**计算机组成原理**

**论文审阅**

学院：计算机科学与技术学院

班级：11403202

学号：1140320206

姓名：霍峻杰

1. **论文审阅所选篇目：**

A Fully Associative, Tagless DRAM Cache

(全相联无标记DRAM缓存)

1. **论文框架结构概述：**

首先，论文进行一个简短的摘要概括，向读者讲述了基本的论文主题及其意义所在。接着，开始进入到正文。正文分为七个主要的部分。

第一部分是一个介绍部分，主要讲述层叠芯片DRAM的当前发展情况及未来前景，指出目前此种TLB和cache tag组合的不足之处，并较概略地引出了自己设计的cache map TLB（cTLB），并在其中，将二者的一些性能（命中率，延迟等）进行比较，使读者对此现有一个了解。

第二部分是一个对背景知识和设计动机的说明。向读者介绍了一些基本的知识，便于对这篇论文的理解有很大的帮助。先介绍了层叠芯片DRAM，其中介绍了两种创造快速封装DRAM的方法，一种是软件管理快速主存，另一种是软件透明最后一级缓存。比较了两种方法的优劣性后（后一种节约空间，使用tag），给出了基于层叠芯片SRAM标签的基线页缓存的访问路径（为了和下面的无标签访问路径进行对比）。分析了这种页缓存的缓存页开销，可见这种开销还是比较大的，并且随着DRAM的发展，这种开销会越来越大。这样就引出了要论述的第三个部分。

第三部分讲的就是论文的主体，作者通过前面的铺垫引出了自己的设计理念。首先，将自己的设计进行一个概述，包括如何实现和开销分析。作者给出自己设计的无标签缓存访问路径。通过上面的对比，我们可以看出，大部分是相同的，而不同的地方在于后者比前者少了一块SRAM标签。实现的思想是将原来的虚拟地址到物理地址，再由物理地址通过标签找到缓存这两部分结合为一个，变为由虚拟地址直接找到缓存。中间不需要标签的参与，这样就可以忽略标签的开销，从而达到无标签的设计思想。接着，作者又讲了支持这种设计一些基本结构支撑，包括页表，cTLB，全局转化页表和空闲队列。介绍完基本的思想，实现流程，和需要的支承结构组件以后，作者给出了一个简单的用例方案进一步说明无标签缓存，然后给出了一个缓存命中和未命中延迟的两者对比分析表，使读者直观地看出二者的效率差别，我们可以看到，在命中时，cTLB占用的时间会更短一些，但是在未命中时，其占用时间会稍长，途中也明确标出了具体的原因。接下来，作者根据不同的请求，列出了一个对比三种方法（基于块，基于页，无标签）的表格，从中，我们可以发现大多数请求的情况下，基于块的方法效果很差，而无标签方法在大多数情况下都展现出很好的效果。总体来看，三种方法的优秀程度为无标签>基于页>基于块。最后，作者对自己的设计进行了一个概述，列出了一个全相联无标签缓存的组成部件及其一些参数。

第四部分进行了实验装置的说明。包括了估计实验结果和效率的工具McSimA+ simulator，和进行组装的基本部件（L3 缓存，堆交叉，SRAM标签，cTLB等）。

第五部分对实验结果进行了评估。首先进行单程序的，比较了指定的SPEC CPU 2006程序的归一化IPC和归一化EDP，并列出了表格数据对比。接下来，给出了SRAM标签和cTLB的平均L3访问延迟的对比图，我们可以看到，cTLB的延迟在任何的几何平均值中，都要小于SRAM标签。进行完单程序的评估，作者为了证明其全面的正确性，还进行了多程序的评估，对比了IPC和EDP，给出了数据对比图。接下来，给出了DRAM缓存大小敏感度对SRAM标签和无标签缓存的影响对比图，我们可以发现，cTLB变化是比较剧烈的。然后，为了进一步说明，作者又进行了多线程的测试，同时也给出了无缓存形式的潜在益处。

第六部分是进行一些其他的讨论，例如，共享页的支持，超级页的支持，多套接字的支持和连接。

第七部分主要讲了一些和本论文相关的科研领域。这些中，可能会用到本文所提出的设计，如封装DRAM缓存，混杂存储系统。

第八部分对论文进行了一个总结，总结自己的设计和发展前景。

最后鸣谢相关指导人员并列出了参考书籍。

1. **论文主体设计思想及过程：**
2. 未命中时，TLB未命中句柄先执行传统的页表搜寻来找到正确的页表入口。如果主存页表最近在缓存中出现过，那就仅仅通过物理地址到缓存的地址映射更新cTLB，如果没有，就分配一个缓存块，并产生一个将页从未封装DRAM拷贝到分配的缓存块的请求，同时更新GIPT。一旦完成上述操作，TLB未命中句柄会用虚拟地址到缓存的地址映射来更新cTLB。
3. 在TLB命中时，cTLB返回缓存地址，而不是原来的物理地址。
4. 如果层叠芯片缓存未命中，缓存地址将被用来访问封装DRAM缓存。这样保证了不需要在已存在的DRAM缓存中在支付检查标签操作的额外代价。
5. 利用空闲的队列同步缓存块的清除。为了在访问缓存时，清除掉回写的开销，无标签缓存确保有少部分的空闲块总是可以相应缓存存储请求。如果一个缓存块被清除，则它的PTE和咨询GIPT得到的物理地址更新到缓存地址。
6. **总体评价：**

这篇论文通过对目前DRAM的创造方法进行了分析，挖掘出其中的不足（待改进）之处，在分析问题后，得到了开销的主要原因，并通过思考，找出了一种减少开销的方法，即本文所主要论述的全相联无标签缓存。通过进一步的思考，确定出了制作此项目需要的部件，完成了对其的搭建，并测试实验验证。在验证的过程中，通过不同的方面对其进行测试（单程序，多程序，多线程），并在每一组数据中给出了相应的表格，更加直观的反映出数据的走势和与之前方法的对比。最后指出了论文的可应用领域，表明此设计有用武之地。综合来看，这是一篇思路清晰，用实验数据说话的论文，值得我们去学习。