# 19道小米网络运维工程师笔试真题。（含答案）

2017年11月23日 [学习路线](http://www.coooz.com/archives/category/zhiye/lines), [开发语言](http://www.coooz.com/archives/category/sa/lang) [暂无评论](http://www.coooz.com/archives/1396" \l "respond) 阅读 1,285 views 次

## 第一部分：Linux基础

**题目1：**

有一百个图片文件，它们的地址都是

http://down.xiaomi.com/img/1.png

http://down.xiaomi.com/img/2.png

…

一直到http://down.xiaomi.com/img/100.png

批量下载这100个图片文件，并找出其中大于500KB的文件。

**题目2：**

一个文本文件info.txt的内容如下：

aa,201

zz,502

bb,1

ee,42

每行都是按照逗号分隔，其中第二列都是数字，请对该文件按照第二列数字从大到小排列。

**题目3：**

查看当前Linux服务器是否监听80端口，如果在监听，请找出其进程ID，并结束该进程。

**题目4：**

使用curl或wget命令获取http服务的header信息。

**题目5：**

关于Linux的用户账号，下面说法正确的有：

A.用户的密码，是以明文形式存储在 /etc/passwd 文件中的

B.用户的密码，是以密文形式存储在 /etc/passwd 文件中的

C.用户的密码，是以密文形式存储在 /etc/shadow 文件中的

D.用户登录的时候，会把用户的密码明文与保存的密码做对比

**题目6：**

对于N块硬盘组成的硬盘阵列，下面的说法哪个是错误的：

A.raid1 与 raid5 相比，读取数据的速度 raid5 更快

B.raid1 与 raid5 相比，raid5 的磁盘空间利用率更高

C.raid1 在 （N-1）块磁盘损坏的情况下，不影响数据的完整性

D.raid0 相比于raid1、raid5，读写速度最快

**题目7：**

负载均衡，你了解的常用软件有哪些？请写出至少三种以上，并评价各自的缺点。

**题目8：**

执行 $ time sleep 2，输出如下：

real    0m2.003s

user   0m0.004s

sys    0m0.000s

请说明 real、user、sys三者具体代表的意思和区别。

**题目9：**

nginx rewrite 规则中 last、break、redirect、permanent 的含义。

**题目10：**

WEB 服务 cookies 和 session 的区别。

**题目11：**

http 1.0 和 http 1.1下有何区别？http 2.0的主要变化或优势有哪些？

## 第二部分：安全部分

**题目12：**

请列举WEB常见安全问题（不少于三项），阐述其原理、危害。

**题目13：**

请列举常见DOS攻击类型，并分别介绍其原理和防御方法。

**题目14：**

服务器遭到入侵，作为安全管理人员，你应做如何处理？

## 第三部分：网络部分

**题目15：**

请简述TCP的三次握手过程。

**题目16：**

发现系统中存在大量TIME\_WAIT，分析原因并提出三条以上优化建议。

## 第四部分：Python部分

**题目17：**

xrange和range的异同。

**题目18：**

列表 A 和 B，使用Python快速获取 A 和 B 中的共有元素。

**题目19：**

有20台服务器，需要在所有机器上 echo “123”，要求同时并行操作5台服务器，请使用Python或shell写出相关代码。

====================================================  
答案:

第一部分：Linux基础

题目1：

#!/bin/bash  
for i in {1..100}  
    do  
        wget http://down.xiaomi.com/img/$i.png  
    done  
find ./ -name "\*.png" -size +500k

题目2：

#如果仅排序第二列  
cat info.txt | cut -d"," -f 2 | sort -rn  
#按数字列排序整个文本  
sort -rnk2 -t, info.txt

题目3：

kill `netstat -tunlp | grep 80| grep -v grep | awk '{ print $NF}' | awk -F"/" '{ print $1}'`

题目4：

curl -I http://www.coooz.com  
wget --server-response http://www.coooz.com

题目5：  
答：C

题目6：  
答：B

题目7：  
nginx 七层代理，协议支持较少  
haproxy 支持Session的保持，多进程模式支持不够好  
lvs 稳定，无流量，四层代理。不能动静分离，成本高

题目8：  
答：  
real时间是指挂钟时间，也就是命令开始执行到结束的时间。这个短时间包括其他进程所占用的时间片，和进程被阻塞时所花费的时间。 user时间是指进程花费在用户模式中的CPU时间，这是唯一真正用于执行进程所花费的时间，其他进程和花费阻塞状态中的时间没有计算在内。 sys时间是指花费在内核模式中的CPU时间，代表在内核中执系统调用所花费的时间，这也是真正由进程使用的CPU时间。

题目9：  
答：  
不写last和break - 那么流程就是依次执行这些rewrite  
1. rewrite break - url重写后，直接使用当前资源，不再执行location里余下的语句，完成本次请求，地址栏url不变  
2. rewrite last - url重写后，马上发起一个新的请求，再次进入server块，重试location匹配，超过10次匹配不到报500错误，地址栏url不变  
3. rewrite redirect – 返回302临时重定向，地址栏显示重定向后的url，爬虫不会更新url（因为是临时）  
4. rewrite permanent – 返回301永久重定向, 地址栏显示重定向后的url，爬虫更新url

题目10：  
Session是在服务端保存的一个数据结构，用来跟踪用户的状态，这个数据可以保存在集群、数据库、文件中；  
Cookie是客户端保存用户信息的一种机制，用来记录用户的一些信息，也是实现Session的一种方式。

题目11：  
区别：  
1，HTTP/1.0协议使用非持久连接,即在非持久连接下,一个tcp连接只传输一个Web对象,；  
2，HTTP/1.1默认使用持久连接(然而,HTTP/1.1协议的客户机和服务器可以配置成使用非持久连接)。在持久连接下,不必为每个Web对象的传送建立一个新的连接,一个连接中可以传输多个对象!  
http2.0的优势：  
1.采用二进制格式传输数据，而非http1.1文本格式，二进制格式在协议的解析和优化扩展上带来了跟多的优势和可能  
2.对消息头采用Hpack进行压缩传输，能够节省消息头占用的网络流量，http1.1每次请求，都会携带大量冗余的头信息，浪费了很多宽带资源。  
3.异步连接多路复用  
4.Server Push，服务器端能够更快的把资源推送到客户端。  
5.保持与HTTP 1.1语义的向后兼容性也是该版本的一个关键

第二部分：安全部分

题目12：  
sql注入： 拖库，撞库，爆破，删除数据  
跨站脚本攻击: 向web页面里插入恶意html代码 用户浏览该页之时，嵌入其中的html代码会被执行  
失效的身份认证和会话管理

题目13：  
Land攻击：向某个设备发送数据包，并将数据包的源IP和目的IP都设置成攻击目标的地址。  
Teardrop攻击：利用OS处理分片重叠报文的漏洞进行攻击。  
DDos分布式拒绝服务攻击：攻击者攻破了多个系统。并利用这些系统去集中攻击其他目标。成千上万的主机发送大量的请求，受害设备因为无法处理而拒绝服务。  
Ping of Death攻击：通过构造出重缓冲区大小异常的ICMP包进行攻击。ping死你  
Smurf攻击：攻击这冒充受害主机的ip地址，向一个大的网络发送echo request 的定向广播包，此网络的许多主机都做出回应，受害主机会收到大龄的echo reply消息。  
SYN Flooding 攻击：利用TCP连接的3次握手过程进行攻击。攻击者主机使用无效的IP地址。与受害主机进行TCP3次握手。在完成第2步后，受害主机就会处于开放会话的请求之中，但会话并未真正完成。受害主机必须等待连接超时，之后才能清楚为完成的会话。在此期间，受害主机将会连续接收到这样的请求，最终因耗尽用来管理会话资源的资源而停止相应。

题目14：

第三部分：网络部分

题目15：  
首先Client端发送连接请求报文，Server段接受连接后回复ACK报文，并为这次连接分配资源。Client端接收到ACK报文后也向Server段发生ACK报文，并分配资源，这样TCP连接就建立了

题目16：  
对于一个处理大量短连接的服务器,如果是由服务器主动关闭客户端的连接,将导致服务器端存在大量的处于TIME\_WAIT状态的socket, 严重影响服务器的处理能力,甚至耗尽可用的socket,停止服务. TIME\_WAIT是TCP协议用以保证被重新分配的socket不会受到之前残留的延迟重发报文影响的机制,是必要的逻辑保证.  
解决方法：  
1、缩短TIME\_WAIT的等待时间  
2、允许将TIME-WAIT sockets重新用于新的TCP连接  
3、开启TIME-WAIT sockets的快速回收  
4、使用长连接

第四部分：Python部分

题目17：  
答：range([start,] stop[, step])，根据start与stop指定的范围以及step设定的步长，生成一个序列。xrange生成的不是一个数组，而是一个生成器。尽量用xrange吧，除非你是要返回一个列表。

题目18：

print set(A) & set(B)

题目19：

#!/bin/bash  
start\_time=`date +%s`  
[ -e /tmp/fd1 ] || mkfifo /tmp/fd1  
exec 3<>/tmp/fd1  
rm -rf /tmp/fd1  
for ((i=1;i<=5;i++))  
do  
    echo >&3  
done  
while read line  
do  
    read -u3  
    {  
    host=`echo $line| awk '{print $1}'`  
    ssh  root@$host 'echo “123”'  
    echo 'success,'$host  
    echo >&3  
    }&  
done < host.txt  
stop\_time=`date +%s`  
echo "time: `expr $stop\_time - $start\_time`"  
exec 3<&-  
exec 3>&-

《19道小米网运维工程师笔试真题，你能通关吗？》

题目转至：马哥Linux运维

**第一部分：Linux基础**

**题目1：**

有一百个图片文件，它们的地址都是

http://down.xiaomi.com/img/1.png

http://down.xiaomi.com/img/2.png

…

一直到http://down.xiaomi.com/img/100.png

批量下载这100个图片文件，并找出其中大于500KB的文件。

参考答案：

#/bin/bash#by xianwei#2017-9-5

echo "downloading the picture......"for i in {1..100}do

wget "http://down.xiaomi.com/img/$i.png"done

echo "download done!"echo "find the file which is big than 500k"find . -type f -size +500k -print

**题目2：**

一个文本文件info.txt的内容如下：

aa,201

zz,502

bb,1

ee,42

每行都是按照逗号分隔，其中第二列都是数字，请对该文件按照第二列数字从大到小排列。

参考答案：

root@vmUbu:/home/dell/shell# cat info.txt

aa,201

zz,502

bb,1

ee,42

root@vmUbu:/home/dell/shell# cat info.txt |sort -t "," -rnk 2

zz,502

aa,201

ee,42

bb,1

root@vmUbu:/home/dell/shell#

考察知识点：

    sort的使用

        -r 逆序

        -n 按数字排序

        -k 指定列

        -t 指定分隔符号

**题目3：**

查看当前Linux服务器是否监听80端口，如果在监听，请找出其进程ID，并结束该进程。

参考答案

 说明：测试使用3306段口为例。

 1）查看端口对应的进程ID和进程名，有两种方法lsof和netstat

   netstat -tanp |grep 3306

   lsof -i:3306

root@vmUbu:/home/dell/shell# netstat -tanp |grep 3306

tcp 0 0 127.0.0.1:3306 0.0.0.0:\* LISTEN 1309/mysqld

root@vmUbu:/home/dell/shell# lsof -i:3306

COMMAND PID USER FD TYPE DEVICE SIZE/OFF NODE NAME

mysqld 1309 mysql 16u IPv4 19621 0t0 TCP localhost:mysql (LISTEN)

  2)过滤出ID号，并kill

方法一：使用for

for i in `netstat -tanp |grep 3306 |awk '{print $NF}' |cut -d "/" -f 1`;do kill -9 $i;done

方法二：使用xargs

netstat -tanp |grep 3306 |awk '{print $NF}' |cut -d "/" -f 1 |xargs -I '{}' kill -9 {}

方法三：使用awk

netstat -tanp |grep 3306 |awk '{print $NF}' |cut -d "/" -f 1 |awk '{print "kill -9 "$0}' |bash

**题目4：**

使用curl或wget命令获取http服务的header信息。

**题目5：**

关于Linux的用户账号，下面说法正确的有：

A.用户的密码，是以明文形式存储在 /etc/passwd 文件中的

B.用户的密码，是以密文形式存储在 /etc/passwd 文件中的

C.用户的密码，是以密文形式存储在 /etc/shadow 文件中的

D.用户登录的时候，会把用户的密码明文与保存的密码做对比

参考答案：C

**题目6：**

对于N块硬盘组成的硬盘阵列，下面的说法哪个是错误的：

A.raid1 与 raid5 相比，读取数据的速度 raid5 更快

B.raid1 与 raid5 相比，raid5 的磁盘空间利用率更高

C.raid1 在 （N-1）块磁盘损坏的情况下，不影响数据的完整性

D.raid0 相比于raid1、raid5，读写速度最快

参考答案：

    raid0 的读速度比raid1,raid5要慢一些，因为只从一个盘中读取数据

**题目7：**

负载均衡，你了解的常用软件有哪些？请写出至少三种以上，并评价各自的缺点。

参考答案：

Nginx的缺点是：

1. Nginx**仅能支持**http、https和Email协议，这样就在适用范围上面小些，这个是它的缺点。

2. 对后端服务器的健康检查，**只支持通过端口**来检测，不支持通过url来检测。**不支持Session**的直接保持，但能通过ip\_hash来解决。

LVS的缺点是：

1. 软件本身**不支持正则表达式**处理（仅仅支持4层负载均衡），**不能做动静分离**；而现在许多网站在这方面都有较强的需求，这个是Nginx/HAProxy+Keepalived的优势所在。

2. 如果是网站应用比较庞大的话，LVS/DR+Keepalived实施起来就**比较复杂**了，特别后面有 Windows Server的机器的话，如果实施及配置还有维护过程就比较复杂了，相对而言Nginx/HAProxy+Keepalived就简单多了。

请参考：http://www.ha97.com/5646.html

**题目8：**

执行 $ time sleep 2，输出如下：

real    0m2.003s

user    0m0.004s

sys     0m0.000s

请说明 real、user、sys三者具体代表的意思和区别。

参考答案：

**real时间**是指挂钟时间，也就是命令开始执行到结束的时间。这个短时间包括其他进程所占用的时间片，和进程被阻塞时所花费的时间。

**user时间**是指进程花费在用户模式中的CPU时间，这是唯一真正用于执行进程所花费的时间，其他进程和花费阻塞状态中的时间没有计算在内。

**sys时间**是指花费在内核模式中的CPU时间，代表在内核中执系统调用所花费的时间，这也是真正由进程使用的CPU时间。

**题目9：**

nginx rewrite 规则中 last、break、redirect、permanent 的含义。

**题目10：**

WEB 服务 cookies 和 session 的区别。

参考答案：

1、cookie数据存放在客户的浏览器上，session数据放在服务器上

2、cookie不是很安全，别人可以分析存放在本地的COOKIE并进行

3、session会在一定时间内保存在服务器上。当访问增多，会比较占用你服务器的性能

4、单个cookie在客户端的限制是3K，就是说一个站点在客户端存放的COOKIE不能大于3K。

出处：http://blog.csdn.net/u010168160/article/details/47128443

**题目11：**

http 1.0 和 http 1.1下有何区别？http 2.0的主要变化或优势有哪些？

参考答案：

    http 1.0使用短链接,http1.1使用长连接

    http2.0 性能和速度有很大提升。

**第二部分：安全部分**

**题目12：**

请列举WEB常见安全问题（不少于三项），阐述其原理、危害。

略

**题目13：**

请列举常见DOS攻击类型，并分别介绍其原理和防御方法。

略

**题目14：**

服务器遭到入侵，作为安全管理人员，你应做如何处理？

参考答案

[http://netsecurity.51cto.com/art/201608/515574.htm](http://netsecurity.51cto.com/art/201608/515574.htm" \t "http://blog.51cto.com/237085/_blank)

**第三部分：网络部分**

**题目15：**

请简述TCP的三次握手过程。

**题目16：**

发现系统中存在大量TIME\_WAIT，分析原因并提出三条以上优化建议。

**第四部分：Python部分**

**题目17：**

xrange和range的异同。

参考答案:

xrange会比range性能优很多，因为不需要一上来就开辟一块很大的内存空间。

range会直接生成一个list对象,而xrange则不会直接生成一个list，而是每次调用返回其中的一个值

**题目18：**

列表 A 和 B，使用Python快速获取 A 和 B 中的共有元素。

参考答案:

for i in A:

if i in B:

print i

**题目19：**

有20台服务器，需要在所有机器上 echo “123”，要求同时并行操作5台服务器，请使用Python或shell写出相关代码。

# encoding:utf-8

import timeimport sysimport osimport threadpool

def cmd(ip):

#需先设置远程无密码执行命令

os.system(‘ssh ip echo "123" 2>&1’)

time.sleep(2)

pool = threadpool.ThreadPool(5) #创建一个线程池，包括线程数10个

requests = threadpool.makeRequests(cmd, ip\_list) #传递函数和参数[pool.putRequest(req) for req in requests] #将请求放入线程池进行处理

pool.wait()