

DOI: 10.3969/j.issn.1672-8874.2009.01.034

使用模拟器加强计算机学科的综合实验教学*

潘晓辉, 司宏伟

(国防科学技术大学 计算机学院, 湖南 长沙 410073)

[摘要] 本文针对现有计算机学科实验教学中的问题, 提出了使用模拟器加强计算机学科综合性实验教育的观点, 给出了一个使用模拟器加强计算机学科综合性实验教育的案例, 指出这种实验教学不仅对学生系统地理解计算机课程有帮助, 同时能够有效锻炼和提升学生分析问题、解决问题的实际能力。

[关键词] 计算机学科; 实验教学; 综合性实验教学; 模拟器

[中图分类号] G642.423 [文献标识码] A [文章编号] 1672-8874(2009)01-0097-02

实验教学在计算机学科的教育中占有重要的地位。针对现有计算机学科实验教学中存在的问题, 运用模拟器加强计算机学科综合性实验教学, 对锻炼和提升学员分析问题、解决问题的能力具有积极的作用。

一、现有计算机学科实验教学中存在的问题

与很多其它门类的科学不同, 计算机科学本身就是一门人工科学。之所以将计算机科学称为人工科学, 是与传统的自然科学对比而言的, 因此计算机技术通常被归类于工业技术领域。在人们的生产活动中, 为了自动化地完成计算和控制, 人们创造了计算机。这一过程可以追溯到巴贝奇创造差分机和分析机。^[1]可以说, 计算机科学的发展就是人们在自然界中生产、生活不断实践的结果。因此, 作为讲授计算机科学的计算机学科自然也不能离开实践。从计算机学科的课程数目上说, 计算机学科本科教学的课程数目就超过 50 多门。此外, 除了传授知识外, 培训技能也是计算机学科教育的重要组成部分。为了让学生们能够真正学懂、学会, 融会贯通计算机学科中的各种知识和技能, 实验教学扮演了非常重要的角色。因此, 可以说, 实验教学在计算机学科教学中非常必要和重要。

由于实验教学在计算机学科教学中的重要地位, 在计算机学科的各门课程中, 实验教学占了很大比例, 然而现有的实验教学中, 存在着如下两方面问题:

1、从目前的课程实验而言, 存在实验课程中“软”比“硬”多, “应用”比“系统”多。即软件课程的实验比硬件多, 应用课程的实验比系统课程多。这种局面有多方面原因。“软”比“硬”多的最主要原因在于硬件相对软件复杂, 不仅由于硬件设计的复杂度较高, 还受限于硬件的实验平台较软件实验缺乏, 因此很难完成相对大型的硬件系统实验, 大多数的硬件实验只能集中于嵌入式课程领域。“应用”比“系统”多的原因也是一样, 但就数据库原理和操作系统两门课程相比, 学生建立一个简单的数据库系统

相对容易, 而让学生实践编写一个操作系统就变得不太可能。

2、虽然很多课程都有实验设计, 然而缺少各门课程的综合性实验。计算机科学是一门系统性科学, 这一点与计算机系统一样, 软件、硬件密不可分。因此, 对计算机学科的教学仅仅通过分立的课程实验是不够的, 难以起到让学生对计算机系统知识融会贯通的目的。然而, 在计算机科学的研究过程中, 比如博士阶段的课题研究中, 从系统的角度上进行思考、实践和实验的环节往往必不可少, 如果在课程学习阶段缺少这方面的学习和训练, 就很不利于学生建立计算机系统的整体概念, 学到的是一块块分立的知识, 等到研究的时候就常难以综合地理解和思考问题。

二、使用模拟器加强计算机学科的综合实验教学

正因为存在上述问题, 我们认为, 应该多使用模拟器加强计算机学科的综合实验教学。模拟器是通过对需要建模的对象进行参数抽象搭建的软硬件系统, 用于对实际对象进行仿真模拟。一般地, 模拟器都要实现精度和速度两方面的权衡。也就是模拟的精度提上去了, 速度就自然会下降, 要想速度快, 精度往往又受到影响。在计算机科学的研究领域, 模拟器得到广泛利用, 通过对复杂系统的抽象建模, 研究人员可以在模拟器上研究各种新的技术和方法。比如计算机体系结构领域的研究就不可或缺各种处理器的模拟器。这些处理器模拟器是能够模拟处理器等硬件平台的软件, 这种使用软件模拟硬件的方式能够大大提升对硬件设计和研究的效率。

模拟器的各方面优点使得利用模拟器进行计算机各门课程的综合性实验教学成为可能。一是模拟器的使用相当于用软件的方式来理解硬件, 原有硬件实验的复杂程度变得简单很多, 不再受很多硬件条件的限制, 使用普通的微机即可。二是可以方便地进行各门功课的综合性实验教学。硬件模拟器上运行的是软件, 这样可以设计各种综合性实

* [收稿日期] 2009-02-26

[作者简介] 潘晓辉 (1977-), 男, 山西阳泉人, 国防科学技术大学计算机学院训练部参谋, 博士。

验项目,从而加深学生对计算机系统的整体理解。

三、使用模拟器的综合性实验教学案例

这一部分给出一个使用模拟器完成的综合性实验教学案例。这个案例的目的是让学生直观、定量地感受计算机领域一直存在的“存储墙”问题,这个实验中需要综合计算机体系结构、编译,以及程序设计等课程的知识,培养学生独立发现和解决问题的能力。

题目设置如下:

图2中展示了一段C程序的源代码,包含两个多重嵌套的循环,第一个双重嵌套循环完成对数组sum和数据a的初始化,第二个三重嵌套循环对数组a进行计算。要求:

```
#define N 2048
int a[N][N];
int sum [N];

void fun() {
    int i, j, k;

    for (i=0; i<N; i++) {
        sum [i] = 0;
        for (j=0; j<N; j++)
            a [i][j] = i + j;
    }

    for (i=0; i<N; i++)
        for (j=0; j<N; j++)
            for (k=0; k<N; k++)
                sum [i] += a[k][j];
    }
```

图2 综合性实验案例的C程序源代码

1、编写程序调用图2中的fun函数,并在微机上编译运行,统计程序的运行时间。

2、使用SimpleScalar模拟器^[2]定量分析这段程序运行的性能瓶颈,给出定量分析结果,并对程序进行改进。

3、将改进后的程序重新在微机上运行,统计程序的运行时间和改进后的性能加速比。

提示:使用SimpleScalar中的sim-cache模拟器分析程序执行时的Cache性能。

本题综合了程序编程、编译和体系结构三方面知识,主要的性能瓶颈在于Cache的利用率。学生通过使用处理器模拟器SimpleScalar模拟分析程序的Cache性能可发现,程序的运行性能瓶颈与Cache的映射机制有关。由于数组a在第二个循环中属于按列访问,导致访问的数据元素均冲突到一个Cache组中,从而导致Cache利用率的严重下降,程序运行性能较低。本题是一道开放性的综合性实验题,可以说,除了Cache利用率下降这一严重原因外,定量分析和改进方法没有标准答案,有很多种改进策略。下面给出本题的一种相对完全的解题思路。

1、通过编写程序调用fun函数,在微机上编译运行,使用操作系统的time命令得到程序的运行时间,也可以通过在程序中调用系统的时间函数计算函数的运行时间。

2、在SimpleScalar上运行该程序,可以发现,程序的规模太大导致模拟的时间过长,一种解决方案是减少数组的规模,按比例调小N。然后,分析程序运行时Cache的性

能,可以为每个Cache块或Cache组增加计数器,很容易就会发现,数组a的访问严重地冲突到一个Cache组中,导致冲突失效的发生,使得Cache的利用率非常低。

3、针对发现的问题,对程序进行改进。方法有很多,一是分析第二个三重循环,发现后两层循环可交换,交换后数组a变成连续地址访问,从而大大提升Cache的利用率;二是使用Padding技术,^[3]将数组a的声明变成int a [N] [N+ ΔN], ΔN的值与Cache块的大小有关,目的是让a按列访问时不冲突到一个Cache组中。

4、在降低Cache冲突失效的基础上对程序进行Tiling,^[4]进一步减少Cache容量失效。

5、将改进后的程序在模拟器上运行,定量记录Cache的运行性能。

6、将改进后的程序在微机上运行,记录运行时间和加速比。

从上面的解决方案可以看出,使用模拟器可以方便地实现传统实验教学中难以实验的硬件部分(本题是Cache),能够定量分析硬件上的性能开销,同时,在本题的解决方案中需要学生的编程和编译知识,这样才能综合分析出程序的性能瓶颈并进行改进。如果没有模拟器的帮助,虽然也能够通过手工分析的方式分析出程序运行的Cache性能瓶颈,但模拟器方式给出的定量模拟结果会让学生更直观的感受存储墙问题和Cache的重要性,对培养学生未来从事研究工作需要的分析问题和解决问题的能力有很好的帮助。

四、结束语

计算机学科是一门实践性非常强的学科,实验教学在计算机学科教育中占有重要的地位。本文提出使用模拟器加强计算机学科综合性实验教学,以弥补现有实验教学中存在的不足。本文给出了一个使用模拟器加强计算机学科综合性实验教学的案例,指出这种实验教学不仅对学生融会贯通、系统地理解多门课程有帮助,同时能够有效锻炼和提升学生分析问题、解决实际问题的能力。

[参考文献]

- [1] 胡守仁. 计算机技术发展史[M]. 湖南: 国防科技大学出版社, 2004.
- [2] Todd Austin, Eric Larson, Dan Ernst, SimpleScala. An Infrastructure for Computer System Modeling, Computer[J]. 2002, (2).
- [3] E. Hernuzo, O. Plata, E. L. Zapata. Using Padding to Optimize Locality in Scientific Applications, INCS Computational Science- ICCS 2008.
- [4] M. E. Wolf, M. S. Lam: A data locality optimizing algorithm, In Proceedings of the ACM SIGPLAN 1991 Conference on Programming Language Design and Implementation.

(责任编辑: 阳仁宇)