

zip文件格式

zip文件由三部分组成：压缩的文件内容源数据、压缩的目录源数据、目录结束标识的结构

①压缩文件内容的源数据：

记录着压缩的所有文件的内容信息，其数据组织结构是对于每个文件都由File header 、File data 、Data descriptor 三部分组成。

File header（文件头）：用于标识该文件的开始，结构说明如下：

偏移	字节	描述
0	4	文件头标识，值固定(0x04034b50)
4	2	解压文件所需 pkware最低版本
6	2	通用比特标志位(置比特0位=加密，详情见后)
8	2	压缩方式（详情见后）
10	2	文件最后修改时间
12	2	文件最后修改日期
14	2	CRC-32校验码
18	4	压缩后的大小
22	4	未压缩的大小
26	4	文件名长度
28	2	扩展区长度
30	n	文件名
30+n	m	扩展区

www.axiaowei.cn

File data(文件数据)：相应压缩文件的源数据。

Data descriptor（数据描述符）：

用于标识该文件压缩结束，该结构只有在相应的local file header中通用标记字段的第3 bit设为1时才会出现，紧接在压缩文件源数据后。这个数据描述符只用在不能对输出的 ZIP 文件进行检索时使用。例如：在一个不能检索的驱动器（如：磁带机上）上的 ZIP 文件中。如果是磁盘上的ZIP文件一般没有这个数据描述符。

偏移	字节	描述
0	4	CRC-32校验码
4	4	压缩后的大小
8	4	未压缩的大小

www.axiaowei.cn

②Central directory 核心目录

记录了压缩文件的目录信息，在这个数据区中每一条记录对应再压缩源文件数据区中的一条数据。

核心目录结构：

偏移	字节	描述
0	4	核心目录文件header标识= (0x02014b50)
4	2	压缩所用的pkware版本
6	2	解压所需pkware的最低版本
8	2	通用位标记
10	2	压缩方法
12	2	文件最后修改时间
14	2	文件最后修改日期
16	4	CRC-32校验码
20	4	压缩后的大小
24	4	未压缩的大小
28	2	文件名长度
30	2	扩展域长度
32	2	文件注释长度
34	2	文件开始位置的磁盘编号
36	2	内部文件属性
38	4	外部文件属性
42	4	本地文件头的相对位移
46	n	目录文件名
46+n	m	扩展域
46+n+m	k	文件注释内容

www.axiaowei.cn

③End of central directory record(EOCD) 目录结束标识

目录结束标识存于整个归档包的结尾，用于标记压缩目录数据的结束。每个压缩文件必须有且只有一个EOCD记录。

偏移	字节	描述
0	4	核心目录结束标记 (0x06054b50)
4	2	当前磁盘编号
6	2	核心目录开始位置的磁盘编号
8	2	该磁盘上所记录的核心目录数量
10	2	核心目录结构总数
12	2	核心目录的大小
16	4	核心目录开始位置相对于archive开始的位移
20	2	注释长度
22	n	注释内容

www.axiaowei.cn

官方文档: <https://pkware.cachefly.net/webdocs/APPNOTE/APPNOTE-6.2.0.txt>