大数据技术之Shell

# 第1章 Shell概述

大数据程序员为什么要学习Shell呢？

1）需要看懂运维人员编写的Shell程序。

2）偶尔会编写一些简单Shell程序来管理集群、提高开发效率。

# 第2章 Shell解析器

（1）Linux提供的Shell解析器有：

[dahua@hadoop101 ~]$ cat /etc/shells

/bin/sh

/bin/bash

/sbin/nologin

/bin/dash

/bin/tcsh

/bin/csh

（2）bash和sh的关系

[dahua@hadoop101 bin]$ ll | grep bash

-rwxr-xr-x. 1 root root 941880 5月 11 2016 bash

lrwxrwxrwx. 1 root root 4 5月 27 2017 sh -> bash

（3）Centos默认的解析器是bash

[dahua@hadoop102 bin]$ echo $SHELL

/bin/bash

# 第3章 Shell脚本入门

1．脚本格式

脚本以#!/bin/bash开头（指定解析器）

2．第一个Shell脚本：helloworld

（1）需求：创建一个Shell脚本，输出helloworld

（2）案例实操：

[dahua@hadoop101 datas]$ touch helloworld.sh

[dahua@hadoop101 datas]$ vi helloworld.sh

在helloworld.sh中输入如下内容

#!/bin/bash

echo "helloworld"

（3）脚本的常用执行方式

第一种：采用bash或sh+脚本的相对路径或绝对路径（不用赋予脚本+x权限）

sh+脚本的相对路径

[dahua@hadoop101 datas]$ sh helloworld.sh

Helloworld

sh+脚本的绝对路径

[dahua@hadoop101 datas]$ sh /home/dahua/datas/helloworld.sh

helloworld

bash+脚本的相对路径

[dahua@hadoop101 datas]$ bash helloworld.sh

Helloworld

bash+脚本的绝对路径

[dahua@hadoop101 datas]$ bash /home/dahua/datas/helloworld.sh

Helloworld

第二种：采用输入脚本的绝对路径或相对路径执行脚本（必须具有可执行权限+x）

（a）首先要赋予helloworld.sh 脚本的+x权限

[dahua@hadoop101 datas]$ chmod 777 helloworld.sh

（b）执行脚本

相对路径

[dahua@hadoop101 datas]$ ./helloworld.sh

Helloworld

绝对路径

[dahua@hadoop101 datas]$ /home/dahua/datas/helloworld.sh

Helloworld

注意：第一种执行方法，本质是bash解析器帮你执行脚本，所以脚本本身不需要执行权限。第二种执行方法，本质是脚本需要自己执行，所以需要执行权限。

3．第二个Shell脚本：多命令处理

（1）需求：

在/home/dahua/目录下创建一个banzhang.txt,在banzhang.txt文件中增加“I love cls”。

（2）案例实操：

[dahua@hadoop101 datas]$ touch batch.sh

[dahua@hadoop101 datas]$ vi batch.sh

在batch.sh中输入如下内容

#!/bin/bash

cd /home/dahua

touch cls.txt

echo "I love cls" >>cls.txt

# 第4章 Shell中的变量

## 4.1 系统变量

在linux中,声明变量没有数据类型

key=value

a=10

1. 常用系统变量

$HOME、$PWD、$SHELL、$USER等

2．案例实操

（1）查看系统变量的值

[dahua@hadoop101 datas]$ echo $HOME

/home/dahua

（2）显示当前Shell中所有变量：set

[dahua@hadoop101 datas]$ set

BASH=/bin/bash

BASH\_ALIASES=()

BASH\_ARGC=()

BASH\_ARGV=()

## 4.2 自定义变量

1．基本语法

（1）定义变量：变量=值

（2）撤销变量：unset 变量

（3）声明静态变量：readonly变量，注意：不能unset

2．变量定义规则

（1）变量名称可以由字母、数字和下划线组成，但是不能以数字开头，环境变量名建议大写。

（2）等号两侧不能有空格

（3）在bash中，变量默认类型都是字符串类型，无法直接进行数值运算。

（4）变量的值如果有空格，需要使用双引号或单引号括起来。

3．案例实操

（1）定义变量A

[dahua@hadoop101 datas]$ A=5

[dahua@hadoop101 datas]$ echo $A

5

（2）给变量A重新赋值

[dahua@hadoop101 datas]$ A=8

[dahua@hadoop101 datas]$ echo $A

8

（3）撤销变量A

[dahua@hadoop101 datas]$ unset A

[dahua@hadoop101 datas]$ echo $A

（4）声明静态的变量B=2，不能unset

[dahua@hadoop101 datas]$ readonly B=2

[dahua@hadoop101 datas]$ echo $B

2

[dahua@hadoop101 datas]$ B=9

-bash: B: readonly variable

（5）在bash中，变量默认类型都是字符串类型，无法直接进行数值运算

[dahua@hadoop102 ~]$ C=1+2

[dahua@hadoop102 ~]$ echo $C

1+2

（6）变量的值如果有空格，需要使用双引号或单引号括起来

[dahua@hadoop102 ~]$ D=I love banzhang

-bash: world: command not found

[dahua@hadoop102 ~]$ D="I love banzhang"

[dahua@hadoop102 ~]$ echo $A

I love banzhang

（7）可把变量提升为全局环境变量，可供其他Shell程序使用

export 变量名

[dahua@hadoop101 datas]$ vim helloworld.sh

在helloworld.sh文件中增加echo $B

#!/bin/bash

echo "helloworld"

echo $B

[dahua@hadoop101 datas]$ ./helloworld.sh

Helloworld

发现并没有打印输出变量B的值。

[dahua@hadoop101 datas]$ export B

[dahua@hadoop101 datas]$ ./helloworld.sh

helloworld

2

## 4.3 特殊变量：$n

1．基本语法

$n （功能描述：n为数字，$0代表该脚本名称，$1-$9代表第一到第九个参数，十以上的参数，十以上的参数需要用大括号包含，如${10}）

2．案例实操

（1）输出该脚本文件名称、输入参数1和输入参数2 的值

[dahua@hadoop101 datas]$ touch parameter.sh

[dahua@hadoop101 datas]$ vim parameter.sh

#!/bin/bash

echo "$0 $1 $2"

[dahua@hadoop101 datas]$ chmod 777 parameter.sh

[dahua@hadoop101 datas]$ ./parameter.sh cls xz

./parameter.sh cls xz

## 4.4 特殊变量：$#

1．基本语法

$# （功能描述：获取所有输入参数个数，常用于循环）。

2．案例实操

（1）获取输入参数的个数

[dahua@hadoop101 datas]$ vim parameter.sh

#!/bin/bash

echo "$0 $1 $2"

echo $#

[dahua@hadoop101 datas]$ chmod 777 parameter.sh

[dahua@hadoop101 datas]$ ./parameter.sh cls xz

parameter.sh cls xz

2

## 4.5 特殊变量：$\*、$@

1．基本语法

$\* （功能描述：这个变量代表命令行中所有的参数，$\*把所有的参数看成一个整体）

$@ （功能描述：这个变量也代表命令行中所有的参数，不过$@把每个参数区分对待）

2．案例实操

（1）打印输入的所有参数

[dahua@hadoop101 datas]$ vim parameter.sh

#!/bin/bash

echo "$0 $1 $2"

echo $#

echo $\*

echo $@

[dahua@hadoop101 datas]$ bash parameter.sh 1 2 3

parameter.sh 1 2

3

1 2 3

1 2 3

## 4.6 特殊变量：$？

1．基本语法

$？ （功能描述：最后一次执行的命令的返回状态。如果这个变量的值为0，证明上一个命令正确执行；如果这个变量的值为非0（具体是哪个数，由命令自己来决定），则证明上一个命令执行不正确了。）

2．案例实操

（1）判断helloworld.sh脚本是否正确执行

[dahua@hadoop101 datas]$ ./helloworld.sh

hello world

[dahua@hadoop101 datas]$ echo $?

0

# 第5章 运算符

1．基本语法

（1）“$((运算式))”或“$[运算式]”

（2）expr + , - , \\*, /, % 加，减，乘，除，取余

注意：expr运算符间要有空格

2．案例实操：

（1）计算3+2的值

[dahua@hadoop101 datas]$ expr 2 + 3

5

（2）计算3-2的值

[dahua@hadoop101 datas]$ expr 3 - 2

1

（3）计算（2+3）X4的值

（a）expr一步完成计算

[dahua@hadoop101 datas]$ expr `expr 2 + 3` \\* 4

20

（b）采用$[运算式]方式

[dahua@hadoop101 datas]# S=$[(2+3)\*4]

[dahua@hadoop101 datas]# echo $S

# 第6章 条件判断

1．基本语法

[ condition ]（注意condition前后要有空格）

注意：条件非空即为true，[ dahua ]返回true，[] 返回false。

2. 常用判断条件

（1）两个整数之间比较

= 字符串比较

-lt 小于（less than） -le 小于等于（less equal）

-eq 等于（equal） -gt 大于（greater than）

-ge 大于等于（greater equal） -ne 不等于（Not equal）

（2）按照文件权限进行判断

-r 有读的权限（read） -w 有写的权限（write）

-x 有执行的权限（execute）

（3）按照文件类型进行判断

-f 文件存在并且是一个常规的文件（file）

-e 文件存在（existence） -d 文件存在并是一个目录（directory）

3．案例实操

（1）23是否大于等于22

[dahua@hadoop101 datas]$ [ 23 -ge 22 ]

[dahua@hadoop101 datas]$ echo $?

0

（2）helloworld.sh是否具有写权限

[dahua@hadoop101 datas]$ [ ! -w helloworld.sh ]

[dahua@hadoop101 datas]$ echo $?

0

（3）/home/dahua/cls.txt目录中的文件是否存在

[dahua@hadoop101 datas]$ [ ! -e /home/dahua/cls.txt ]

[dahua@hadoop101 datas]$ echo $?

1

（4）多条件判断（&& 表示前一条命令执行成功时，才执行后一条命令，|| 表示上一条命令执行失败后，才执行下一条命令）

[dahua@hadoop101 ~]$ [ condition ] && echo OK || echo notok

OK

[dahua@hadoop101 datas]$ [ condition ] && [ ] || echo notok

notok

# 第7章 流程控制（重点）

## 7.1 if 判断

1．基本语法

if [ 条件判断式 ];then

程序

fi

或者

if [ 条件判断式 ]

then

程序

fi

注意事项：

（1）[ 条件判断式 ]，中括号和条件判断式之间必须有空格

（2）if后要有空格

2．案例实操

（1）输入一个数字，如果是1，则输出banzhang zhen shuai，如果是2，则输出cls zhen mei，如果是其它，什么也不输出。

[dahua@hadoop101 datas]$ touch if.sh

[dahua@hadoop101 datas]$ vim if.sh

#!/bin/bash

if [ $1 -eq "1" ]

then

echo "banzhang zhen shuai"

elif [ $1 -eq "2" ]

then

echo "cls zhen mei"

fi

[dahua@hadoop101 datas]$ chmod 777 if.sh

[dahua@hadoop101 datas]$ ./if.sh 1

banzhang zhen shuai

## 7.2 case 语句

1．基本语法

case $变量名 in

"值1"）

如果变量的值等于值1，则执行程序1

;;

"值2"）

如果变量的值等于值2，则执行程序2

;;

…省略其他分支…

\*）

如果变量的值都不是以上的值，则执行此程序

;;

esac

注意事项：

1. case行尾必须为单词“in”，每一个模式匹配必须以右括号“）”结束。
2. 双分号“**;;**”表示命令序列结束，相当于java中的break。
3. 最后的“\*）”表示默认模式，相当于java中的default。

2．案例实操

（1）输入一个数字，如果是1，则输出banzhang，如果是2，则输出cls，如果是其它，输出renyao。

[dahua@hadoop101 datas]$ touch case.sh

[dahua@hadoop101 datas]$ vim case.sh

!/bin/bash

case $1 in

"1")

echo "banzhang"

;;

"2")

echo "cls"

;;

\*)

echo "renyao"

;;

esac

[dahua@hadoop101 datas]$ chmod 777 case.sh

[dahua@hadoop101 datas]$ ./case.sh 1

1

## 7.3 for 循环

1．基本语法1

for (( 初始值;循环控制条件;变量变化 ))

do

程序

done

2．案例实操

（1）从1加到100

[dahua@hadoop101 datas]$ touch for1.sh

[dahua@hadoop101 datas]$ vim for1.sh

#!/bin/bash

s=0

for((i=0;i<=100;i++))

do

s=$[$s+$i]

done

echo $s

[dahua@hadoop101 datas]$ chmod 777 for1.sh

[dahua@hadoop101 datas]$ ./for1.sh

“5050”

3．基本语法2

for 变量 in 值1 值2 值3…

do

程序

done

4．案例实操

（1）打印所有输入参数

[dahua@hadoop101 datas]$ touch for2.sh

[dahua@hadoop101 datas]$ vim for2.sh

#!/bin/bash

#打印数字

for i in $\*

do

echo "ban zhang love $i "

done

[dahua@hadoop101 datas]$ chmod 777 for2.sh

[dahua@hadoop101 datas]$ bash for2.sh cls xz bd

ban zhang love cls

ban zhang love xz

ban zhang love bd

（2）比较$\*和$@区别

（a）$\*和$@都表示传递给函数或脚本的所有参数，不被双引号“”包含时，都以$1 $2 …$n的形式输出所有参数。

|  |
| --- |
| [dahua@hadoop101 datas]$ touch for.sh  [dahua@hadoop101 datas]$ vim for.sh  #!/bin/bash  for i in $\*  do  echo "ban zhang love $i "  done  for j in $@  do  echo "ban zhang love $j"  done  [dahua@hadoop101 datas]$ bash for.sh cls xz bd  ban zhang love cls  ban zhang love xz  ban zhang love bd  ban zhang love cls  ban zhang love xz  ban zhang love bd |

（b）当它们被双引号“”包含时，“$\*”会将所有的参数作为一个整体，以“$1 $2 …$n”的形式输出所有参数；“$@”会将各个参数分开，以“$1” “$2”…”$n”的形式输出所有参数。

[dahua@hadoop101 datas]$ vim for.sh

#!/bin/bash

for i in "$\*"

#$\*中的所有参数看成是一个整体，所以这个for循环只会循环一次

do

echo "ban zhang love $i"

done

for j in "$@"

#$@中的每个参数都看成是独立的，所以“$@”中有几个参数，就会循环几次

do

echo "ban zhang love $j"

done

[dahua@hadoop101 datas]$ chmod 777 for.sh

[dahua@hadoop101 datas]$ bash for.sh cls xz bd

ban zhang love cls xz bd

ban zhang love cls

ban zhang love xz

ban zhang love bd

## 7.4 while 循环

1．基本语法

while [ 条件判断式 ]

do

程序

done

2．案例实操

（1）从1加到100

[dahua@hadoop101 datas]$ touch while.sh

[dahua@hadoop101 datas]$ vim while.sh

#!/bin/bash

s=0

i=1

while [ $i -le 100 ]

do

s=$[$s+$i]

i=$[$i+1]

done

echo $s

[dahua@hadoop101 datas]$ chmod 777 while.sh

[dahua@hadoop101 datas]$ ./while.sh

5050

# 第8章 read读取控制台输入

1．基本语法

read(选项)(参数)

选项：

-p：指定读取值时的提示符；

-t：指定读取值时等待的时间（秒）。

参数

变量：指定读取值的变量名

2．案例实操

（1）提示7秒内，读取控制台输入的名称

[dahua@hadoop101 datas]$ touch read.sh

[dahua@hadoop101 datas]$ vim read.sh

#!/bin/bash

read -t 7 -p "Enter your name in 7 seconds " NAME

echo $NAME

[dahua@hadoop101 datas]$ ./read.sh

Enter your name in 7 seconds xiaoze

xiaoze

# 第9章 函数

## 9.1 系统函数

1．basename基本语法

basename [string / pathname] [suffix] （功能描述：basename命令会删掉所有的前缀包括最后一个（‘/’）字符，然后将字符串显示出来。

选项：

suffix为后缀，如果suffix被指定了，basename会将pathname或string中的suffix去掉。

2．案例实操

（1）截取该/home/dahua/banzhang.txt路径的文件名称

[dahua@hadoop101 datas]$ basename /home/dahua/banzhang.txt

banzhang.txt

[dahua@hadoop101 datas]$ basename /home/dahua/banzhang.txt .txt

banzhang

3. dirname基本语法

dirname 文件绝对路径 （功能描述：从给定的包含绝对路径的文件名中去除文件名（非目录的部分），然后返回剩下的路径（目录的部分））

4．案例实操

（1）获取banzhang.txt文件的路径

[dahua@hadoop101 ~]$ dirname /home/dahua/banzhang.txt

/home/dahua

## 9.2 自定义函数

1．基本语法

[ function ] funname[()]

{

Action;

[return int;]

}

funname

2．经验技巧

（1）必须在调用函数地方之前，先声明函数，shell脚本是逐行运行。不会像其它语言一样先编译。

（2）函数返回值，只能通过$?系统变量获得，可以显示加：return返回，如果不加，将以最后一条命令运行结果，作为返回值。return后跟数值n(0-255)

3．案例实操

（1）计算两个输入参数的和

[dahua@hadoop101 datas]$ touch fun.sh

[dahua@hadoop101 datas]$ vim fun.sh

#!/bin/bash

function sum()

{

s=0

s=$[ $1 + $2 ]

echo "$s"

}

read -p "Please input the number1: " n1;

read -p "Please input the number2: " n2;

sum $n1 $n2;

[dahua@hadoop101 datas]$ chmod 777 fun.sh

[dahua@hadoop101 datas]$ ./fun.sh

Please input the number1: 2

Please input the number2: 5

7

作业:

使用shell 脚本编写 99乘法表.