073)

【Abstract】 Aiming at the problem that memory system on chip becomes a bottleneck of improving the performance of multi-core processor as the speed distinction between CPU and memory increases. This paper analyses the design characters of memory system in multi-core processor, such as Non-Uniform Cache Access(NUCA) caused by delay, constraint to access performance of connection between core and cache, and complexity of on-chip cache design.

【Key words】 multi-core; memory system; Non-Uniform Cache Access(NUCA)

1 概述

在多核处理器中能够方便地进行任务划分和有效的线程调度，从而很好地满足开发线程级并行性的需求。与传统超标量和超长指令字结构的单核处理器相比，多核处理器已经在 Web 服务和联机事务处理等线程级并行性较高的商业服务领域表现出明显的竞争优势。

多核处理器中数据处理速度和存储器访问速度之间的不匹配会导致多核资源利用的不平衡，阻碍多核系统吞吐率的提高。本文从目前国内外多核处理器研究现状出发，分析几款典型的多核处理器片上存储系统的设计特点，对多核处理器片上访存系统的研究热点和面临挑战进行分析和讨论。

2 多核处理器研究现状

国外研究机构和知名商业公司对片上多核处理器的研究起步较早，许多公司和科研单位都已经有成熟的商用多核处理器和原型芯片问世。

Standford 大学早在 1996 年就设计了 CMP 结构的 Hydra 处理器[1]。它在片内集成 4 个 MIPS 内核，单核私有哈佛结构的一级 cache，4 核共享 L2-cache，L3-cache 置于片外，主存与最低一级的 L3-cache 采用 128 位总线连接。

2007 年，MIT 和 Tilera 公司合作研制了一款基于 tile 体系结构的片内集成 64 个核的多核处理器——Tile64，使用二维网格(Mesh)将 64 个独立的处理器核挂接在网络节点上，核间通信和数据路由部件进行了专门设计以提高数据通信能力和访存带宽，在网络服务、数字视频和通信等计算密集型应用领域能够取得非常好的性能。

Sun 公司着眼于当前线程级并行度越来越高的商业应用程序和工作集负载大的服务器应用，先后推出了 UltraSPARC T1 和 T2 处理器。此外，ROCK 处理器也是 Sun 公司的典型产品代表。IBM, Intel, AMD 和 Fujitsu 等公司也已推出自己的商用多核处理器并在市场上占据相当份额。

我国具有独立自主知识产权的多核处理器研究技术发展相对缓慢，与国外相比还有较大的差距，多家研究单位对多核处理器体系结构的研究和实验已经处于积极开展之中[2]。

3 典型的多核处理器片上存储系统

3.1 UltraSPARC T2 和 ROCK

在 Sun UltraSPARC T2[3]处理器中，单核内设计了 5.7 KB 的整数寄存器文件和 0.3 KB 的浮点寄存器文件，每核私有独立的 16 KB 一级指令 cache 和 8 KB 的数据 cache。片上集成了 4 MB 的 L2-cache，划分为 8 个 Bank，16 路组相联，采用伪随机替换策略，L2-cache 访问采用 9 栈流水。另外，片内集成了 4 个存储访问控制单元来处理 L2-cache 与片外 DRAM 之间的数据交互。

为了提高多个处理器核对 L2-cache 的访问效率，T2 设计了专门的 Crossbar 结构将 8 个 SPARC 内核与 L2-cache 的 8 个 Bank 以及 I/O 访问单元相连。从数据源发出的请求信号在 Crossbar 中排队后发送至目的单元，由于多个数据源可能访问同一个目的单元，因此为了解决访问冲突，在 Crossbar 中集成了一个仲裁部件，按照访问次序对请求进行排队以保

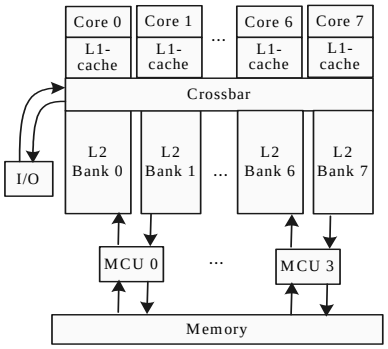
基金项目：国家自然科学基金资助项目(60873016)；国家“863”计划基金资助项目(2008AA01Z147, 2007AA01Z102)

作者简介：黄安文(1983－)，男，博士研究生，主研方向：微处理器体系结构；高 军，讲师、博士研究生；张民选，教授、博士生导师

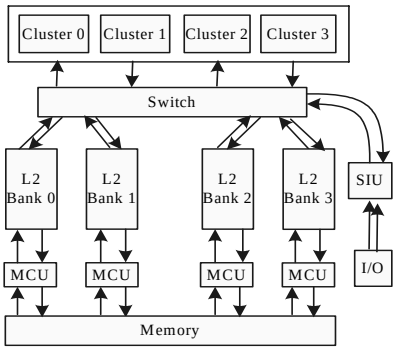
收稿日期：2009-08-22 **E-mail：**anwenhuang@hotmail.com

证公平性和顺序性。随着片上核数和 L2-cache Bank 数目的增加，Crossbar 的规模迅速膨胀，多核对 cache 的访问延迟随之增大。Crossbar 结构面临的可扩展性较差问题给多核访问 L2-cache 造成了极大挑战。

在进行 ROCK 处理器的设计时，Sun 公司对其交叉开关结构进行了调整，如图 1 所示。ROCK 单片集成了 16 个处理器核，采用分组/分层的思想将每 4 核构成一个 Cluster，L2-cache 划分为 4 个 Bank，通过 switch 与 4 个 Cluster 相连，开关阵列的规模缩小为 4×4。另外，片上集成 4 个存储控制单元，负责 L2-cache 和外存之间的数据交互。



(a)T2



(b)ROCK

图 1　UltraSPARC T2 及 ROCK 片上存储结构对比

3.2　基于 PIM 体系结构的 VIRAM

在传统的体系结构设计中，处理器和存储单元(内存)通常在独立的芯片上实现，系统性能受限于两者速度的差异和远程存储器访问延迟。PIM(Processor-In-Memory)技术则打破了传统的设计思路，为了缩小 CPU 和存储器之间的速度差异，充分利用当前超大规模集成电路工艺上的优势，把处理器核与 DRAM 集成在同一个芯片上，以弥补两者之间的速度差异，为处理器访问存储器提供足够带宽，减小了访存延迟。

VIRAM 实际上也可以看作一款单片集成 MIPS 核与向量协处理器的异构双核处理器，该芯片由加州大学伯克利分校开发[4]，如图 2 所示。VIRAM 基于 Processor-in-Memory 的思想，将一个单发射的 MIPS 核与一个 256 位的向量协处理器集成在同一片内，MIPS 核具有私有的直接映射一级数据 cache 和指令 cache，大小均为 8 KB，向量协处理器可以根据不同应用划分为若干向量处理单元，为向量协处理器专门设计了 8 KB 的向量寄存器文件。它没有为片上处理器核设计相应的二级 cache，却将大容量 DRAM 划分为 8 个 Bank 集成在片上，MIPS 核与向量协处理器通过 Memory Crossbar 与片上 DRAM 相连，负责与外部通信的 I/O 单元也挂接在 Crossbar 上。事实上，VIRAM 可以像使用主存一样使用片上 DRAM 存储器，DRAM 已经取代了大容量 cache 的作用。

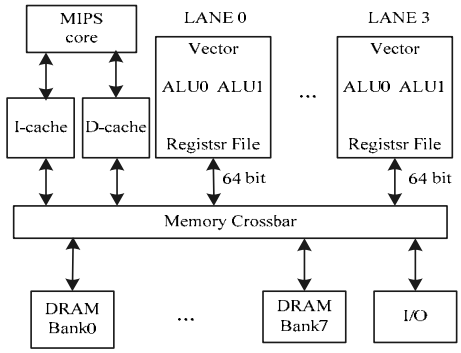


图 2　VIRAM 片上存储系统及互连

3.3　Larrabee

Larrabee 处理器是 Intel 公司迈入众核领域的里程碑标志，它打破了以往“酷睿”的设计思路，在体系结构设计上专门针对高性能并行处理进行了优化[5]。Larrabee 将 8 个核集成在片上，并以 Ring 环形总线的形式进行互连以实现高速数据通信，单核内划分为标量和向量处理 2 个部分，每核私有 32 KB 的指令 cache 和数据 cache，为实现运算部件与存储单元的低延迟数据交互提供了保证。

在 Larrabee 结构中设置了容量为 4 MB 的 L2-cache，根据片上多核的数目划分为相应大小的子集，各核访问自己的本地子集，各个子集通过 Ring 双向环与处理器核相连。CPU 核能够利用 Ring 环进行高速的本地存储访问，并以此来简化数据共享和同步机制的实现，在进行 L2-cache 访问时提供非常高的带宽。得益于 Larrabee 的 Ring 环结构，片上多核的不同 L2-cache 子集间的数据访问变得更加简单。L2-cache 的一致性也可以通过 Ring 环结构得到保证。

4　片上多核访存系统研究热点及面临的挑战

集成电路工艺水平和体系结构设计能力的不断提高，为研究面向不同应用领域的高性能、低功耗、低延迟多核处理器提供了技术保障，业界和研究领域不断涌现出结构各异的新型片上多核系统。多核系统的结构复杂化和规模扩大化趋势使访存系统的设计与优化面临许多前所未有的挑战，主要包括：

(1)线延迟造成的 NUCA 问题日益突出

在多核系统中，各个处理器经常需要对片上集成的大容量共享 cache 进行读写访问来实现核间通信和交互。由于 cache 容量不断增大，传统的 H-tree 结构已经不能满足访存的延迟和时序要求。为了减小存储访问延迟，在多核共享大容量 cache 的结构中一般根据具体需要将 cache 划分为多个 Bank 分布在片上，各个 Bank 与多核之间通过内部互连网络相连接，由于 cache 容量较大，加之内部互连结构复杂度和规模的影响，因此线延迟成了影响 cache 访问的主要因素，核对 cache 中各个 Bank 之间的访问请求会因在互连结构上移动距离长度不同而不同，从而造成非一致 cache 访问现象(Non-Uniform Cache Access, NUCA)。随着片内可集成 cache 规模的不断增大，NUCA 问题的解决将变得更加棘手。为了更好地解决 NUCA 问题，还需要对片上存储层次间的组织形式、数据迁移策略等细节问题进行分析和处理[6]。

(2)核与 cache 互连形式对访存性能的束缚更加明显

在片上集成大容量 cache 的多核结构中，特别是目前常用的多 Bank 结构中，处理器核对 cache 的访问速度很大程度

上依赖于两者之间的互连形式，多核与 cache Bank 之间的互连将成为制约片上存储系统访问速度和带宽的主要因素之一。无论是 Sun UltraSPARC T2 中的 Crossbar 结构还是 Intel Larrabee 中采用的 Ring 结构，都是为了解决 core 对 cache 的高效访问而专门设计的。在进行核与 cache 或者核与多个 cache Bank 体之间互连设计时，需要在存储访问速度与带宽之间作出权衡。总线形式、交叉开关以及网格都是普遍采用的方式，三者的特点对比如表 1 所示。由于核与片上存储资源之间的互连形式很大程度上影响了多核处理器访存性能的优化，设计者需要对两者之间的互连形式进行优化，以满足低延迟、高带宽、低功耗的存储访问需求。随着片内核数与片上存储容量的增大，核与存储单元之间的互连形式还呈现出规模扩大化、形式多样化、结构复杂化的特点，需要根据具体体系结构并针对特定的应用对其进行优化设计以缓解片上存储访问的延迟问题。

表 1　核与片上存储单元互连形式的分析与对比

形式	典型例子	优缺点分析
总线	Larrabee (Intel)	拓扑结构简单，功耗开销较小； 分时复用会造成资源竞争，访问延迟较大
交叉开关	T1, T2, ROCK (Sun)	吞吐率较高，速度较快，可同时满足多核对多个 cache Bank 的访问请求 核数与 cache 体数目增加时会迅速影响开关阵列规模，功耗和面积开销较大
Mesh 网格	Trips (Texas 大学)	拓扑简单，可扩展性好 网络直径增大时传输延迟挑战会变得严峻

(3)片上 cache 设计更加复杂化

cache 能够充分利用指令和数据的局部性特征，将所需的指令和数据放在距离处理器最近的位置来减小访存延迟。目前的体系结构设计能力和工艺水平已经能够将大容量 cache 甚至 DRAM(如 VIRAM 处理器)集成至片内以缓解访存带宽和速度对系统性能的影响。随着多核与 cache 组织形式的多样化，片内 cache 在设计时面临许多新的问题：1)cache 一致性协议的维护变得更加复杂，尤其是引入多线程机制以后，不仅要考虑单核内的数据相关，还需要考虑多核间数据交互时造成的相关，给 cache 的一致性维护带来较大困难。2)其次，cache 失效处理带来的延迟问题更加严重。由于多核间的数据交互和通信是通过共享 cache 完成的，因此多核系统中 Cache 失效处理机制变得更加复杂，特别是引入多线程思想以后，处理 cache 失效所需的时间代价不容忽视。另外，多核 cache 层次的组织形式、私有/共享机制的选择、替换策略、划分机制都有可能根据具体的多核结构和访存特性进行相应调整，以便在低访问延迟和高命中率之间进行权衡与折中。

(4)模拟器在访存系统设计中的作用不容忽视

由于多核条件下片上访存系统设计复杂性越来越高，因此需要功能更加强大的模拟器进行功能和时序模拟。目前，除了对传统的 SimpleScalar 模拟器进行扩展使其能够支持多核系统的仿真模拟外，已经出现了若干具有发展潜力的针对多核系统的模拟仿真平台。Wisconsin 大学开发了一款名为 GEMS(General Execution-driven Multiprocessor Simulator)的多核模拟器[7]，结构如图 3 所示。它的核心模块是存储系统模型 Ruby，能够模拟 cache 及 cache 控制器、主存及存储控制器、互连网络等模块的时序特性，在进行多处理器存储系统模拟时，Random Testor, MicroBenchmark 和 Opal 以及 Simics 等驱动源会发出对 Ruby 的访存请求并与之交互。

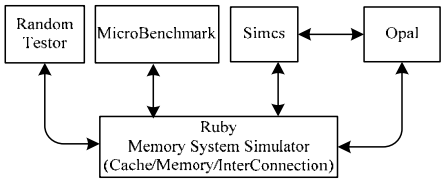


图 3　GEMS 结构

SimOS 是另一款具有代表性的多处理器系统模拟器，由 Standford 大学开发。由于该处理器采用执行驱动模拟的方式，对目标机的模拟性能依赖于主机的硬件结构，因此可移植性有待进一步提高。对于片上 cache，在以前单核条件下经常借助 CACTI(Cache Access and Cycle Time Information)工具进行性能模拟和评估。该工具的最新版本 CACTI 6.0 已经针对多核共享大容量 cache 出现的 NUCA 等情况进行了扩展和改进，能够较好地模拟多核情况下非一致 cache 访问的情况。

随着片上多核系统设计的复杂化，开发功能更加强大、模拟速度更快、精度更高的多核模拟器并完善相应的访存模拟工具，必将为研究新型多核体系结构、优化访存系统性能提供便利。

5　结束语

不断提高的微体系结构设计能力和迅猛发展的集成电路工艺水平为多核处理器的发展提供了技术保证，片上多核系统作为未来商用微处理器和高性能应用领域的主要发展方向之一，已经表现出很强的竞争力。随着片上多核计算能力和访存速度之间矛盾的加剧，迫切要求研究界和工业界寻找行之有效的方法来弥补两者之间的速度差异。如何根据具体应用对多核处理器片上存储系统进行设计和优化，将成为多核体系结构设计领域必须考虑的重要问题。

参考文献

[1] Hammond L. The Standford Hydra[J]. IEEE Micro, 2000, 20(2): 71-84.

[2] 何　军, 王　飙. 多核处理器的结构设计研究[J]. 计算机工程, 2007, 33(16): 208-210.

[3] Sun Microsystems, Inc.. OpenSPARC T2 Core Microarchitecture Specification[Z]. 2007.

[4] Nguyen T P Q, Zakhor A, Yelick K. Performance Analysis of an H.263 Video Encoder for VIRAM[C]//Proc. of IEEE International Conference on Image Processing. [S. l.]: IEEE Press, 2000: 98-101.

[5] Seiler L, Carmean D, Sprangle E, et al. Larrabee: A Many-core x86 Architecture for Visual Computing[J]. ACM Transactions on Graphics, 2008, 27(3): 18-26.

[6] Muralimanohar N, Balasubramonian R, Jouppi NP. Architecting Efficient Interconnects for Large Caches with CACTI 6.0[J]. IEEE Micro, 2008, 28(1): 69-79.

[7] Martin M M K, Sorin D J, Beckmann B M, et al. Multifacet's General Execution-driven Multiprocessor Simulator(GEMS) Toolset[J]. ACM SEGARCH Computer Architecture News, 2005, 33(4): 92-99.

编辑　张正兴