

Linux系统下实现大数据采集（linux大数据采集）

随着社会的发展，大数据已成为我们日常生活中不可或缺的一部分。几乎每个行业都利用大数据实现信息化管理。在使用这些信息之前，我们必须首先为它采集数据，这就涉及到Linux系统下实现大数据采集的问题了。

Linux系统采集大数据主要分为以下几步：

- 1.数据采集：在Linux系统中，可以使用采集工具来采集指定的数据，常见的方式有shell脚本、crontab定时任务等；
- 2.数据解析：在Linux系统中，可以使用多种解析工具来完成数据解析，包括Python、sed、awk等；
- 3.数据存储：Linux系统中可以使用多种方式来存储采集的数据，包括Mongodb、MySQL等；
- 4.数据分析：在Linux系统中可以使用多种分析工具来对采集的数据进行分析，包括pySpark、hive等。

以上是在Linux系统中完成大数据采集需要完成的步骤，下面我们以shell脚本采集MySQL数据库数据为例，来实现Linux系统下实现大数据采集：

- 1.编写shell脚本：

```
#!/bin/bash
```

```
mysql -uroot -ppassword -e "select * from database.table" -B  
| sed 's/\t/"/g;s/^"/";s/$/"/;s/\n//g' >output.csv
```

2.实现定时任务：

```
echo "0 0 * * * root /path/to/script.sh" > /etc/crontab
```

3.让定时任务生效：

```
service cron reload
```

以上就是Linux系统下实现大数据采集的具体操作，大家在实际操作之前，可以查看相关的文档和文章，实践之后再加以改进，这样才可能实现高效率、节省时间和费用的采集。

[Linux](#)[Linux教程](#)[Linux资讯](#)

■ 我想要获取技术服务或软件

服务范围：MySQL、ORACLE、SQLSERVER、
MongoDB、PostgreSQL、程序问题

服务方式：远程服务、电话支持、现场服务，沟通指定方
式服务

技术标签：数据恢复、安装配置、数据迁移、集群容灾、
异常处理、其它问题

沟通购买：[👤QQ咨询](#) [淘宝咨询](#) [微信咨询](#) [淘宝店
铺](#)

■ 版权申明及联系

本站部分文章参考或来源于网络，如有侵权请联系站长。本站提供
相关远程技术服务，有需要可联系[👤QQ](#)

数据库远程运维 » Linux系统下实现大数据采集 (linux大数据采集)

分享到：



上一篇

Oracle练习题，锻炼数据分析能力的好选择（oracle练习题）

下一篇

型Oracle 浮点型数据类型: 当高精度数据很重要时（oracle浮点）

相关推荐



thumbnail

编程深入浅出UDF Linux编程实战（udflinux）



thumbnail

简单易懂！服务器如何安装Linux桌面系统（服务器安装linux桌面）



thumbnail

Linux下全盘格式化教程（linux格式化整个硬盘）



thumbnail

Linux RAID配置：构建安全、可靠的数据环境（linuxraid配置）



thumbnail

Linux C编程：一步一步开启新的世界（linuxc编程入门）



thumbnail

Linux FIFO管道：实现进程间的数据交换（linuxfifo管道）



thumbnail

Linux 自启动程序：解决Linux开机自动加载问题（linux自启动程序）



thumbnail

root用户如何让Linux非root用户使用if语句（linuxif非）

随机文章



thumbnail

Linux C语言实现读取一行文本（linuxc读一行）



thumbnail

Linux下的GCC：提高编程效率的利器（linux是否gcc）



thumbnail

令人惊叹的Linux反向代理技术（linux反向代理）



thumbnail

展示你的艺术：使用Linux系统图片功能创作（linux系统图片）



thumbnail

权限Linux用户权限分配全面概览 (linux用户有哪些)



thumbnail

法Linux VI输入法：轻松解决文本编辑问题 (linuxvi输入)

标签

Linux Linux教程 Linux资讯 MacOS MacOS教程
MacOS资讯 MongoDB MongoDB教程 MongoDB资讯
MSSQL MSSQL错误 MySQL mysql教程 MySQL维护
MySQL资讯 Neo4j Neo4j教程 Neo4j资讯 ORACLE
Oracle优化 oracle内部视图 oracle参数 oracle开发
oracle异常修复 oracle故障处理 oracle教程 oracle维护
oracle视图 ORACLE资讯 oracle远程维护 ORA错误码
Redis Redis教程 Redis资讯 SQLServer SQLServer报错
SQLServer教程 SQLServer资讯 SQL修复 SQL异常
SQL远程处理 Windows 技术文档 操作系统 数据库

本站部分资源来自互联网收集,仅供用于学习和交流,本站一切资源不代表本站立场,如有侵权、后门、不妥请联系本站站长删除

© 2020 Theme by - 数据服务 . All rights reserved 苏ICP备15021567号