

5 款可替代du 命令的工具！

良許Linux 2021-12-15 07:33

以下文章來源於杰哥的IT之旅，作者JackTian



杰哥的IT之旅

一名致力於在技術道路上的終身學習者、實踐者、分享者，一位忙起來又偶爾偷懶的原...



良許Linux

技術分享 | 資料共享 | 英語交流

后台回复【进群】，带你进入高手如云交流群



在Linux 環境下，du 命令大家都不陌生，它可以對Linux 系統中的文件和目錄空間使用情況進行計算和匯總。

當文件佔用磁盤空間80 % 且在暫時無法擴容空間的情況下，我們可以找出系統中佔比磁盤空間較大的具體目錄下的文件進行清理，只需執行如下命令即可查看文件夾佔用空間大小。

```
# du -sh
```

進入佔用空間較大的文件夾執行如下命令可以將當前目錄下的文件從大到小排序。

```
# du -sh * |sort -rh
```

將輸出後佔用空間較大的文件根據實際情況判斷是否可以正常刪除，從而來釋放磁盤空間並進行觀察磁盤使用率。

當文件佔用磁盤空間較多時，du 就會執行很慢，每次會反復執行來找出大文件，效率較低，此時 `ncdu` 就可以來解決該問題。

一、ncdu

ncdu 是一個使用C 語言和ncurses 接口開發的簡易快速的磁盤用量分析器，可以用來查看目錄或文件在本地或遠程系統上佔用磁盤空間的情況。

Ubuntu 21.10 安裝ncdu：

```
# sudo apt-get install ncdu
```

ncdu 參數選項：

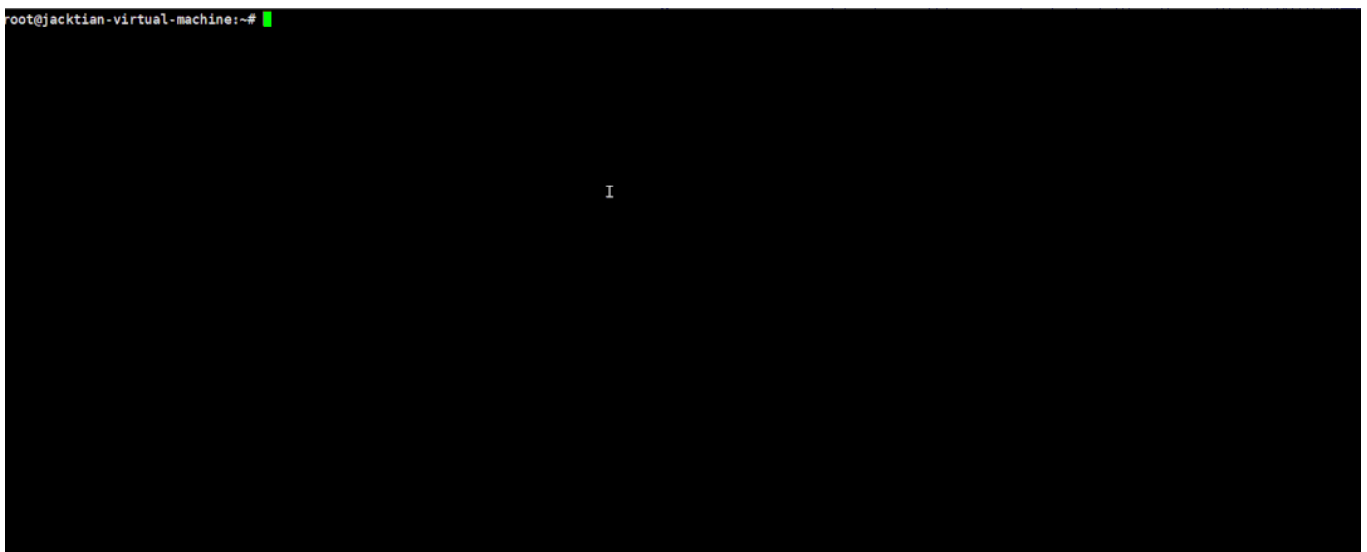
- `-h`：幫助信息
- `-q`：靜音模式，刷新間隔為2 秒
- `-v`：打印版本
- `-x`：相同的文件系統
- `-e`：啟用擴展信息
- `-r`：只讀
- `-o FILE`：將掃描目錄導出為文件
- `-f FILE`：從文件中導入掃描目錄
- `-0,-1,-2`：掃描時使用的UI (0=none,2=full ncurses)
- `--si`：使用基數10 (SI) 前綴而不是基數 2
- `--exclude PATTERN`：排除匹配PATTERN 的文件
- `-X, --exclude-from FILE`：排除與文件中任何模式匹配的文件
- `-L, --follow-symlinks`：按照符號鏈接（不包括目錄）
- `--exclude-caches`：排除包含 CACHEDIR.TAG 的目錄
- `--exclude-kernfs`：排除 Linux 偽文件系統 (procfs、sysfs、cgroup...)

- `--confirm-quit`：确认退出
- `--color SCHEME`：设置配色方案

如果不使用任何参数选项，可直接在需要分析的目录下执行命令 `ncdu`，等待 `ncdu` 扫描完毕后，此时就会弹出一个交互式的终端界面。

可以清晰的看出每个目录的大小，使用方向键 或 `Enter` 键可进入或退出所选的具体目录。

当发现某个目录或文件占用空间较大时，可根据上述使用键盘情况找到那个文件，切合实际情况确认如果可以删除的话，无需退出 `ncdu`，选中要删除的目录或文件后，按 `d` 键即可删除，若删除完毕，按 `q` 键即可退出。



二、dust

`dust` (`du` + `rust` = `dust`) 是用 `rust` 编写的一个免费、开源、更直观的 `du` 工具，也是一个提供文件类型和元数据的工具。

Ubuntu 21.10 安装 dust:

`du-dust_0.7.1_amd64.deb` 是 `dust` GitHub 存储库中的最新版本，从 `dust` 的发布页面下载 `.deb` 文件，可执行如下命令进行下载和安装。

```
# wget https://github.com/bootandy/dust/releases/du-dust_0.7.1_amd64.deb
# dpkg -i du-dust_0.7.1_amd64.deb
```

对于其他系统，可以参考 GitHub 上 `dust` 的其他操作系统安装方式。

dust 用法：

```
dust [FLAGS] [OPTIONS] [--] [inputs]...
```

dust 参数选项：

- **-f**：目录“大小”是子文件/目录的数量，而不是磁盘大小
- **-s**：使用文件长度而不是块
- **-p**：子目录的路径不会缩短
- **-h**：帮助信息
- **-i**：不显示隐藏文件
- **-x**：仅计算与提供的目录位于同一文件系统上的文件和目录
- **-b**：不会显示百分比条或百分比
- **-c**：不会打印颜色（通常最大的目录是彩色的）
- **-r**：打印树倒置（最大的最高）
- **-t**：仅显示这些文件类型
- **-V**：打印版本信息
- **-d**：显示深度
- **-e**：仅包含与此正则表达式匹配的文件。对于 png 文件类型：-e ".png\$"
- **-X**：排除任何具有此名称的文件或目录
- **-v**：排除与此正则表达式匹配的文件。忽略 png 文件类型：-v ".png\$"
- **-n**：要显示的输出行数。这是高度，（但 h 有帮助）[默认值：23]
- **-w**：指定输出的宽度覆盖终端宽度的自动检测

dust 使用：

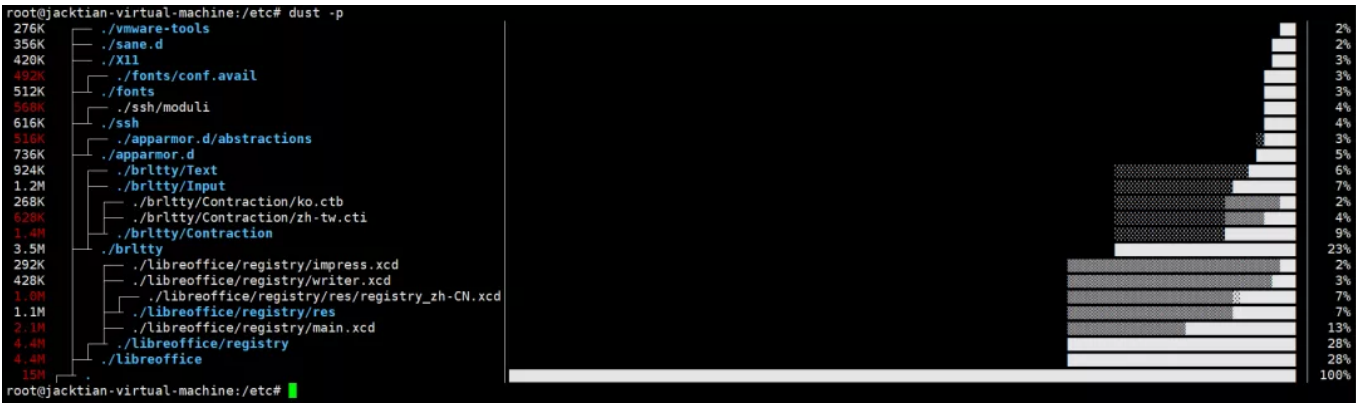
查看当前目录及所有子目录下的文件大小，可以执行以下命令：

```
# dust
```



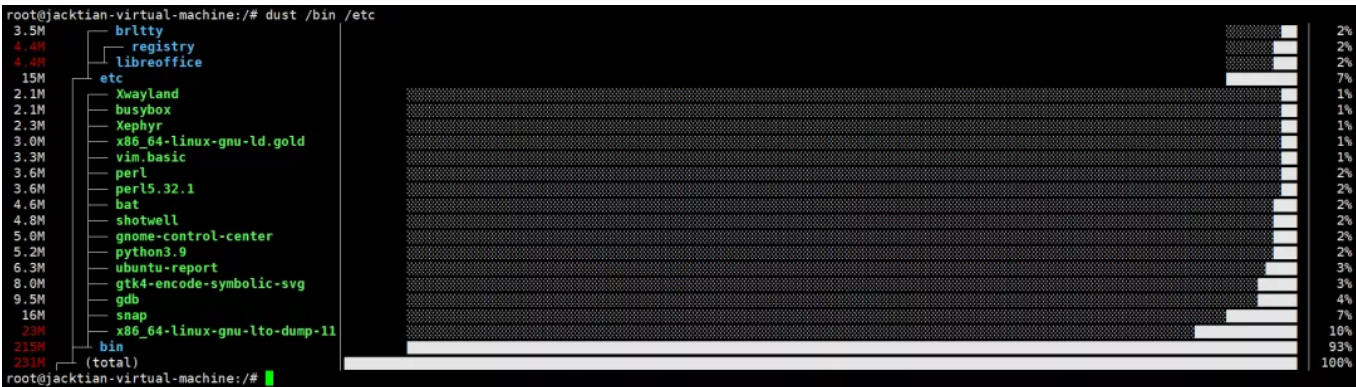
-p 参数可以按照从当前目录起始的完整目录进行显示

```
# dust -p
```



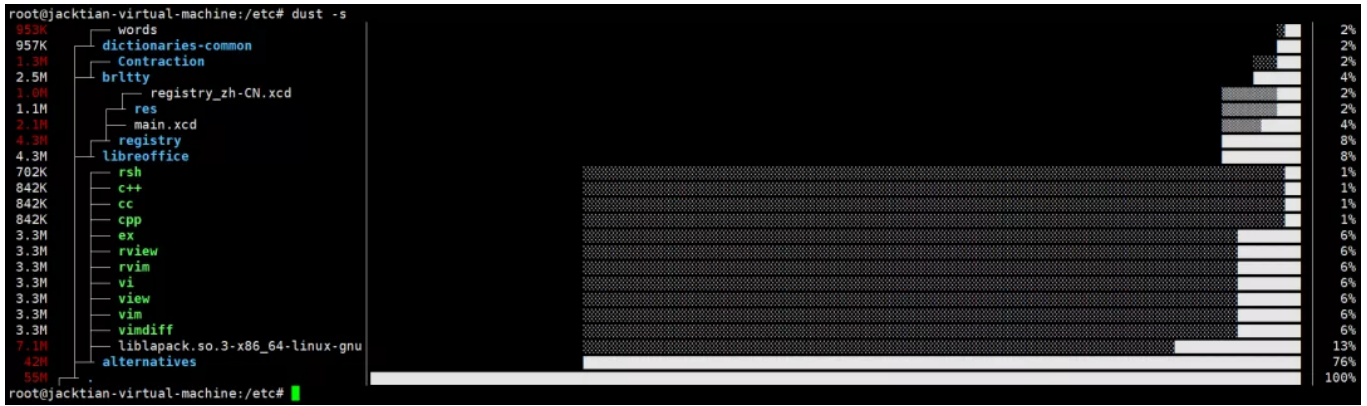
如果只查看多个目录的大小，只需要同时列出这些目录，并用空格分隔开即可：

```
# dust /bin /etc
```



显示文件的长度：

```
# dust -s
```



只显示 10 个目录：

```
# dust -n 10
```



三、duu

duu 是 Directory Usage Utility（目录使用实用程序）的缩写，基于 Python 语言编写的查看指定目录大小的工具，具有跨平台特性，可以在 Windows、Linux 和 MacOS 操作系统上使用，它是以千字节为单位显示目录磁盘使用情况。

Ubuntu 21.10 安装 duu：

安装 duu 工具之前需要先确保系统有没有安装 Python 3 环境，没有的话需要安装一下，有的话忽略即可，不过目前在大多数 Linux 发行版中默认软件仓库都是有 Python 3 的。

v2.22.tar.gz 是 duu GitHub 存储库中最新版本的源码包，从 duu 的发布页面下载 `.tar.gz` 文件，可执行如下命令进行下载和安装。

```
# wget https://github.com/jftuga/duu/archive/refs/tags/v2.22.tar.gz
# tar -zxvf v2.22.tar.gz
```

解压完毕后，进入 duu-2.22 文件夹，查看 `duu.py` 文件是否存在。

```
# ls | grep duu
duu-2.22
```

```
# cd duu-2.22/  
# ls  
duu.py  LICENSE  README.md
```

duu 用法：

```
duu.py [-h] [-b] [-e] [-q] [-s STATUS] [-n] [-N] [-f] [-S] [-H] [-T THREADS] [-x EXCLUDE]
```

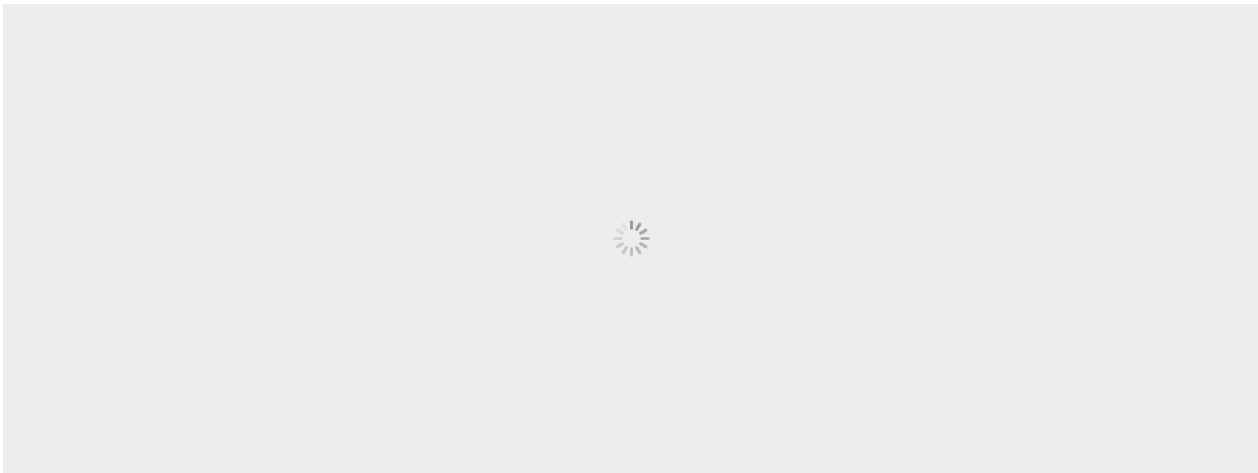
duu 参数选项：

- **-h**：帮助信息
- **-b**：不打印摘要或统计数据，专门用于排序时使用
- **-e**：总结文件扩展名
- **-q**：不显示单个目录
- **-s**：将处理状态发送到 STDERR，每个 STATUS 目录数
- **-n**：跳过以"."开头的目录
- **-N**：不递归
- **-f**：显示每个目录中的文件数
- **-S**：显示均值、中值、众数和标准差文件统计信息
- **-H**：以更易读的格式显示数字
- **-T**：并发线程数，考虑 SAN
- **-x**：排除不分大小写字符串的冒号分隔列表
- **-X**：排除不分大小写正则表达式的冒号分隔列表
- **-o**：输出到 CSV 文件

duu 使用：

查看当前目录的大小，只需要执行如下命令：

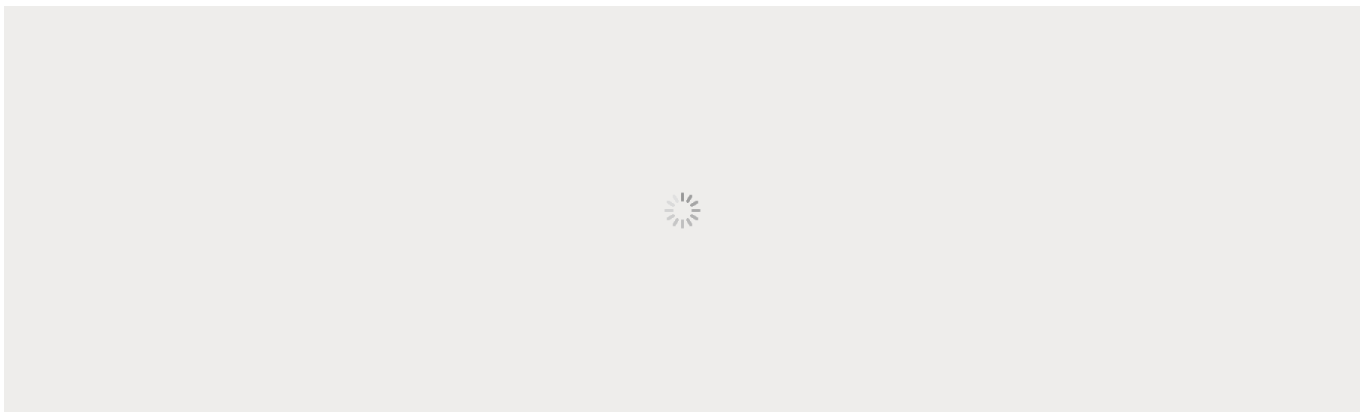
```
# python3 duu.py
```

如上图，可以看出 duu 会显示当前目录下文件和目录的数量并按照 Bytes、KB、MB 为单位显示这些文件的总大小，以及每个文件的大小。

查看某个目录的大小，只需要把目录的绝对路径加上即可：

```
# python3 duu.py /etc/init.d/
```



四、diskus

diskus 也是基于 Rust 编写的一个小型、快速、可替代 `du -sh` 命令的开源工具，diskus 会计算当前目录下所有文件的大小，执行命令 `diskus` 效果和 `du -sh`、`du -sh --bytes` 是一样的。

根据 diskus GitHub 存储库的开发者表示，在他的 8 核笔记本电脑上，以 hyperfine 命令行基准测试工具将中等大小的文件夹（15GB、100k 目录、400k 文件）进行测试，对于较小的文件夹测试实际上是没有意义的，因为所有程序都会在合理的时间内完成，而不会中断您的工作流程，从而得出的结论是：diskus 比 du 冷磁盘缓存快十倍左右，热磁盘缓存快三倍多。

以下是冷磁盘缓存和热磁盘缓存基于 `diskus`、`du -sh`、`sn p -d0 -j8` 和 `dust -d0` 命令测试出的结果对比。

冷磁盘缓存

Command	Mean [s]	Min [s]	Max [s]	Relative
diskus	1.746 ± 0.017	1.728	1.770	1.00
du -sh	17.776 ± 0.549	17.139	18.413	10.18
sn p -d0 -j8	18.094 ± 0.566	17.482	18.579	10.36
dust -d0	21.357 ± 0.328	20.974	21.759	12.23

热磁盘缓存

Command	Mean [s]	Min [s]	Max [s]	Relative
diskus	500.3 ± 17.3	472.9	530.6	1.00
du -sh	1098.3 ± 10.0	1087.8	1122.4	2.20
sn p -d0 -j8	1122.2 ± 18.2	1107.3	1170.1	2.24
dust -d0	3532.1 ± 26.4	3490.0	3563.1	7.06

Ubuntu 21.10 安装 diskus:

diskus_0.7.0_amd64.deb 是 diskus GitHub 存储库中的最新版本，从 diskus 的发布页面下载 `.deb` 文件，可执行如下命令进行下载和安装。

```
# wget https://github.com/sharkdp/diskus/releases/download/v0.7.0/diskus_0.7.0_amd64.deb
# dpkg -i diskus_0.7.0_amd64.deb
```

对于其他系统，可以参考 GitHub 上 diskus 的其他操作系统安装方式进行安装。

diskus 用法:

```
diskus [OPTIONS] [path]...
```

diskus 参数选项:

- `-j`：设置线程数（默认：3 x num cores），文件大小的输出格式（十进制：MB，二进制：MiB）[默认：十进制][可能的值：十进制、二进制]
- `-v`：不隐藏文件系统错误
- `-b`：计算表观大小而不是磁盘使用量
- `-h`：帮助信息
- `-v`：版本信息

diskus 使用：

使用 `du -sh 路径` 的命令可以查看某个目录的大小，这里的 `-s` 参数表示显示总大小。

```
# du -sh /etc
16M    /etc
```

如果使用 `diskus` 可直接显示当前目录的总大小。

```
# diskus
```

如果你想查看某个目录的大小，可以在 `diskus` 后面加绝对路径。

```
# diskus /etc
16.66 MB (16,662,528 bytes)
```

五、tin-summer

tin-summer (sn) 是基于 Rust 语言编写的开源工具，也是 `du` 命令的替代品，可以用于查找占用磁盘空间的文件，以更好的输出、更清晰的命令和默认值，而且由于多线程，它甚至可以在计算大目录下的大小时会比 `du` 命令运行得更快，tin-summer 与 `du` 命令之间的区别是前者读取文件的大小，而后者则读取磁盘使用情况。

此外，tin-summer 的开发者也针对 `du` 和 `sn` 做了鲜明的使用理由对比。

使用 du 的理由：

- 可读取磁盘使用情况，而不仅仅是文件大小
- 可选的取消引用符号链接

- 可在小目录上稍微快一点
- 稳定且支持良好

使用 sn 的理由：

- 在大目录下更快
- 使用正则表达式进行排除，使其在与 --exclude 标志一起使用时比 du 快得多。
- 默认为可读的输出
- 彩色输出
- 通过获得更好的帮助
- 提供排序输出
- 查找构建工件
- 读取文件大小，而不是磁盘使用情况
- 在 Rust 中可扩展

同时，也做了基准测试，从而来对比 sn 和 du 执行不同目录大小的时间结果。

Directory Size	Tool	Command	Time
600MB	sn	sn p	60.74 ms
600MB	sn	sn d	99.92 ms
600MB	du	du -hacd2	88.28 ms
4GB	sn	sn p	185.2 ms
4GB	sn	sn d	271.9 ms
4GB	du	du -hacd2	195.5 ms
700MB	sn	sn p	91.05 ms
700MB	sn	sn d	176.3 ms
700MB	du	du -hacd2	153.8 ms
7MB	sn	sn p	19.48 ms
7MB	sn	sn d	12.72 ms

Directory Size	Tool	Command	Time
7MB	du	du -hacd2	10.13 ms

安装 tin-summer:

```
# curl -LSfs https://japaric.github.io/trust/install.sh | sh -s -- --git vmchale/tin-sum
```

也可以使用 cargo 软件包管理器安装 tin-summer，前提是需要要在系统上安装 Rust。如果 Rust 已安装好的话，可执行如下命令：

```
# cargo install tin-summer
```

如果上述两种安装方式都尚未成功安装 tin-summer，还可以通过 tin-summer GitHub 存储库的发布区下载最新版本的二进制文件进行手动安装。

tin-summer 的使用：

查看当前目录文件大小，无需添加任何参数，tin-summer 也是默认以易读的格式向用户进行输出，可执行如下命令：

```
# sn f
```

查看某个特定目录的文件大小，可执行如下命令：

```
# sn f <path-to-the-directory>
```

在当前目录中搜索包含构建工件的目录，可执行如下命令：

```
# sn ar
```

查找占用超过 200MB 磁盘空间的工件或包含工件的目录，可执行如下命令：

```
# sn ar -t200M
```

获取 \$DIR 中 10 个最大目录的排序列表，可执行如下命令：

```
# sn sort $DIR -n12
```

sn 实际在操作大目录的时候速度比较快，在操作小目录的时候，速度会相对比较慢一些。

参考

dust: <https://github.com/bootandy/dust/>

duu: <https://github.com/jftuga/duu>

diskus: <https://github.com/sharkdp/diskus/>

tin-summer: <https://github.com/vmchale/tin-summer>

本公众号全部博文已整理成一个目录，请在公众号里回复「m」获取！

推荐阅读：

再见Xshell！这个开源的终端工具更酷炫！

手把手教你给 SSH 启用二次身份验证

全中国一共有多少IP地址？

5T技术资源大放送！包括但不限于：C/C++，Linux，Python，Java，PHP，人工智能，单片机，树莓派，等等。在公众号内回复「1024」，即可免费获取！！



良许Linux

良许，自学转行IT并顺利进入500强外企担任Linux开发工程师。公众号分享大量Linux...
210篇原创内容



公众号

阅读原文

喜欢此内容的人还喜欢

再见Xshell！这个开源的终端工具更酷炫！

良许Linux



满背纹身 满背关公纹身 关羽纹身 关二哥纹身 满背纹身手稿
纹身手稿图



夏延軍已任河北省委統戰部部長
中國婦女報

