本节重点：

> - 学习进程创建,fork/vfork
> - 学习到进程等待
> - 学习到进程程序替换, 微型shell，重新认识shell运行原理
> - 学习到进程终止,认识$?

进程创建

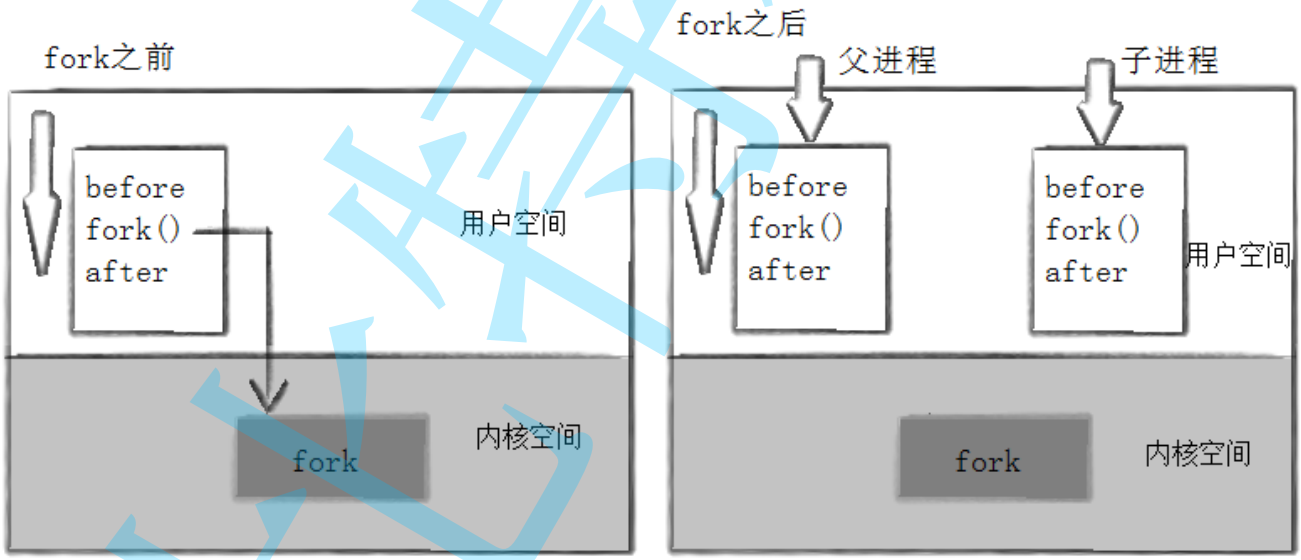
fork函数初识

在linux中fork函数时非常重要的函数，它从已存在进程中创建一个新进程。新进程为子进程，而原进程为父进程。

```
#include <unistd.h>
pid_t fork(void);
返回值：自进程中返回0，父进程返回子进程id，出错返回-1
```

进程调用fork，当控制转移到内核中的fork代码后，内核做：

> - 分配新的内存块和内核数据结构给子进程
> - 将父进程部分数据结构内容拷贝至子进程
> - 添加子进程到系统进程列表当中
> - fork返回，开始调度器调度

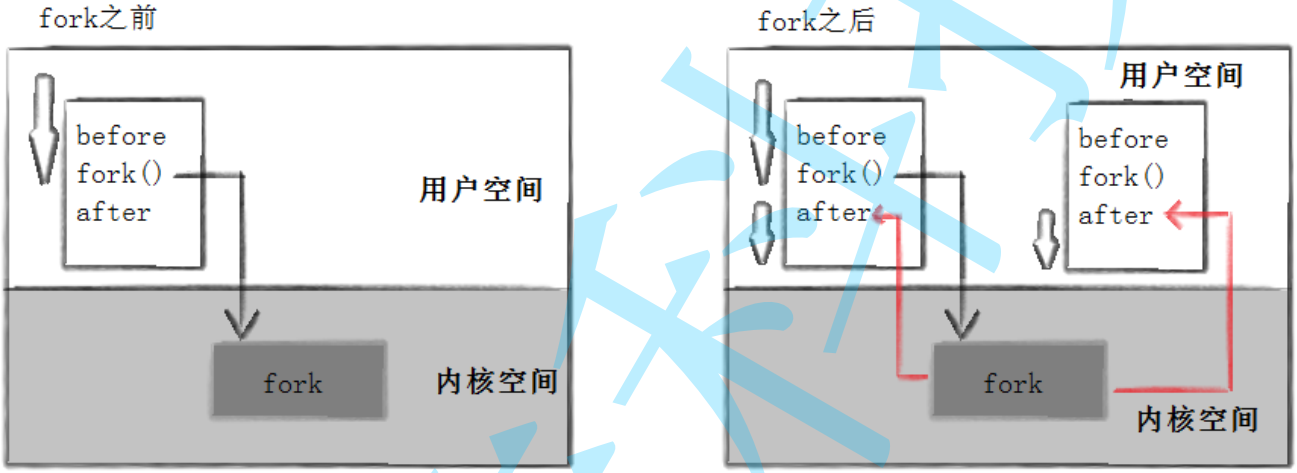


当一个进程调用fork之后，就有两个二进制代码相同的进程。而且它们都运行到相同的地方。但每个进程都将可以开始它们自己的旅程，看如下程序。

```
int main( void )
{
   pid_t pid;


   printf("Before: pid is %d\n", getpid());
```

```
    if ( (pid=fork()) == -1 )perror("fork()"),exit(1);
    printf("After:pid is %d, fork return %d\n", getpid(), pid);
     sleep(1);
    return 0;
}

运行结果:
[root@localhost linux]# ./a.out
Before: pid is 43676
After:pid is 43676, fork return 43677
After:pid is 43677, fork return 0
```

这里看到了三行输出，一行before，两行after。进程43676先打印before消息，然后它有打印after。另一个after消息有43677打印的。注意到进程43677没有打印before，为什么呢？如下图所示



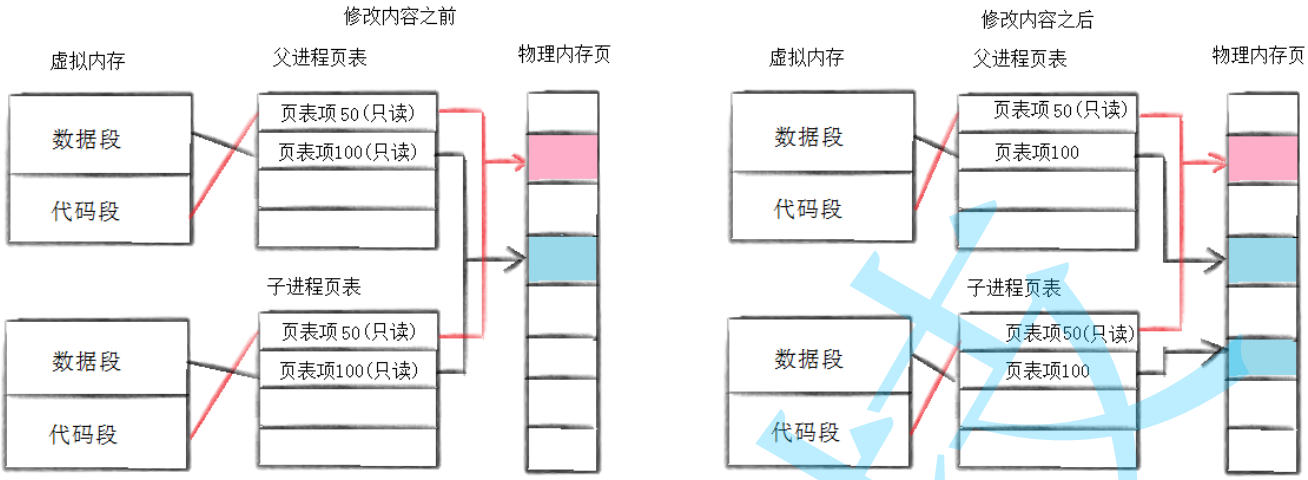
所以，fork之前父进程独立执行，fork之后，父子两个执行流分别执行。注意，fork之后，谁先执行完全由调度器决定。

fork函数返回值

> - 子进程返回0，
> - 父进程返回的是子进程的pid。

写时拷贝

通常，父子代码共享，父子再不写入时，数据也是共享的，当任意一方试图写入，便以写时拷贝的方式各自一份副本。具体见下图:



fork常规用法

> - 一个父进程希望复制自己，使父子进程同时执行不同的代码段。例如，父进程等待客户端请求，生成子进程来处理请求。
> - 一个进程要执行一个不同的程序。例如子进程从fork返回后，调用exec函数。

fork调用失败的原因

> - 系统中有太多的进程
> - 实际用户的进程数超过了限制

进程终止

进程退出场景

> - 代码运行完毕，结果正确
> - 代码运行完毕，结果不正确
> - 代码异常终止

进程常见退出方法

正常终止（可以通过 `echo $?` 查看进程退出码）：

> 1. 从main返回
> 2. 调用exit
> 3. _exit

异常退出：

> - ctrl + c，信号终止

_exit函数

```
#include <unistd.h>
void _exit(int status);

参数：status 定义了进程的终止状态，父进程通过wait来获取该值
```

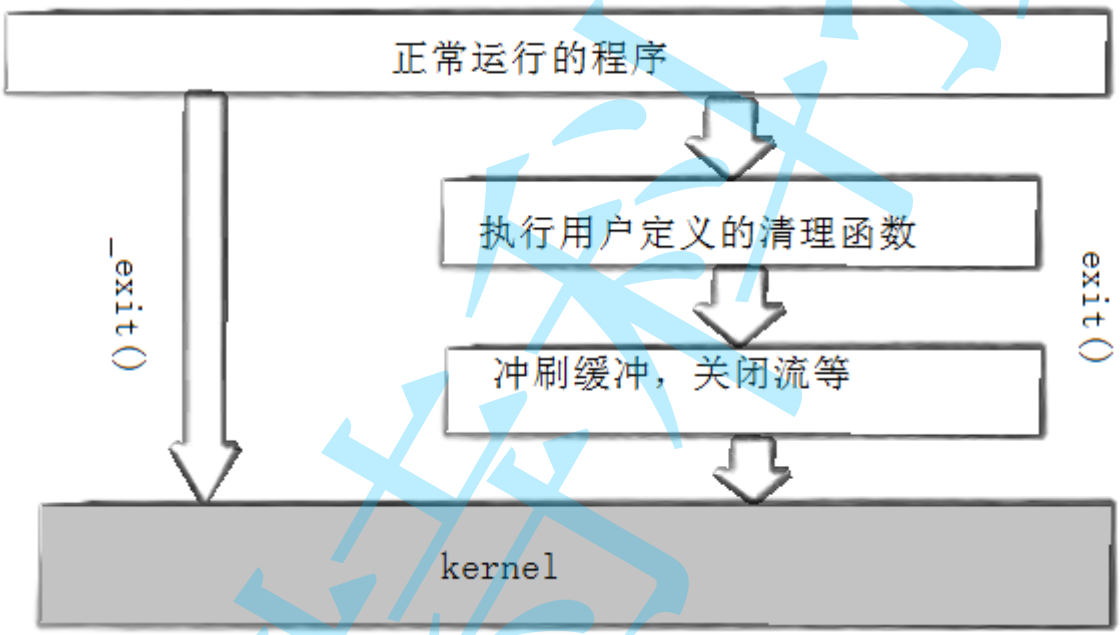
- 说明：虽然status是int，但是仅有低8位可以被父进程所用。所以_exit(-1)时，在终端执行$?发现返回值是255。

exit函数

```
#include <unistd.h>
void exit(int status);
```

exit最后也会调用*exit, 但在调用*exit之前，还做了其他工作：

1. 执行用户通过 atexit或on_exit定义的清理函数。
2. 关闭所有打开的流，所有的缓存数据均被写入
3. 调用_exit



实例：

```
int main()
{
    printf("hello");
    exit(0);
}
运行结果：
[root@localhost linux]# ./a.out
hello[root@localhost linux]#

int main()
{
    printf("hello");
    _exit(0);
}
运行结果：
[root@localhost linux]# ./a.out
[root@localhost linux]#
```

return退出

return是一种更常见的退出进程方法。执行return n等同于执行exit(n),因为调用main的运行时函数会将main的返回值当做 exit的参数。

进程等待

进程等待必要性

> - 之前讲过，子进程退出，父进程如果不管不顾，就可能造成‘僵尸进程’的问题，进而造成内存泄漏。
> - 另外，进程一旦变成僵尸状态，那就刀枪不入，“杀人不眨眼”的kill -9 也无能为力，因为谁也没有办法杀死一个已经死去的进程。
> - 最后，父进程派给子进程的任务完成的如何，我们需要知道。如，子进程运行完成，结果对还是不对，或者是否正常退出。
> - 父进程通过进程等待的方式，回收子进程资源，获取子进程退出信息

进程等待的方法

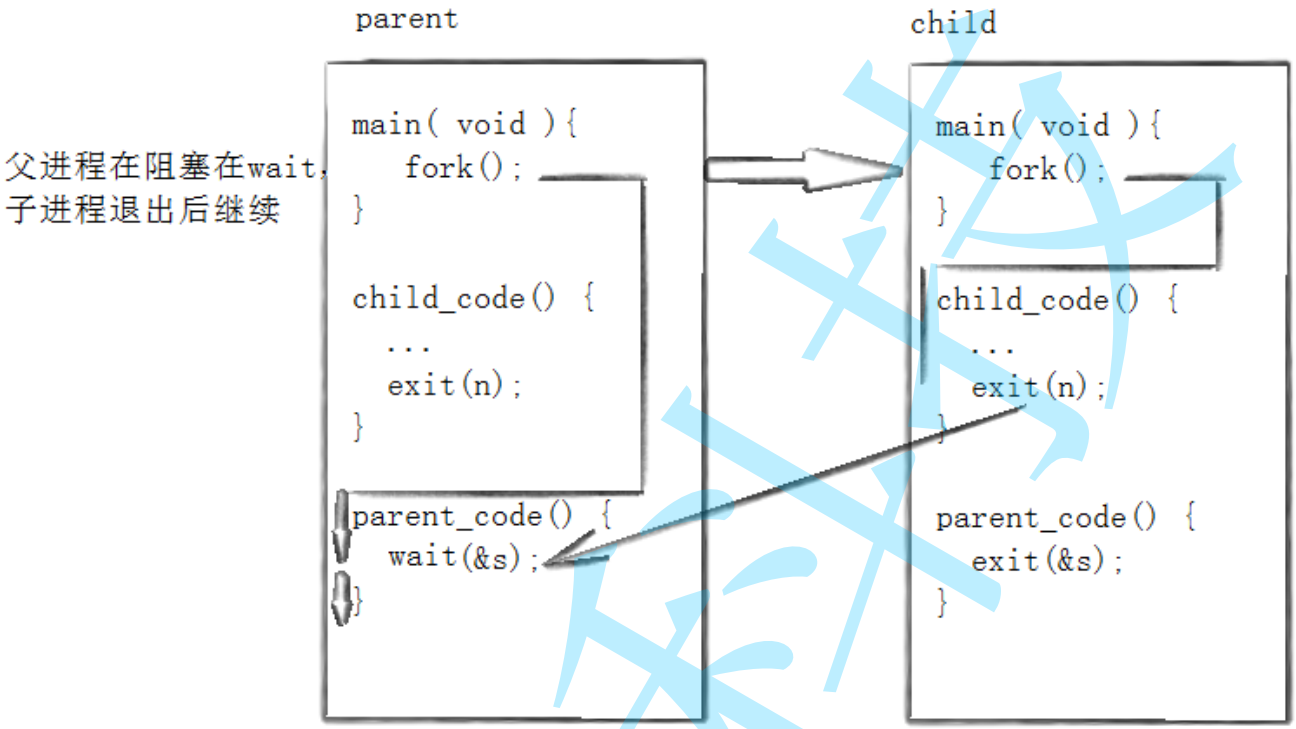
wait方法

```
#include<sys/types.h>
#include<sys/wait.h>

pid_t wait(int*status);

返回值:
    成功返回被等待进程pid，失败返回-1。
参数:
    输出型参数，获取子进程退出状态,不关心则可以设置成为NULL
```

waitpid方法

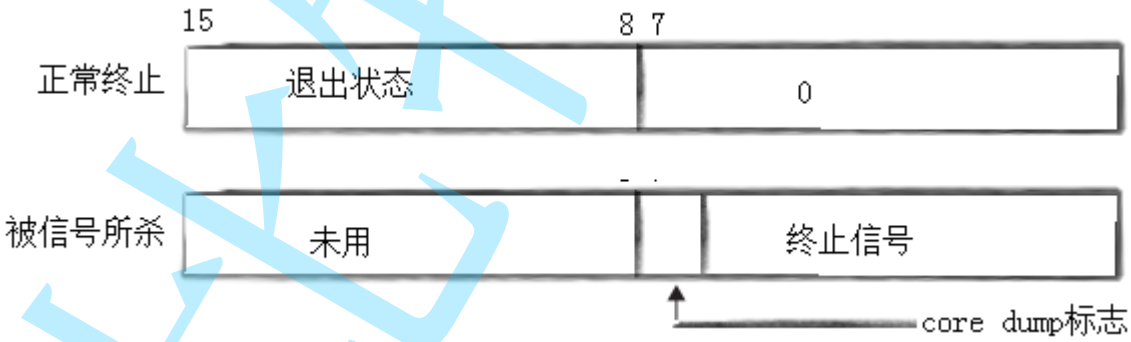
```
pid_ t waitpid(pid_t pid, int *status, int options);
返回值:
    当正常返回的时候waitpid返回收集到的子进程的进程ID;
    如果设置了选项WNOHANG,而调用中waitpid发现没有已退出的子进程可收集,则返回0;
    如果调用中出错,则返回-1,这时errno会被设置成相应的值以指示错误所在;
参数:
    pid:
        Pid=-1,等待任一个子进程。与wait等效。
        Pid>0.等待其进程ID与pid相等的子进程。
    status:
        WIFEXITED(status): 若为正常终止子进程返回的状态，则为真。（查看进程是否是正常退出）
        WEXITSTATUS(status): 若WIFEXITED非零，提取子进程退出码。（查看进程的退出码）
    options:
        WNOHANG: 若pid指定的子进程没有结束，则waitpid()函数返回0，不予以等待。若正常结束，则返回该子进程的ID。
```

- 如果子进程已经退出，调用wait/waitpid时，wait/waitpid会立即返回，并且释放资源，获得子进程退出信息。
- 如果在任意时刻调用wait/waitpid，子进程存在且正常运行，则进程可能阻塞。
- 如果不存在该子进程，则立即出错返回。



获取子进程status

- wait和waitpid，都有一个status参数，该参数是一个输出型参数，由操作系统填充。
- 如果传递NULL，表示不关心子进程的退出状态信息。
- 否则，操作系统会根据该参数，将子进程的退出信息反馈给父进程。
- status不能简单的当作整形来看待，可以当作位图来看待，具体细节如下图（只研究status低16比特位）：



测试代码：

```
#include <sys/wait.h>
#include <stdio.h>
#include <stdlib.h>
#include <string.h>
#include <errno.h>

int main( void )
```

```
{
    pid_t pid;

    if ( (pid=fork()) == -1 )
        perror("fork"),exit(1);

    if ( pid == 0 ){
        sleep(20);
        exit(10);
    } else {
        int st;
        int ret = wait(&st);

        if ( ret > 0 && ( st & 0X7F ) == 0 ){ // 正常退出
            printf("child exit code:%d\n", (st>>8)&0XFF);
        } else if( ret > 0 ) {  // 异常退出
            printf("sig code : %d\n", st&0X7F );
        }
    }
}

测试结果:
    [root@localhost linux]# ./a.out #等20秒退出
    child exit code:10
    [root@localhost linux]# ./a.out #在其他终端kill掉
    sig code : 9
```

具体代码实现

> - 进程的阻塞等待方式:

```
int main()
{
    pid_t pid;
    pid = fork();
    if(pid < 0){
        printf("%s fork error\n",__FUNCTION__);
        return 1;
    } else if( pid == 0 ){ //child
        printf("child is run, pid is : %d\n",getpid());
        sleep(5);
        exit(257);
    } else{
        int status = 0;
        pid_t ret = waitpid(-1, &status, 0);//阻塞式等待，等待5S
        printf("this is test for wait\n");
        if( WIFEXITED(status) && ret == pid ){
            printf("wait child 5s success, child return code is :%d.\n",WEXITSTATUS(status));
        }else{
            printf("wait child failed, return.\n");
            return 1;

        }
```

```
    }
    return 0;
}

运行结果:
[root@localhost linux]# ./a.out
child is run, pid is : 45110
this is test for wait
wait child 5s success, child return code is :1.
```

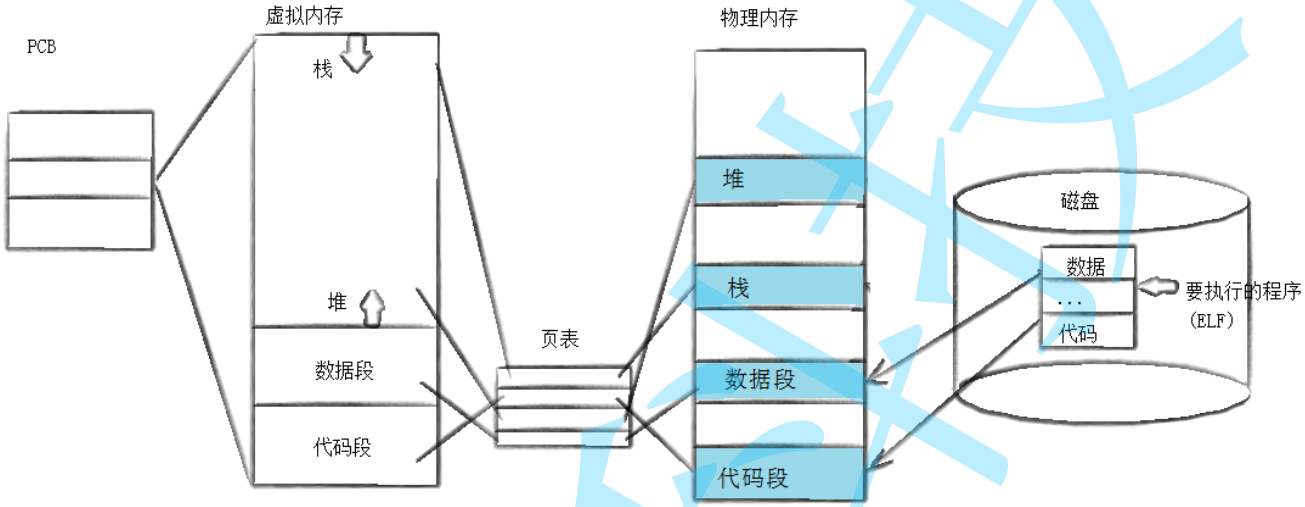
- 进程的非阻塞等待方式:

```
#include <stdio.h>
#include <unistd.h>
#include <stdlib.h>
#include <sys/wait.h>

int main()
{
    pid_t pid;

    pid = fork();
    if(pid < 0){
        printf("%s fork error\n",__FUNCTION__);
        return 1;
    }else if( pid == 0 ){ //child
        printf("child is run, pid is : %d\n",getpid());
        sleep(5);
        exit(1);
    } else{
        int status = 0;
        pid_t ret = 0;
        do
        {
            ret = waitpid(-1, &status, WNOHANG);//非阻塞式等待
            if( ret == 0 ){
                printf("child is running\n");
            }
            sleep(1);
        }while(ret == 0);

        if( WIFEXITED(status) && ret == pid ){
            printf("wait child 5s success, child return code is :%d.\n",WEXITSTATUS(status));
        }else{
            printf("wait child failed, return.\n");
            return 1;
        }
    }
    return 0;
}
```

进程程序替换

替换原理

用fork创建子进程后执行的是和父进程相同的程序(但有可能执行不同的代码分支),子进程往往要调用一种exec函数以执行另一个程序。当进程调用一种exec函数时,该进程的用户空间代码和数据完全被新程序替换,从新程序的启动例程开始执行。调用exec并不创建新进程,所以调用exec前后该进程的id并未改变。



替换函数

其实有六种以exec开头的函数,统称exec函数:

```
#include <unistd.h>`

int execl(const char *path, const char *arg, ...);
int execlp(const char *file, const char *arg, ...);
int execle(const char *path, const char *arg, ...,char *const envp[]);
int execv(const char *path, char *const argv[]);
int execvp(const char *file, char *const argv[]);
```

int execve(const char *path, char *const argv[], char *const envp[]);

函数解释

> - 这些函数如果调用成功则加载新的程序从启动代码开始执行,不再返回。
> - 如果调用出错则返回-1
> - 所以exec函数只有出错的返回值而没有成功的返回值。

命名理解

这些函数原型看起来很容易混,但只要掌握了规律就很好记。

> - l(list) : 表示参数采用列表
> - v(vector) : 参数用数组
> - p(path) : 有p自动搜索环境变量PATH
> - e(env) : 表示自己维护环境变量

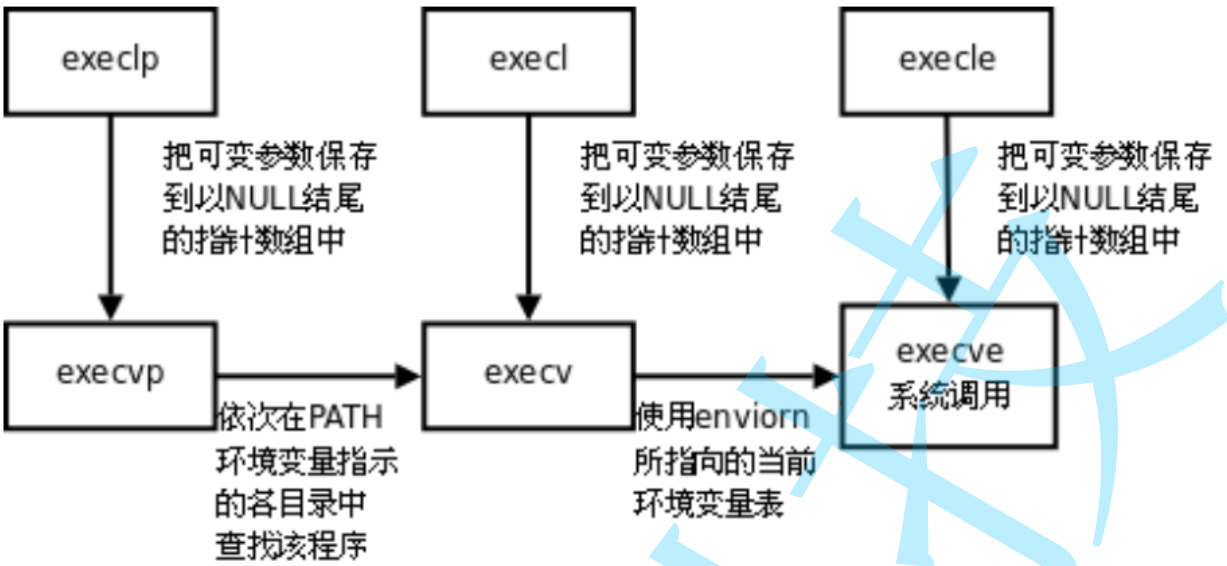
函数名	参数格式	是否带路径	是否使用当前环境变量
execl	列表	不是	是
execlp	列表	是	是
execle	列表	不是	不是，须自己组装环境变量
execv	数组	不是	是
execvp	数组	是	是
execve	数组	不是	不是，须自己组装环境变量

exec调用举例如下:

```
#include <unistd.h>

int main()
{
    char *const argv[] = {"ps", "-ef", NULL};
    char *const envp[] = {"PATH=/bin:/usr/bin", "TERM=console", NULL};

    execl("/bin/ps", "ps", "-ef", NULL);

    // 带p的，可以使用环境变量PATH，无需写全路径
    execlp("ps", "ps", "-ef", NULL);

    // 带e的，需要自己组装环境变量
    execle("ps", "ps", "-ef", NULL, envp);

    execv("/bin/ps", argv);

    // 带p的，可以使用环境变量PATH，无需写全路径
    execvp("ps", argv);

    // 带e的，需要自己组装环境变量
    execve("/bin/ps", argv, envp);

    exit(0);
}
```

事实上,只有execve是真正的系统调用,其它五个函数最终都调用 execve,所以execve在man手册 第2节,其它函数在man手册第3节。这些函数之间的关系如下图所示。

下图exec函数族 一个完整的例子:

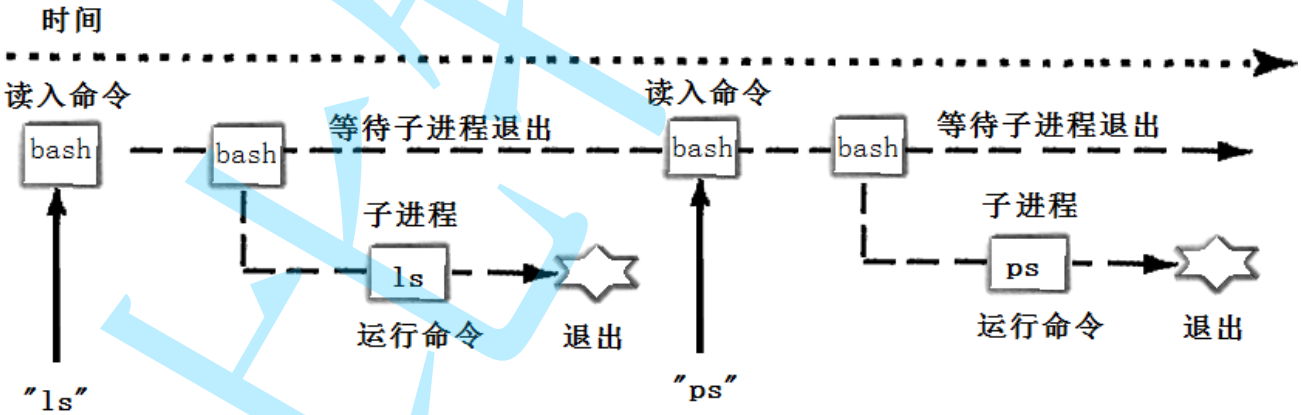


我们可以综合前面的知识，做一个简易的shell

考虑下面这个与shell典型的互动：

```
[root@localhost epoll]# ls
client.cpp  readme.md  server.cpp  utility.h
[root@localhost epoll]# ps
  PID TTY          TIME CMD
 3451 pts/0    00:00:00 bash
 3514 pts/0    00:00:00 ps
```

用下图的时间轴来表示事件的发生次序。其中时间从左向右。shell由标识为sh的方块代表，它随着时间的流逝从左向右移动。shell从用户读入字符串"ls"。shell建立一个新的进程，然后在那个进程中运行ls程序并等待那个进程结束。



然后shell读取新的一行输入，建立一个新的进程，在这个进程中运行程序 并等待这个进程结束。
所以要写一个shell，需要循环以下过程：

> 1. 获取命令行
> 2. 解析命令行
> 3. 建立一个子进程（fork）
> 4. 替换子进程（execvp）

> 5. 父进程等待子进程退出（wait）

根据这些思路，和我们前面的学的技术，就可以自己来实现一个shell了。

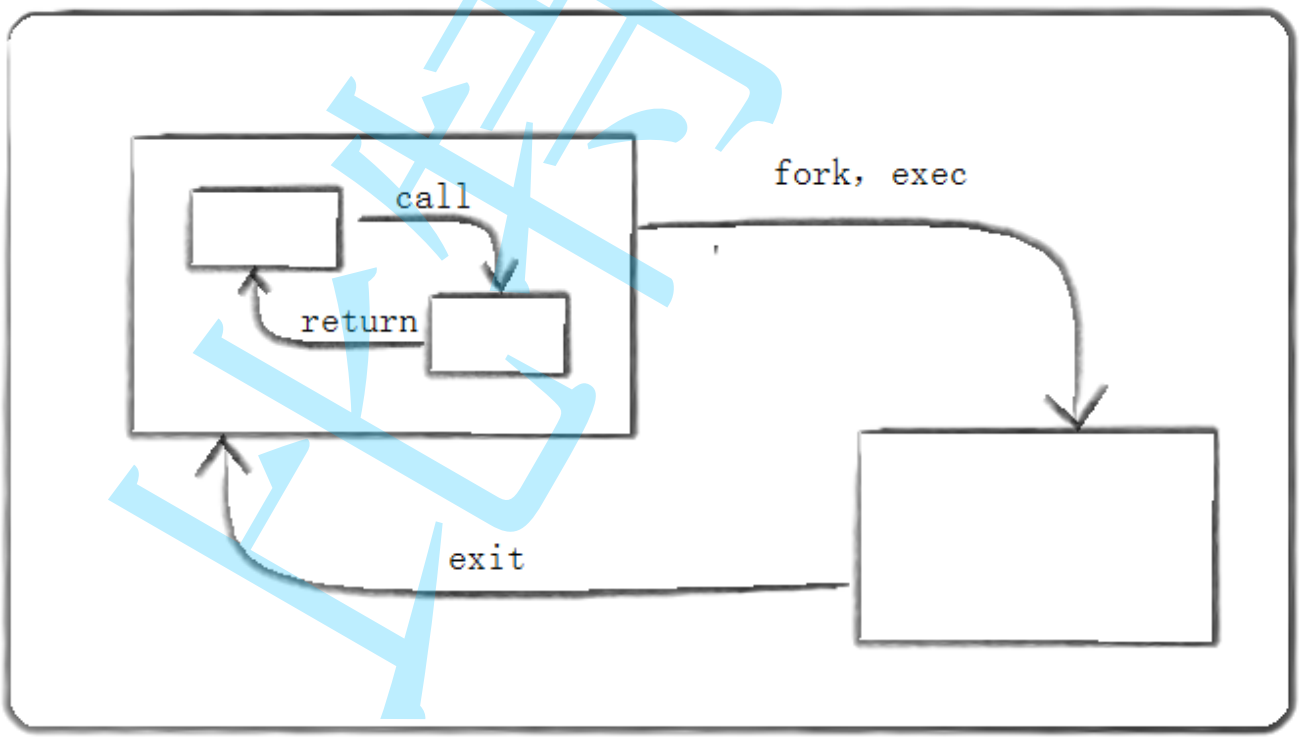
实现代码:

```
#include <stdio.h>
#include <stdlib.h>
#include <unistd.h>
#include <string.h>
#include <fcntl.h>

#define MAX_CMD 1024
char command[MAX_CMD];

int do_face()
{
    memset(command, 0x00, MAX_CMD);
    printf("minishell$ ");
    fflush(stdout);
    if (scanf("%[^\n]%*c", command) == 0) {
        getchar();
        return -1;
    }
    return 0;
}
char **do_parse(char *buff)
{
    int argc = 0;
    static char *argv[32];
    char *ptr = buff;

    while(*ptr != '\0') {
        if (!isspace(*ptr)) {
            argv[argc++] = ptr;
            while((!isspace(*ptr)) && (*ptr) != '\0') {
                ptr++;
            }
        }else {
            while(isspace(*ptr)) {
                *ptr = '\0';
                ptr++;
            }
        }
    }
    argv[argc] = NULL;
    return argv;
}
int do_exec(char *buff)
{
    char **argv = {NULL};


    int pid = fork();
```

```
    if (pid == 0) {
        argv = do_parse(buff);
        if (argv[0] == NULL) {
            exit(-1);
        }
        execvp(argv[0], argv);
    }else {
        waitpid(pid, NULL, 0);
    }

    return 0;
}
int main(int argc, char *argv[])
{
    while(1) {
        if (do_face() < 0)
            continue;
        do_exec(command);
    }
    return 0;
}
```

在继续学习新知识前，我们来思考函数和进程之间的相似性

exec/exit就像call/return
一个C程序有很多函数组成。一个函数可以调用另外一个函数，同时传递给它一些参数。被调用的函数执行一定的操作，然后返回一个值。每个函数都有他的局部变量，不同的函数通过call/return系统进行通信。
这种通过参数和返回值在拥有私有数据的函数间通信的模式是结构化程序设计的基础。Linux鼓励将这种应用于程序之内的模式扩展到程序之间。如下图



一个C程序可以fork/exec另一个程序，并传给它一些参数。这个被调用的程序执行一定的操作，然后通过exit(n)来返回值。调用它的进程可以通过wait（&ret）来获取exit的返回值。