

Hadoop vs HDFS 命令



从你的描述和输出信息来看，这段 `INFO` 级别日志是 Hadoop 在读取 HDFS 文件时生成的运行日志，它并不是文件内容的一部分，因此也不会出现在本地文件中。

为什么会看到这条日志？

这条日志信息来源于 Hadoop 的日志系统，而不是你本地文件中的内容：

1. 日志级别：

- Hadoop 默认会输出一些 `INFO` 级别的日志，包括客户端和服务端之间的数据传输、加密信任等状态。
- 在 `cat` 命令操作时，Hadoop 会记录它访问 HDFS 的行为，因此显示了这条日志。

2. 内容解读：

plaintext

```
INFO sasl.SaslDataTransferClient: SASL encryption trust check: localhostTrusted = false, remoteHostTrusted = false
```

- `INFO`：表明这是普通信息，而非警告或错误。
- `SASL encryption trust check`：指 Hadoop 检查了 SASL（简单身份验证和安全层）加密是否被信任。
- `localhostTrusted = false, remoteHostTrusted = false`：
 - 本地主机和远程主机的信任未建立，说明你未启用 Kerberos 或其他加密机制。
 - 这种默认配置在开发或测试环境下是常见的，对正常操作没有影响。

3. 文件内容无修改：

- 这段日志是在 HDFS 操作时生成的，并未改变文件的实际内容。
- 如果你用 `vi` 或 `cat` 检查本地文件内容，这些日志信息是不会出现在本地文件中的。

如何确认文件内容未变？

运行以下命令直接查看本地文件内容：

bash

```
cat /home/hadoop/myLocalFile.txt
```

如果显示内容为：

```
plaintext

Hadoop
Spark
XMU DELAB
```

则文件内容未受到任何影响。

如何隐藏这些日志？

方法 1：更改 Hadoop 的日志级别

Hadoop 默认输出 `INFO` 级别日志，可以将日志级别降低到 `ERROR`，以隐藏这些信息。

1. 打开 Hadoop 的日志配置文件：

```
bash

vi /usr/local/hadoop/etc/hadoop/log4j.properties
```

2. 找到以下内容：

```
properties

log4j.rootLogger=INFO, console
```

3. 修改为：

```
properties

log4j.rootLogger=ERROR, console
```

4. 保存并退出，然后重新启动 Hadoop 服务。

方法 2：重定向日志输出

在执行命令时，将日志输出重定向到 `/dev/null`，从而只显示文件内容：

```
bash

./bin/hdfs dfs -cat /user/hadoop/input/myLocalFile.txt 2>/dev/null
```

此命令会隐藏所有日志，只显示文件的实际内容。

总结

- 你看到的 `INFO` 日志来自 Hadoop 的运行日志，**不是文件内容的一部分**。
- 文件内容未被修改，你可以通过本地 `cat` 或编辑工具检查确认。
- 如果想隐藏这些日志，可以调整 Hadoop 的日志级别或通过命令重定向日志输出。