**HDF5(Hierarchical Data Format)**

层次数据格式是一种针对超过内存的大量数据进行组织和存储的文件格式，一般以.h5或者.hdf5作为后缀

HDF5文件结构中有两个基本对象： Dataset和Group

Dataset类似于Numpy数组

Group是像文件夹一样的容器，类似于Python中的字典，也有自己的键和值，”键”就是组成员的名称，”值”就是组成员对象本身(组或者数据集)

**h5py模块**

用于处理HDF5文件的模块

import h5py

x = np.arange(100)

with h5py.File('test.h5','w') as f:

dataset = f.create\_dataset('test\_numpy',data=x)*#创建dataset*

group = f.create\_group('subgroup')*#创建group*

subdataset = group.create\_dataset('test\_numpy',data=x)#*在group中创建dataset*

subgroup = group.create\_group('subsub')#*在group中创建group*

subsubdataset = subgroup.create\_dataset('test\_numpy',data=x) #*在group的group中创建dataset*