



Linux 下的远程备份（镜像）工具 rsync

（陈云川 yc2084@163.com UESTC,CD 2007 年 11 月 7 日星期三）

0 简介

传统上，在 Unix 系统下使用的备份工具主要有 tar、dump (rdump)、restore (rrestore)、dd、cpio、rcp 等工具。这些工具中大多数适合于本地备份 (native backup)，倘若要执行远程备份 (remote backup) 的话，往往显得力不从心。在 Linux 系统下提供了一个工具——rsync，据其文档说明，它是 rcp 工具的一个替代品，但是更快捷、更灵活。rsync 采用的 remote-update 协议允许 rsync 通过网络链路传输两组文件的增量，而不是文件本身，据此 rsync 可以提高传输效率。

rsync 所具有的功能特色主要有：

- 支持对连接、设备、用户、组、权限的拷贝；
- 支持类似 GNU tar 工具的 exclude 和 exclude-from 选项；
- 类似 CVS 的 exclude 模式：忽略相同的文件；
- 可以使用多种远程 shell 工具，比如 rsh 或者 ssh；
- 不需要 root 权限即可运行；
- 将文件传输流水线化以减少延迟开销；
- rsync 服务器允许以匿名或者认证方式登录（特别适合于做镜像）。

和大多数网络工具一样，rsync 也是以服务器/客户端的方式工作的。但较为特别的一点是：rsync 服务器和客户端都使用同一个程序——rsync（如果按常见的做法，则应该是服务器程序为 rsyncd，客户端程序为 rsync）。启动服务器的时候需要指明--daemon 选项，以表明启动的是服务器。而在启动客户端程序的时候则不应开启--daemon 选项。

在下文中，假设需要备份的主机为 192.168.0.100，其上运行 rsync 服务器，将该主机称之为 rsync-server；存放备份数据的主机为 192.168.0.200，其上运行 rsync 客户端程序，将该主机称之为 rsync-client。

1 服务器端的配置

rsync 的配置文件为/etc/rsyncd.conf，这个配置文件控制着 rsync 服务器的认证、访问、logging、可用模块等信息。关于这个配置文件的详细说明请参见 rsyncd.conf(5)。在本文中，仅对将要使用到的几个配置选项进行说明。需要提醒的是，默认情况下 Fedora Core 系统的 /etc 目录下没有 rsyncd.conf 这个文件，自行创建该文件即可。rsync 服务器默认运行在 TCP 端口 873 上。



在本文的示例中，要备份的数据统统位于 rsync-server 主机上的/home 目录下，rsync-server 主机上的配置文件/etc/rsyncd.conf 的内容如下所示：

```
[home]
path=/home
uid=0
gid=0
```

配置文件 rsyncd.conf 的写法非常类似于 samba 服务器配置文件。rsyncd.conf 文件主要包括两类配置信息：全局参数和模块参数。全局参数配置包含着用来控制 rsync 服务器运行方式的信息，比如日志文件的位置、套接口选项等，全局参数应该出现在第一个模块配置之前。在全局参数之后，是模块参数配置。不同的模块（module）以模块名区分，模块名包含在方括号之中，模块名之后的行都是针对该模块的配置信息，直到遇到下一个模块为止。rsyncd.conf 配置文件是基于行的——也就是说一个文本行就是一条配置信息。每一行中出现的第一个等号是重要的，在这个等号之前是参数名，等号之后是参数值。也就是说，每一个参数都是以 name=value 的形式出现在每一行上的。在第一个等号之前或之后的空格都将被忽略。以#开始的行将被视为注释行而被忽略，如果参数信息太长以至于在一行上写不下，那么可以使用\将其延续到下一行上。参数值分为两种：字符参数和布尔参数，如果是布尔参数，那么可以使用的值为：yes/no，1/0，true/false。布尔参数值不用关心字符的大小写，但是字符参数值却是区分大小写的。

有了上述说明之后，就可以来看一下上面的示例配置文件的含义了：

- [home]
创建一个名为 home 的模块；
- path=/home
home 模块要导出的路径为/home。每一个模块都必须具有 path 参数；
- uid=0
uid 参数指定了以何种用户权限访问该模块（即文件传输时的用户权限）；
- gid=0
gid 参数指定了以何种组权限访问该模块（即文件传输时的组权限）；

除了上面这些模块参数之外，对于模块还有其他很多参数可以选用，比如，可以指定访问控制选项以控制哪些主机可以访问 rsync 服务器，哪些主机不可以；可以指定哪些文件应当排除在被传输的文件之外（exclude）；或者可以指定哪些文件应当被包括在被传输的文件集合之内（include）；可以指定 rsync 客户端的超时时间；可以指定是否采用压缩传输；等等，详见 rsyncd.conf(5)。在指定了 rsync 服务器配置选项之后，应当启动 rsync 服务器，采用如下命令即可：

```
[root@localhost etc]# rsync --daemon
[root@localhost etc]# ps auxw | grep rsync
root      4988  0.0  0.0 1688   4 ?        S    Nov07   0:00 rsync --daemon
root      7427  0.0  0.2 3544  552 pts/1   S    01:08   0:00 grep rsync
```



rsync 服务器在启动之后即转入后台运行。ps 命令显示的结果表明 rsync 服务器程序进入了睡眠状态。当有客户端与之建立连接时，rsync 服务器将被唤醒。

2 运行客户端

rsync 客户端有六种运行方式：

- 拷贝本地文件。调用方式为：

```
rsync [OPTION]... SRC [SRC]... DEST
```

注：SRC 和 DEST 都是本地文件（或目录）。

- 通过远程 shell 从本地主机拷贝文件到远程主机。调用方式为：

```
rsync [OPTION]... SRC [SRC]... [USER@]HOST:DEST
```

注：SRC 是本地文件，DEST 是远程主机上的文件。

- 通过远程 shell 从远程主机拷贝文件到本地主机。调用方式为：

```
rsync [OPTION]... [USER@]HOST:SRC DEST
```

注：SRC 是远程主机上的文件，DEST 是本地文件。

- 从远程 rsync 服务器拷贝文件到本地主机。调用方式为：

```
rsync [OPTION]... [USER@]HOST::SRC [DEST]
```

或者

```
rsync [OPTION]... rsync://[USER@]HOST[:PORT]/SRC [DEST]
```

注：SRC 是远程主机上的文件，DEST 是本地文件。

- 从本地主机拷贝文件到远程 rsync 服务器。调用方式为：

```
rsync [OPTION]... SRC [SRC]... [USER@]HOST::DEST
```

注：SRC 是本地文件，DEST 是远程主机上的文件。

- 列出远程主机上的文件列表。调用方式为：

```
rsync [OPTION]... rsync://[USER@]HOST[:PORT]/SRC
```

注：SRC 是远程主机上的文件（或目录）。

在上述六种运行模式中，除了列表模式（最后一种）外，其余模式都要求 SRC 和 DEST 中至少一个是本地文件（或目录）。

rsync 的传输既可通过远程 shell 通道完成，也可以直接以服务器/客户端的方式完成。如果是后者，实际上就是通过套接口实现的数据传输。在本文的示例中，采用服务器/客户端的方式，不采用远程 shell 通道作为传输接口。

本文的示例中，在 rsync-client 上如下运行 rsync 客户端程序：

```
[root@localhost backup_31_64]# rsync -v --archive --recursive --progress --delete --t
```



```
imes --perms --compress rsync://rsync-server:/home /home/backup_rsync_server/
```

该命令将把 rsync-server 主机上的 /home 目录下的所有数据同步到本地主机的 /home/backup_rsync_server/ 目录下。在这个命令中使用的几个选项的含义如下：

- -v
输出详细信息。
- --archive
采用归档模式。在这种模式下，几乎所有东西都将被保留（如用户、组、权限、时间等）。
- --recursive
这个选项告诉 rsync 要递归地拷贝目录。如果不指定该选项，rsync 不会拷贝目录。
- --progress
这个选项告诉 rsync 打印传输的进展情况。通常与 -v 联合使用。
- --delete
这个选项告诉 rsync 删除任何存在于接收端但不存在于发送端的文件。
- --times
这个选项告诉 rsync 将文件的更改时间与文件一起传输，并在远程系统上更新它们。
- --perms
这个选项命令 rsync 确保远程系统上的文件权限与本地系统上的文件权限相同。
- --compress
这个选项告诉 rsync 将数据压缩之后再传输。该选项在低速链路上非常有用。

可以看到，这里采用的是 rsync 六种工作方式中的第四种工作方式，rsync:// 是一个 URL 标记。以上列举的是 rsync 客户端程序常用的选项，rsync 命令完整的选项列表请参考 rsync(1)。

发出上述命令之后，rsync 开始从 rsync-server 主机同步数据到本地主机的 /home/backup_rsync_server/ 目录下。同步过程中的文件传输等信息将在控制台上显示出来。

3 结论

本文展示了在 Linux 系统下如何通过 rsync 工具完成备份工作的基本流程。尽管只就使用 rsync 工具的一种简单情形做了说明，但是在此基础上很容易扩展到 rsync 的其它工作模式上去。rsync 工具灵活的选项、增量更新的备份方式、多样的功能模式使得它非常适合于在多台主机之间进行远程备份。