进程（Process）和线程（Thread）都是操作系统中的基本概念，它们之间有一些优劣和差异。

# 进程

## 概述

进程是程序执行时的一个实例，是系统进行资源分配的基本单位。

所有与该进程有关的资源，都被记录在进程控制块(PCB)中。以表示该进程拥有这些资源或正在使用它们。

另外，进程也是抢占处理机的调度单位，它拥有一个完整的虚拟地址空间。当进程发生调度时，不同的进程拥有不同的虚拟地址空间，而同一进程内的不同线程共享同一地址空间。

进程与程序：

程序是存放在磁盘文件中的可执行文件。

进程时程序的执行实例。

## 进程环境

### main函数和参数

#### main函数

格式：

int main(int argc,char \*\*argv)

初启函数：

认为：C程序从main开始执行

实际：真正启动点是系统提供的一个初启函数

位置：位于库文件中(crt0.o)

装配：ld在链接时候（加到main函数前面），形成可执行文件

作用：参数组织，调用main，中止进程

实例：

#include <stdio.h>

int main(int argc,char \*\*argv)

{

return 0;

}

#### 命令行参数

格式：

以空格分开的字符串

例如：$./a.out –f foo test

POSIX命令行参数语法约定：

选项、选项值（选项操作数）、操作数（命令操作数）

短横线开头，选项

多个选项，无选项值，可以放在一起

选项先于操作数

两个短横线终止所有选项

多个选项可任意顺序出现，可多次出现

扫描命令行中的选项：

选项如何提取：未知选项

选项值如何提取：选项如何对应选项值，缺少选项值

操作数如何获取：

报错：

扫描命令行中的选项：getopt函数

实例：

#include <unistd.h>

#include <stdlib.h>

#include <stdio.h>

int main(int argc, char \*argv[])

{

int flags, opt;

int nsecs, tfnd;

int i;

nsecs = 0;

tfnd = 0;

flags = 0;

while ((opt = getopt(argc, argv, "nt:")) != -1) {

switch (opt) {

case 'n':

flags = 1;

break;

case 't':

nsecs = atoi(optarg);

tfnd = 1;

break;

default: /\* '?' \*/

fprintf(stderr, "Usage: %s [-t nsecs] [-n] name\n",

argv[0]);

exit(EXIT\_FAILURE);

}

}

printf("flags=%d; tfnd=%d; optind=%d\n", flags, tfnd, optind);

if (optind >= argc) {

fprintf(stderr, "Expected argument after options\n");

exit(EXIT\_FAILURE);

}

for(i=optind;i<argc;i++)

{

printf("name argument = %s\n", argv[i]);

}

/\* Other code omitted \*/

exit(EXIT\_SUCCESS);

}

#### 环境变量参数

argv为程序的某一次特定运行参数

不常改变或系统共享信息

标准环境变量名：env和export命令设置查看

环境表

访问环境



实例：

#include <stdio.h>

extern char \*\*environ;

int main(int argc,char \*\*argv)

{

char \*\* p = NULL;

for(p=environ;NULL!=\*p;p++)

{

printf("%s\n",\*p);

}

return 0;

}

进程中环境变量操作：

#include<stdio.h>

char \*getenv(const char \*name);

返回：指向与name关联的value的指针，若未找到则返回NULL

int putenv(char \*str);

int setenv(const char \*name, const char \*value, int rewrite);//第3个参数表示是否覆盖原有环境变量

int unsetenv(const char \*name);

返回：成功返回0，出错返回非0

实例：

#include <stdio.h>

#include <stdlib.h>

#include <unistd.h>

extern char \*\*environ;

void showenv()

{

int i = 0;

char \*str;

while((str = environ[i]) != NULL)

{

printf("%s\n",str);

i++;

}

}

int main(void)

{

showenv();

printf("----------------------------\n");

printf("PATH:%s\n",getenv("PATH"));

setenv("PATH",".:/bin:/usr/bin",1);

printf("PATH:%s\n",getenv("PATH"));

putenv("NAME=jiekexueyuan");

printf("NAME:%s\n",getenv("NAME"));

return 0;

}

### 终止进程

进程终止流程：

注：\_exit通知内核，把进程占据的资源释放，而exit是用户态的终止函数，通知终止处理函数，IO清理函数及内核。

如果直接调用\_exit则标准IO未清除，一些资源没有在终止处理函数中释放，则有可能存在问题。

#### 正常终止

从main返回

调用exit（标准库提供）

调用\_exit或\_Exit（内核提供）

最后一个线程从其启动例程返回

最后一个线程调用pthread\_exit

##### 终止函数

exit和\_exit区别：

flush I/O：exit是先刷新缓冲区（调用atexit函数刷新缓冲区），然后结束进程，而\_exit是不刷新缓冲区，直接从进程退出

实例：

#include <stdio.h>

#include <stdlib.h>

#include <unistd.h>

#include <string.h>

#include <errno.h>

int main(int argc,char \*argv[])

{

if(argc < 3)

{

fprintf(stderr,"usage : %s file return | exit | \_exit\n",argv[0]);

exit(1);

}

FILE \*fp = fopen(argv[1],"w");

char \*str = "helloworld";

fprintf(fp,"%s",str);

if(!strcmp(argv[2],"return"))

{

return 0;

}else if(!strcmp(argv[2],"exit"))

{

exit(0);

}else if(!strcmp(argv[2],"\_exit"))

{

\_exit(0);

}

else

{

printf("process error\n");

}

return 0;

}

##### 出口状态

有两种方式：exit(0)和return 0;

通常返回值0/1采用宏定义SUCCESS或FILED。

##### 终止前的清理

实例：

#include <stdio.h>

#include <stdlib.h>

void fun(void)

{

//printf("clean something!!!\n");

}

int main(int argc,char \*\*argv)

{

printf("hello world");

atexit(fun);//注册函数

atexit(fun);//注册函数

exit(EXIT\_SUCCESS);

return 0;

}

#### 异常终止

调用abort（与信号有关）

接受到一个信号并终止（和信号有关）

最后一个线程对取消请求作出响应

##### abort函数

实例：

#include <stdio.h>

#include <stdlib.h>

void fun(void)

{

printf("clear someting!!!\n");

}

int main(int argc,char \*\*argv)

{

printf("hello world");

atexit(fun);

abort();

return 0;

}

#include <stdio.h>

#include <stdlib.h>

void fun(void)

{

printf("clean something!!!\n");

}

int main(int argc,char \*\*argv)

{

printf("hello world");

atexit(fun);//注册函数

atexit(fun);//注册函数

\_exit(EXIT\_SUCCESS);

return 0;

}

##### core文件

##### 信号

### 进程存储空间

#### 存储结构



#### 进程地址空间



#### 内存分配/释放

堆上分配

边界对齐

分配大小合适

内存泄露

实例：

#include <stdio.h>

#include <stdlib.h>

#define MAX 100

int main(int argc,char \*\*argv)

{

char \*buf = NULL;

//申请100字节内存，堆

buf = malloc(MAX);

if ( NULL == buf )

{

perror("malloc");

return -1;

}

while(1)

{

//memset 将内存清0

memset(buf,'\0',MAX);

fgets(buf,MAX,stdin);

if(0==strncmp("exit",buf,4))

{

exit(0);

}

printf("%s\n",buf);

}

free(buf);

buf = NULL;

}

### 非局部转移

非局部跳转应注意的问题：

1. 自动，寄存器和易失变量

寄存器中的变量在非局部跳转时可以恢复原始值。

1. 自动变量的潜在问题

#### setjmp

int setjmp(jmp\_buf env);

返回：直接调用返回0，若从longjump调用返回，则返回非0值

#### longjmp

void longjmp(jmp\_buf env, int val);

实例：

#include <stdio.h>

#include <setjmp.h>

static jmp\_buf jmpbuf;

int a = 0;

void fun(int a,int b)

{

printf("in fun a is %d, b is %d\n",a,b);

longjmp(jmpbuf,3);

}

int main(int argc,char \*\*argv)

{

int b = 1;

if( 3 == setjmp(jmpbuf) )

{

printf("a is %d\nb is %d\n",a,b);

exit(0);

}

a = 2;

b = 3;

fun(a,b);

return 0;

}

### 进程的资源

#### 资源限制

Linux中可用的资源限制：

RLIMIT\_AS：进程可用存储区大小

RLIMIT\_CORE：core文件最大字节数

RLIMIT\_CPU：CPU时间最大值

RLIMIT\_DATA：数据段最大长度

RLIMIT\_FSIZE：可创建文件的最大长度

RLIMIT\_LOCKS：文件锁的最大数

RLIMIT\_MEMLOCK：使用mlock能否在存储器中锁定的最长字节数

RLIMIT\_NOFILE：能打开的最大文件数

RLIMIT\_NPROC：每个用户ID可拥有的最大子进程数

RLIMIT\_RSS：最大驻内存集的字节长度

RLIMIT\_STACH：栈的最大长度

#### 查看资源

##### 虚拟空间的最大字节

##### core文件的最大字节

##### CPU时间最大值

##### 数据段空间的最大字节

##### 创建的最大文件的字节数

##### 在物理内存中上锁的最大虚拟存储字节数

##### 最多能够打开的文件个数

##### 创建的子进程个数

##### 最大物理内存页数

##### 栈的最大字节数

##### 不同资源的个数

实例：

#include <stdio.h>

#include <stdlib.h>

#include <sys/time.h>

#include <sys/resource.h>

int main(int argc,char \*\*argv)

{

struct rlimit limit;

getrlimit(RLIMIT\_CORE,&limit);

printf("soft limit core is %d\n",limit.rlim\_cur);

printf("hard limit core is %d\n",limit.rlim\_max);

limit.rlim\_cur = limit.rlim\_max;

setrlimit(RLIMIT\_CORE,&limit);

abort();

return 0;

}

#### 设置资源

进程资源限制：

#include<sys/resource.h>

int getlimit(int resource, struct rlimit \*rlptr);

int setlimit(int resource, const struct rlimit \*rlptr);

返回：成功返回0，出错返回非0

struct rlimit{

rlim\_t rlim\_cur;/\*soft limit,current limit\*/

rlim\_t rlim\_max;/\*hard limit,maxium limit\*/

}

资源限制的修改规则：

1. 任何一个进程可以降低或者提升其软资源限制、但必须小于等于其硬资源限制
2. 任何一个进程可以降低其硬件资源限制，但必须大于其软限制，普通用户不可逆此操作
3. 只有root用户可以提升硬件资源限制

Linux中进程资源的初始值由进程0建立，并被后续进程继承。

实例：

#include <sys/resource.h>

#include <stdio.h>

#include <errno.h>

#include <stdlib.h>

#include <string.h>

#define LIMIT(resource)\

query\_limit(#resource,resource)

void query\_limit(char \*name,int res)

{

struct rlimit limit;

if(getrlimit(res,&limit))

{

fprintf(stderr,"getrlimit error %s\n",strerror(errno));

exit(1);

}

printf("%15s ",name);

if(limit.rlim\_cur == RLIM\_INFINITY)

{

printf(" soft (INFINITY)");

}else

{

printf(" soft (%8ld)",limit.rlim\_cur);

}

if(limit.rlim\_max == RLIM\_INFINITY)

{

printf(" hard (INFINITY)");

}else

{

printf(" hard (%8ld)",limit.rlim\_max);

}

printf("\n");

}

int main(void)

{

LIMIT(RLIMIT\_AS);

LIMIT(RLIMIT\_CORE);

LIMIT(RLIMIT\_CPU);

LIMIT(RLIMIT\_DATA);

LIMIT(RLIMIT\_FSIZE);

LIMIT(RLIMIT\_LOCKS);

LIMIT(RLIMIT\_MEMLOCK);

LIMIT(RLIMIT\_NOFILE);

struct rlimit lt;

lt.rlim\_cur = 4;

lt.rlim\_max = 1024;

if(setrlimit(RLIMIT\_NOFILE,&lt))

{

fprintf(stderr,"setrlimit error %s\n",strerror(errno));

exit(1);

}

printf("----------------------------\n");

LIMIT(RLIMIT\_NOFILE);

FILE \*fp1 = fopen("a.txt","r");

if(fp1 == NULL)

{

fprintf(stderr,"file:%s\n",strerror(errno));

exit(1);

}else

{

printf("open sucessed\n");

}

FILE \*fp2 = fopen("a.txt","r");

if(fp2 == NULL)

{

fprintf(stderr,"file:%s\n",strerror(errno));

exit(1);

}else

{

printf("open sucessed\n");

}

return 0;

}

#### 统计资源

### 用户信息

#### 获取用户名

#### 用户数据库管理

#### 组数据库管理

实例：

#include <stdio.h>

#include <unistd.h>

int main(int argc,char \*\*argv)

{

printf("%s\n",getlogin());

return 0;

}

### 进程身份及调整

#### 身份凭证

#### 调整身份凭证

## 进程调度

任何进程要想占有CPU，从而真正处于执行状态，就必须经由进程调度。进程调度机制主要涉及到调度方式、调度时机和调度策略。

### 调度方式

Linux内核的调度方式基本上采用“抢占式优先级”方式，即当进程在用户模式下运行时，不管是否自愿，在一定条件下(如时间片用完或等待I/O)，核心就可以暂时剥夺其运行而调度其它进程进入运行。但是，一旦进程切换到内核模式下运行，就不受以上限制而一直运行下去，直至又回到用户模式之前才会发生进程调度。

Linux系统中的调度策略基本上继承了Unix的以优先级为基础的调度。就是说，核心为系统中每个进程计算出一个优先权，该优先权反映了一个进程获得CPU使用权的资格，即高优先权的进程优先得到运行。核心从进程就绪队列中挑选一个优先权最高的进程，为其分配一个CPU时间片，令其投入运行。在运行过程中，当前进程的优先权随时间递减，这样就实现了“负反馈”作用：经过一段时间之后，原来级别较低的进程就相对“提升”了级别，从而有机会得到运行。当所有进程的优先权都变为0时，就重新计算一次所有进程的优先权。

### 调度策略

Linux系统针对不同类别的进程提供了三种不同的调度策略，即SCHED\_FIFO、SCHED\_RR及SCHED\_OTHER。

SCHED\_FIFO适合于实时进程，它们对时间性要求比较强，而每次运行所需的时间比较短，一旦这种进程被调度开始运行后，就要一直运行到自愿让出CPU，或者被优先权更高的进程抢占其执行权为止。

SCHED\_RR对应“时间片轮转法”，适合于每次运行需要较长时间的实时进程。一个运行进程分配一个时间片(如200毫秒)，当时间片用完后，CPU被另外进程抢占，而该进程被送回相同优先级队列的末尾。SCHED\_OTHER是传统的Unix调度策略，适合于交互式的分时进程。这类进程的优先权取决于两个因素，一个因素是进程剩余时间配额，如果进程用完了配给的时间，则相应优先权为0；另一个是进程的优先数nice，这是从Unix系统沿袭下来的方法，优先数越小，其优先级越高。

nice的取值范围是19－20。用户可以利用nice命令设定进程的nice值。但一般用户只能设定正值，从而主动降低其优先级；只有特权用户才能把nice的值置为负数。进程的优先权就是以上二者之和。核心动态调整用户态进程的优先级。这样，一个进程从创建到完成任务后终止，需要经历多次反馈循环。当进程再次被调度运行时，它就从上次断点处开始继续执行。对于实时进程，其优先权的值是(1000＋设定的正值)，因此，至少是1000。所以，实时进程的优先权高于其它类型进程的优先权。另外，时间配额及nice值与实时进程的优先权无关。如果系统中有实时进程处于就绪状态，则非实时进程就不能被调度运行，直至所有实时进程都完成了，非实时进程才有机会占用CPU。

后台命令（在命令末尾有&符号，如gcc f1.c& )对应后台进程（又称后台作业），后台进程的优先级低于任何交互（前台）进程的优先级。所以，只有当系统中当前不存在可运行的交互进程时，才调度后台进程运行。后台进程往往按批处理方式调度运行。

### 调度时机

核心进行进程调度的时机有以下几种情况：

(1)、当前进程调用系统调用nanosleep( )或pause( )使自己进入睡眠状态，主动让出一段时间的CPU使用权；

(2)、进程终止，永久地放弃对CPU的使用；

(3)、在时钟中断处理程序执行过程中，发现当前进程连续运行的时间过长；

(4)、当唤醒一个睡眠进程时，发现被唤醒的进程比当前进程更有资格运行；

(5)、一个进程通过执行系统调用来改变调度策略或降低自身的优先权(如nice命令)，从而引起立即调度。

### 调度算法

进程调度的算法应该比较简单，以便减少频繁调度时的系统开销。Linux执行进程调度时，首先查找所有在就绪队列中的进程，从中选出优先级最高且在内存的一个进程。如果队列中有实时进程，那么实时进程将优先运行。如果最需要运行的进程不是当前进程，那么当前进程就被挂起，并且保存它的现场所涉及的一切机器状态，包括计数器和CPU寄存器等，然后为选中的进程恢复运行现场。

## 僵尸进程

进程停止后，该进程就会从进程表中移除。但是，有时候有些程序即使执行完了也依然留在进程表中。那么，这些完成了生命周期但却依然留在进程表中的进程，我们称之为“僵尸进程”。

### 产生

当你运行一个程序时，它会产生一个父进程以及很多子进程。所有这些子进程都会消耗内核分配给它们的内存和CPU资源。

这些子进程完成执行后会发送一个exit信号然后死掉。这个exit信号需要被父进程所读取。父进程需要随后调用 wait 命令来读取子进程的退出状态，并将子进程从进程表中移除。

若父进程正确第读取了子进程的exit信号，则子进程会从进程表中删掉。

但若父进程未能读取到子进程的exit信号，则这个子进程虽然完成执行处于死亡的状态，但也不会从进程表中删掉。

### 危害

由于僵尸进程并不做任何事情，不会使用任何资源也不会影响其它进程， 因此存在僵尸进程也没什么坏处。不过由于进程表中的退出状态以及其它一些进程信息也是存储在内存中的，因此存在太多僵尸进程有时也会是一些问题。

你可以想象成这样：

“你是一家建筑公司的老板。你每天根据工人们的工作量来支付工资。 有一个工人每天来到施工现场，就坐在那里， 你不用付钱， 他也不做任何工作。 他只是每天都来然后呆坐在那，仅此而已！”

这个工人就是僵尸进程的一个活生生的例子。但是， 如果你有很多僵尸工人， 你的建设工地就会很拥堵从而让那些正常的工人难以工作。

### 查找

打开终端并输入下面命令:

ps aux | grep Z

会列出进程表中所有僵尸进程的详细内容。

### 杀死

如何杀掉僵尸进程？

正常情况下我们可以用 SIGKILL 信号来杀死进程，但是僵尸进程已经死了， 你不能杀死已经死掉的东西。 因此你需要输入的命令应该是

kill -s SIGCHLD pid

将这里的 pid 替换成父进程的进程 id，这样父进程就会删除所有以及完成并死掉的子进程了。

你可以把它想象成：

“你在道路中间发现一具尸体，于是你联系了死者的家属，随后他们就会将尸体带离道路了。”

不过许多程序写的不是那么好，无法删掉这些子僵尸（否则你一开始也见不到这些僵尸了）。 因此确保删除子僵尸的唯一方法就是杀掉它们的父进程。

# 线程

## 概述

线程，有时也被称为轻量级进程，是程序执行流的最小单元，是进程中的一个实体，是被系统独立调度和分派的基本单位。与进程不同，线程与资源分配无关，线程自己不拥有系统资源，它属于某一个进程，并与进程内的其他线程一起共享进程的资源。线程只由相关堆栈（系统栈或用户栈）寄存器和线程控制表TCB组成。

在许多类Unix系统中，如Linux、FreeBSD、Solaris等，进程一直都是操作系统内核调用的最小单位，也都采用多进程模型。后来引入了线程概念，有以下两种概念的线程：

## 用户级线程

用户级线程(User-Level Thread，ULT)。由应用进程利用线程库创建和管理，不在内核中实现线程，只在用户态中模拟出多线程，不依赖于核心，操作系统内核完全不知道多线程的存在。

## 内核线线程

内核线线程(Kernel-Level Thread，KLT)，又称为内核支持的线程或轻量级进程。是在核心空间实现的，内核为每个线程在核心空间中设置了一个线程控制块，用来登记该线程的线程标识符、值、状态、优先级等信息，所有对线程的操作，如创建、撤销和切换都是通过系统功能调用由内核中的相应处理完成，内核维护进程及线程的上下文切换以及线程切换，类系统中一般通过修改进程的实现方式来实现，可以使用不完全的进程创建方式创建共享数据空间的进程，在 Linux下这种系统调用为clone()，而在FreeBSD下它为rfork()。

# 联系

通常在一个进程中可以包含若干个线程，它们可以利用进程所拥有的资源。但是，一个线程只属于一个进程。进程间相互独立，同一进程的各线程间共享。某进程内的线程在其它进程不可见。而且需要注意的是，线程不是一个可执行的实体。

# 区别

进行和线程之间的差异可以从下面几个方面来阐述：

调度 ：在引入线程的操作系统中，线程是调度和分配的基本单位 ，进程是资源拥有的基本单位 。把传统进程的两个属性分开，线程便能轻装运行，从而可 显著地提高系统的并发程度 。在同一进程中，线程的切换不会引起进程的切换；在由一个进程中的线程切换到另一个进程中的线程时，才会引起进程的切换。

并发性 ：在引入线程的操作系统中，不仅进程之间可以并发执行，而且在一个进程中的多个线程之间亦可并发执行，因而使操作系统具有更好的并发性，从而能 更有效地使用系统资源和提高系统吞吐量。

拥有资源 ：不论是传统的操作系统，还是设有线程的操作系统，进程都是拥有资源的一个独立 单位，它可以拥有自己的资源。一般地说，线程自己不拥有系统资源（只有一些必不可少的资源，但它可以访问其隶属进程的资源。

系统开销：由于在创建或撤消进程时，系统都要为之分配或回收资源，因此，操作系统所付出的开销将显著地大于在创建或撤消线程时的开销。进程切换的开销也远大于线程切换的开销。

通信：进程间通信IPC，线程间可以直接读写进程数据段（如全局变量）来进行通信——需要进程同步和互斥手段的辅助，以保证数据的一致性，因此共享简单。但是线程的数据同步要比进程略复杂。

# 选择

需要频繁创建销毁的优先使用线程;因为对进程来说创建和销毁一个进程代价是很大的。

线程的切换速度快，所以在需要大量计算，切换频繁时用线程，还有耗时的操作使用线程可提高应用程序的响应

因为对CPU系统的效率使用上线程更占优，所以可能要发展到多机分布的用进程，多核分布用线程;

并行操作时使用线程，如C/S 的服务器端并发线程响应用户的请求;

需要更稳定安全时，适合选择进程;需要速度时，选择线程更好。

# 总结

从上面的分析可以看到，似乎线程有很多优势，比如，数据共享效率高，可应对并发操作，有效利用等待时间等等，但是多线程的编程比多进程要复杂，同时，多进程的可靠性较好，因为进程间不会相互影响。实际情况还是需要自己分析拿捏的。但是一般来说，实际应用中常常采用“进程+线程”结合的方式，而不是非此即彼，因为它们两者没有绝对的好与不好，而是适合于不同场景。

4.1 用户模式+内核模式

一般说来，一个进程在CPU上运行可以有两种运行模式，既可在用户模式下运行，又可在内核模式下运行（即进程分别工作在用户态和内核态，在内核态工作仍旧是这个进程，除非进行了进程的切换）。通常操作系统把虚拟地址空间划分为用户空间和内核空间，例如x86平台的Linux系统虚拟地址空间是0x00000000~0xffffffff，前3GB（0x00000000~0xbfffffff）是用户空间，后1GB（0xc0000000~0xffffffff）是内核空间。用户加载到用户空间，在用户模式下执行，不能访问内核中的数据，也不能跳转到内核代码中执行。这样可以保护内核，如果一个进程访问了非法地址，顶多这一个进程崩溃，而不会影响到内核和整个系统的稳定性。Cpu在产生中断或异常时不仅会跳转到中断或异常服务城西，还会自动切换模式，从用户模式切换到特权模式，因此从中断或异常程序可以跳转到内核代码中执行。事实上，整个内核就是由各种中断和异常处理程序组成的。即，正常情况下处理器在用户模式执行用户程序，在中断或异常情况下处理器切换到特权模式执行内核程序，处理完中断或异常之后再返回用户模式继续执行用户程序，例如，用户进程A调用了内核系统调用来获取当前的时钟滴答数，在执行用户进程A中的系统调用指令时会保存当前用户进程的IP，CS等当前状态，然后再跳转到内核空间（即内核代码区域）去执行像应的系统调用函数，获取当前的时钟滴答数。执行完后再通过IRET指令返回到进程A中（就是将进入时保存的信息再复位到相应的寄存器中），再接着从CS：EIP地址开始执行A进程的指令

进程在创建的时候除了创建进程的控制块之外，在内核里还创建了进程的内核栈，进程通过系统调用（例如fopen()或者open()）进入内核后，此时处理器处于特权级最高的（0级）内核代码中执行，当进程处于内核态时，执行的内核代码会使用当前进程的内核栈，是指向在进程的上下文上的，

内核模式的权限高于用户模式的权限。

用户级。系统用户可以与进行交互操作，如运行应用和系统命令，用户级通过系统调用接口访问内核级；内核级。操作系统自动运行一些功能，它们主要对硬件进行操作