

# notify+rsync实时同步 百万级别文件

原创 ronon77 最后发布于2016-01-27 09:12:03 阅读数 395 ☆ 收藏

## 背景

我们公司在用notify+rsync做实时同步，来解决分布式集群文件一致性的问题。但当web文件越来越多(百万级数量html,jpg等小文件)，同步就越来越慢，根本做不到照网上的调优方法都尝试过，问题根本没有解决。经过我一翻细致研究，终于把慢的核心问题研究明白，先总结一句 inotifywait响应不会有延迟，rsync也很快。大家的烦恼，那是因为网上的inotify+rsync的教程都是坑。下面我们来分析。

## inotifywait 单独分析

```
1 /usr/local/bin/inotifywait-mrq--format'%Xe %w%f'-emodify,create,delete,attrib/data/
```

执行上面命令，是让inotifywait监听/data/目录，当监听到有发生modify,create,delete,attrib等事件发生时，按%Xe %w%f的格式输出。

在/data/目录touch几个文件

```
1 touch/data/{1..5}
```

观看inotify输出

```
1 ATTRIB/data/1 --表示发生了ATTRIB事件路径为/data/1
2 ATTRIB/data/2
3 ATTRIB/data/3
4 ATTRIB/data/4
5 ATTRIB/data/5
```

知道上面的输出效果之后 我们应该想得到，可以用rsync获取inotifywait监控到的文件列表来做指定的文件同步，而不是每次都由rsync做全目录扫描来判断文件是否存在。

## 网上的inotify+rsync分析

我们来看网上的教程，我加了注释。(网上所有的教程基本都一模一样，尽管写法不一样，致命点都是一样的)

```
1 #!/bin/bash
2 /usr/bin/inotifywait-mrq--format'%w%f'-ecreate,close_write,delete/backup|while read file
3 #把发生更改的文件列表都接收到file 然后循环，但有什么鬼用呢？下面的命令都没有引用这个$file 下面做的是全量rsync
4 do
5     cd/backup&&rsync-az--delete/backup/rsync_backup@192.168.24.101::backup/--password-file=/etc/rsync.password
6 done
```

#注意看 这里的rsync 每次都是全量的同步(这就坑爹了)，而且 file列表是循环形式触发rsync，等于有10个文件发生更改，就触发10次rsync全量同步(简直就是噩梦)，直接写个死循环的rsync全量同步得了。

#有很多人会说 日志输出那里明明只有差异文件的同步记录。其实这是rsync的功能，他本来就只会输出有差异需要同步的文件信息。不信你直接拿这句rsync来跑试试。

#这种在需要同步的源目录文件量很大的情况下，简直是不堪重负。不仅耗CPU还耗时，根本不可以做到实时同步。

## 改良方法

要做到实时，就必须减少rsync对目录的递归扫描判断，尽可能的做到只同步inotify监控到已发生更改的文件。结合rsync的特性，所以这里要分开判断来实现一个目改变对应的操作。

脚本如下

1 | #!/bin/bash



ronon77

TA的个人主页 >

原创 748 粉丝 36 获赞 35 评论 10 访问 12万+

等级: 博客 4 周排名: 6万+ 积分: 878 总排名: 9万+

关注

私信

## 亿速云服务器免备案CN2商

亿速云香港服务器, CN2高速连接, 低可免换IP, 安全稳定, 技术团队在线稳定无忧

打开

## 最新文章

- python 类装饰器的四种场景写法
- monacmonaco-editor使用
- python logging 重复写日志问题
- CSS 黑魔法小技巧, 让你少写不必要的JS, 代码更优雅
- Redis 数据类型及应用场景

## 分类专栏

- 监控与安全、性能 70篇
- 管理哲学、业余提升 8篇
- 前端负载、开发、db 138篇
- 系统集成、架构 114篇
- Trouble Shooting 45篇

展开

## 归档

- 2018年1月 12篇
- 2017年12月 12篇
- 2017年11月 5篇
- 2017年10月 5篇
- 2017年9月 19篇
- 2017年7月 26篇

ync --daemon 发布的名称, rsync --daemon这里就不做介绍了, 网上搜一下, 比较简单。  
# rsync验证的密码文件

1

2

on定义的验证用户名

ync同步的特性, 这里必须要先cd到源目录, inotify再监听 ./ 才能rsync同步后目录结构一致, 有兴趣的同学可以进行各种尝试观看其效果

t'%Xe %n%f'-emodify,create,delete,attrib./|while read file # 把监控到有发生更改的"文件路径列表"循环

)# 把inotify输出切割 把事件类型赋值给INO\_EVENT

# 把inotify输出切割 把文件路径赋值给INO\_FILE

-----'

件的操作, 即使是新建目录, 要同步的也只是一个空目录, 不会影响速度。

ENT=~'MODIFY']])# 判断事件类型

client.pass\${INO\_FILE}\${user}@\${ip1}:::\${des}&&# INO\_FILE 变量代表路径哦

client.pass\${INO\_FILE}\${user}@\${ip2}:::\${des}

FILE}变量 即每次只针对性的同步发生改变的文件 然后用-R参数把源的目录结构递归到目标后面 保证目录结构一致性

/rsync-client.pass\$(dirname\${INO\_FILE})\${user}@\${ip1}:::\${des}&&

/rsync-client.pass\$(dirname\${INO\_FILE})\${user}@\${ip2}:::\${des}

\_FILE}会报no such or directory错误 所以这里同步的源是被删文件或目录的上一级路径, 加上--delete来删除目标上有而源中没有的文件, 则同步的目录月多, 同步删除的操作就越花时间。这里有更好方法的同学, 欢迎交流。

等操作

则不同步, 因为同步目录会发生递归扫描, 等此目录下的文件发生同步时, rsync会顺带更新此目录。

client.pass\${INO\_FILE}\${user}@\${ip1}:::\${des}&&



首页 博客 学院 下载 论坛 问答 活动 专题 招聘 APP VIP会员 搜博主文章



创作



2017年6月

22篇

展开

热门文章

Logstash由于时区导致8小时时差解决方案

阅读数 3677

Python实战mongodb第3篇：Pymongo的分页查询

阅读数 2531

用threading 解决 gunicorn worker timeout

阅读数 1916

python时区设置的两种方式

阅读数 1859

Channels ——django实时推送系统

阅读数 1755

最新评论

Nginx设置Header头信息 ...

weixin\_44050791：前期如果使用了yum install 安装了nginx，后期可以试试 不用编译 安装 该模 ...

国际物流大屏(echarts)+a...

WL19980330：有代码仓库地址吗？

python 版工作流设计

weixin\_39077457：博主可以分享一下源码吗？

n种elasticsearch按照...

Nerd\_F：请问yaml文件是放在elasticsearch的confi g文件下面吗

Sanic框架实战

chf1142152101：整理了一个sanic的中文文档:http ://wiki.mediastack.cn

亿速云服务器免备案CN2高速直连

亿速云香港服务器，CN2高速连接，ping值低可免费换IP,安全稳！技术团队24小时在线稳定无忧

QQ客服

kefu@csdn.net

客服论坛

400-660-0108

工作时间 8:30-22:00

关于我们 招聘 广告服务 网站地图

京ICP备19004658号 经营性网站备案信息

公安备案号 11010502030143

©1999-2020 北京创新乐知网络技术有限公司

网络110报警服务

北京互联网违法和不良信息举报中心

中国互联网举报中心 家长监护 版权申诉

move

create

期间的文件发生更改，他是不知道的，所以这里每2个小时做1次全量同步，防止各种意外遗漏，保证目录一致。

```
sync-client.pass/data/root@192.168.0.18::data&&rsync-avz--password-file=/etc/rsync-ta
```

同步了。

的文件系统监控机制，\*&####&\*\_0\_\*&####&\*内核从2.6.13起，加入Inotify可以监控文件系统中添加、删除、修改移动文件的各种变化情况。

参数说明	
	始终保持事件监听状态
	递归查询目录
	只打印监控事件的信息
	排除文件或目录时，不区分大小写
	超时时间
	指定时间输出格式
	指定时间输出格式
	后面指定删、增、改等事件
事件说明	
	读取文件或目录内容
	修改文件或目录内容
	文件或目录的属性改变
	修改真实文件内容
	文件或目录被打开
	文件或目录移动到
	文件或目录从移动
	移动文件或目录移动到监视目录
	在监视目录下创建文件或目录

举报

delete	删除监视目录下的文件或目录
delete_self	
unmount	卸载文件系统

### 优化 Inotify

# 在/proc/sys/fs/inotify目录下有三个文件，对inotify机制有一定的限制

```
1 [root@web~]# ll /proc/sys/fs/inotify/
2 总用量0
3 -rw-r--r--1root root09月923:36max_queued_events
4 -rw-r--r--1root root09月923:36max_user_instances
5 -rw-r--r--1root root09月923:36max_user_watches
```

-----

max\_user\_watches #设置inotifywait或inotifywatch命令可以监视的文件数量(单进程)

max\_user\_instances #设置每个用户可以运行的inotifywait或inotifywatch命令的进程数

max\_queued\_events #设置inotify实例事件(event)队列可容纳的事件数量

-----

```
1 [root@web~]# echo 50000000>/proc/sys/fs/inotify/max_user_watches -- 把他加入/etc/rc.local就可以实现每次重启都生效
2 [root@web~]# echo 50000000>/proc/sys/fs/inotify/max_queued_events
```

### 作者信息

广州 - 小朱

点赞 收藏 分享 ...

ronon77

发布了748 篇原创文章 · 获赞 35 · 访问量 12万+

私信



### 常用erp软件有哪些

系统erp

评论



想对作者说点什么

- #### rsync实时同步

阅读数 634

假设有如下需求：假设两个服务器：192.168.0.1 源服务器 有目录 /opt/test/192.168.0.2 目标服务器 有目录 /opt/bak/t... 博文 来自： weixin\_33994444...
- #### 基于RSYNC的实时目录同步

阅读数 772

应用场景搭建多台web服务器实现负载均衡，因此多台服务器的网站文档内容必须实时同步。本实例中使用两台web... 博文 来自： Stephen\_Curry11...
- #### Rsync实时同步

阅读数 2445

一、rsync接收端安装 1、查看rsync安装包# rpm -qa rsync2、安装rsync # yum install rsync -y3、添加rsync服务的用... 博文 来自： 张存的博客

阅读数 60

博文 来自: [weixin\\_34341229...](#)



阅读数 1025

博文 来自: 嘻哈王大头

阅读数 21

博文 来自: [weixin\\_30421809...](#)

阅读数 63

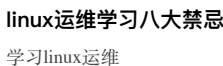
博文 来自: [qq 15140133的博客](#)

阅读数 51

博文 来自: [weixin 33894992...](#)

阅读数 81

博文 来自: [weixin\\_30915951...](#)



阅读数 146

博文 来自: [迷茫的IT男孩](#)

阅读数 6652

博文 来自: [冰恋云的专栏](#)



weixin\_33994444

4649篇文章

[关注](#)

排名:千里之外



Stephen\_Curry11

44篇文章

[关注](#)

排名:千里之外



北上广不相信眼泪

76篇文章

[关注](#)

排名:千里之外



weixin\_34341229

4552篇文章

[关注](#)

排名:千里之外

## 10-11

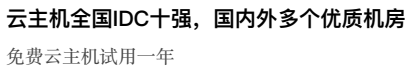
下载

阅读数 123

博文 来自: [lilygg的博客](#)

阅读数 184

博文 来自: [weixin\\_34068198...](#)



阅读数 440

博文 来自: [陈即彪的博客](#)

阅读数 535

博文 来自: [搁浅在PHP沙滩上...](#)

<b>rsync_inotify实现文件双向同步</b>	阅读数 56
2019独角兽企业重金招聘Python工程师标准>>> .....	博文 来自: weixin_33895695...
<b>Rsync+Inotify实时同步</b>	阅读数 21
#实现两服务器文件实时同步rsync: 负责传输同步文件inotify: 负责监测文件变动 (服务端) # yum -y install rsync上...	博文 来自: 口区人的博客
<b>nginx读写分离+rsync+notify</b>	阅读数 133
先配置一个简单的基于nginx的读写分离, 为了方便都用过二进制安装实验环境: nginx: 192.168.122.94web1 (写) ...	博文 来自: kerry2018的博客
 <b>linux运维学习八大禁忌</b> 学习linux运维	广告
<b>Linux下inotify+rsync文件实时同步</b>	阅读数 199
使用背景与传统的cp、tar备份方式相比, rsync具有安全性高、备份迅速、支持增量备份等优点, 通过rsync可以解决...	博文 来自: Linux Story
<b>linux命令之rsync同步速度分析</b>	阅读数 8235
rsync是类unix系统下的数据镜像备份工具——remote sync。一款快速增量备份工具 Remote Sync, 远程同步 支持...	博文 来自: zhibin的程序日记
<b>rsync传输引起的网络延迟</b>	阅读数 5
故障现象: 用户无法访问网站, 具体表现为网页打不开或打开极慢, 体验极差排查过程: 用户调查和模拟用户角色, ...	博文 来自: anjiji1666的博客
<b>读写大文件在linux下占用大量内存且不会释放? 【急】</b>	
代码如下: public void copyFile(String oldPath, String newPath) throws Exception { int bytesum = 0; int b	论坛
<b>linux以rsync对文件目录进行备份, 利用exclude跳过指定类型文件</b>	阅读数 3186
有时会有备份某些文件(例如脚本)的需求, 但又不想把脚本执行的日志或一些结果文件备份, 可以使用rsync中的excl...	博文 来自: Lu_Luke的博客
<b>通过rsync+inotify实现服务器之间数据的实时备份</b>	阅读数 1838
在上一篇博客中, 我们只使用了rsync来实现数据的备份, 但却不能做到实时, 因为守护进程触发和执行有一段的时...	博文 来自: turtle的博客
<b>rsync主动推送</b>	阅读数 183
本文出处参考: http://www.jbxue.com/article/10493.html一.安装步骤(服务器端192.168.0.5)1) 安装rsync 点击下载rsy...	博文 来自: weixin_34297704...
<b>Rsync+inotify 搭建企业实时备份数据服务器</b>	阅读数 20
一. Rsync简介二. Rsync优点与不足三. Inotify简介四. Rsync+Inotify组合实验案例一. Rsync简介Rsync (remote ...	博文 来自: weixin_33766168...
<b>rsync速度问题求解</b>	
测试的时候, 发现rsync速度有点异常, 如下, 求问高手, 什么原因可能导致公网同机房rsync速度这么慢, 或者求排查方法...	论坛
<b>Linux: RSync实现文件备份同步图解及代码</b>	阅读数 675
目录: 1、需求2、图解3、代码备注: 参考RSync命令 http://blog.csdn.net/silentwolfyh/article/details/546139411、需...	博文 来自: 北京小辉
<b>rsync+inotify同步备份MYSQL数据</b>	阅读数 1583
备份源10.16.77.93备份端10.16.77.95原理: 利用inotify监控mysql数据库数据目录: /usr/local/mysql/data一. 备份端...	博文 来自: https://github.com/l...
<b>CentOS下(rsync+inotify)监听式实现服务器间的备份实时同步</b>	阅读数 1027
背景:前几天, 整理了一篇 (rsync+crontab) 实现文件同步的文章, 不过从网上得知, 网上很多言论认为在优势上不如...	博文 来自: u011415782的专栏
<b>rsync+inotify实现实时同步、增量备份</b>	阅读数 672
rsync+inotify实现开机自启实时同步、增量备份。含优化脚本下载!	博文 来自: 朝锡相处
<b>NFS搭建以及rsync+inotify实时备份</b>	阅读数 116
搭建NFS假设有2台Server, 分别是Server1和Server2。现在要在Server1上安装NFS, 在Server2上挂载该目录。1. ...	博文 来自: Clay's Coding Life

## 通过rsync+inotify实现数据的实时备份

阅读数 17

在前面的博文中，我讲到过利用rsync实现数据的镜像和备份，但是要实现数据的实时备份，单独靠rsync还不能... 博文 来自: [weixin\\_34337381...](#)

## Centos rsync+inotify 实现数据同步备份

阅读数 15

最近公司做了一... 博文 来自: [weixin\\_34279579...](#)

## rsync+inotify目录文件实时同步备份

阅读数 12

说明: 192.168.1.2(rsync+inotify)-----FTP192.168.1.3(rsync)-----backup实验目的: 实时自动同步:... 博文 来自: [weixin\\_34240657...](#)

## Rsync+Inotify实时监听备份

阅读数 15

说明，下面的inotify是建立在rsync的配置过程大前提是rsync daemon 配置成功，rsync配置看上一遍博文，在客户端... 博文 来自: [weixin\\_34332905...](#)

## 完整利用Rsync实现服务器/网站数据增量同步备份

阅读数 169

我们在选择VPS、服务器架设项目之后，所有的项目、网站数据都需要我们自行备份和维护，即便有些服务商有提供... 博文 来自: [weixin\\_30287169...](#)

## Rsync+Sersync实时同步

阅读数 34

Rsync+Sersync实时同步一、基本介绍1.什么是Rsync? Rsync (Remote Synchronize) 是一款开源的、快速的、多... 博文 来自: [weixin\\_33938733...](#)

## rsync + Inotify 实现数据的实时同步更新

阅读数 18

1.检查rsync是否安装rpm -q rsync2.定义配置文件/etc/rsyncd.confuid=sunliyangid=sunliyanuserchroot=nomaxconne... 博文 来自: [weixin\\_34249678...](#)

[python](#) [json](#) [java](#) [mysql](#) [pycharm](#) [android](#) [linux](#) [json格式](#)