Data Details

From IllustrisWiki

The complete Illustris data set (all snapshots, group catalogs, and merger trees) are available on Odyssey.

This page describes the raw HDF5 snapshot files and their access. For derived data products, see Halo Catalogs, Merger Trees and Post-Processed Catalogs, respectively.

The snapshot and subbox snapshot descriptions apply to all three resolution levels of the full physics ("FP") runs.

Contents

Corrupted Data

Please be aware that **for the Illustris-1 (1820FP) run:**

- **snapshot=53 (z=4.17) is corrupt** (multiple sub-chunks).
- **snapshot=55 (z=3.94) is corrupt** (multiple sub-chunks).
  - *The likely best course of action for analysis is to simply pretend that this snapshot does not exist, e.g. that the snapshot spacing was slightly worse at this time. However, care is perhaps requiring for using tracer quantities across this snapshot, or using merger trees across this snapshot.*

- **the following subbox snapshots are missing some (or all) files:**
  - subbox0 - 2806, 3026, 3533.
  - subbox1 - 462, 2806, 3026, 3526, 3533.
  - subbox2 - 462, 2806, 3026.
  - subbox3 - 462, 2806, 3026.

- **the following subbox snapshots contain corrupt files:**
  - subbox0 - 2353, 2989, 3311.
  - subbox1 - TBD
  - subbox2 - TBD
  - subbox3 - 3619.

Please be aware that **for the Illustris-2 (910FP) run:**

- **the following subbox snapshots are missing some (or all) files:**
  - subbox2 - 2252 (however, the 'grouporderedd' snapshot files are complete, so use those).

Incorrect Data

Please be aware that **for the Illustris-1 (1820FP) run:**

- Stellar metallicities in the group catalogs are incorrect for some galaxies, due to a bug in the subfind_properties function.
- The total number of particles in FOF groups (Nids_Total in the group catalogs) is incorrect at snapshots>=58.

Other Important Notes

Please be aware that **for the Illustris-1 (1820FP) run:**

- Starting at snapshot=72, Coordinates are written in double (64 bit). In previous snapshots they are in float (32 bit).

Data Format

Snapshot Organization

There are 136 snapshots stored for every run, the complete table of spacings and redshifts is here:

- Snapshot Spacing and Redshift

A truncated version is:

Snap	Scalefactor	Redshift

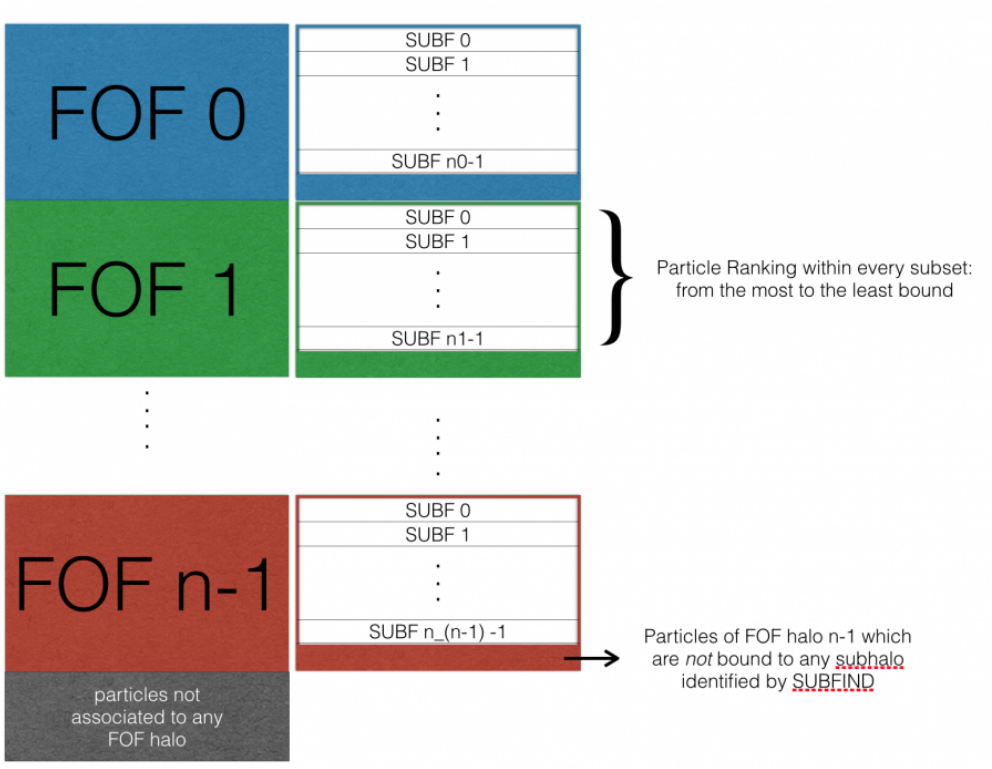
0	0.020932	46.773
32	0.090937	9.9966
45	0.14264	6.0108
49	0.16678	4.9959
54	0.19968	4.0079
60	0.24949	3.0081
68	0.33311	2.002
85	0.50068	0.9973
103	0.66531	0.50305
135	1	0

Every snapshot is stored in a series of "Chunks", i.e. more manageable, smaller-size files. The number of chunks per snapshots is different for the different runs:

Run	**aka**	**Total NumPart (DM)**	**Chunks per Snapshot**	**Snapshot Size**	**Total Data Volume**
L75n455FP	Illustris-3	94,196,375	32	22 GB	3 TB
L75n455NR	Illustris-Adiabatic-3	94,196,375	8	9 GB	1 TB
L75n455DM	Illustris-Dark-3	94,196,375	8	3.2 GB	0.4 TB
L75n910FP	Illustris-2	753,571,000	256	176 GB	24 TB
L75n910NR	Illustris-Adiabatic-2	753,571,000	32	57 GB	8 TB
L75n910DM	Illustris-Dark-2	753,571,000	32	26 GB	3.5 TB
L75n1820FP	Illustris-1	6,028,568,000	512	1.5 TB	204 TB
L75n1820DM	Illustris-Dark-1	6,028,568,000	128	203 GB	28 TB

Note: The snapshot data is **not** organized according to spatial position. Rather, particles within the snapshot files are sorted according to their group/subgroup memberships, according to the FoF/Subfind algorithms. Within each particle type, the sort order is: GroupNumber, SubgroupNumber, BindingEnergy, where particles belonging to the group but not to any of its subgroups ("fuzz") are included after the last subgroup.

This is a schematic view of the particle organization within a snapshot, for **one particle type**:



Note: The truncation of a snapshot in chunks is arbitrary, thus halos may happen to be stored across multiple, subsequent chunks. Similarly, the different particle types of a halo can be stored in different sets of chunks.

Snapshot Contents

Every HDF5 snapshot contains a "Header" and 5 additional groups named "PartTypeX", corresponding to the following particle types (the DM only runs have a single PartType1 group):

- PartType0 - GAS
- PartType1 - DM
- PartType2 - unused
- PartType3 - TRACERS
- PartType4 - STARS + WIND PARTICLES (Winds have negative ages)
- PartType5 - BLACK HOLES

The most important fields of the header are:

Header

Field	Dimensions	Units	Description
BoxSize	1	$ckpc/h$	Spatial extent of the periodic box.
MassTable	6	$10^{10} M_\odot/h$	Masses of particle types which have a constant mass (only DM).
NumPart_ThisFile	6		Number of particles (of each type) included in this (sub-)file.
NumPart_Total	6		Total number of particles (of each type) across all (sub-)files of this snapshot, modulo 2^{32}.
NumPart_Total_HighWord	6		Total number of particles (of each type) across all (sub-)files of this snapshot, divided by 2^{32} and rounded downwards.
Omega0	1		The cosmological density parameter for matter.
OmegaLambda	1		The cosmological density parameter for the cosmological constant.
Redshift	1		The redshift corresponding to the current snapshot.
Time	1		The scale factor (=1/(1+z)) corresponding to the current snapshot.

These are the snapshot fields for a Full Physics (FP) run:

PartType0 (gas)

Field	Dimensions	Units	Description
Coordinates	N,3	$ckpc/h$	Spatial position within the periodic box of size 75000 ckpc/h. *Comoving coordinate*.
Density	N	$(10^{10} M_\odot/h)/(ckpc/h)^3$	Mass density of cell (calculated as mass/volume).
ElectronAbundance	N		
GFM_AGNRadiation	N		
GFM_CoolingRate	N		
GFM_Metallicity	N	-	The ratio M_Z/M_{total} where M_Z is the total mass all metal elements (above He). Is NOT in solar units. To convert to solar metallicity, divide by 0.0127 (the primordial solar metallicity).
GFM_Metals	N,9	-	Individual abundances of nine species: H, He, C, N, O, Ne, Mg, Si, Fe.
GFM_WindDMVelDisp	N		
InternalEnergy	N		
Masses	N	$10^{10} M_\odot/h$	Gas mass in this cell. Refinement/derefinement attempts to keep this value within a factor of two of the targetGasMass for every cell.
NeutralHydrogenAbundance	N		
NumTracers	N	-	The number of child MC tracers residing within this gas cell.
ParticleIDs	N	-	The unique ID (64bit) of this gas cell. Constant for the duration of the simulation. May cease to exist (as gas) in a future snapshot due to conversion into a star/wind particle, or accretion into a BH.
Potential	N		
SmoothingLength	N	$ckpc/h$	

StarFormationRate	N	$M_\odot/yr$	Converted into these normal units when the snapshots are written.
SubfindDensity	N		
SubfindHsml	N	$ckpc/h$	The radius of the sphere enclosing the X (20?64?) nearest neighbors (of this type only, or of DM?).
SubfindVelDisp	N		
Velocities	N,3	$km/s/\sqrt{a}$	Spatial velocity. The peculiar velocity is obtained by multiplying this value by $\sqrt{a}$.
Volume	N		

PartType1 (dm)

Field	Dimensions	Units	Description
Coordinates	N,3	$ckpc/h$	Spatial position within the periodic box of size 75000 ckpc/h. *Comoving coordinate*.
ParticleIDs	N	-	The unique ID (64bit) of this DM particle. Constant for the duration of the simulation.
Potential	N		
SubfindDensity	N		
SubfindHsml	N	$ckpc/h$	The radius of the sphere enclosing the X (20?64?) nearest neighbors (of this type only, or of DM?).
SubfindVelDisp	N		
Velocities	N,3	$km/s/\sqrt{a}$	Spatial velocity. The peculiar velocity is obtained by multiplying this value by $\sqrt{a}$.

PartType3 (MC tracers)

Field	Dimensions	Units	Description
FluidQuantities	N,13	Various	Thirteen auxiliary quantities stored for each tracer with differing significance. See #Tracer Quantities below.
ParentID	N	-	The unique ID (64bit) of the parent of this tracer. *Could be a gas cell, star, wind phase cell, or BH*.
TracerID	N	-	The unique ID (64bit) of this tracer. Constant for the duration of the simulation.

PartType4 (stars / wind phase gas)

Field	Dimensions	Units	Description
Coordinates	N,3	$ckpc/h$	Spatial position within the periodic box of size 75000 ckpc/h. *Comoving coordinate*.
GFM_InitialMass	N	$10^{10} M_\odot/h$	Mass of this star particle when it was formed (will subsequently decrease due to stellar evolution).
GFM_Metallicity	N	-	See entry under PartType0. Inherited from the gas cell spawning/converted into this star, at the time of birth.
GFM_Metals	N,9	-	Individual abundances of nine species: H, He, C, N, O, Ne, Mg, Si, Fe.
GFM_StellarFormationTime	N	-	The "BirthTime" field (scalefactor when this star was created). **Only differentiation between a real star (>=0) and a wind phase gas cell (<0) is the sign of this quantity.**
GFM_StellarPhotometrics	N,8		Magnitudes in eight bands: U, B, V, K, g, r, i, z
Masses	N	$10^{10} M_\odot/h$	Mass of this star or wind phase cell.
NumTracers	N	-	Number of child tracers belonging to this star/wind phase cell.
ParticleIDs	N	-	The unique ID (64bit) of this star/wind cell. Constant for the duration of the simulation.
Potential	N		

SubfindDensity	N		
SubfindHsml	N	$ckpc/h$	The radius of the sphere enclosing the X (20?64?) nearest neighbors (of this type only, or of DM?).
SubfindVelDisp	N		
Velocities	N,3	$km/s/\sqrt{a}$	Spatial velocity. The peculiar velocity is obtained by multiplying this value by $\sqrt{a}$.

PartType5 (black holes)

Field	Dimensions	Units	Description
BH_CumEgyInjection_QM	N		
BH_CumMassGrowth_QM	N		
BH_Density	N		
BH_Hsml	N		
BH_Mass	N		
BH_Mass_bubbles	N		
BH_Mass_ini	N		
BH_Mdot	N	internal mass-over-time code units: $(10^{10} M_\odot/h)/(0.978 Gyr/h)$ - multiply by 10.22 to get $M_\odot/yr$	The mass accretion rate onto the black hole, instantaneous.
BH_Pressure	N		
BH_Progs	N		
BH_U	N		
Coordinates	N,3	$ckpc/h$	Spatial position within the periodic box of size 75000 ckpc/h. *Comoving coordinate*.
HostHaloMass	N	$10^{10} M_\odot/h$	
Masses	N	$10^{10} M_\odot/h$	Mass of this black hole. Monotonically increases with time according to the accretion prescription, starting from the seed mass.
NumTracers	N	-	The number of child MC tracers residing within this BH.
ParticleIDs	N	-	The unique ID (64bit) of this black hole. Constant for the duration of the simulation. May cease to exist in a future snapshot due to a BH merger.
Potential	N		
SubfindDensity	N		
SubfindHsml	N	$ckpc/h$	The radius of the sphere enclosing the X (20?64?) nearest neighbors (of this type only, or of DM?).
SubfindVelDisp	N		
Velocities	N,3	$km/s/\sqrt{a}$	Spatial velocity. The peculiar velocity is obtained by multiplying this value by $\sqrt{a}$.

Tracer Quantities

Each Monte Carlo tracer particle stores 13 auxiliary values. These are updated every timestep where the tracer parent is active. Many are reset to zero immediately after they are written out to a snapshot, such that their recording duration is precisely the timesteps between two successive snapshots. Some are only relevant when the tracer resides within a parent of a specific particle type (e.g. gas or star).

Number	Name	Reset Each Snapshot?	Units	Description
0	TMax	Y	Kelvin	The maximum past temperature of the parent gas cell, back to the previous snapshot. *Only updated when parent is a gas cell*.
1	TMax_Time	Y	-	Scalefactor of the above TMax event. *Only updated when parent is a gas cell*.
				Density of the parent gas cell when the

2	TMax_Time_Rho	Y	$(10^{10} M_\odot/h)/(ckpc/h)^3$	most recent TMax was recorded. *Only updated when parent is a gas cell.*
3	RhoMax	Y	$(10^{10} M_\odot/h)/(ckpc/h)^3$	Maximum past density of the parent gas cell, back to the previous snapshot. *Only updated when parent is a gas cell.*
4	RhoMax_Time	Y	-	Scalefactor of the above RhoMax event. *Only updated when parent is a gas cell.*
5	MachMax	Y	-	Maximum past mach number of the parent gas cell, as set by `get_mach_numbers()` in the Riemann solver. *Only updated when parent is a gas cell.*
6	EntMax	Y	$P/(\rho/a^3)^\gamma$	Maximum past entropy of the parent gas cell, back to the previous snapshot. *Only updated when parent is a gas cell.* Note slightly strange units, where P and ρ are pressure and density, as in the snapshots.
7	EntMax_Time	Y	-	Scalefactor of the above EntMax event. *Only updated when parent is a gas cell.*
8	Last_Star_Time	N	-	Scalefactor, set only when this tracer exchanges from a star/wind to a gas, or from a gas to a star/wind. These four cases respectively set `LST = { a, -a, a+1, a+2 }`.
9	Wind_Counter	N	int32	Integer counter initialized to zero, increased by one each time this tracer is moved from a gas cell to a wind particle.
10	Exchange_Counter	N	int32	Integer counter initialized to zero, increased by one each time this tracer is exchanged, regardless of parent type.
11	Exchange_Distance	N	$ckpc/h$	Cumulative sum of the spatial distance over which this tracer has moved due to Monte Carlo exchange between gas cells. In particular, the sum of the parent gas cell radii when either the originating parent or destination parent is of gas type.
12	Exchange_Distance_Error	N	$ckpc/h$	Cumulative sum of $r_{\text{cell}} \times (\sqrt{N_{\text{exch}}} - \sqrt{N_{\text{exch}} - 1})$, when either the originating or destination parent is of gas type.

Note: Tracers are exchanged (and can therefore change their parents) in the following ways:

- Gas -> Gas *(finite volume fluxes, refinement, derefinement)*
- Gas -> Stars *(star formation, both spawning new stars and converting cells into stars)*
- Stars -> Gas *(stellar mass return)*
- Gas -> Wind *(galactic scale stellar winds)*
- Wind -> Gas *(recoupling stellar wind)*
- Gas -> BHs *(blackhole accretion)*
- BHs -> BHs *(blackhole mergers)*

Subboxes

Four separate "subbox" cutouts exist, for each FP run. These are spatial cutouts of fixed comoving size and fixed comoving coordinates. They are output at each highest timestep, that is, their time resolution is significantly better than the main snapshots. This may be useful for some types of analysis or particular science questions, or for making movies.

- Caution: The time spacing of the subboxes is not uniform in scale factor or redshift, but scales with the time integration hierarchy of the simulation, and is thus variable, with some discrete factor-2 jumps at several points during the simulations. The output times can be found in files named **.../postprocessing/subboxes_times.txt** that exist for each of the simulations.

- Note: The subboxes, unlike the full box, are not periodic.

Run	Number of Subbox Snapshots	Chunks per Snap	Time Resolution (at z=6)	(at z=2)	(at z=0)
455_FP	1426	1	~7 Myr	~12 Myr	~33 Myr
910_FP	2265	16	~4 Myr	~6 Myr	~17 Myr
1820_FP	3976	512	~2 Myr	~3 Myr	~8 Myr

Subbox #	Environment	Ω_m^{sub}	XYZ Center	BoxSize	Volume Fraction
0	Crowded, including a ~5e13Msun halo	1.47	(9000, 17000, 63000)	7.5 cMpc/h	0.1%
1	Less crowded, including several >1e12Msun halos	0.16	(43100, 53600, 60800)	8.0 cMpc/h	0.12%
2	Less crowded, including several >1e12Msun halos	0.29	(37000, 43500, 67500)	5.0 cMpc/h	0.03%
3	Least crowded, including several ~1e12Msun halos	0.25	(64500, 51500, 39500)	5.0 cMpc/h	0.03%

The four subboxes sample four different areas of the large box, roughly described by the Environment above. Subbox snapshots have a format identical to the main snapshots. They (should, at least in theory) have group catalogs and merger trees also in formats identical to those for the main snapshots.

Odyssey Data Access

Available at:

```
/n/hernquistfs1/Illustris/Runs/
```

For examples of reading snapshot data, see Data Access Examples.

HITS Data Access

What data is available at HITS? todo
