# Hadop Offline image Viewer&&Offline Edits Viewer

**1.Offline Image Viwer**

Offline Image Viewer是将fsimage文件转换为可以阅读的工具，并提供Web HDFS API以进行离线分析及Hadoop集群的检查。该工具是对image文件进行快速的离线分析，输出多种格式有多种：

* Web，默认的输出进程，启动HTTP Server，提供WebHDFS API，用户可以通过REST API查询NameSpace
* XML, 生成Fsimage的XML格式文件，包含fsimage文件中的所有信息，可以使用XML工具来分析
* FileDsitribution，用于通过image来分析文件的大小分布，输出文件的格式用tab分割
* Delimited，生成text文件，每行内容包含inodes、inodes-under-construction等信息，默认用\t隔开。

使用Offline Image Viewer收集命名空间中文件数据，使用这些信息可以匹配文件系统或者使用任意的特定规则进行分析。分割处理器可以输出进一步处理的格式如Pig。

1. 执行命令的格式如下：

*Usage:*

*bin/hdfs oiv [OPTIONS] -i INPUTFILE -o OUTPUTFILE*

其参数描述如下：

|  |  |
| --- | --- |
| 参数 | 描述 |
| -i | -inputFile | 定义要处理的fsimage文件 |
| -o| -outputFile | 定义输出文件 |
| -p| -processor | 定义使用image文件的处理器，默认Web |
| -addr | 定义Web的host和port，默认localhost:5978 |
| -maxSize | FileDistribution Procesor可以处理的文件大小范围 |
| -step | 文件分布钟定义文件大小的粒度 |
| -delimiter | Delimited处理器的间隔Str，默认/t |
| -t | -temp | 临时目录 |
| -h| -help | 帮助信息，使用说明 |

最简单方式是通过-i和-o设定输入及输出文件，使用示例如下：

*hdfs oiv -p XML -i fsimage\_0000000000000028703 -o fsimage.xml*

在当前目录中，生成fsimage.xml文件，内如如下：

*<?xml version="1.0"?>*

*<fsimage>*

*<NameSection>*

*<genstampV1>1000</genstampV1>*

*<genstampV2>1002</genstampV2>*

*<genstampV1Limit>0</genstampV1Limit>*

*<lastAllocatedBlockId>1073741826</lastAllocatedBlockId>*

*<txid>37</txid>*

*</NameSection>*

*<INodeSection>*

*<lastInodeId>16400</lastInodeId>*

*<inode>*

*<id>16385</id>*

*<type>DIRECTORY</type>*

*<name></name>*

*<mtime>1392772497282</mtime>*

*<permission>theuser:supergroup:rwxr-xr-x</permission>*

*<nsquota>9223372036854775807</nsquota>*

*<dsquota>-1</dsquota>*

*</inode>*

*...remaining output omitted...*

XML文件中包含了fsimage中所有信息，比如inodeid、type、name、修改时间、权限等

1. **Offline Edit Viewer**

Offline Edits Viewer是解析Edit日志文件的工具，用于实现Edits文件实现不同格式之间的转，目前包括二进制文件和XML，XML比二进制格式更易于编辑。其提供三种输出处理器：

* binary，Hadoop内部使用的原生二进制格式
* xml，XML格式
* stats，输出统计信息

执行命令：

*bin/hdfs oev [OPTIONS] -i INPUT\_FILE -o OUTPUT\_FILE*

参数描述如下：

|  |  |
| --- | --- |
| 参数 | 描述 |
| -h | 使用帮助信息 |
| -i | -inputFile | 定义要处理的fsimage文件 |
| -o| -outputFile | 定义输出文件 |
| -p| -processor | 定义使用image文件的处理器，默认Web |
| -v | 版本信息 |

使用示例如下：

*hdfs oev -i edits\_0000000000000052566-0000000000000052567 -o edits.xml*

edits.xml格式如下：

*<?xml version="1.0" encoding="UTF-8"?>*

*<EDITS>*

*<EDITS\_VERSION>-63</EDITS\_VERSION>*

*<RECORD>*

*<OPCODE>OP\_START\_LOG\_SEGMENT</OPCODE>*

*<DATA>*

*<TXID>52566</TXID>*

*</DATA>*

*</RECORD>*

*<RECORD>*

*<OPCODE>OP\_END\_LOG\_SEGMENT</OPCODE>*

*<DATA>*

*<TXID>52567</TXID>*

*</DATA>*

*</RECORD>*

*</EDITS>*

http://hadoop.apache.org/docs/r2.7.2/hadoop-project-dist/hadoop-hdfs/HdfsEditsViewer.html

http://hadoop.apache.org/docs/r2.7.2/hadoop-project-dist/hadoop-hdfs/HdfsImageViewer.html