



Not *HDFS*
(Sorry)



HDF5 and h5py

Thomas Kluyver
PyData Edinburgh, June 2019

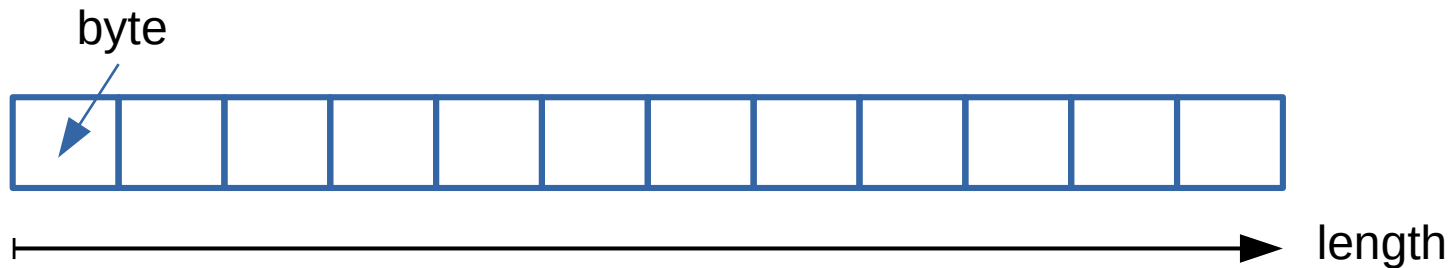
Filesystem in a file

```
[takluyver@localhost scratch]$ h5glance demo.h5
```

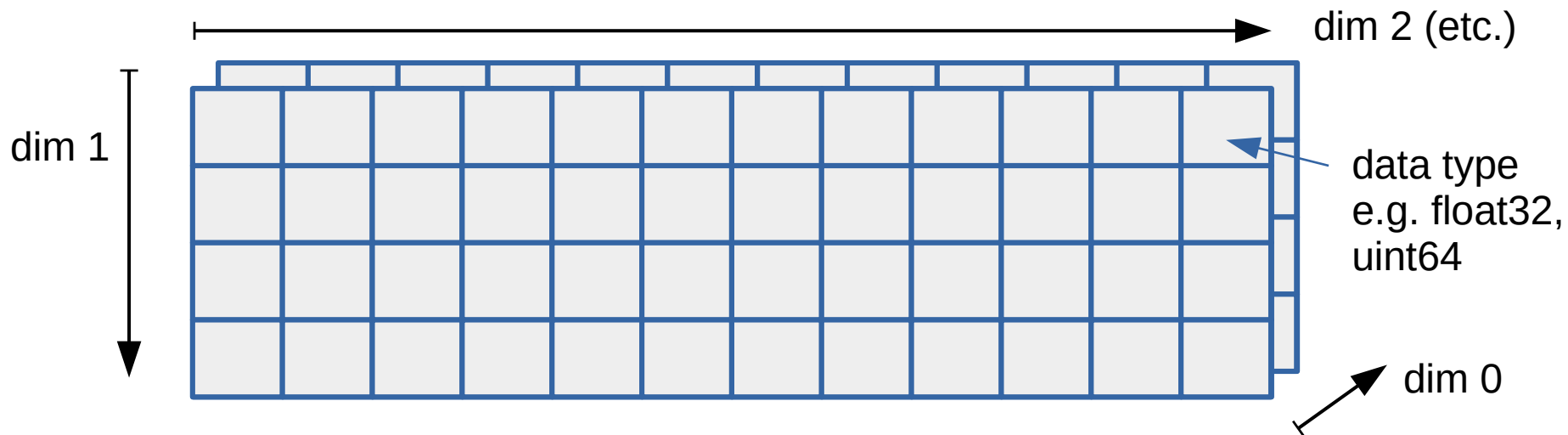
demo.h5

```
|group
|  |subgroup
|    |data1      [uint8: 20]
|  link_external -> other_file.h5/path/to/object
|  link_internal -> group/subgroup
```

File:



HDF5 Dataset:



Drawbacks

- Complex file format
 - Mostly dependent on libhdf5
 - Be wary of untrusted files
- Gaps take up space
- Alternative: *exdir* (real filesystem)

h5py

```
f = h5py.File('demo2.h5', 'w')  
f['grp/data'] = numpy.zeros((10, 20))  
f['grp/data'][:] # Read
```