|  |  |
| --- | --- |
|  |  |
|  |  |

Diagram

Description automatically generated

程序调试实践

Program Debug Practice

王浩

（软件工程师）

公司：安波福电子（苏州）有限公司

部门：AS&UX CAM&TEL TEAM

邮箱：luis.wang@aptiv.com

导师：

日期：2022/10/30

**安波福电子**

**论文原创性声明**

本人郑重声明：所提交的论文是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已经注明引用的内容外，本论文不含其他个人或集体已经发表或撰写过的研究成果。对本文的研究作出重要贡献的个人和集体，均已在文中以明确方式标明。

论文作者签名： 年 月 日

程序调试实践

中文摘要

随着计算机行业的不断发展和信息技术的继续迭代，越来越多的行业在计算机和互联网的帮助下来提升自身的效率。这对于软件工程来说意味着会有更复杂更多样的需求，而且这些需求也会使得代码变得复杂，程序更加庞大。而对于程序员们来说，这些变化意味着他们需要有更优秀的编程能力和代码调试能力。

本文主要介绍当程序出了不符合预期的表现时，也就是当程序出了BUG时，程序员应当如何去找到出现这个问题的原因。当然为了解释清楚BUG的来龙去脉，本文来将会介绍与之有关的计算机的必要知识，这些知识主要集中在操作系统和程序计算机中的如何编译运行的。

由于目前开发的环境大多集中于UNIX体系的操作系统中，所以本文着重介绍的还是在Linux的操作系统下程序调试相关内容，其中最主要涉及的就是Linux下关于调试方法的研究，比如GDB等调试工具的使用和在调试过程的中应用等。

**关键词：软件调试、Linux、GDB、BUG、DEBUG**

**Program Debug Practice**

**Abstract**

With the development of computer industry and information technology grows, more and more industries are making themselves more efficient with the help of internet and computer. it means that there are more complex and diversified requirement for software engineering, which will make code become sophisticated and make program become huge. Meanwhile, these changes will require programmer with perfect coding skills and excellent code debug ability to cover these solutions.

This paper mainly introduces how to find out the cause of problem for programmer when the program appears non-expected performance, in other words, the program appears with a bug. Of course, in order to better explain the context of the bug, this paper also will introduce some related necessary knowledge about computer and program, and this knowledge mainly center on operation system and how does the program compile and run in the computer.

So far, because of the development environment mainly fasten-on class of the UNIX operation system, so this article focuses on the introduction of the relevant content of program debug technique under Linux operation system, in which the most of important of all is about research on debugging methods in Linux, such as the use of debugging tools and application during debugging process.

**Keywords:** **software、Linux、GDB、BUG、DEBUG**

Table of Contents

[第一章 绪 论 1](#_Toc93417709)

[1.1 程序调试研究的目的 1](#_Toc93417710)

[1.2 程序调试研究的意义 1](#_Toc93417711)

[第二章 程序调试的介绍 1](#_Toc93417712)

[2.1 程序调试的背景 1](#_Toc93417713)

[2.1.1 什么是程序调试 1](#_Toc93417714)

[2.1.2 什么是BUG 1](#_Toc93417715)

[2.1.3 程序调试和程序测试的区别 1](#_Toc93417716)

[2.2 程序调试的目的 1](#_Toc93417717)

[2.2.1 为什么要进行程序调试 1](#_Toc93417718)

[2.2.2 什么时候进行程序调试 1](#_Toc93417719)

[2.2.3 程序调试的步骤 1](#_Toc93417720)

[第三章 调试前的知识准备 1](#_Toc93417721)

[3.1 操作系统相关知识 1](#_Toc93417722)

[3.1.1 字节序 1](#_Toc93417723)

[3.1.2 寄存器 1](#_Toc93417724)

[3.1.3 地址 1](#_Toc93417725)

[3.1.4 数据类型 1](#_Toc93417726)

[3.1.5 堆栈 1](#_Toc93417727)

[3.2 程序运行相关知识 2](#_Toc93417728)

[3.2.1 程序在计算机中的存储结构 2](#_Toc93417729)

[3.2.2 程序的编译链接过程 2](#_Toc93417730)

[3.2.3 程序的运行过程 2](#_Toc93417731)

[3.3 汇编语言相关知识 2](#_Toc93417732)

[3.3.1 汇编语言的常用语法 2](#_Toc93417733)

[3.3.2 汇编语言的使用分析 2](#_Toc93417734)

[第四章 调试的知识准备 2](#_Toc93417735)

[4.1 调试平台差异介绍 2](#_Toc93417736)

[4.2 调试工具介绍 2](#_Toc93417737)

[4.3 GDB基础命令 2](#_Toc93417738)

[4.4 GDB调试步骤 2](#_Toc93417739)

[第五章 简单的程序调试实践 3](#_Toc93417740)

[5.1 简单的main程序调试实践 3](#_Toc93417741)

[5.2 带有函数的程序调试实践 3](#_Toc93417742)

[5.3 加载静态库的程序调试实践 3](#_Toc93417743)

[5.4 加载动态库的程序调试实践 3](#_Toc93417744)

[第六章 多进程程序调试实践 3](#_Toc93417745)

[6.1 多进程程序编译运行 3](#_Toc93417746)

[6.2 多进程程序调试实践 3](#_Toc93417747)

[第七章 多线程程序调试实践 3](#_Toc93417748)

[7.1 多线程程序调试实践 3](#_Toc93417749)

[7.2 多线程程序调试实践 3](#_Toc93417750)

[第八章 内核有关的调试实践 3](#_Toc93417751)

[8.1 与程序崩溃有关的调试实践 3](#_Toc93417752)

[8.2 与程序停止响应的调试实践 3](#_Toc93417753)

[8.3 与系统运行缓慢的调试实践 3](#_Toc93417754)

[8.4 与CPU负载过高的调试实践 3](#_Toc93417755)

[第九章 总结 3](#_Toc93417756)

[参考文献 4](#_Toc93417757)

[致谢 4](#_Toc93417758)

# 绪 论

## 程序调试研究的原因

随着社会科学技术的发展，各行各业对于计算机的需求也越来越旺盛，随之而来的就是计算机的从业的人员也越来越多，计算机行业也在快速的迭代和发展。计算机行业中对于编码方面的书籍和论文数不胜数，但是对于程序调试方法的研究却是寥寥无几。

在一个程序的实现过程中，程序的编码无疑是占了一个很重要的地位，但是程序调试的地位也丝毫不逊色于程序编码的重要性。甚至程序就是在写代码和改代码的过程中，逐步的去满足最终的需求。程序的调试过程甚至于要比编码在项目周期中占用的时间更多。所以对程序调试进行研究这一工作是一件很有必要的事。

虽然程序调试在程序开发过程中占有很重要的位置，但是目前来说，程序员大多只能靠经验来支撑这一部分的工作，大多数人只能通过不断在工作中踩坑来积累经验，并且市面上对于调试部分的研究论文和书籍也是寥寥无几。并且各种语言和项目又千差万别，不能一概而论的现象又加剧了这种情况的延续。

所以这篇文章就是基于以上的理由，开始研究关于在Linux系统下关于C/C++程序的调试的一些方法和步骤，特别是一些Linux自带的工具甚至也可能涉及到一些第三方工具的使用，都是为了更好的对调试做一些介绍和研究。

## 程序调试研究的目的

这篇文章的研究目的主要有两个：

其一：旨在希望当有需要的人在读完这篇文章以后，能帮助他们快速上手关于C/C++程序在Linux系统下的调试方法。为什么要强调适用语言，主要是各个语言对于本身的调试调试支持，如果不是限制到该语言中，有些方法和工具是需要具体分析的，有些语言对于调试支持的比较好会有自己的输出逻辑和框架。此外说是在Linux系统下，其实相关的Unix下，关于C/C++语言的调试都是可以使用到的，因为我们需要用到的调试工具GDB在Linux系列的系统下大都支持，并且如果在GDB工具不支持使用的情况下，还有一些更好用的第三方工具可以选择来帮助到调试。

其二：旨在帮助学习这篇文章的人可以借着学习调试的过程中来更深入的学习计算机底层的知识，构建计算机操作系统的知识网络。因为C/C++是一门和计算机底层联系比较密切的语言，有C/C++编译出来的代码很容易就和操作系统做了很深的交互，如果对底层知识可以很熟悉的运用和了解的话，对于一些调试的方法可以起到事半功倍的效果。此外其他语言虽然可能有自己的调试模块来帮助程序员做对应的调试工作，但是这些部分最终都是需要走到计算机操作系统的调用。所以对于学习操作系统这件事来说，对于程序员提升自己的计算机素养，很有帮助且，很有必要。

## 程序调试研究的意义

进行程序调试的细致研究，主要是为了通过对程序调试的前因后果的过程探索，梳理调试过程中涉及到的知识体系，可以让程序员掌握计算机更深层次的知识，更重要的是通过此次研究过程，可以让程序员能构建出对程序调试这一过程中的知识网络。

现阶段，大部分程序员对于程序在计算机中的操作系统等知识一直都是一知半解，但是随着互联网的继续发展下去，计算机会和各行各业都进行一个赋能结合的过程，这也就意味着程序员除了需要有过硬的编程能力之外，还需要了解计算机的底层知识和对程序的Debug能力，这样才能应对在整个程序开发周期过程中面临的主要的编码和调试问题。

而通过对程序调试的研究，不仅可以帮助程序员了解和梳理计算机底层知识，还能学习在面对的程序出现不符合预期的现象时，如何去快速准确追踪溯源的找到该问题发生的原因，并且恰到好处的解决问题。

除此之外，通过研究程序调试的方法，要求程序员去掌握和了解在操作系统层次下关于程序的调试工具的使用方法，和培养当没有直接输出时如何去模拟程序运行环境下程序调试的能力。

总而言之，程序调试研究的意义主要是想让程序员掌握计算机底层的知识以及面对没有直接调试信息输出时，或者只有调试信息如何去准确的确定问题并解决问题的能力。

# 程序调试的介绍

## 2.1 程序调试的背景

### 2.1.1 什么是程序调试

每一个程序都是为了解决一个问题。所以在一个完整的，规范的项目实现过程中，程序的最初来源是需求定义，然后是根据需求，进行程序的总体设计，然后再是根据总体设计制定详细设计，再经过程序员的编码，系统测试步骤，程序测试调试等步骤，最终完成一个程序来解决开头提出的那个问题。当然这些步骤的能实现的前提是项目立项，需求确定。

当需求以文档确定好以后，程序员会拆分整个需求，构建程序架构，然后分配各个模块到每个组，或者每个人手中。程序在编码过程中会有自测过程，当一个模块完成以后会进行模块或者单元测试，验证程序是否像需求文档描述那样符合预期的运转。自测通过以后，程序会整理打包交到测试人员手中进行正式的系统测试，发现问题。通过几轮测试修改，程序运行结果如需求文档要求后，就可以正式对外发布正式版本，交付最终产品。

在这个过程中，程序员在开发时进行的自测过程，也算是程序调试。当测试人员发现整体某一点和需求文档有出入的话，会通过约定的方式通知开发人员这处错误，开发人员会继续修改这个问题，这个过程也需要进行程序调试。

程序调试的目的就是为了解决程序中不符合预期现象，而方法就是通过追踪程序中流转的数据，找到数据在哪一部分出现不符合预期的错误。而这个方法就是我们接下来需要重点讨论的话题，并且程序调试的方式多种多样，但是最根本的就是通过种种手段来跟踪程序数据的流转和处理过程。这些手段主要分为系统或者第三方工具来追踪程序堆栈或者是使用者自己使用输出日志形式观察程序运行过程的数据并且保存下来。

总结就是，运行程序检查程序是否符合预期的这一过程就叫做程序调试，其中进行程序调试的时机主要有两个，一个当程序员根据需求初步完成编码时需要进行自测，二是当整体模块编写完成并且通过时，程序被以正式版本的形式交给专业的测试人员进行系统测试。

### 2.1.2 什么是BUG

BUG原意为虫子或者臭虫，在现代的计算机行业中通常被视为在程序中的缺陷或者问题。

BUG从原意到现在的引申义之间是有一个故事的，最初的计算机都是由数量庞大的真空管组合而成，并且还需要用庞大的电力使得真空管发光。当时的计算机体积约有一间屋子大小，1945年，在计算机工作时，一只小虫子飞蛾被计算机的电和光所吸引，落在了计算机内部一个继电器的触点上，导致计算机无法继续工作。然后这种导致计算机无法正常工作的现象就被称为BUG，与之相对应的解决问题的过程则被称为DEBUG。

广义上的BUG是由电脑系统的硬件、系统软件、或者应用软件出错；而狭义的BUG指的是程序的漏洞和错误，除此之外BUG还包括由测试工程师或者用户所发现的和提出的软件可改进的细节、或者与需求文档存在差异的功能实现等。

BUG出现的原因则有1）对各种流程分支考虑的不够充分；2）对于边界处理有错误；3）编码的错误导致；任何软件在发布的时候都不可能是零BUG的状态，并且现在的软件复杂程度早就超过了一般人能控制的范围了，所以如何减少BUG的出现以及DEBUG软件开发过程中一个重要的部分。

### 2.1.3 程序调试和程序测试的区别

程序调试是由开发人员解决BUG的过程；而程序测试是测试测试人员找到BUG的过程。这是这两者的根本区别，前者的目标是解决问题，而后者的目标是发现问题，通过不停的对程序进行测试和DEBUG的过程，使程序能如需求文档那样顺利流畅的运行，最后进行软件发布。所以程序调试和程序测试的最终目的就是为了使开发的软件能正常顺利的运行发布。

一般来说，非正式测试以后（确认程序可以正常的编译运行）就是测试和调试过程，但是BUG发现的时机不仅仅是在正式的测试阶段由测试工程师测试出来，它也会在平常的使用过程中，被开发人员，或者其他人员发现，而发现的方式也是多种多样。

而无论是由谁，以任何方式方式发现的问题，最总都会转回到负责该模块的开发人员手中，由开发人员通过程序调试的方式来找到这个问题出现的原因，并且解决掉这个问题，然后该问题的修复结果会以相同的方式返回到发现人的手里进行验证。如果这个问题是由测试人员发现的，他们会以一个正式的项目管理的方式将这个问题通知给开发人员，而开发人员在解决掉该问题以后放在正式的版本中，然后再通知测试人员验证该问题是否继续存在。

这就是程序测试和程序调试的联系，程序测试如果出现问题，则紧接着就会有程序调试来解决问题。

## 2.2 程序调试的目的

### 2.2.1 为什么要进行程序调试

程序员在完成一个项目时，大致会经历需求定义、软件设计、程序编码、程序测试、和程序调试这几个阶段，如图：



程序调试的根本原因就是为了找到程序出现的错误，然后解决掉。而程序出现错误的原因有很多，可能是主观上程序员没有注意到某个细节；可能是程序过于复杂，程序没有通盘考虑到这个情况；也可能是外界的环境不同导致；也可能是代码版本管理出了问题；凡此种种有客观的，主观的原因，都会导致问题的出现。所以想要解决这些问题，就需要根据程序表现的现象，一步步追踪到根本原因处，然后采用合适的方案来解决。

为什么要程序调试，就是为了准确的找到问题出现的原因，并且修正。BUG在整个软件开发的过程中都是一直被伴随着的状态，所以程序调试也是贯穿整个开发周期的始终。

### 2.2.2 什么时候进行程序调试

发现BUG的时机，下一步就是进行程序调试的时候。根据上图所示，BUG发现的时机是在程序可以正确启动以后，和程序测试阶段，然后接下来就会转回到程序的开发人员手中进行程序调试过程。但是BUG发现的时机不单单只有上述的两个阶段才会发生，只有该程序是在使用的状态过程中，都有可能被自己人，或者是其他任何人，以各种各样的方式发现。

但是在项目的正式发布之前，使用该软件的只有相应的开发人员和测试人员，此时程序调试的时机就在开发与测试过程中发现BUG后进行程序调试。

### 2.2.3 程序调试的步骤

如果开发人员被告知有一个问题，那么他首先应该做些什么去解决这些BUG。做些什么的这个过程，其实就是程序调试的步骤。

程序调试在软件开发过程中的流程大概为下图所示：

****

无论BUG是开发人员自己自测出来的，还是其他人发现的，其主要的调试步骤如下：

1. 复现BUG；
2. 程序调试
3. 确认程序的执行过程（测试）
4. 程序行为与预期相同，则停止，程序行为与预期不同，跳转到步骤2.

而本文就是来介绍步骤2中的各种调试方法和经验。

# 调试前的知识准备

## 3.1 操作系统相关知识

### 3.1.1 字节序

字节序，顾名思义就是字节的排序顺序。

首先解释一下什么叫字节。大多数计算机都使用8位的块来作为最小的可寻址单位，这个8位的块，就被称为字节。从C语言角度来说，就是每8位来表示一个ASCII码，然后以此为单位组成更多的内容。

如果一个多字节的对象保存在机器的内存中，机器如何判断哪个字节在前哪个字节在后，这个问题产生了字节序。比如在32bit的操作系统中，int类型的数据占了4字节，假设一个值为0x12345678的int型变量x地址为 0x100，那么x的值就会被存储到地址为 0x100,0x101,0x102,0x103 这四个字节的内存中，但是从哪里开始放呢，是将 12、34、56、78 依次放到 0x100、0x101、0x102、0x103 地址的内存中还是依次放到 0x103、0x102、0x101、0x100 的四个地址中呢。

以上两种存储多字节对象的方式，以最低有效位（0x78）在前（低地址-0x100）为小端法，在后（高地址：0x103）大端法。

对于大多数程序员来说，机器使用的字节顺序其实是无需关心的。只要写入的字节规则和读取的字节规则保持一致，换句话说，只要同一台机器或者是使用相同字节序的机器都不会因为字节序的不同产生问题。但是当两台字节序不同的机器进行数据传输，那么对于同样的内存的数据就会因为字节序的原因产生错误。一个最常见的问题就是当一个采用小端法的机器发送的数据被大端法的机器收到时，就会出现字节反序的错误。所以为了避免这类问题，网络应用程序都需要按照已建立好的字节顺序进行转换。

### 3.1.2 寄存器

寄存器是中央处理单元内部能够存储参与运算的数据和结果的时序逻辑电路。计算机本质是一个存储数据并且计算处理数据的机器。其中存储数据的功能就在磁盘、内存和寄存器中体现。当CPU需要进行数据计算时，会预先将数据从硬盘读取到内存中，然后再把即将要使用的数据从内存读取到寄存器中。寄存器，又被称为缓存，是容量最小，读写速度最快的存储数据的设备，但是它的造价是最昂贵的。内存和磁盘对应的读写速度和造价依次下降，容量依次增大。

寄存器分为通用寄存器组、指针和变址寄存器、段寄存器、指令指针寄存器、标志寄存器。

**3.1.2.1 通用寄存器**

通用寄存器组包括AX、BX、CX、DX 4 个寄存器，其中每个寄存器根据CPU迭代的版本，%ax 是16位的寄存器，%eax 是32位的寄存器，%rax 是64位的寄存器；BX、CX、DX系列的寄存器同理也有16、32和64位的寄存器迭代版本。

AX：累加器，主要存储返回值；16位的是%ax，有AH、AL的高8位和低8位区分；32位的是%eax；64位的是%rax；

BX：基址存储器，主要被调用者保存；16位的是%bx，有BH、BL的高8位和低8位区分；32位的是%ebx；64位的是%rbx；

CX：基数寄存器，保存函数调用的第四个参数，16位的%cx；32位的%ecx;64位的%rcx；

DX：数据寄存器，保存函数调用的第三个参数；16位的%dx；32位的%edx;64位的%rdx；

**3.1.2.2 指针和变址寄存器**

BP（BASE Pointer Register）：基址指针寄存器，被调用者保存

SP（Stack Pointer Register）：堆栈指针寄存器，保存栈指针；

SI（Source Index Register）：源变址寄存器，保存函数调用的第一个参数；

DI（Destination Index Register）：目的变址寄存器，保存函数调用的第二个参数。

**3.1.2.3 段寄存器**

CS：代码段寄存器

DS：数据段寄存器

SS：堆栈段寄存器

ES：附加段寄存器

**3.1.2.4 指令指针寄存器**

%rip：程序计数器，通常被称为PC，记录的是将要执行的下一条指令在内存中的地址；

**3.1.2.5 标志寄存器**

标志位寄存器也被称为程序状态位寄存器；标志位寄存器是16bit的寄存器，其中用来存放状态标志和控制标志的位数有9位。6位状态标志，分别是CF、PF、AF、ZF、SF和OF，用来记录程序运行的状态信息。控制位有3位，IF、DF、TF，用来CPU的控制。

**3.1.2.6 工作原理**

一个触发器能存放一个二进制代码， 若干个触发器和门电路组成一个寄存器。其中若干个的数量可以分为8，16，32，64。通常一个 x86-64 的中央处理单元包含一组16个能存储64位值的通用目的寄存器，而寄存器主要是用来存储整数数据和指针。

寄存器具有接受数据、存放数据和输出数据的功能。当寄存器收到“存入脉冲”时接受数据，得到“读出”指令时，将保存的数据输出。

寄存器存取数据有并行和串行两种方式，将n位二进制一次写入或者读出的方式被称为并行方式；将n位二进制以每次1位的分为n次存入寄存器或者从寄存器中读出的方式被称为串行方式。并行方式只需要一个时钟脉冲就可以完成工作，但是需要n条输入和输出数据线，穿行方式需要n个时钟才能完成n次的写入或者读取操作，但是只需要一根输入或者输出数据线，传输线少，适用于远距离传输。

### 3.1.3 地址空间

操作系统将内存视为一个非常大的字节数组，这个字节数组就被称为虚拟内存。内存的每一个字节都是由唯一的一个数字，该标识被称为地址。

地址是一个非负整数地址的游戏集合。在一个带虚拟内存的系统中，如果该系统是64bit，那么CPU就从一个有N=2^64个地址的地址空间中生成虚拟地址。这个地址空间被称为虚拟地址空间。一个系统还有一个物理地址空间，对应的大小是物理空间的字节数。

操作系统会将虚拟地址空间分割为虚拟页，作为磁盘内容拷贝到内存中的存储单元，这样每次从磁盘拷贝数据到内存中，直接拷贝一整页内容，而无需每次需要什么就去磁盘找然后拷贝。因为去磁盘拷贝数据的过程是非常耗费时间的。为了实现这个拷贝一个虚拟页的想法，物理空间地址也会遵循虚拟页的大小，分割成一个个物理页。操作系统通过内存管理单元将物理页和虚拟页进行映射，进行页表替换等操作。

### 3.1.4 数据类型

基本的数据类型包括字节（8bit）、字（16bit）、双字（32bit）、四字（64bit）。

在C/C++中，

### 3.1.5 堆栈

## 3.2 程序运行相关知识

### 3.2.1 程序在计算机中的存储结构

### 3.2.2 程序的编译链接过程

### 3.2.3 程序的运行过程

## 3.3 汇编语言相关知识

### 3.3.1 汇编语言的常用语法

### 3.3.2 汇编语言的使用分析

# 调试的知识准备

### 4.1 操作系统介绍

### 4.2 调试工具介绍

### 4.3 GDB基础命令

### 4.4 GDB调试步骤

# 简单的程序调试实践

### 5.1 简单的main程序调试实践

### 5.2 带有函数的程序调试实践

### 5.3 加载静态库的程序调试实践

### 5.4 加载动态库的程序调试实践

### 5.5 总结

# 多进程程序调试实践

### 6.1 多进程程序编译运行

### 6.2 多进程程序调试实践

# 多线程程序调试实践

### 7.1 多线程程序编译运行

### 7.2 多线程程序调试实践

# 内核有关的调试实践

### 8.1 与程序崩溃有关的调试实践

### 8.2 与程序停止响应的调试实践

### 8.3 与系统运行缓慢的调试实践

### 8.4 与CPU负载过高的调试实践

# 总结

# 参考文献

# 致谢