

内核工匠

码龄11年🏆 暂无认证

100

9433

1万+

14万+



原创

周排名

总排名

访问

等级

1988

829

188

75

1050

积分

粉丝

获赞

评论

收藏



私信

关注

搜博文文章🔍

热门文章

FUSE文件系统🔗 7335

纯干货, PSI 原理解析与应用🔗 6893

GKI改造原则、机制和方法🔗 6075

UFS特性介绍🔗 5730

Android ART dex2oat 浅析🔗 5341

最新评论

Linux SPI 驱动

紫荆鱼: 不错不错, 期待回访👍

cpufreq schedutil原理剖析

weixin_38786831: 请问, CPU最大算力是什么指标

Linux thermal governor之IPA分析

Tae - Cho: 谢谢博主, 对我的帮助很大

盘点只读压缩文件系统

烤红薯的红薯: 为什么部分解压的数据需要重复使用呀?

Load_balance函数情景分析

jrglinux: 能像贵单位这么沉浸是分析linux内核实现并产出高质量文章且愿公开分享的...

您愿意向朋友推荐“博客详情页”吗?











强烈不推荐

不推荐

一般般

推荐

强烈推荐

Linux overlayfs文件系统介绍

原创内核工匠🕒 于 2021-10-15 17:00:00 发布👁 243🌟 收藏 3

文章标签:linuxjavapython数据库大数据

版权

OverlayFS, 顾名思义是一种堆叠文件系统, 可以将多个目录的内容叠加到另一个目录上。OverlayFS并不直接涉及磁盘空间结构, 看起来像是将多个目录的文件按照规则合并到同一个目录。且对多个源目录具体使用文件系统类型没有要求, 即使各个源目录的文件系统类型不同也不影响使用。

一、挂载OverlayFS

下面就让我们来看看如何挂载一个OverlayFS文件系统:

```
mount -t overlay -o lowerdir=/lower,upperdir=/upper,workdir=/work overlay /merged
```

上面的命令可以将"lowerdir"和"upper"目录堆叠到/merged目录, "workdir"工作目录要求是和"upperdir"目录同一类型文件系统的空目录。

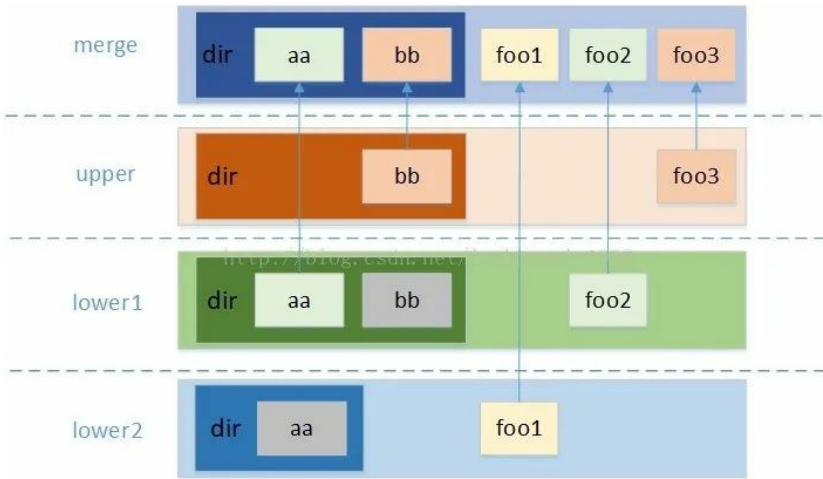
也可以省略upperdir和workdir参数, 但/merged为只读属性了:

```
mount -t overlay -o lowerdir=/upper:/lower overlay /merged
```

也可支持多lowerdir目录堆叠:

```
mount -t overlay -o lowerdir=/lower1:/lower2:/lower3,upperdir=/upper,workdir=/workoverlay /merged
```

lowerdir的多层目录使用":"分隔开, 其中层级关系为/lower1> /lower2 > /lower3。



在使用如上mount进行OverlayFS合并之后, 遵循如下规则:

- lowerdir和upperdir两个目录存在同名文件时, lowerdir的文件将会被隐藏, 用户只能看到upperdir的文件。
- lowerdir低优先级的同目录同名文件将会被隐藏。
- 如果存在同名目录, 那么lowerdir和upperdir目录中的内容将会合并。

目录

- 一、挂载OverlayFS
- 三、创建文件和目录
- 四、rename目录
- 五、Android中的应用
- 六、小结
- 参考文献

最新文章

JVM 垃圾回收算法与ART CC回收器实现概述

FreeRTOS内存机制

Android 设备树的匹配流程

2022年	3篇	2021年	47篇
2020年	54篇	2019年	1篇

- 当用户修改mergedir中来自upperdir的数据时，数据将直接写入upperdir中原来目录中，删除文件也同理。
- 当用户修改mergedir中来自lowerdir的数据时，lowerdir中内容均不会发生任何改变。因为lowerdir是只读的，用户想修改来自lowerdir数据时，overlayfs会首先拷贝一份lowerdir中文件副本到upperdir中（这也被称作OverlayFS的copy-up特性）。后续修改或删除将会在upperdir下的副本中进行，lowerdir中原文件将会被隐藏。
- 如果某一个目录单纯来自lowerdir或者lowerdir和upperdir合并，默认无法进行rename系统调用。但是可以通过mv重命名。如果要支持rename，需要CONFIG_OVERLAY_FS_REDIRECT_DIR。

一般lowerdir为只读文件系统，upperdir为可写文件系统，这形成了一个有趣的机制，似乎我们可以修改lowerdir下的文件或目录，lowerdir看上去变成了一个可读写的文件系统。

二、删除文件和目录

为了支持rm和rmdir而又不修改lower文件系统，需要在upper文件系统中记录文件或目录已经被删除。OverlayFS引入了whiteout文件的概念。如果需要删除lower层的文件或目录，需要在upper层创建一个whiteout文件。

```
test@ubuntu:overlay$ sudo mount -t overlay -o lowerdir=lower,upperdir=upper,workdir=work overlay merged
sudo: unable to resolve host engineer
[sudo] password for :
test@ubuntu:overlay$ mount | grep overlay
overlay on /work/ /lab/overlay/merged type overlay (rw,relatime,lowerdir=lower,upperdir=upper,workdir=work)
test@ubuntu:overlay$ ls -lni merged/
total 8
1053684 -rw-rw-r-- 1 1031 1031 3 Oct 14 10:22 aa
1053683 -rw-rw-r-- 1 1031 1031 0 Oct 13 10:34 bb
1053685 -rw-rw-r-- 1 1031 1031 0 Oct 14 10:22 cc
1089625 drwxrwxr-x 1 1031 1031 4096 Oct 14 10:21 dir
test@ubuntu:overlay$ ls -lni upper/
total 8
1053684 -rw-rw-r-- 1 1031 1031 3 Oct 14 10:22 aa
1053685 -rw-rw-r-- 1 1031 1031 0 Oct 14 10:22 cc
1089616 drwxrwxr-x 2 1031 1031 4096 Oct 14 10:21 dir
test@ubuntu:overlay$ ls -lni lower/
total 8
1053682 -rw-rw-r-- 1 1031 1031 4 Oct 13 10:35 aa
1053683 -rw-rw-r-- 1 1031 1031 0 Oct 13 10:34 bb
1089625 drwxrwxr-x 2 1031 1031 4096 Oct 13 11:00 dir
test@ubuntu:overlay$ cat merged/aa
up
test@ubuntu:overlay$ cat upper/aa
up
test@ubuntu:overlay$ cat lower/aa
low
test@ubuntu:overlay$ rm -rf merged/*
test@ubuntu:overlay$ ls -lni merged/
total 0
test@ubuntu:overlay$ ls -lni upper/
total 0
1053687 c----- 1 0 0 0, 0 Oct 14 10:28 aa
1053684 c----- 1 0 0 0, 0 Oct 14 10:28 bb
1050344 c----- 1 0 0 0, 0 Oct 14 10:28 dir
test@ubuntu:overlay$ ls -lni lower/
total 8
1053682 -rw-rw-r-- 1 1031 1031 4 Oct 13 10:35 aa
1053683 -rw-rw-r-- 1 1031 1031 0 Oct 13 10:34 bb
1089625 drwxrwxr-x 2 1031 1031 4096 Oct 13 11:00 dir
test@ubuntu:overlay$
```

可以看到删除merged目录下的文件或目录后，在upper层新建了aa、bb、dir三个whiteout文件，whiteout文件不是普通文件，而是主/次设备号都是0的字符设备。只存在于upper的文件cc直接删除就可以了。

三、创建文件和目录

创建操作与删除操作类似，都是在upper层进行修改。创建文件直接在upper层新增文件即可，如果upper层存在对应的whiteout文件，先删除whiteout文件再创建文件。创建目录与创建文件类似，区别在于upper层存在whiteout文件时，删掉whiteout文件创建目录，如果就此结束，lower层对应目录（因为有whiteout文件）的文件就被显示到merged目录了，所以还需要将目录的"trusted.overlay.opaque"属性设为"y"（所以这也就需要upper层所在的文件系统支持xattr扩展属性），OverlayFS在读取上下层存在同名目录的目录项时，如果upper层的目录被设置了opaque属性，它将忽略这个目录下层的所有同名目录中的目录项，以保证新建的目录是一个空的目录。

```
test@ubuntu:overlay$ rm -rf merged/*
test@ubuntu:overlay$ ls -lni merged/
total 0
```

```
test@ubuntu:overlay$ ls -lni upper/
total 0
1053687 c----- 1 0 0 0, 0 Oct 14 10:28 aa
1053684 c----- 1 0 0 0, 0 Oct 14 10:28 bb
1050344 c----- 1 0 0 0, 0 Oct 14 10:28 dir
test@ubuntu:overlay$ ls -lni lower/
total 8
1053682 -rw-rw-r-- 1 1031 1031 4 Oct 13 10:35 aa
1053683 -rw-rw-r-- 1 1031 1031 0 Oct 13 10:34 bb
1089625 drwxrwxr-x 2 1031 1031 4096 Oct 13 11:00 dir
test@ubuntu:overlay$
test@ubuntu:overlay$
test@ubuntu:overlay$ touch merged/ee
test@ubuntu:overlay$ mkdir merged/dir
test@ubuntu:overlay$ ls -lni merged/
total 4
1089616 drwxrwxr-x 2 1031 1031 4096 Oct 14 10:52 dir
1053685 -rw-rw-r-- 1 1031 1031 0 Oct 14 10:52 ee
test@ubuntu:overlay$ ls -lni merged/dir
total 0
test@ubuntu:overlay$ ls -lni upper/
total 4
1053687 c----- 1 0 0 0, 0 Oct 14 10:28 aa
1053684 c----- 1 0 0 0, 0 Oct 14 10:28 bb
1089616 drwxrwxr-x 2 1031 1031 4096 Oct 14 10:52 dir
1053685 -rw-rw-r-- 1 1031 1031 0 Oct 14 10:52 ee
test@ubuntu:overlay$ ls -lni lower/
total 8
1053682 -rw-rw-r-- 1 1031 1031 4 Oct 13 10:35 aa
1053683 -rw-rw-r-- 1 1031 1031 0 Oct 13 10:34 bb
1089625 drwxrwxr-x 2 1031 1031 4096 Oct 13 11:00 dir
test@ubuntu:overlay$ ls -lni lower/dir/
total 4
1053686 -rw-rw-r-- 1 1031 1031 5 Oct 13 11:00 dd
test@ubuntu:overlay$ sudo getfattr -n "trusted.overlay.opaque" upper/dir
sudo: unable to resolve host engineer
[sudo] password for :
# file: upper/dir
trusted.overlay.opaque="y"
test@ubuntu:overlay$
```

四、*rename*目录

当我们想重命名一个在lower层的目录，OverlayFS有两种处理方式：

1. 返回EXDEV错误码：rename系统调用试图穿过文件系统边界移动一个文件或目录时返回这个错误。这个是默认行为。
2. 当使能"redirect_dir"特性后，rename操作成功，在upper层产生一个副本目录。

有以下几种方式控制"redirect_dir"特性：

1. KernelConfig Options：
 - OVERLAY_FS_REDIRECT_DIR
 - OVERLAY_FS_REDIRECT_ALWAYS_FOLLOW

使能后，redirect_dir特性默认打开。

2. sys文件系统：

参照KernelConfig设置：

/sys/ [module](#)  /overlay/parameters/redirect_dir

/sys/module/overlay/parameters/redirect_always_follow

/sys/module/overlay/parameters/redirect_max

3. MountOptions：

redirect_dir=on/off/follow/nofollow

五、*Android*中的应用

OverlayFS文件系统可以类似达到把只读文件系统改为可写文件系统的效果，这一特性在Android开发的场景下得到应用，userdebug模式下我们adb remount后似乎就可以往/system/目录下push内容了，查看remount前后的mount信息，可以看到/system/目录被重新挂载成可读写的OverlayFS文件系统了：

remount前：

```
/dev/block/dm-1 on / type ext4 (ro,seclabel,relatime)
/dev/block/dm-8 on /vendor type ext4 (ro,seclabel,relatime)
/dev/block/dm-3 on /product type ext4 (ro,seclabel,relatime)
/dev/block/dm-2 on /system_ext type ext4 (ro,seclabel,relatime)
```

remount后：

```
overlay on /system type overlay (rw,seclabel,noatime,lowerdir=/system,upperdir=/cache/overlay/system/upper,workdir=/cache/overlay/system/work,override_creds=off)
overlay on /vendor type overlay (rw,seclabel,noatime,lowerdir=/vendor,upperdir=/cache/overlay/vendor/upper,workdir=/cache/overlay/vendor/work,override_creds=off)
overlay on /product type overlay (rw,seclabel,noatime,lowerdir=/product,upperdir=/cache/overlay/product/upper,workdir=/cache/overlay/product/work,override_creds=off)
overlay on /system_ext type overlay (rw,seclabel,noatime,lowerdir=/system_ext,upperdir=/cache/overlay/system_ext/upper,workdir=/cache/overlay/system_ext/work,override_creds=off)
```

重启：

```
overlay on /system type overlay (ro,seclabel,noatime,lowerdir=/system,upperdir=/cache/overlay/system/upper,workdir=/cache/overlay/system/work,override_creds=off)
overlay on /vendor type overlay (ro,seclabel,noatime,lowerdir=/vendor,upperdir=/cache/overlay/vendor/upper,workdir=/cache/overlay/vendor/work,override_creds=off)
overlay on /product type overlay (ro,seclabel,noatime,lowerdir=/product,upperdir=/cache/overlay/product/upper,workdir=/cache/overlay/product/work,override_creds=off)
overlay on /system_ext type overlay (ro,seclabel,noatime,lowerdir=/system_ext,upperdir=/cache/overlay/system_ext/upper,workdir=/cache/overlay/system_ext/work,override_creds=off)
```

system、vendor、product等目录是以ext4文件系统方式挂载的，remount后以OverlayFS挂载，之后重启也会以OverlayFS方式挂载，以使之前的修改生效。

system和vendor等的upperdir都在/cache可写文件系统中，往/system目录push东西实际上都存放在/cache/overlay/system/upper目录中了。实际的system分区并没有被修改，修改的文件全部存放在cache分区了。

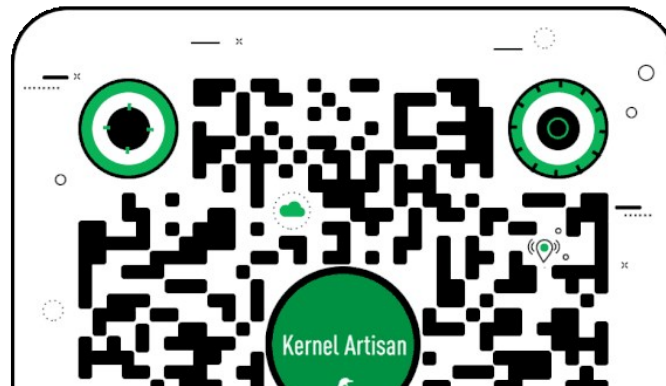
OverlayFS也被应用在把多个不同分区的目录堆叠到一个目录下面，可以更好做到软件系统的组件解耦，不同特性的组件内容分别放到不同分区，最后通过OverlayFS堆叠到一个目录下，提升软件的可维护性。

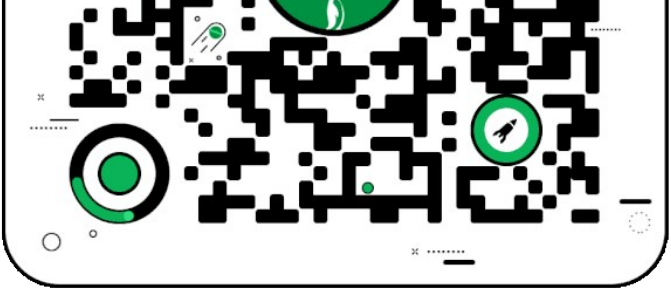
六、小结

OverlayFS以其独特的优势正得到越来越广泛的应用，Androiduserdebug/eng模式中使用其实现对系统分区的修改。OpenWRT系统也利用OverlayFS减少擦写闪存次数，延长闪存的使用寿命。OverlayFS更多的应用在云平台的容器镜像，基础的容器镜像通常不应被改变，新功能叠加可以通过OverlayFS实现且可以再次分发。

参考文献

1. <https://www.kernel.org/doc/html/latest/filesystems/overlayfs.html>
2. <https://blog.csdn.net/luckyapple1028/article/details/78075358>





长按关注

内核工匠微信

Linux 内核黑科技 | 技术文章 | 精选教程

深入理解overlayfs（一）：初识

My Linux Journey 2万+

Overlayfs是一种类似aufs的一种堆叠文件系统，于2014年正式合入Linux-3.18主线内核，目前其功能已经基本稳定（虽然还存在一些特性尚未实现）且被...

overlayfs简介及使用

chuanzhilong的博客 9424

overlayfs是目前使用比较广泛的层次文件系统，实现简单，性能较好。可以充分利用不同或则相同overlay文件系统的page cache，具有上下合并同名遮盖...



请发表有价值的评论，博客评论欢迎灌水，良好的社区氛围需大家一起维护。

抢沙发



评论

深入理解overlayfs(二):使用与原理分析_My Linux Journ...

1-10

由于目前主线内核对overlayfs正在不断的开发和完善中,因此不同的内核版本改动可能较大,本文尽量与最新的内核版本保持一致,但可能仍会存在细微的出...

Linux overlayFS文件系统_baynk的博客_linux overlay

1-8

Linux overlayFS文件系统 本文是在学习docker时,为了加强对docker的了解进行的知识的扩展。 OverlayFS是一种堆叠文件系统,它依赖并建立在其它的...

overlay文件系统浅析

热门推荐

esion23011的专栏 1万+

overlayfs文件系统类似于aufs，相比aufs，overlay实现更简洁，很早就合入了linux主线，合入主线后overlayfs修改为overlay。docker使用overlay文件...

Linux内核异常调试工具与方法

最新发布

weixin_39247141的博客 331

常见调试通道（工具） 串口调试通道 通过串口shell查看/修改寄存器: 查看一个寄存器空间内容： a.1 启动linux内核，进入shell界面 a.2 执行“cd /sys/class...

linux overlay目录作用,把玩overlay文件系统_weixin_39...

1-12

作为UnionFS之一的OverlayFS就是这么一种文件系统,它于2014年被合并到3.18内核,顾名思义其主要特性就是“覆盖”,可以结合下图进行理解(图非原创)。 ...

linux文件系统详解_weixin_30352645的博客

1-2

auto fs4 ceph debug fs efs ext4 fuse hfs jbd2 lockd nfs notify overlayfs quota sys fs xfs #硬盘设备识别 MBR,全称Master Boot Record,可译为硬盘主引导...

使用Squashfs和Overlayfs提高嵌入式Linux文件系统可靠性

toradexsh的博客 1141

By Toradex胡珊逢 在使用嵌入式Linux系统的时，会出现由于设备意外断电引起文件系统损坏而最终使该设备无法启动的现象。为了应对这种情况，通常会...

转载两篇，深入理解overlayfs（一）：初识

hziarm的博客 452

Overlayfs是一种类似aufs的一种堆叠文件系统，于2014年正式合入Linux-3.18主线内核，目前其功能已经基本稳定（虽然还存在一些特性尚未实现）且被...

深入理解overlayfs（二）：使用与原理分析

linyonghui1213的专栏 2253

在初步了解overlayfs用途之后，本文将介绍如何使用overlayfs以及理解该文件系统所特有的一些功能特性。由于目前主线内核对overlayfs正在不断的开发...

linux overlay目录作用,配置overlay(overlayfs)根目录挂载构建deb

weixin_32379547的博客 986

配置overlay(overlayfs)根目录挂载构建deb发布时间：2020-02-27 09:47:21来源：51CTO阅读：9628作者：btvnlue据wikipedia描述overlayfs在2014年融...

Linux层次文件系统overlayfs

发呆健将的博客 1290

简介 OverlayFS是一种目前广泛应用的层次文件系统，openwrt、树莓派、docker等上都有应用。它于2014年被合并到Linux3.18内核中，顾名思义其主要...

浅谈overlayfs

happyAnger6的专栏 396

overlayfs overlayfs试图在其它文件系统之上提供一个联合的文件系统视图 Upper and Lower overlayfs组合了2个文件系统--Upper文件系统和Lower文件...	
Docker存储驱动之overlay新镜像存储的实现和inode耗尽问题 镜像是按层下载和管理的，新镜像下载的文件临时存放在/var/lib/docker/tmp，文件命名方式是GetImageBlobxxx（xxx是一串随机数字），这些临时文件时...	idwtwt的专栏 6264
overlay文件系统解析 一个 overlay 文件系统包含两个文件系统，一个 upper 文件系统和一个 lower 文件系统，是一种新型的联合文件系统。overlay是“覆盖…上面”的意思...	guizaijianchic的博客 5747
linux内核的overlayfs,linux kernel Documentation filesystems overlayfs Written by: Neil BrownPlease see MAINTAINERS file for where to send questions.Overlay Filesystem=====This document describes a ...	weixin_39562615的博客 139
玩转开发板-几种常见的根文件系统的介绍 前言 根文件系统首先是一种文件系统，该文件系统不仅具有普通文件系统的存储数据文件的功能，但是相对于普通的文件系统而言它还是内核启动时...	PipiAvenger的博客 944
Linux overlayFS文件系统 本博文是在学习docker时，为了加强对docker的了解进行的知识的扩展。OverlayFS是一种堆叠文件系统，它依赖并建立在其它的文件系统之上（例如ext...	baynk的博客 926
linux内核的overlayfs,docker底层原理———overlayFS介绍与实践 8种机械键盘轴体对比本人程序员，要买一个写代码的键盘，请问红轴和茶轴怎么选？概述在docker底层原理的学习过程中，接触到了UnionFS的概念。U...	weixin_39708557的博客 62
overlay文件系统切换为根文件系统 内核配置 CONFIG_OVERLAY_FS 配置后编译烧写，可见支持了overlay文件系统 /tmp # cat /proc/filesystems nodev sysfs nodev rootfs nodev ramfs nod...	TSZ0000的博客 556
/etc 可读写内核 overlay 内核配置 CONFIG_OVERLAY_FS=y 文件系统挂载 # cd /data # mkdir etc data # mount -t overlay overlay -o lowerdir=/etc,upperdir=/data/etc,workdir=/d...	YuZhuQue的博客 123
linux内核开启overlay,linux-overlay启动挂载为只读或写文件异常 背景路由器在使用过程中，偶像启动挂载文件系统异常，查看打印后发现挂载overlay异常，打印错误代码-39，跟踪代码发现由于work目录非空导致清除...	weixin_30744857的博客 213
©2022 CSDN 皮肤主题：大白 设计师：CSDN官方博客 返回首页	
关于我们 招贤纳士 广告服务 开发助手 400-660-0108 kefu@csdn.net 在线客服 工作时间 8:30-22:00 公安备案号11010502030143 京ICP备19004658号 京网文〔2020〕1039-165号 经营性网站备案信息 北京互联网违法和不良信息举报中心 网络110报警服务 中国互联网举报中心 家长监护 Chrome商店下载 ©1999-2022北京创新乐知网络技术有限公司 版权与免责声明 版权申诉 出版物许可证 营业执照	

内核工匠

关注

👍 0

💬 0

🌟 3

📄

🔗

