

tar2h5: Small Files Packer for Machine Learning Tasks

Gerd Heber (HDFGroup), Dawei Mu (NCSA)

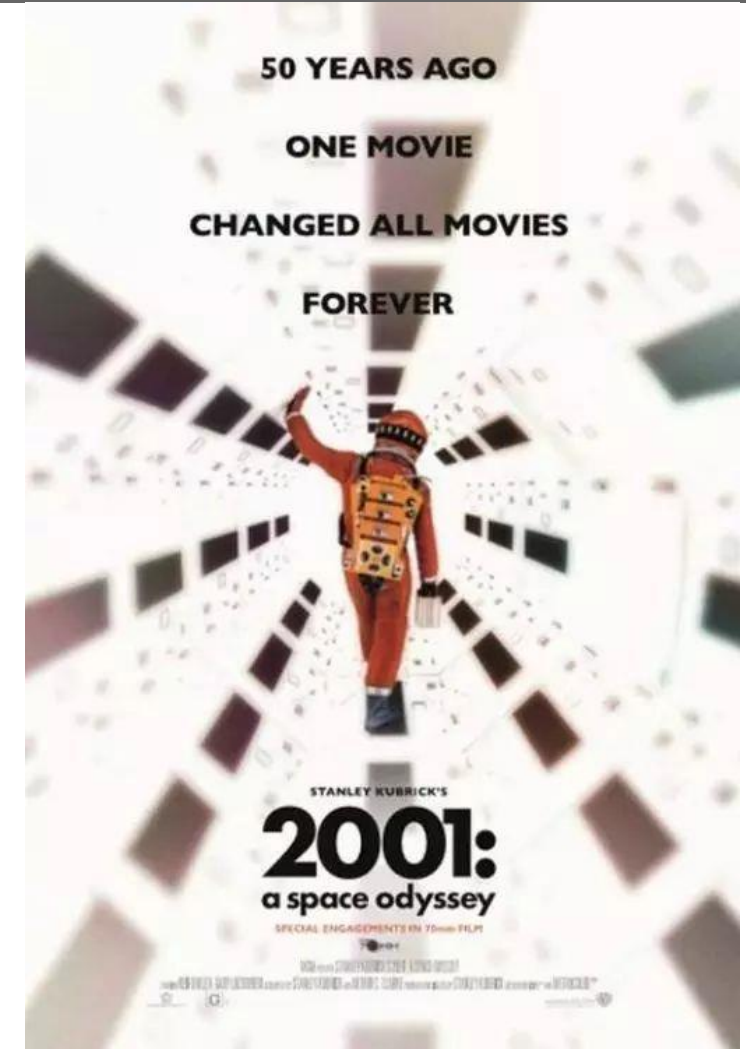
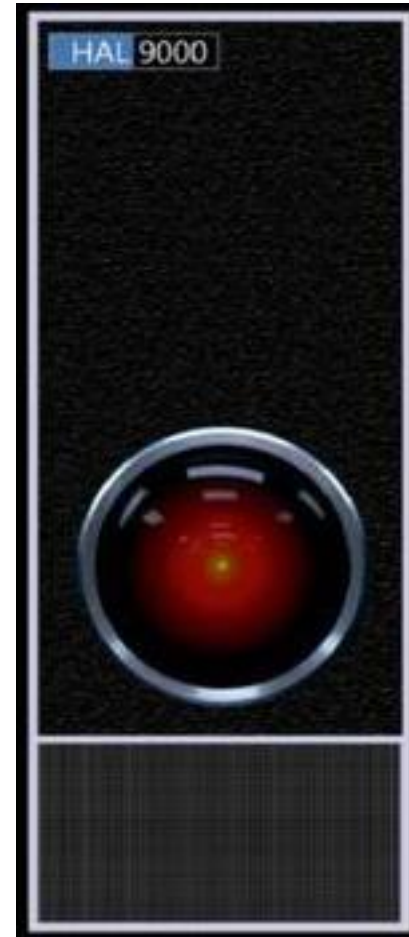


ILLINOIS

NCSA | National Center for
Supercomputing Applications

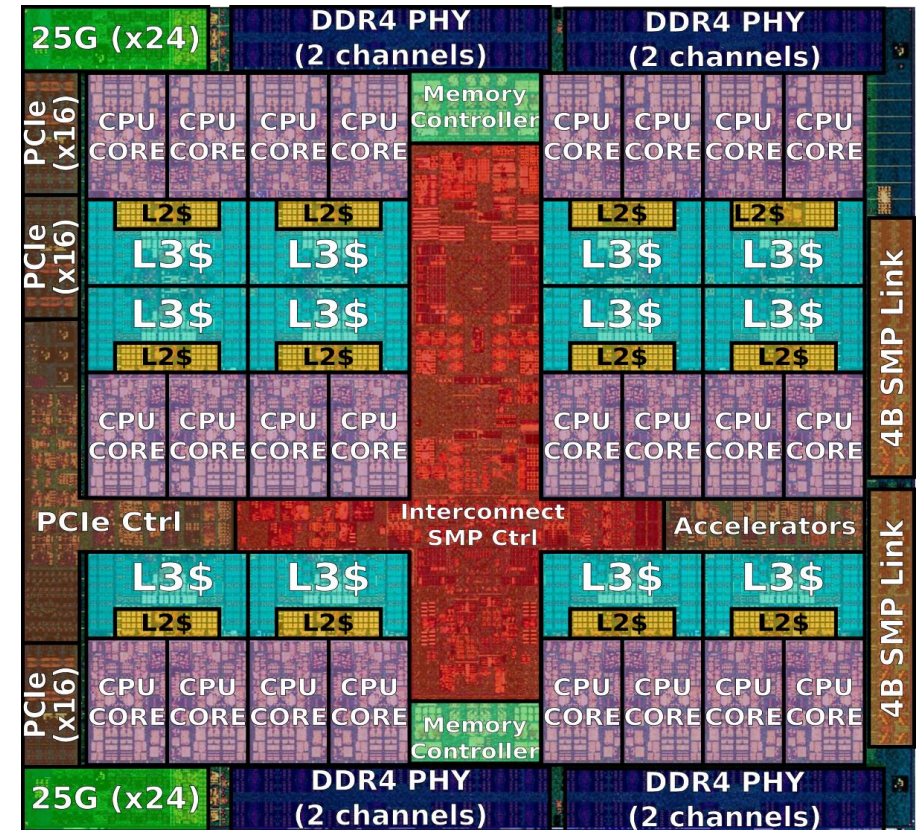
HAL System Overview

- **NSF-funded IBM cluster for Deep Learning applications**
 - 16x compute nodes,
 - 640x physical CPU cores,
 - 64x Nvidia V100 GPUs
 - 224 TB of All-Flash Storage
- **The Origin of Machine Name**
 - 2001: a space odyssey
 - Early concept of an artificial intelligence system
 - Didn't end well and we decided to give "him" a second chance



HAL System Overview

- **IBM POWER9 CPUs**
 - 14nm finFET semiconductor
 - Stronger Thread Performance – **SMT**
 - POWER ISA 3.0 Architecture
 - Enhanced Cache Hierarchy
 - NVIDIA **NVLink 2.0**
 - I/O System – **PCIe Gen4**
- **2x 20 Cores with SMT4**
 - Map to OS as 160 CPUs per node



HAL System Overview

- **NVIDIA V100 GPUs**

- Peak **7.8 TFLOP/s** (double-precision).
- Peak **15 TFLOP/s** (single-precision).
- SM / Cores : **80 / 5120**.
- HBM2 Memory 16 GB : **900 GB/s**.
- Config up to **128 KB** L1 Cache per SM.
- Config up to **96 KB** Shared Memory per SM.
- Constant memory 64 KB.
- 65536 32-bit Registers per SM.
- Clock Frequency : 1530 MHz



HAL System Overview

- **DDN GS400NVE Flash Arrays Server**
 - 224 TB usable GPFS
 - Balance between bandwidth and IOPs



HAL Software Overview

- **HAL Software**

- OS : CentOS Linux 7.7
- Compilers :
 - GNU 4.8.5
 - Advance Toolchain 12.0
 - IBM XL 16.1.1
 - CUDA 10.2.89
 - PGI 2019.10
- Tools :
 - PowerAI 1.7.0 (Watson Machine Learning Community Edition)
 - OpenMPI 4.0.3
 - CMake 3.14.0
 - Singularity 3.5.3

The IO Challenge

- **The catalyst of the rise of machine learning - Datasets**
 - dataset composed by millions of small files
 - dominated random access pattern
- **Researchers new to this region**
 - expertise in domain science
 - who has not enough HPC experience
 - using small files as dataset
 - overwhelming workload to shared storage
 - we used to reduce compute nodes to ease the I/O pressure

Tar2h5: Small Files Packer

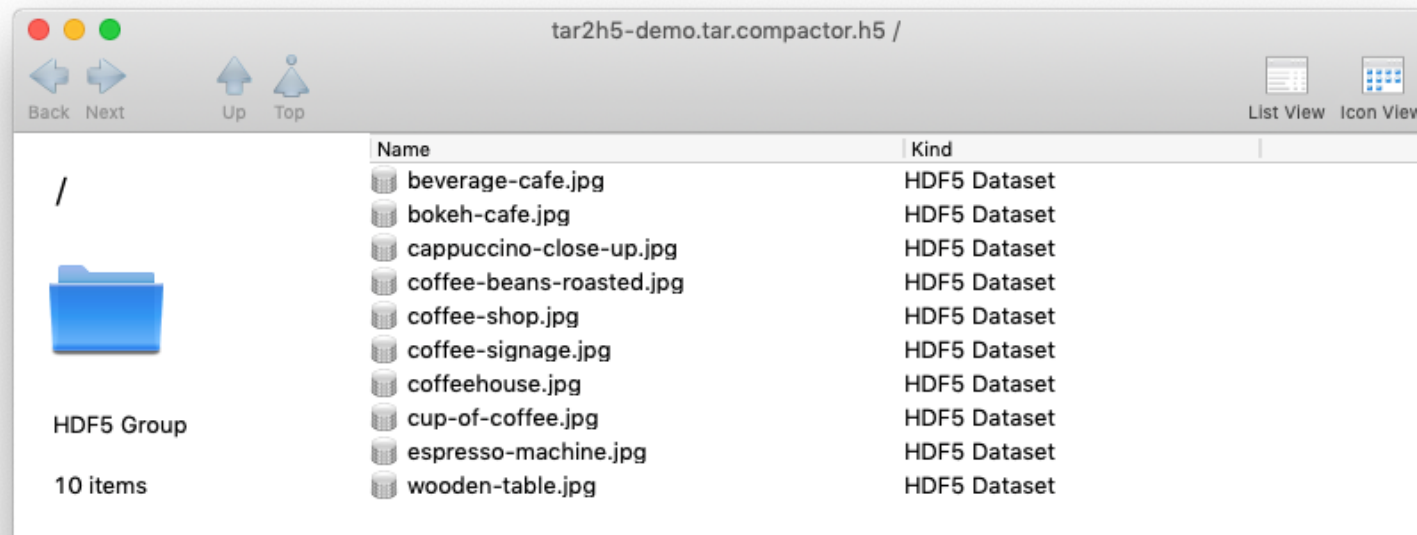
- **Convert Tape ARchives to HDF5 files**
 - easy to use
 - solutions for different scenarios
- **Functions**
 - archive checker
 - h5compactor
 - h5shredder

archive checker

- archive_checker
 - check how many files can be extracted from the input tar file.
- archive_checker_64k
 - check if any files within input tar files larger than 64 KB.

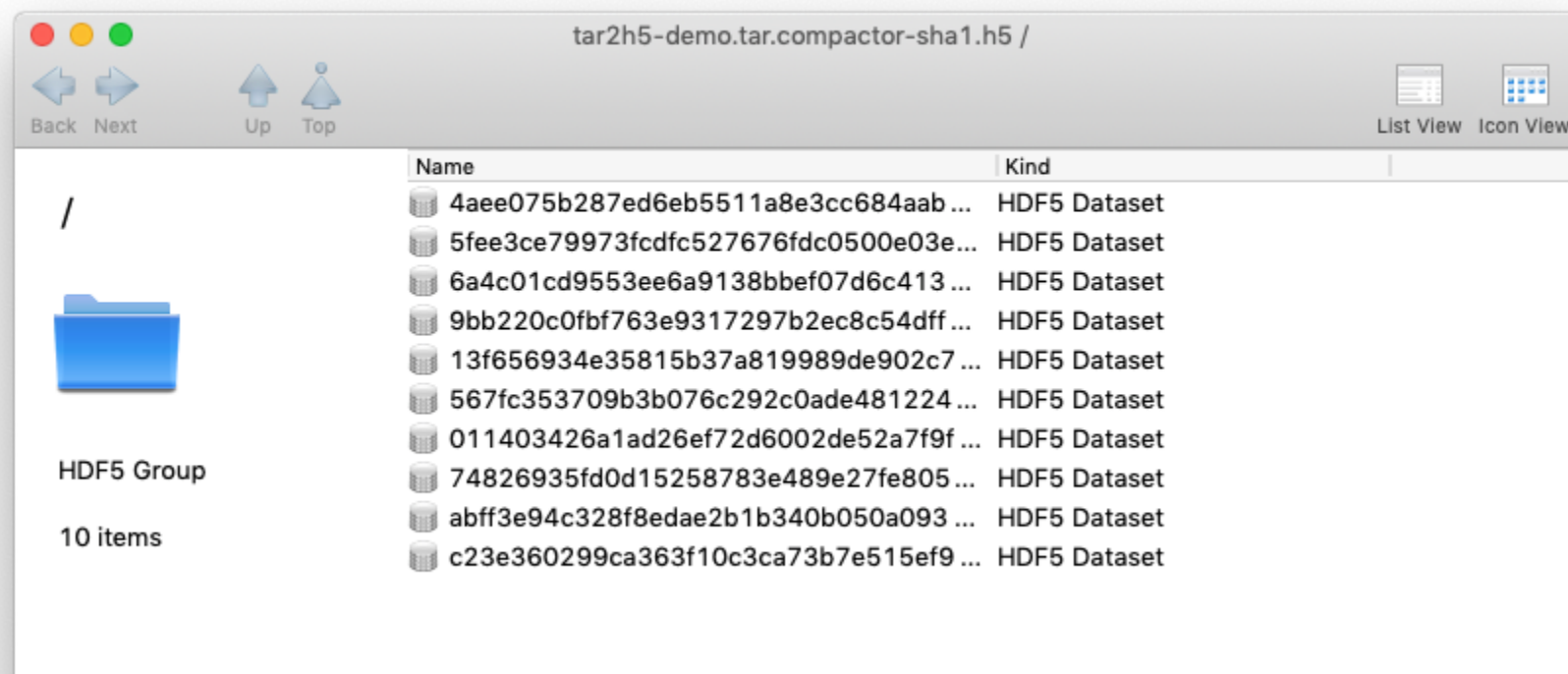
h5compactor

- h5compactor
 - convert input tar file into hdf5 file, all files within tar file should be smaller than 64KB, using small file names as dataset names.



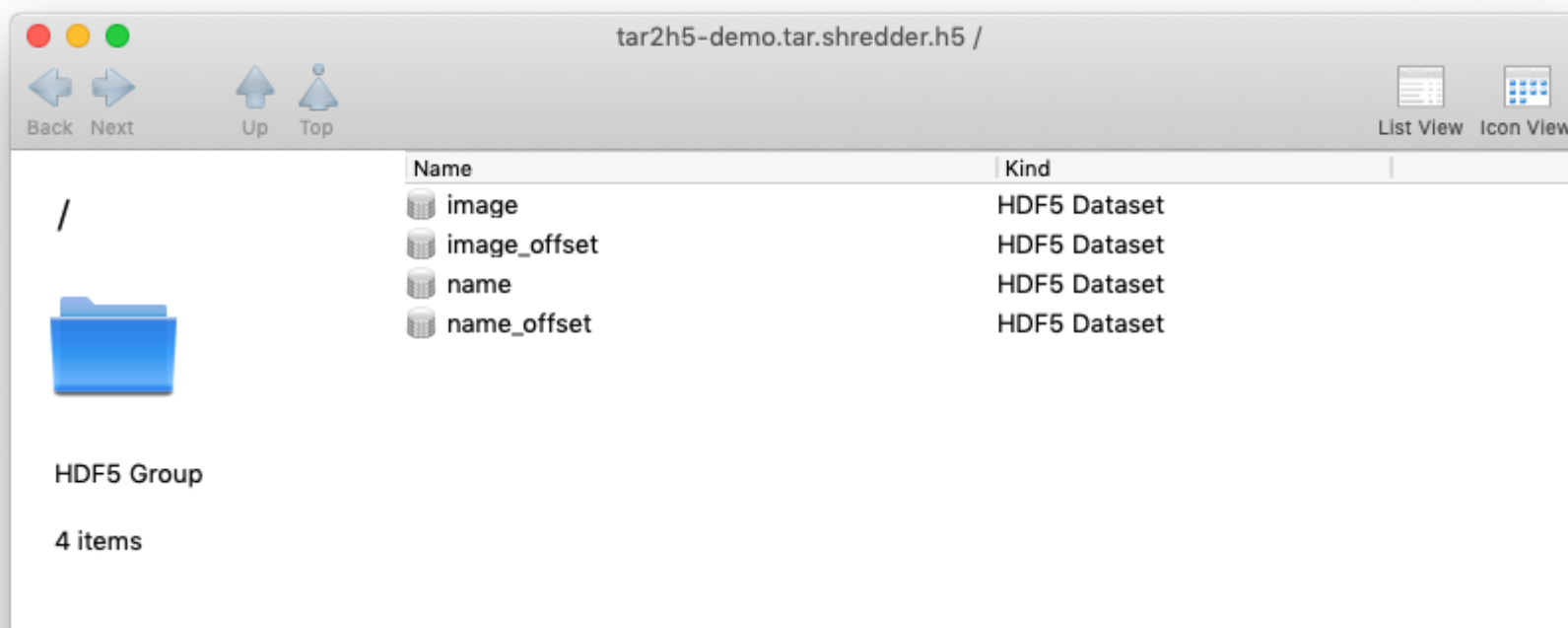
- h5compactor-sha1

- convert input tar file into hdf5 file, all files within tar file should be smaller than 64KB, using small files sha1 values as dataset names.



- h5shredder

- convert input tar file into hdf5 file, no file size limitation, concatenate data and offsets into 4 separate arrays for better randomized access.



User Case

- **Project Title**
 - Efficient Large-Scale Video Generation with GANs
- **HAL User**
 - CS Research Assistant: Daniel B. McKee
- **Data Info**
 - large Kinetics-400 video dataset composed of around 240k videos
 - the number of files to around 25million
 - total size of jpg dataset around 125GB as a tar file
- **Performance Comparison**
 - Loading from the compact hdf5 file made training about 5x faster

tar2h5 Open Source

- The tar2h5 is now available at GitHub
 - <https://github.com/HDFGroup/tar2h5>

The screenshot displays the GitHub repository page for `HDFGroup/tar2h5`. The repository is currently on the `master` branch. The file browser shows a list of files and folders, including `demo`, `img`, `src`, `.gitignore`, `CMakeLists.txt`, `README.md`, `float-my-boat.org`, `tar-to-hdf5.html`, `tar-to-hdf5.org`, `tar2h5.html`, and `tar2h5.org`. The `README.md` file is selected, showing the following content:

tar2h5

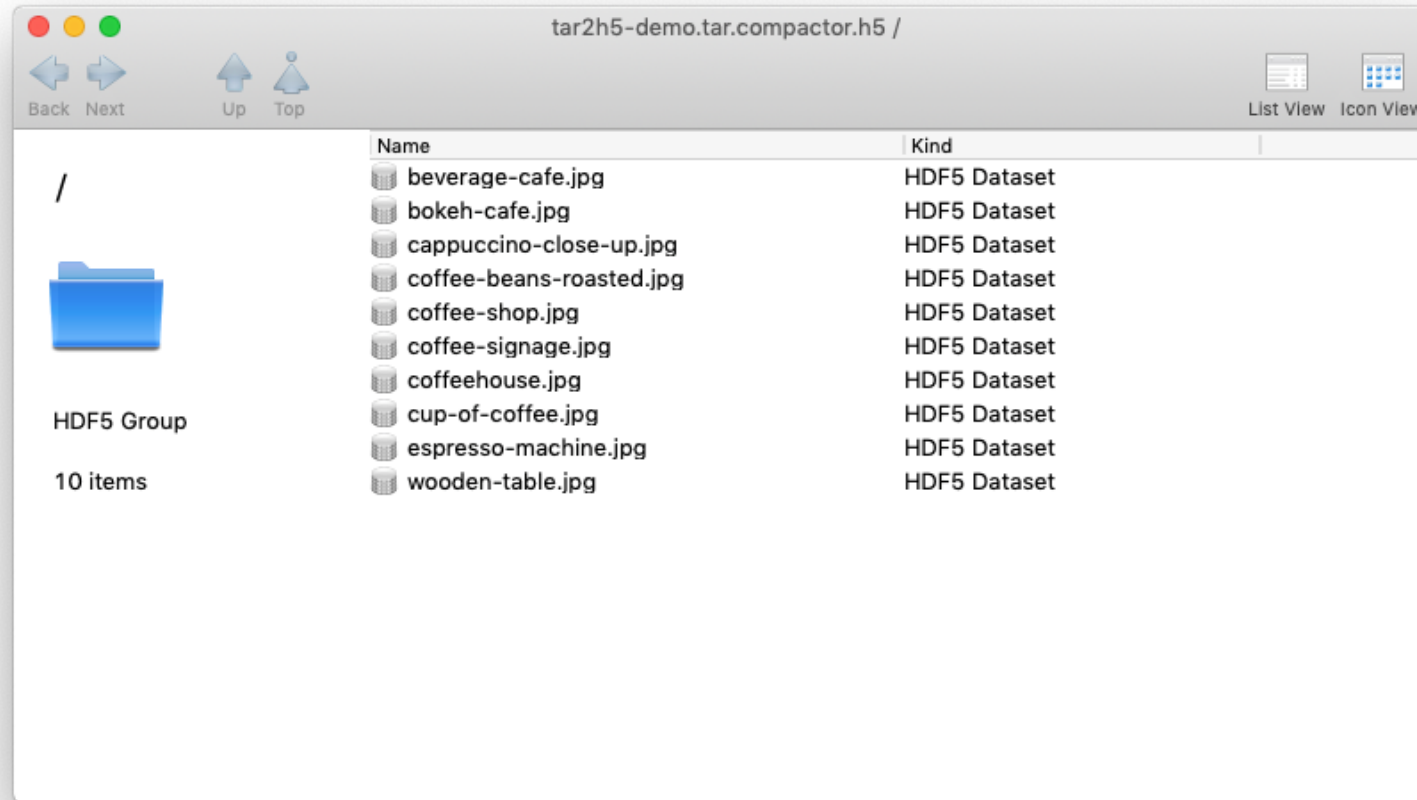
Convert Tape ARChives to HDF5 files

- `archive_checker` check how many files can be extracted from the input tar file.
- `archive_checker_64k` check if any files within input tar files larger than 64 KB.
- `h5compactor` convert input tar file into hdf5 file, all files within tar file should smaller than 64KB, using small files name as dataset names.

The right sidebar shows the repository's statistics: 6 Unwatch, 0 Stars, and 0 Forks. It also includes sections for Releases, Packages, Contributors (gheber Gerd Heber, superdavidxp Dawei Mu), and Languages (C 87.1%, CMake 12.9%).

Visualization with HDFCompass

- <https://support.hdfgroup.org/projects/compass/>



Future Work

- Mixed Precision Support
 - Float32
 - BFloat16
 - TFloat32



THANK YOU FOR YOUR TIME !



ILLINOIS

NCSA | National Center for
Supercomputing Applications