### segments.gen,segments\_N

文件名如: segment\_N为段文件，保存的是存储提交点的信息。相关的类名是IndexFileNames来实现。SegmentInfos的write(Directory directory)。

1. 当首次提交段时，生成pending\_segments\_1文件。
2. 首先写入**IndexHeader:**

**Header**: CODEC\_MAGIC, codec(编码格式识别文件)，version(format version).lucene5.0后为***VERSION\_50***

**Id**: 通过*StringHelper.randomId()*生成的*ID\_LENGTH*长度的id，唯一识别这个特别文件实例。

**Suffix：**文件名后缀，通过*Long.toString(nextGeneration, Character.MAX\_RADIX)。*首先写后缀占用的bit长度，再写内容。

1. **version：**Counts how often the index has been changed，当初始化IndexWriter的构造方法里算一次。

**Counter：**Used to name new segments

**size()：**segments的个数，

**SegmentCommitInfo(遍历输入每个):**

**SegmentInfo.name** ：当为首个是为\_0,

**segmentID：**segmentInfo的id如果为null则写一位0，否则写一位1.再写入segmentId的值。

**SegmentInfo.getCodec().getName()：**这个段对应的编码

**SegmentCommitInfo.getDelGen():**Generation number of the live docs file (-1 if there are no deletes yet)

**SegmentCommitInfo.getDelCount():** How many deleted docs in the segment:

**SegmentCommitInfo.getFieldInfosGen():** Generation number of the FieldInfos (-1 if there are no updates)

**SegmentCommitInfo.getDocValuesGen():**Generation number of the DocValues (-1 if there are no updates)

**SegmentCommitInfo.getFieldInfosFiles():**先写size，再遍历写入这个set，track the fieldInfos update files

**SegmentCommitInfo.getDocValuesUpdatesFiles()：**先写map的size,再依次写key-value。Track the per-field DocValues update files。

**SegmentInfos.userData:**用户可以在自己设置。

**writeFooter：**FOOTER\_MAGIC, 0(AlgorithmID,表示checksum使用的算法,0 for zlib-crc32), checksum

### .si

文件名为.si为段信息文件，存储段的元数据,由Lucene50SegmentInfoFormat的write方法写入。

1. 首先写入**IndexHeader:**

**Header**: CODEC\_MAGIC, codec(编码格式识别文件，这个为Lucene50SegmentInfo)，version

**Id**: 通过*StringHelper.randomId()*生成的*ID\_LENGTH*长度的id，唯一识别这个特别文件实例。

**Suffix：**文件名后缀，通过*Long.toString(nextGeneration, Character.MAX\_RADIX)。*首先写后缀占用的bit长度，再写内容

1. **Version：**Write the Lucene version that created this segment, since 3.1

**version.major：**Major version, the difference between stable and trunk

**version.minor：**minor version, incremented within the stable branch

**version.bugfix：**Bugfix number, incremented on release branches

**SegmentInfo.getDocCount():**number of docs in seg

**SegmentInfo.getUseCompoundFile():**这个段是否是合并。

**SegmentInfo.diagnostics：**创建私有的，作为debug的辅助。对于每个创建的段，它包含一些元数据，比如:Lucene版本，操作系统，Java的版本。segment创建的原因(merge, flush,addIndexes)

**SegmentInfo.setFiles：**这个段对应的所有文件。

**writeFooter：**FOOTER\_MAGIC, 0(AlgorithmID,表示checksum使用的算法,0 for zlib-crc32), checksum

### .fdm

文件名为.fdm。域文件，存储field信息,通过Lucene50FieldInfosFormat:

1. 首先写入**IndexHeader:**

**Header**: CODEC\_MAGIC, codec(编码格式识别文件，这个为*Lucene50FieldInfos)*，version

**Id**: 通过*segmentInfo.getId()*;*ID\_LENGTH*长度的id，唯一识别这个特别文件实例。

**Suffix：**文件名后缀，通过*Long.toString(nextGeneration, Character.MAX\_RADIX)。*首先写后缀占用的bit长度，再写内容

1. **infos.size():** the number of fields

**FieldInfos:** 遍历写入所有的fieldInfo

**FieldInfo.name:** Field's name

**FieldInfo.number:** Internal field number,内部域的次序。

**一个byte:** 由fi的storeTermVector，omitNorms，storePayloads从低位到高位组成。

**IndexOptions：**DOCS\_ONLY为1，DOCS\_AND\_FREQS为2，DOCS\_AND\_FREQS\_AND\_POSITIONS为3，DOCS\_AND\_FREQS\_AND\_POSITIONS\_AND\_OFFSETS为4

**DocValuesType：**the docValues. this may be null if the field has no docvalues，pack the DV type and hasNorms in one byte。

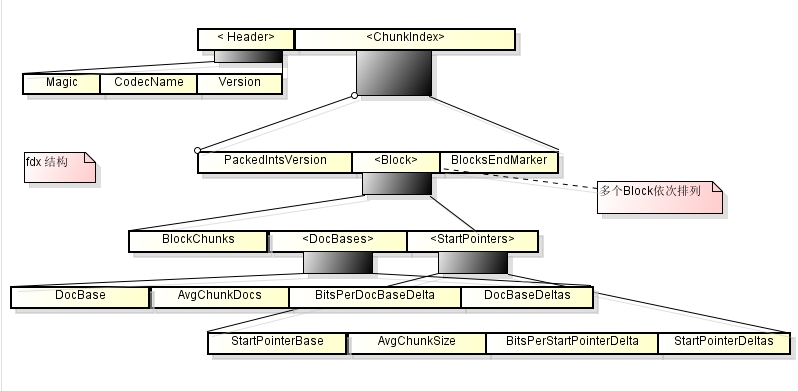
**DocValuesGen：**the docValues generation of this field, or -1 if no docValues updates exist for it

**attributes ：**internal codec attributes map. May be null if no mappings exist

**writeFooter：**FOOTER\_MAGIC, 0(AlgorithmID,表示checksum使用的算法,0 for zlib-crc32), checksum

### .fdx

文件名为.fdx域指针文件，包含域数据数据的指针



1,BlockChunks表示当前Block中Chunk的个数；

2,<DocBases>表示当前Block中每个Chunk的doc个数，可以看作一个数组;

3,<StartPointers>表示当前Block中每个Chunk在fdt文件中的起始位置，其结构与<DocBases>相同。

### .fdt

