

Hadoop vs HDFS 命令



./bin/hadoop和./bin/hdfs这两条命令有什么区别？



在 Hadoop 系统中，`./bin/hadoop` 和 `./bin/hdfs` 是两种不同的命令入口，它们的功能和作用各不相同：

1. `./bin/hadoop`

- **功能：**
这是 Hadoop 的通用命令入口，用于执行 Hadoop 提供的所有工具和服务的通用管理任务。
- **常见用途：**
 - 提供对整个 Hadoop 生态系统组件的访问，比如 `yarn`、`dfsadmin` 等。
 - 提供通用的命令帮助信息，通过 `hadoop help` 查看支持的子命令列表。
 - 运行 MapReduce 作业，比如通过 `hadoop jar` 提交作业。
- **示例：**

```
bash
```

```
./bin/hadoop jar my-job.jar com.example.MyJob input output  
./bin/hadoop fs -ls /user/hadoop
```

- **适用场景：**
当你需要管理整个 Hadoop 生态系统，或者需要与 HDFS 交互时，使用这个命令。

2. `./bin/hdfs`

- **功能：**
专用于管理和操作 **HDFS（Hadoop 分布式文件系统）** 的命令入口。
- **常见用途：**
 - 操作 HDFS 中的文件和目录，例如创建、删除、查看文件。
 - 管理 HDFS 服务，例如启动、关闭 NameNode 或 DataNode。
- **示例：**

```
bash
```

```
./bin/hdfs dfs -ls /  
./bin/hdfs dfs -put localfile /user/hadoop  
./bin/hdfs namenode -format
```

- **适用场景：**
如果操作仅涉及 HDFS 文件系统或服务管理，使用这个命令更直接。

核心区别：

特性	<code>./bin/hadoop`</code>	<code>./bin/hdfs`</code>
作用范围	整个 Hadoop 生态系统	仅限于 HDFS 文件系统
支持的功能	包括 HDFS、YARN、MapReduce 等	仅用于 HDFS 的文件和服务操作
常用子命令	<code>fs`</code> 、 <code>jar`</code> 、 <code>yarn`</code> 、 <code>distcp`</code> 等	<code>dfs`</code> 、 <code>namenode`</code> 、 <code>datanode`</code> 等
适用场景	综合管理和任务提交	专注于 HDFS 文件系统管理

总结：

- 如果你不确定具体需要哪个命令，一般可以从 `./bin/hadoop`` 开始。
- 如果明确要管理或操作 HDFS 文件系统，直接使用 `./bin/hdfs`` 更加合适。