

一种基于 ACM 程序设计竞赛在线评测系统解决方案

车明珠¹, 纪洪波²

(1. 延边大学 信息化中心, 吉林 延吉 133002 ; 2. 通化师范学院 计算机科学系, 吉林 通化 134001)

摘要:分析了一种基于 ACM 程序设计竞赛在线评测系统的基本原理及系统的构成,重点阐述了程序性能评判原理及实现方法。对通信安全、资源占用等关键问题给出了相应的解决方案,主要模块给出了相应代码。

关键词:评测系统; ACM; 程序性能

中图分类号: TP 393

文献标识码: A

An online evaluation system solution based on ACM programming contest

CHE Ming Zhu¹, JI Hong Bo²

(1. Info.Center, Yanbian University, Yanji 133002, China;

2. Computer Science Department, Tonghua Normal University, Tonghua 134001, China)

Abstract:This study analyses the basic principles and system's structure which based on ACM international collegiate programming contest online evaluation system, puts emphases on the the principle and implementation of program performance evaluation. The corresponding solution of some key problems such as communications security and resources occupy are presented. Moreover, the main module are given the corresponding code.

Key words: judge system; ACM; program performance

ACM 国际大学生程序设计竞赛 (ACM/ICPC 或 ICPC) 是由美国计算机协会 (ACM) 主办、旨在展示大学生创新能力、团队精神和在压力下编写程序、分析和解决问题能力的年度竞赛。经过近 30 年的发展,ACM 国际大学生程序设计竞赛已经发展成为最具影响力的大学生计算机竞赛。竞赛规模的迅速扩大对阅卷工作的自动化、高效性、合理性和公正性提出了更高的要求,建立一套准确、高效的程序评测系统成为非常迫切的需求。基于以上原因,本文给出了一种网络自动化的程序性能分析评价系统——ACM 国际大学生程序设计竞赛在线评测系统的实现方案。

1 评测系统简介

本系统提供了对 C、C++、Java 三种语言所编写的程序进行处理的功能。用户按照竞赛题目的要求,通过对问题的分析并给出解决方案后,就可以向系统提交解决问题的源代码程序。系统可以根据用户提供的源代码,采取相应的程序语言环境编译、运行。在编译与运行正常以后,系统按照题目要求来判断该程序结果是否正确,同时给出程序的运行时间、内存的开销等情况,根据程序性能信息表^[1-2]来判断各个用户的得分情况。

《微型机与应用》2010 年第 4 期

系统运行于 Linux 操作系统,借助 Linux 操作系统所提供的多任务、多线程功能及其稳定的性能和免费开源的内核,从而大大减少程序员书写系统代码的工作量,也提高了系统的运行效率。代码是用 C 语言编写的,通过自己编写函数或调用 C 语言内部函数来访问系统所使用的资源。

2 评测系统的体系结构

评测系统网络体系结构采用先进的 B/S 模式的三层架构,如图 1 所示,这种结构符合面向服务 (SOA) 的模块设计,模块耦合性低,使得业务逻辑的添加、修改较为容易,并且网络通信安全性能方面也大为提高。用户通过浏览器连接评测系统,表示层 UI 是系统呈现给用户的使用界面,用户通过 UI 进行登录、浏览、提交代码等操作。业务逻辑层 BL 负责实现根据性能信息表数据评分等除测试部分以外的各模块的功能。本系统将评测部分(主要是黑盒测试)从业务逻辑层中独立了出来,成为一个专门的测试模块,这种做法的优点是对测试模块的修改、增删都很容易进行。数据库 DB 主要用来存储用户提交的程序性能信息表及用户信息表等,程序性能信息表中主要包括程序的各种性能信息,例如是否通过

欢迎网上投稿 www.pcachina.com 11

编译、是否超出最大时间、是否正常结束、语言类别、结果是否正确、代码运行时间、内存开销、CPU 占用量、代码长度等。测试模块从数据库取得未测试代码,进行评测,将程序性能信息表中相关项的评测结果返回给数据库,业务逻辑则从数据库取得程序性能信息表,并根据表中各项数据评分,对于未通过编译、超出最大时间、非正常结束、结果不正确的程序按不合格处理,对于上面各项都合格的则按程序性能信息表中其余项进行综合评分。

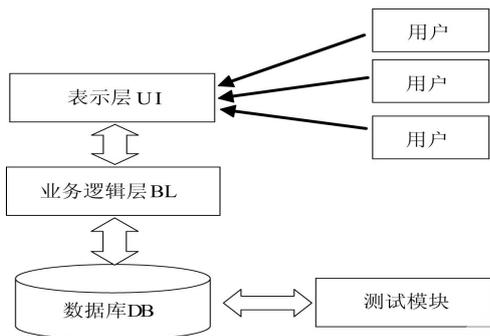


图 1 评测系统体系结构

3 评判原理

对程序编译与运行过程中出现的错误,系统会根据错误的不同给出相应的提示,例如是编译错误或者是运行错误。如果编译和运行都正确,系统将进一步通过相应的题目要求、对程序要求给出的输入和程序输出的结果,来判断该程序是否正确。

在进行程序评判时,先启动服务器评测进程,在侦听到有提交记录时,系统立即对提交记录的相关题目进行评判。每个问题均对应一组数据(输入和输出文件),系统设计时采用 Linux 管道(pipe(int filedes^[5]))来处理,评判进程启动一个子进程(fork())来编译运行用户提交的程序,读入该题目的输入数据,评判程序每次从管道读入一个字符与标准输出数据比较,如果两个文件完全一样,则表示程序正确;如发现只是相差空格、Tab、回车,则给出该程序为输出格式错误信号(Presentation Error);如果发现其他字符的不匹配,则立即中止该程序,则给出该程序为结果错误的信号 Wrong Answer。从程序运行时刻开始计时,在时间允许范围之内,得到了正确答案,则此提交程序符合题目要求并通过系统时间和内存的限制,给出接受的信号(Accepted)。反之,如果答案错、或者表达错、运行错等,则向该进程发出 SIGKILL 信号,强行中止该进程,对数据库记录进行相应的修改。运行流程如图 2 所示。

4 程序核心内容分析

为了能有效地解决和判断用户程序对系统资源的使用情况,系统定义了 rusage 结构体。该结构体具体信息为:

```
struct rusage {
```

12

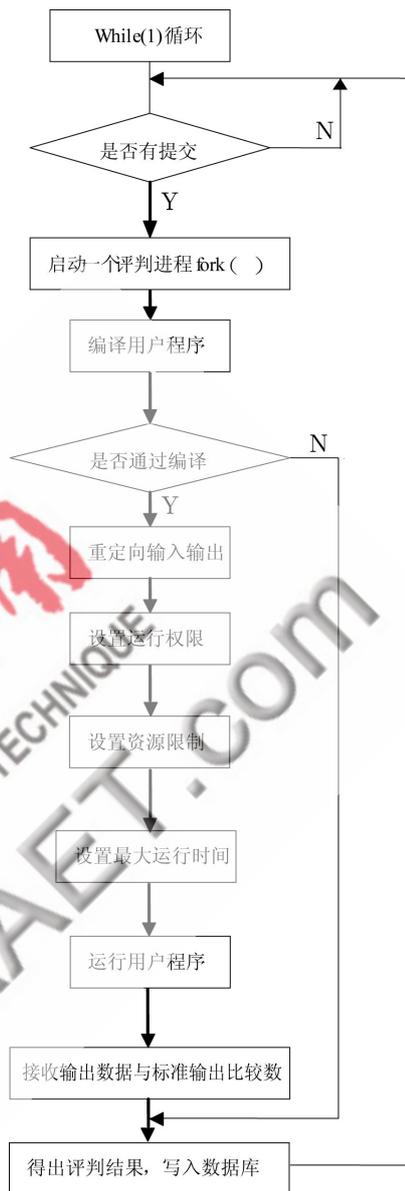


图 2 评测主程序流程

```
struct timeval ru_utime;          /* user time used */
struct timeval ru_stime;         /* system time used */
long ru_minflt;                 /* page reclaims */
long ru_majflt;                 /* page faults */
};
```

ru_utime 和 ru_stime 成员变量包含了在用户模式和系统模式中执行时间的总和。其结构都为 timeval 结构。ru_minflt 成员指不需要 I/O 的页缺失数。页缺失发生在内核需要得到一个内存页以供进程访问。

ru_majflt 值指需要 I/O 的页缺失数。页缺失发生在内核需要得到一个内存页以供进程访问时。

有时,一个进程会被调出内存,以提供空间给其他进程使用。ru_nswap 指的就是一个进程被调出内存的次数。

通过该结构体就可以统计出各用户程序对资源的

《微型机与应用》2010 年第 4 期

软件天地

Software Technology

占用情况,通过调用函数 `int getrusage (int who, struct rusage *rusage)` 对用户程序运行时间和内存占用情况进行分析统计。

统计程序运行时间:

```
passtime = usage.ru_utime.tv_sec + usage.ru_stime.tv_sec +
(float) (usage.ru_stime.tv_usec + usage.ru_utime.tv_usec)/
1000000;
```

统计程序所占内存的开销:

```
usedMemory = usage.ru_minflt*4
```

通过统计该两项,能计算出程序运行时间和内存使用情况的数据,为结果排序做准备。

5 系统安全处理

用户提交的程序在服务器上运行,为了保证系统安全运行、减少或杜绝故障的发生,在解决安全故障方面主要有以下几个方面。

(1) 文件操作

为了让用户提交上来的程序不破坏本机的文件系统,本设计采用了管道技术,在执行用户程序之前,先把输入流定向到标准输入文件,然后使用 `chroot` 命令改变用户程序的执行目录,让其只能在一个临时文件夹下面做操作。

(2) 网络操作

为了不让用户程序影响评判系统的设置,保证评判系统的公正性,必须禁止用户对网络的访问。系统通过对用户实行权限限制机制,判断系统在用户程序运行之前,通过调用 Linux 系统中 `setuid(99)` 和 `setgid(99)` 两个函数^[4],让用户程序的拥有者权限降到最低,达到防止用户对网络使用的目的。

(3) 资源管理

资源管理的作用是对客户进程所使用的资源进行管理, Linux 系统提供的资源限制函数 `setrlimit()`,可以避免用户进程在运行的过程中过多地占用系统资源,以影响服务器系统的稳定性,最终有可能导致系统的崩溃。

为了防止用户进程占用过多的资源,影响服务器系统的稳定,对用户进程在占用资源方面进行限制,主要

考虑程序运行时间和内存占用两个方面,在程序运行之前,通过调用系统函数 `setrlimit()` 来限制程序对系统资源的占用。函数 `setrlimit(RLIMIT_AS,&memlim)` 和 `setrlimit(RLIMIT_CPU,&timlim)` 分别用来限制用户程序对 CPU 和内存的占用情况,如果某用户的程序超过了系统对 CPU 和内存的最大设置值,该进程就会被自动中止,标志用户的程序不满足比赛要求。

通过对基于 ACM 国际大学生程序设计竞赛在线评测系统的分析研究,设计了相应的 ACM 测试软件,习题的测试难度可由指导老师自己设计。该软件也可以作为计算机专业学生算法测试工具,通过对该测试系统的使用,经过一定题量的练习,可以增强学生在算法方面设计的技能,为学生奠定雄厚的编程基础。

参考文献

- [1] ESPINOSA A, MARGALEF T, LUQUE E. Automatic performance evaluation of parallel programs. The Sixth Euromicro Workshop on Parallel and Distributed Processing, 1998(21):43-49.
- [2] ESPINOSA A, MARGALEF T, LUQUE E. Automatic detection of parallel program performance problems. Lecture Notes in Computer Science 1999(1573):365-377.
- [3] STEVENS W R, RAGO S A 著. UNIX 环境高级编程(第 2 版)[M]. 尤晋元,张来英,戚正伟,译.北京:人民邮电出版社,2006.
- [4] SARWAR S M, KORETSKY R, SARWAR S A 著. Linux 教程[M].李善平,施韦,林欣,译.北京:清华大学出版社,2005.

(收稿日期:2009-11-09)

作者简介

车明洙,男,1975 年生,实验师,硕士,主要研究方向:数据库原理、算法设计;

纪洪波,男,1976 年生,讲师,硕士,主要研究方向:算法设计、优化技术及统计分析。