

Loader 运维管理

www.huawei.com





目标

- 学完本课程后，您将能够：
 - 掌握**Loader**服务状态检查方法
 - 掌握**Loader**服务基本性能参数调优
 - 熟悉**Loader**常见故障恢复方法



目录

1. 目录介绍
2. 服务检查
3. 日志介绍
4. 常用性能调优参数
5. 常见问题及定位手段

目录介绍

目录	用途
/opt/huawei/Bigdata	各组件安装目录
/opt/huawei/Bigdata/FusionInsight/ FusionInsight-Sqoop-1.99.3	Loader 组件安装目录
/var/log/Bigdata/loader	Loader 日志存储目录
/opt/huawei/Bigdata/FusionInsight/ FusionInsight-Sqoop- 1.99.3/FusionInsight-Sqoop- 1.99.3/server/webapps/loader/WEB- INF/ext-lib	Loader 关系型数据库 JDBC 驱动存放路径



目录

1. 目录介绍
2. 服务检查
3. 日志介绍
4. 常用性能调优参数
5. 常见问题及定位手段

服务检查

一、FusionInsight Manager的“服务管理”页面，查看Loader状态

- Loader操作状态：已启动
- Loader健康状态：良好
- Loader配置状态：已同步

系统概览 服务管理 主机管理 告警管理 审计管理 租户管理 系统设置

服务 >

集群 (r2c60)


+ 添加服务 下载客户端 更多操作 ▼

服务	操作状态	健康状态	配置状态
DBService	✓ 已启动	✓ 良好	✓ 已同步
HBase	✓ 已启动	✓ 良好	✓ 已同步
HDFS	✓ 已启动	✓ 良好	✓ 已同步
KrbServer	✓ 已启动	✓ 良好	✓ 已同步
LdapServer	✓ 已启动	✓ 良好	✓ 已同步
Loader	✓ 已启动	✓ 良好	✓ 已同步
Mapreduce	✓ 已启动	✓ 良好	✓ 已同步
Yarn	✓ 已启动	✓ 良好	✓ 已同步
ZooKeeper	✓ 已启动	✓ 良好	✓ 已同步

服务检查

二、“告警管理”页面中，无Loader相关告警

系统概览

服务管理

主机管理

告警管理

审计管理

租户管理

系统设置

告警管理 > 告警

告警 事件

 导出全部 清除告警当前页: 0 总页数: 0 |     | 跳转到 [确定](#)

服务检查

三、在“**Loader**服务状态”页面，执行“更多操作 > 启动服务健康检查”，显示

- 服务状态：健康
- 告警信息：无告警信息

服务健康检查

导出报告

	项目	开始时间	进度	结束时间
-	1. Loader	2016-04-06 20:03:53	100%	2016-04-06 20:03:53
服务状态：良好				✓
告警信息：无告警信息				✓

✓ 操作成功。
单击完成继续。

完成



目录

1. 目录介绍
2. 服务检查
3. 日志介绍
4. 常用性能调优参数
5. 常见问题及定位手段

日志介绍——安装日志

Loader安装日志的存储路径为：
/var/log/Bigdata/loader/scriptlog

日志文件名	分析描述
postInstall	Loader 服务安装前的准备工作日志。
preStart.log	Loader 服务启动前的准备工作日志。

日志介绍——运行日志

Loader运行日志的存储路径为：
/var/log/Bigdata/loader/runlog

日志文件名	分析描述
loader.log	Loader 服务端运行日志

日志介绍——审计日志

Loader审计日志的存储路径为：
/var/log/Bigdata/loader/audit

日志文件名	分析描述
default.audit	Loader 安全审计日志。



目录

1. 目录介绍
2. 服务检查
3. 日志介绍
4. 常用性能调优参数
5. 常见问题及定位手段

常用性能调优参数

一、**Loader**作业的配置跟性能息息相关，针对不同的场景，指定不同的配置项值，可以提高性能。

数据源	作业配置项	场景	取值和要求
文件系统 (SFTP/FTP)	“输入设置” - “文件分割方式”	大量小文件导入	SIZE
		单个大文件导入	SIZE
关系型数据库 (RDB)	“输入设置” - “分区列名”	任何	“分区列名”的选择， 需要保证经过该字段 分区后数据分布均匀， 优先使用表的主键， 如果表不存在主键， 优先选择整型字段

常用性能调优参数

二、关联的组件的参数会影响作业的性能。

关联组件	配置项	缺省值	调优场景
Yarn/MapReduce	yarn.app.mapreduce.am.resource.mb	1536	AppllicationMaster 的内存限制。单位： MB
Yarn/MapReduce	mapreduce.reduce.memory.mb	4096	每个 reduce 任务占有用的内存大小，适当调小可以增加 map 数，提高并发。 单位： MB
Yarn/MapReduce	mapreduce.map.memory.mb	4096	每个 map 任务占有用的内存大小，适当调小可以增加 map 数提高并发。 单位： MB
HDFS	dfs.blocksize	134217728	HDFS 块大小，适应调节可以增加 HDFS 读写速度。 单位： MB



目录

1. 目录介绍
2. 服务检查
3. 日志介绍
4. 常用性能调优参数
5. 常见问题及定位手段

常见问题及定位方法-举例1

- 问题描述

从**SFTP**导入至**HDFS**，作业执行失败，提示**authorization negotiation failed**。

- 问题背景

XX局点，新搭建了一套集群，使用**Loader**，从**SFTP**导入至**HDFS**，作业连接器连接正常，提交也成功了，但一直执行失败，提示**authorization negotiation failed**。

- 问题现象

环境一切正常，配置作业从**SFTP**导入至**HDFS**，作业运行过程中，**SFTP**服务器在写入文件的时候，一直提示**authorization negotiation failed**。

常见问题及定位方法-举例1

- 问题分析

从提示信息中，可以看出，是**SFTP**服务认证授权的加密算法验证失败，需要修改它的加密算法。

- 解决办法

修改**SSH**服务器守护进程配置文件**/etc/ssh/sshd_config**，其中**Ciphers**属性用于指定**SSH**允许使用的认证加密算法，多个算法之间使用逗号分隔。

- 具体操作

打开当前作业配置中**SFTP**节点，执行指令**vi /etc/ssh/sshd_config**修改文件，找到**Ciphers**属性，修改成支持的加密算法，值为**Ciphers aes128-cbc,aes192-cbc,aes256-cbc,aes128-ctr,aes192-ctr,aes256-ctr**，保存后执行**service sshd restart**命令，重启**ssh**服务即可。

常见问题及定位方法-举例2

- 故障：Loader服务异常

- 现象描述

Loader服务界面显示状态为“故障”。

- 可能原因

- Loader服务依赖的内部服务异常。

- ZooKeeper服务异常。

- HDFS服务异常。

- DBService服务异常。

- Yarn服务异常。

- Mapreduce服务异常。

- 环境故障：网络异常，Loader服务无法与其依赖的内部服务通信，无法提供服务。

- 软件故障：Loader服务无法正常运行。

常见问题及定位方法-举例2

- 修复思路

故障的原因多种多样，没有固定的模式来解决所有故障问题，需要根据具体情况具体分析，思路如下：

- 1、检查依赖的组件状态是否良好。
- 2、检查**Loader**和依赖组件之间的网络连接。
- 3、检查**Loader**浮动IP地址是否冲突。
- 4、收集故障信息和日志，查看**Loader**自身软件问题。

Thank you

www.huawei.com