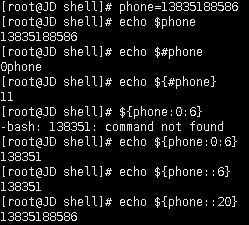
**字符串处理**

子串截取

子串截取的方法

方法一，使用${}表达式

格式：${var:起始位置：长度}（起始位置，编号从0开始，可省略）



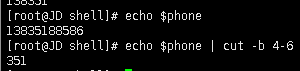
方法二，使用expr substr

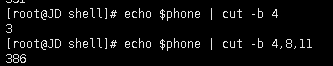
格式：expr substr “$var” 起始位置 长度（起始位置，编号从1开始）



方法三，使用cut工具

格式：echo $var|cut -b 起始位置-结束位置





路径分割

取目录位置

格式：dirname “字符串”

取文档的基本名称

格式：basename “字符串”

子串替换

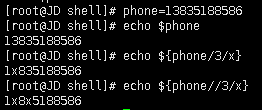
子串替换的方法

只替换第1个匹配结果

格式：${var/old/new}

替换全部匹配结果

格式：${var//old/new}



按条件掐头去尾

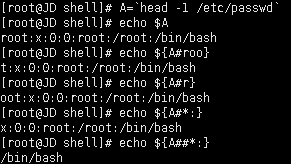
字符串掐头

从左向右，最短匹配删除

格式：${变量名#\*关键字}（#用来删除头部，\*通配）

从左向右，最长匹配删除

格式：${变量名##\*关键字}（#用来删除头部，\*通配）



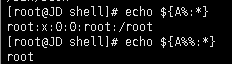
字符串去尾

从右向左，最短匹配删除

格式：${变量名%关键字\*}

从右向左，最长匹配删除

格式：${变量名%%关键字\*}



变量的初始值处理

初值的检测及设置

取值，${var:-word}

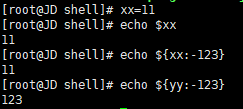
若变量var已存在且非NULL，则返回$var的值

否则返回字串“word”，变量var值不变

赋值，${var:=word}

若变量var已存在且非null，则返回$var的值

否则返回字串“word”，并赋值给var



**扩展脚本的技巧**

shell数组（一个变量存多个值）

关于变量的类型

shell对变量类型的管理比较松散

变量的值默认均视为文本

用在数学运算中时，自动将其转换为整数

定义/赋值数组

方法一，整体赋值：

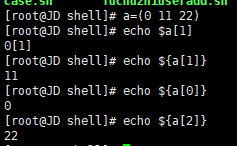
格式：数组名=（值1 值2 ... 值N）

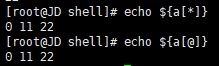
实例：SVRS=（www ftp mail club）

方法二：为单个元素赋值

格式：数组名[下标]=值（从0开始）

实例：fqdns[0]=www.tarena.com







输出数组元素

获得单个数组元素

格式：${数组名[下标]}

获得所有数组元素

格式：${数组名[@]}

获取数组元素个数

格式：${#数组名[@]}

获取连续的多个数组元素

格式：${数组名[@]:起始下标:元素个数}

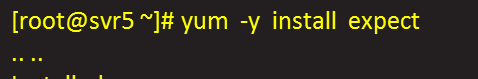
expect预期交互

最后一行不执行

简介

基于TCL编写的自动交互式程序

可以用在shell脚本中，为交互式过程自动输送预先准备的文本或指令，而无需人工干预





**正则表达式**

正则表达式概述

什么是正则表达式

使用“一串符号”来描述共同属性的数据

egrep过滤工具

文本处理顺序

以行为单位，逐行进行处理

默认只输出与表达式相匹配的文本行

基本用法

格式1：egrep [选项] ‘正则表达式’ 文件...

格式2：前置命令 | egrep [选项] ‘正则表达式’

等同于grep -E，表示允许使用扩展的正则表达式

常用命令选项

-i：忽略字母大小写

-v：条件取反

-c：统计匹配的行数

-q：静默、无任何输出，一般用于检测

-n：显示出匹配结果所在的行号

--color：标记显示匹配字串

基本元字符

行首尾及单字匹配

|  |  |  |  |
| --- | --- | --- | --- |
| 类型 | 含义 | 事例 | 说明 |
| ^ | 匹配行首 | ^abc | 以abc开头的行 |
| ^# | 以#号开头的行（比如注释行） |
| $ | 匹配行尾 | abc$ | 以abc结尾的行 |
| ^$ | 空行 |
| . | 单个字符 | . | 除换行符\n以外的任意单个字符 |

未定匹配次数

|  |  |  |  |
| --- | --- | --- | --- |
| 类型 | 含义 | 示例 | 说明 |
| + | 最少匹配一次 | a+ | 一个或多个连续的a |
| （abc）+ | 一个或多个连续的abc |
| ？ | 最多匹配一次 | a? | 0个或1个a |
| （abc）? | 0个或1个abc |
| \* | 匹配任意次数 | a\* | 0个或多个连续的a |
| （abc）\* | 0个或多个连续的abc |
| .\* | 匹配前一个字符的任意次数（0次或者多次）\*不允许单独使用 |

｛｝限定次数

限定表达式的匹配次数{n}、{n,m}、{n,}

|  |  |  |  |
| --- | --- | --- | --- |
| 类型 | 含义 | 示例 | 说明 |
| \{n\} | 匹配前一个字符n次（包含） | （ab）\{3\} | 匹配ababab |
| \{n,m\} | 匹配前一个字符n-m次（包含） | (ab)\{1,3\} | 匹配ab、abab、ababab |
| \{n,\} | 匹配前一个字符至少n次（包含） | (ab)\{2,\} | 匹配2个及以上连续的ab |

其他元字符

[]范围内匹配单个字符

匹配指定字符集合内的任意一个字符

[]内加^可取反

|  |  |
| --- | --- |
| 示例 | 说明 |
| [alc45\_?] | 匹配a、l、c、4、5、\_、？ |
| [a-z] | 匹配任意小写字母 |
| [A-Z] | 匹配任意大写字母 |
| [0-9] | 匹配任意数字 |
| [a-Z0-9] | 匹配任意字母或数字 |
| [^A-Z] | 匹配非大写字母的行 |
| ^[^a-z] | 匹配不以小写字母开头的行 |
| [] | 集合，匹配范围内的任意单个字符 |

整体及边界匹配

|  |  |  |  |
| --- | --- | --- | --- |
| 类型 | 含义 | 示例 | 说明 |
| () | 组合为整体 | ab{1,3} | 匹配ab、abb、abbb |
|  |  | (ab){1,3} | 匹配ab、abab、ababab |
| | | 或者 | root|bin | 匹配root、bin |
| \b | 单词边界 | \broot\b | 匹配单词root，不匹配keroot、rooty、brooty等字符串 |
| \< | 单词的开头 | \<th | 匹配以th开头的单词 |
| \> | 单词的结束 | \<root\> | 作用与\broot\b相同 |

基本正则表达式

|  |  |
| --- | --- |
| 正则符号 | 描述 |
| ^ | 匹配行首 |
| $ | 匹配行尾 |
| [] | 集合，匹配集合中的单个任意字符 |
| [^] | 对集合取反 |
| . | 匹配任意单个字符 |
| \* | 匹配前一个字符任意次（不允许单独使用） |
| \{n,m\} | 匹配前一个字符n到m次 |
| \{n\} | 匹配前一个字符n次 |
| \{n,}\ | 匹配前一个字符n次以上 |
| \(\) | 保留 |

扩展正则表达式

|  |  |
| --- | --- |
| 正则符号 | 描述 |
| + | 最少匹配一次 |
| ？ | 最多匹配一次 |
| {n,m} | 匹配n到m次 |
| () | 组合为整体，保留，相当于复制，\序号粘贴 |
| | | 或者 |
| \b | 单词边界 |