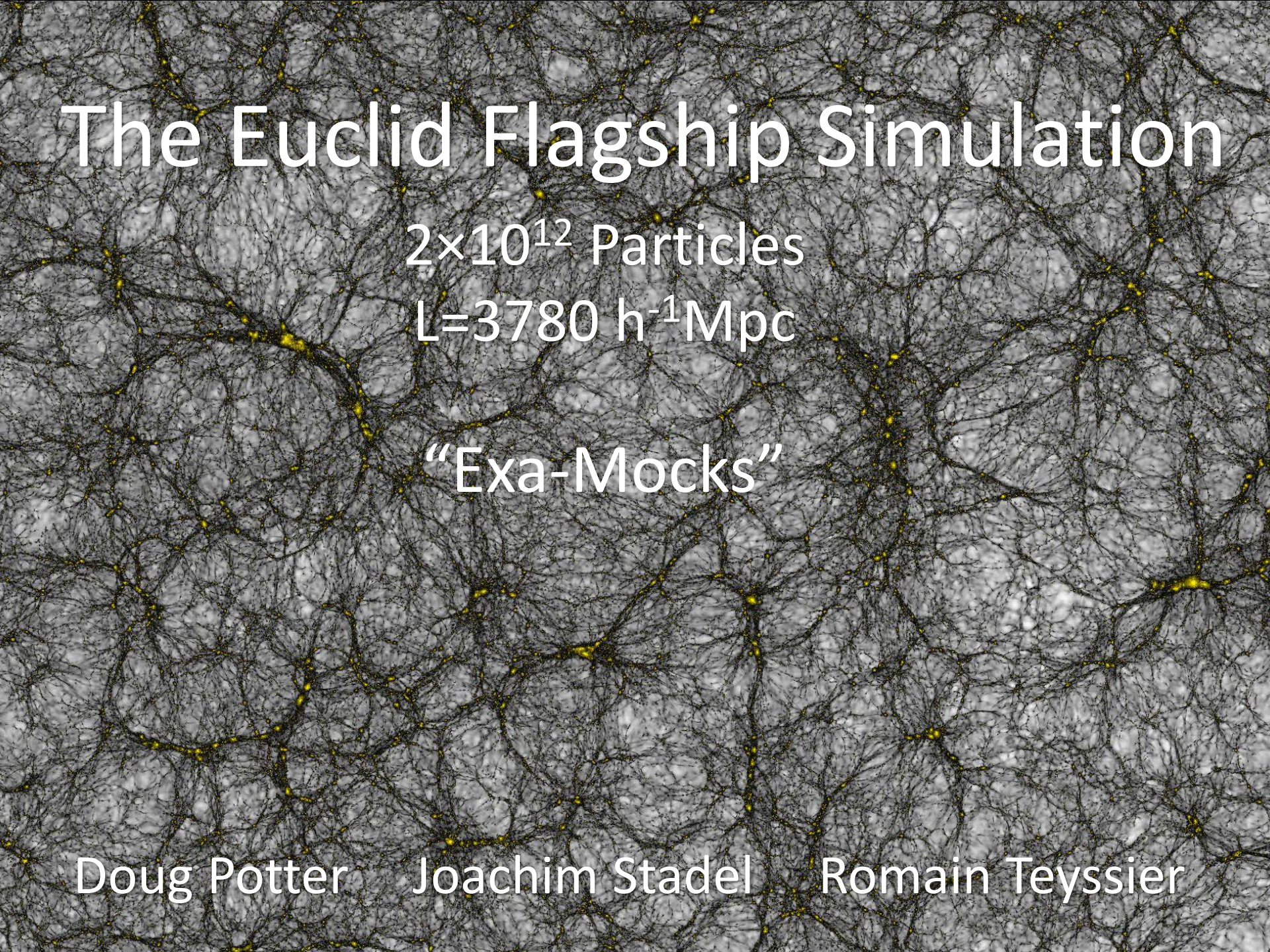
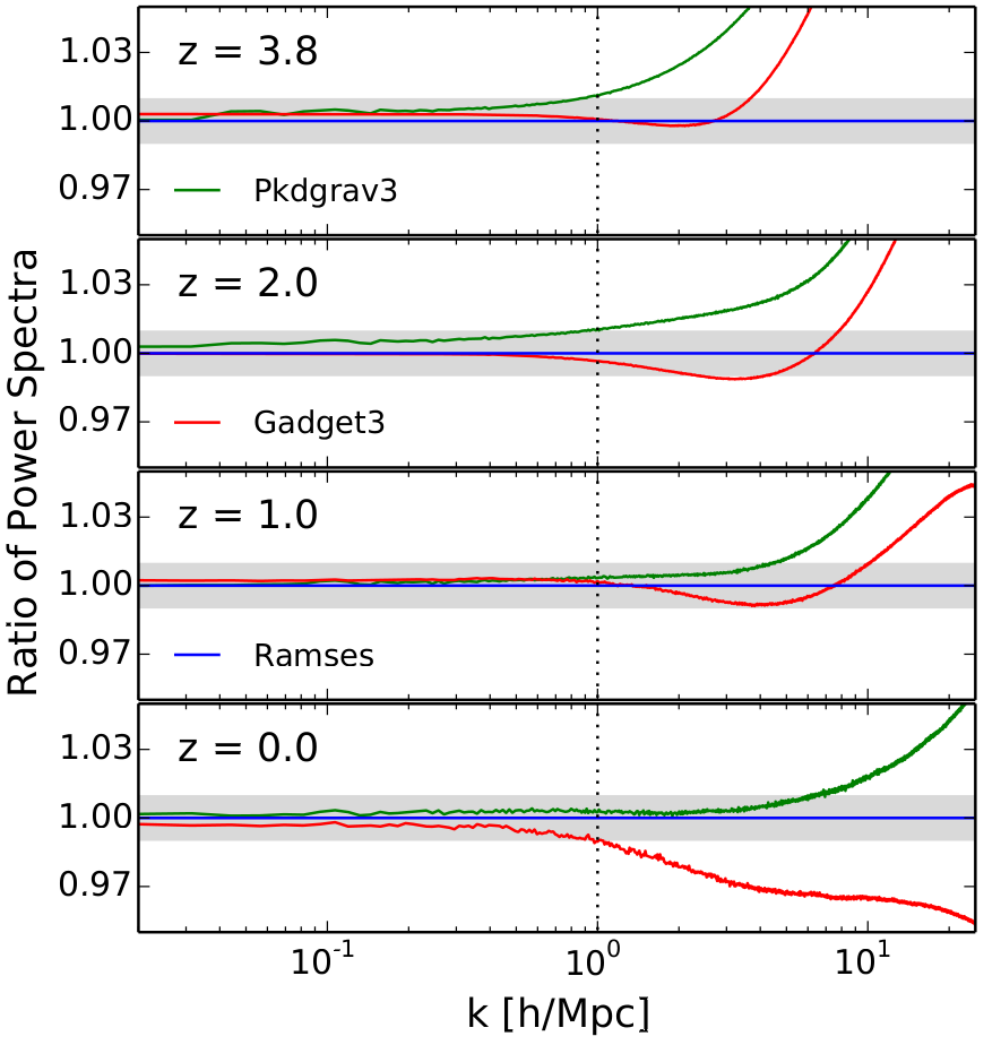
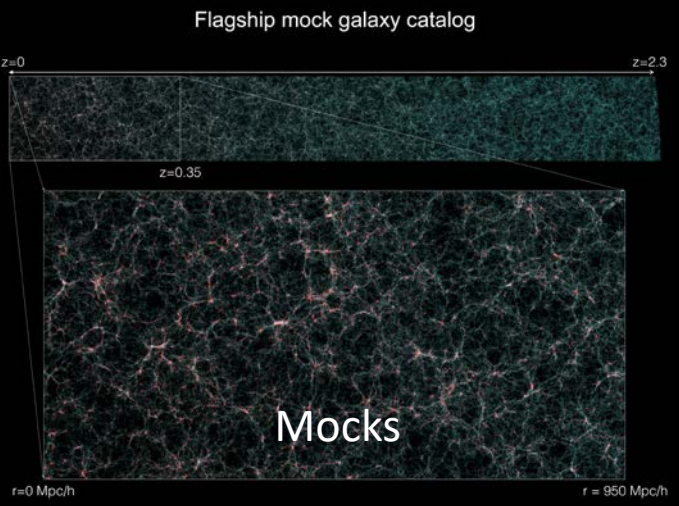

The Euclid Flagship Simulation
2×10^12 Particles
L=3780 h^-1Mpc
“Exa-Mocks”
Doug Potter
Joachim Stadel
Romain Teyssier

“Accurate” Theory?

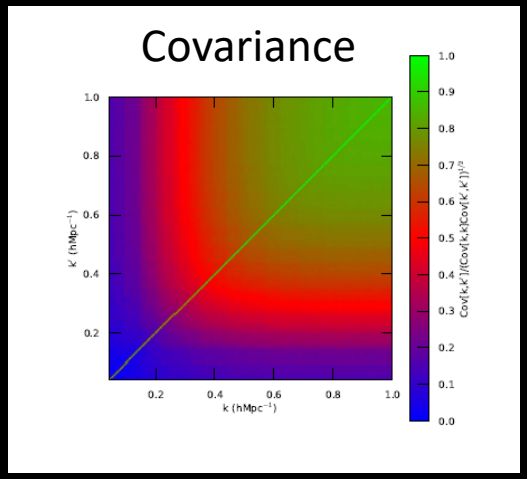


Note: all 3 codes have very different Poisson solvers and integration methods!

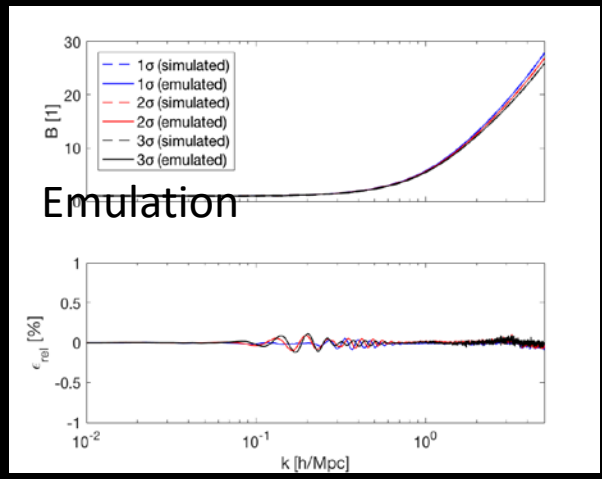


N-Body

Klypin & Prada 17, Blot+ 15



Knabenhans+ (prep)



Piz Daint – over 5000 GPU Nodes

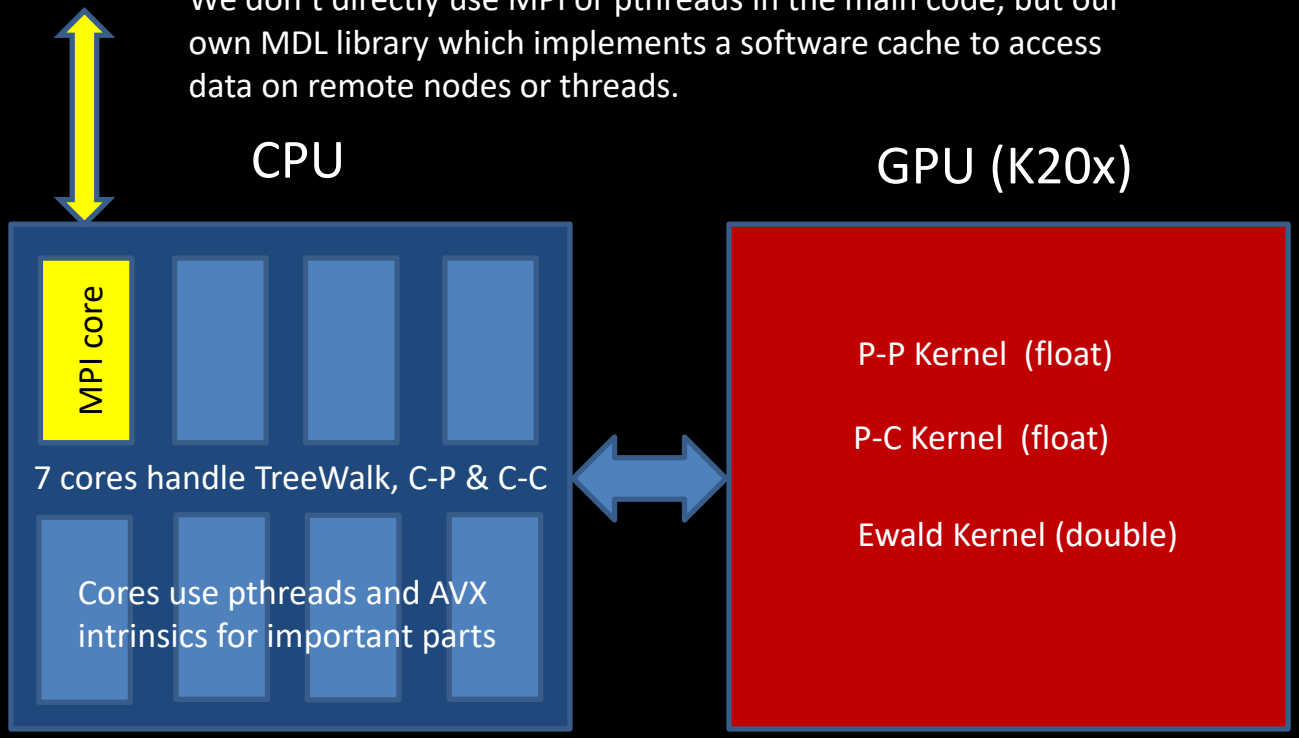
Swiss National Computing Center (CSCS) in Lugano, Switzerland

3rd Fastest Computer in the World Now!

GPU Hybrid Computing

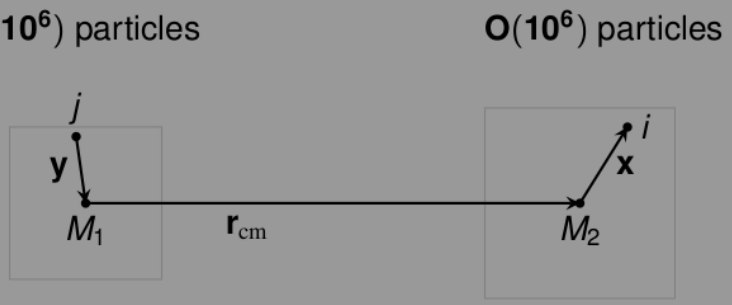
Piz Daint 2016 Flagship 1.0

We don’t directly use MPI or pthreads in the main code, but our own MDL library which implements a software cache to access data on remote nodes or threads.



Fast Multipole Method - O(N)

Quick explanation of FMM



Direct **O**(**10**^{12}) interactions to calculate! $O(N^2)$ code.

Tree Use a multipole approximation for the mass at M_2 to calculate the force at each j: **O**(**10**^{6}) interactions to calculate. $O(N \log N)$ code.

FMM Use a multipole approx for the mass at M_2 to approximate the "potential landscape" at M_1 (n^{th} order gradients of the potential): **O**(**1**) interaction to calculate. $O(N)$ code!

Memory Usage in pkdgrav3

1 billion particles can fit on a 64 Gbyte Node like Piz Daint

28 bytes persistent

6 bits: old rung 24: group id	
pos[0]	int32_t
pos[1]	int32_t
pos[2]	int32_t
vel[0]	float
vel[1]	float
vel[2]	float

<28 bytes / particle

Tree Cells
Binary Tree

4th order
Multipoles
(float prec)

~5 bytes / particle

Cache/Buffers

0-8 bytes ephemeral

Group finding
Other analysis

ClAoS is used for the particle and cell memory which makes moving particles around simple

AoSoA is used for all interaction lists which are built by the TreeWalk algorithm.

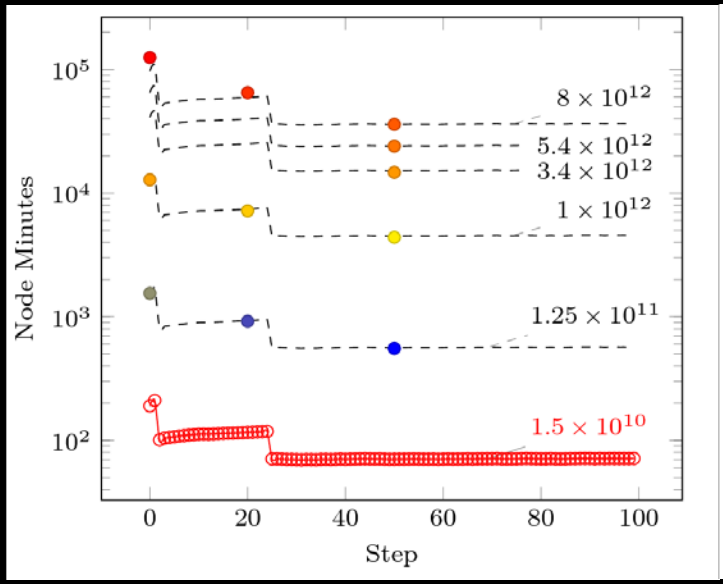
Reducing memory usage increases the capability of existing machines, but also increases performance somewhat. Simulations are limited more by memory footprint.

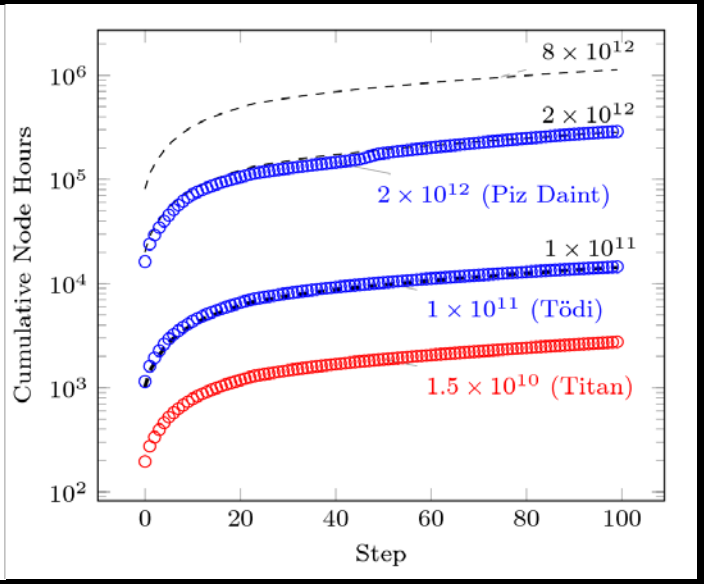
see Potter+ 2017 for details

Benchmarking on Titan and Piz Daint

Nearly Perfect Weak Scaling makes performance prediction very accurate for these simulations.
120 seconds for an all N gravity solve!

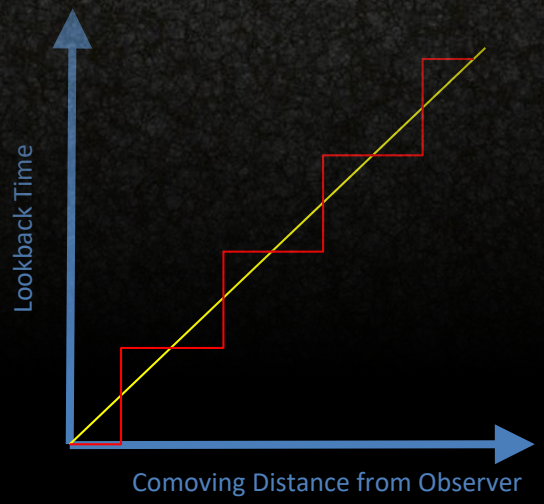
We show that it is quite feasible to run 8 trillion particles on Titan with a little over 1 million node hours. **10 PFlops**



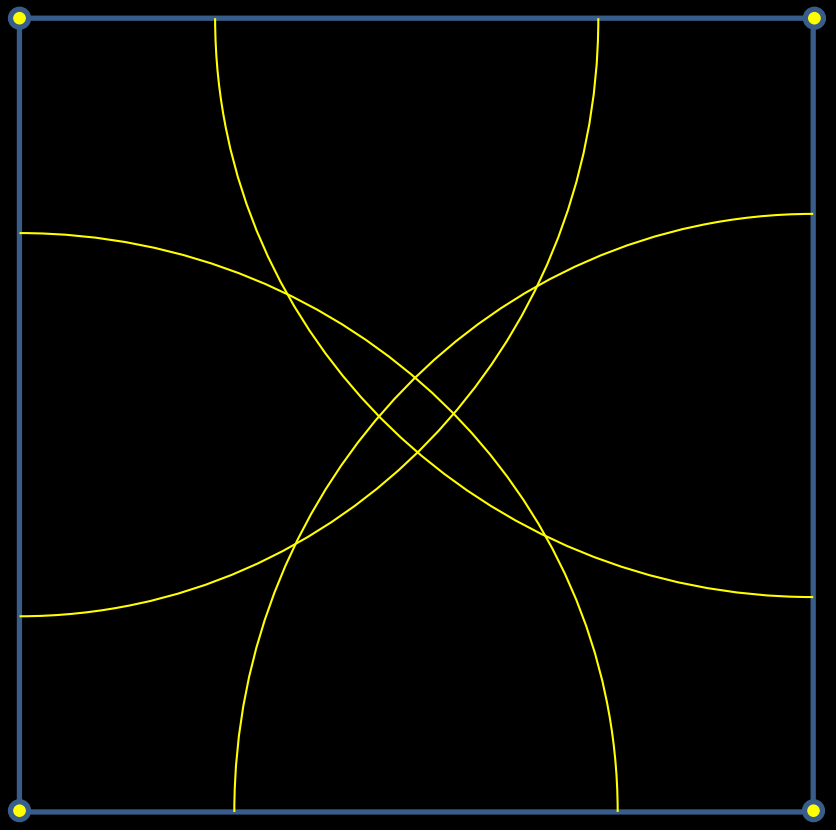


see Potter+ 2017 for details

For the Flagship Mock we produced a light cone without discontinuous jumps in the observed structure.



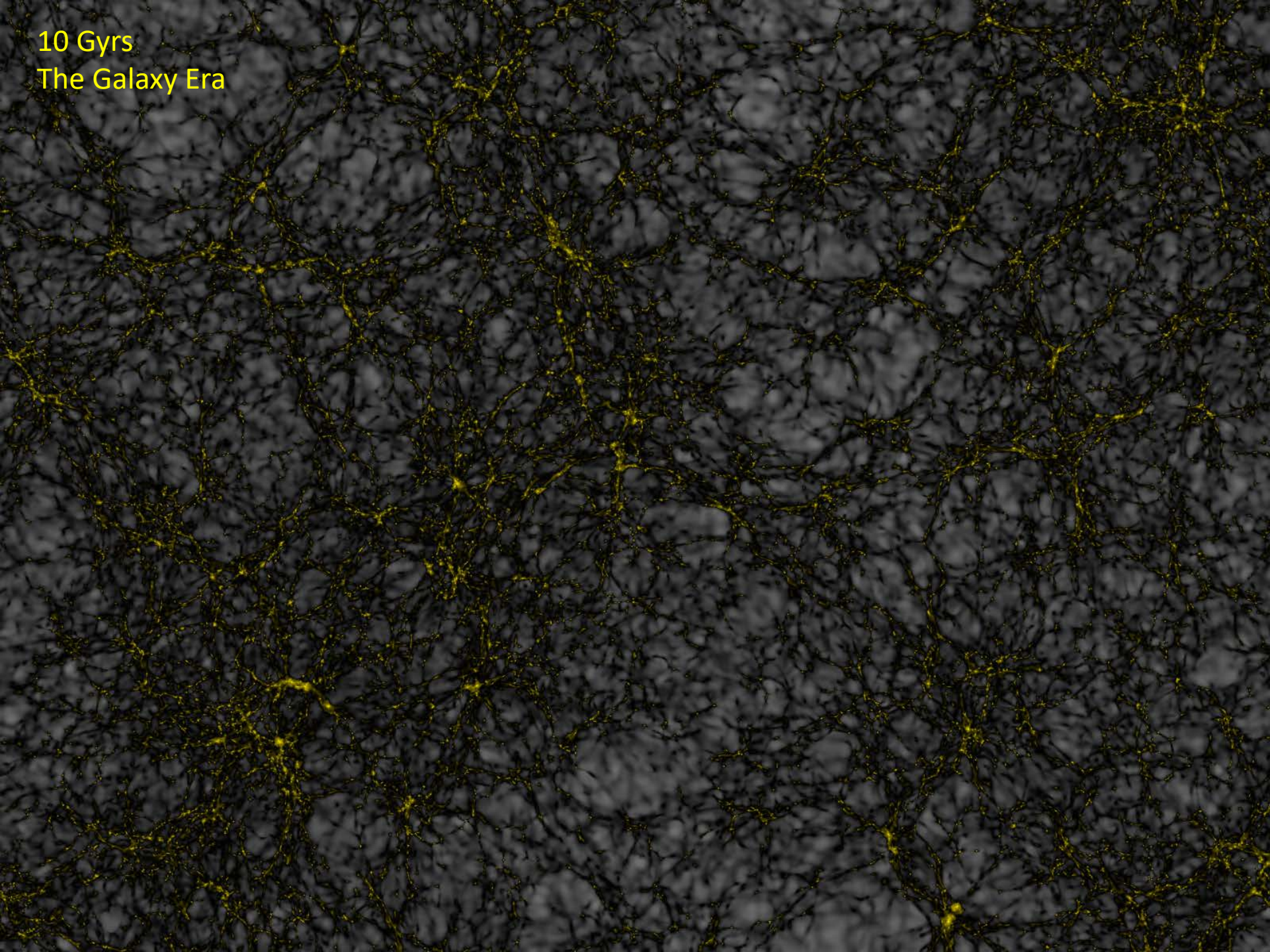
Continuous Light Cone Generation

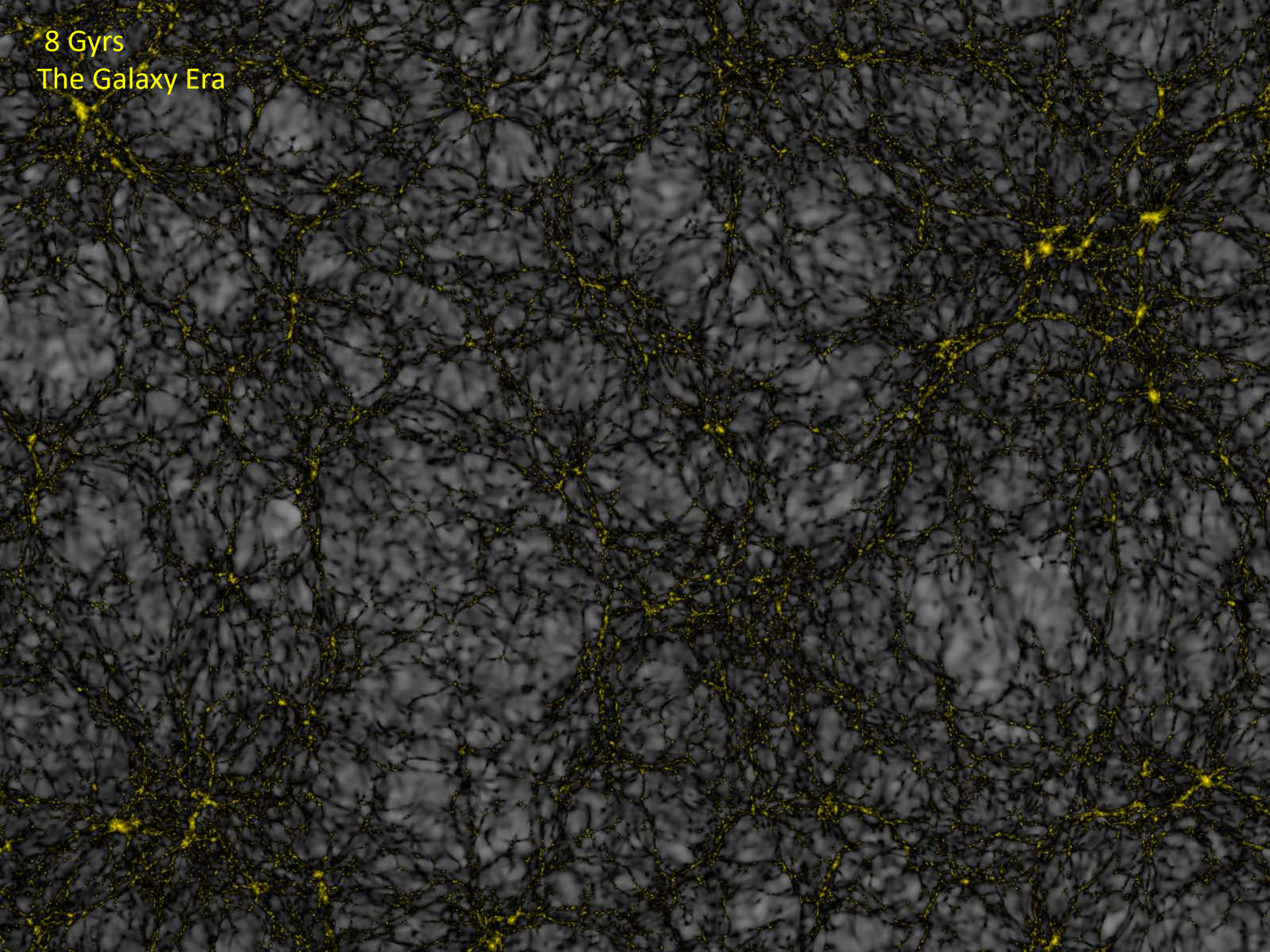


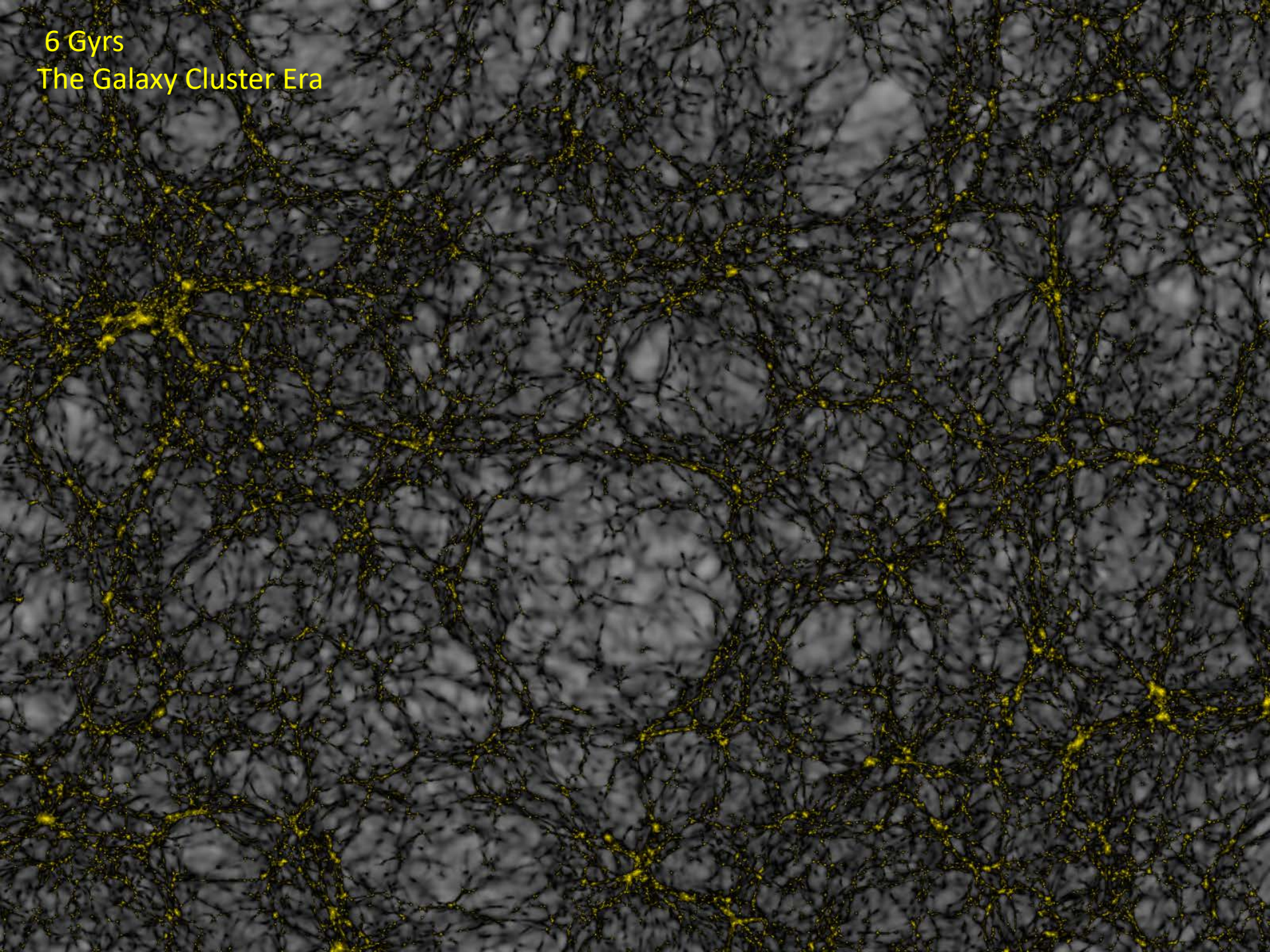
Solve for the precise time at which the light surface intersects a given particle, even when this particle is between 2 of the smallest sub-steps in the simulation.

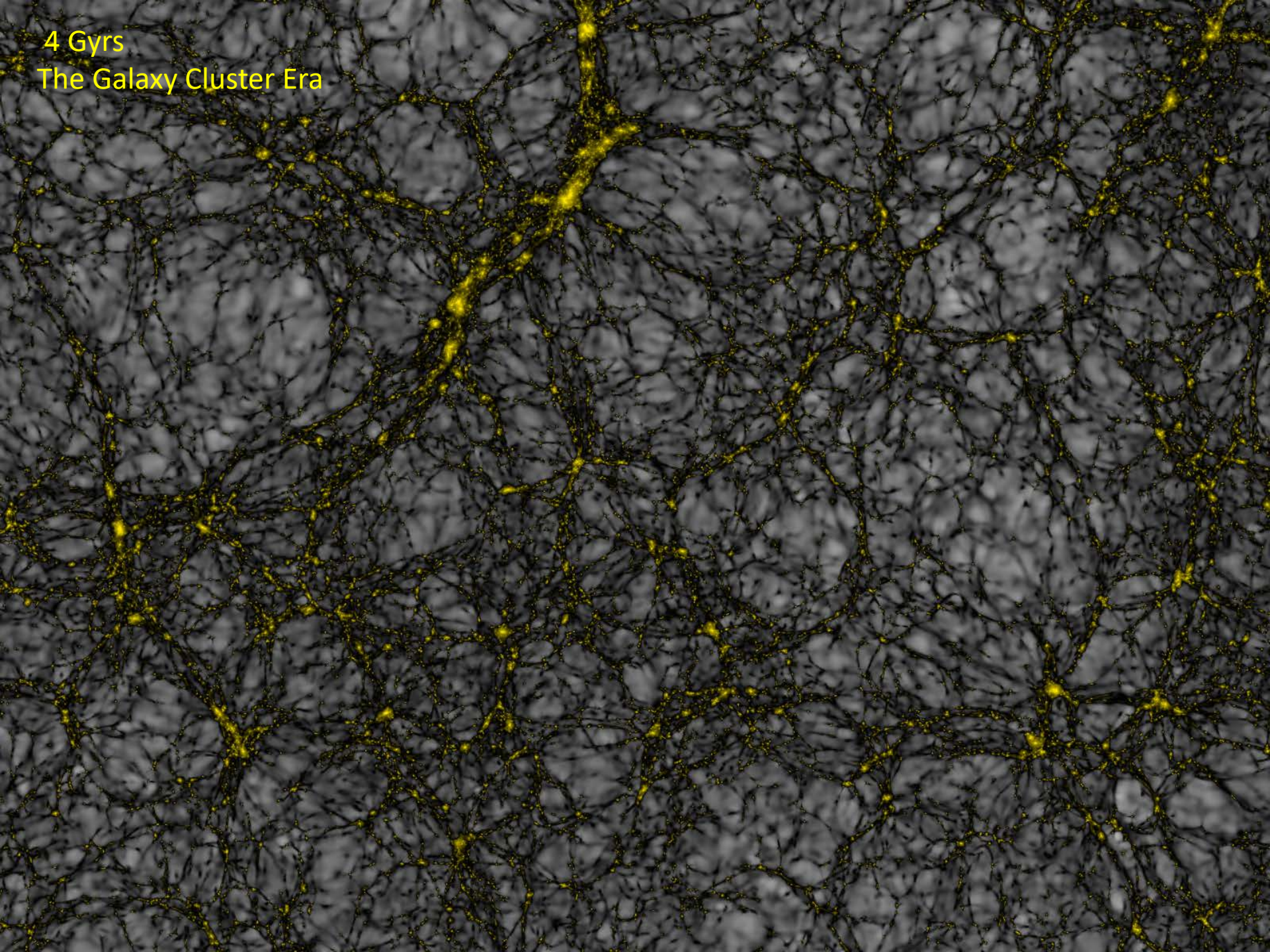
Speed this test up by considering each of the 8 corners of the light cone using AVX vector operations.

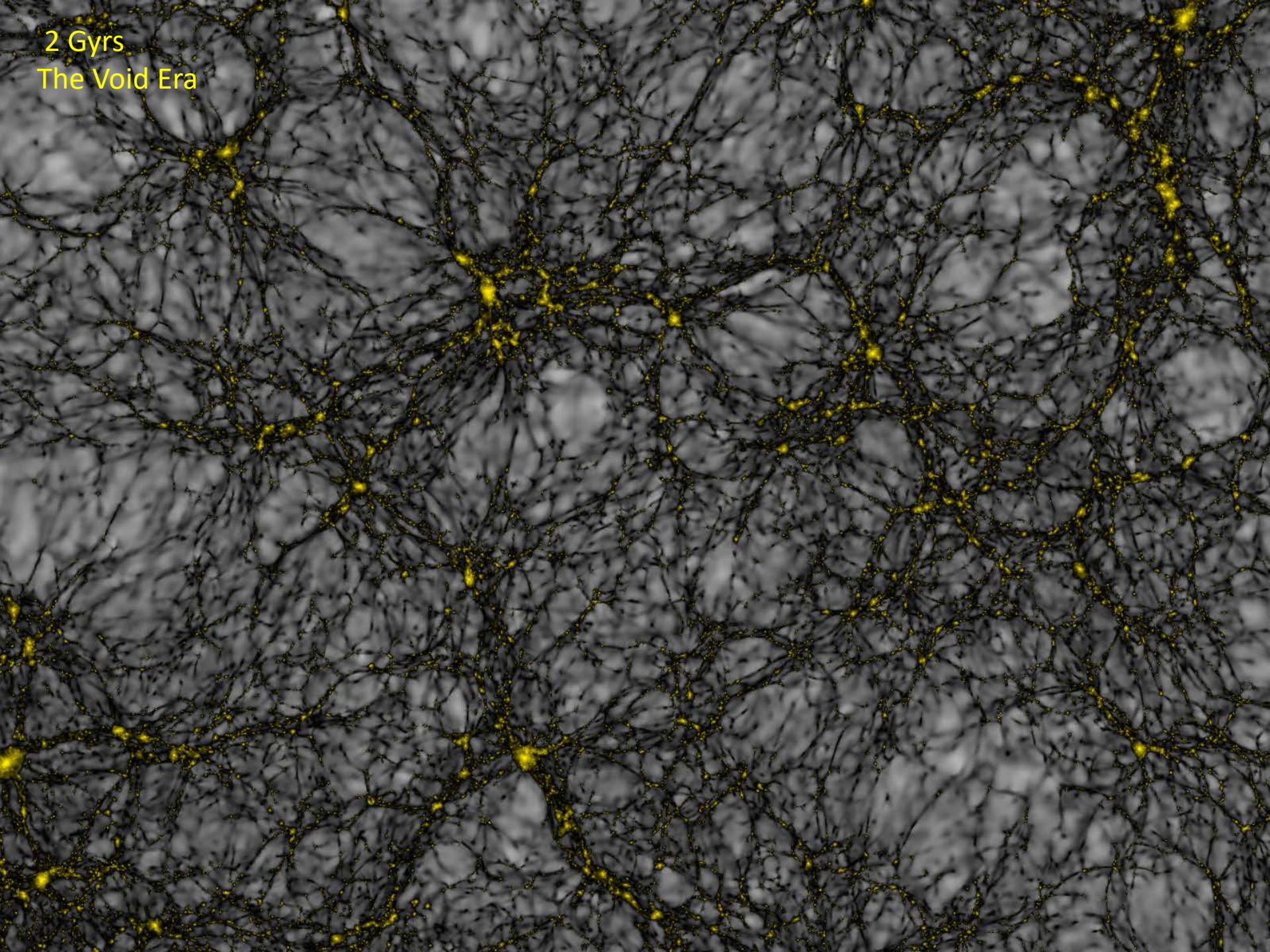
For 2 replicas there are 64 tests and for 3 replicas there are 184 tests.

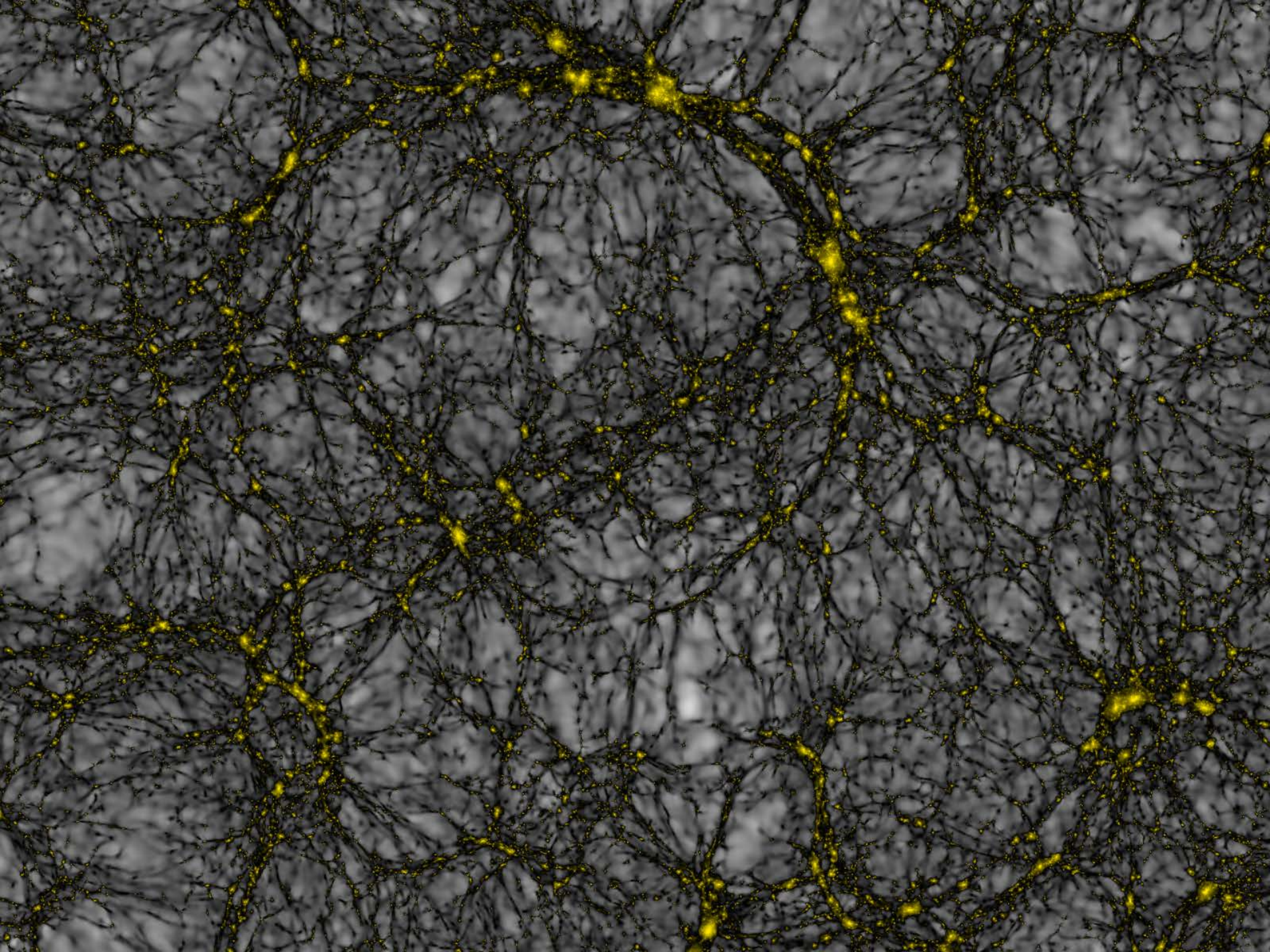

10 Gyrs
The Galaxy Era


8 Gyrs
The Galaxy Era


6 Gyrs
The Galaxy Cluster Era


4 Gyrs
The Galaxy Cluster Era


2 Gyrs
The Void Era



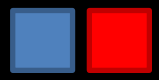
Evolution of Systems
All 3 here have O(5000) Nodes

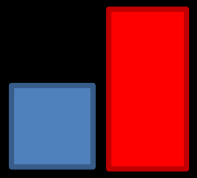
Flagship

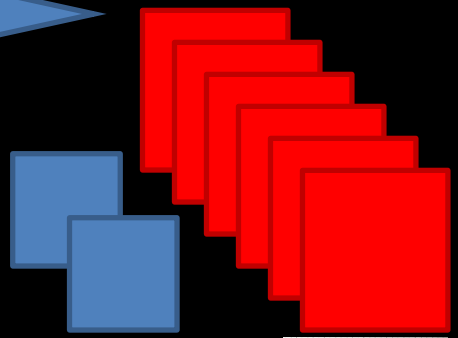
Flagship v2

or?

Future Mock







8 core v2 cpu
Kepler gpu

~2 Tflops/Node

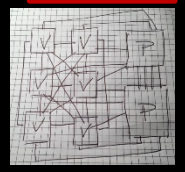
Piz Daint
2016

12 core v3 cpu
Pascal gpu

~5 Tflops/Node

Piz Daint
2017

2 x 24 core
Power9 cpus
6 x Volta gpus



~40 Tflops/Node

Summit
2019

Flagship 2.0 (2019) Summary

- $1x10^{9}\ h^{-1}\ M_{\odot}$ in $4000\ h^{-1}$ Mpc $\cong 4x10^{12}$ particles
- Support for SHAM/SAM Mocks
  - Merger Trees, for time-slice halos and light-cone halos
  - "Orphaned Galaxies" tracked thoughout
  - Several on-the-fly halos finders (Rockstar-*like*, FoF, others?)
  - More halo properties
- Massive neutrinos (0.05 ev) Linear treatment included
- Supporting beyond Born approximation/ray-tracing lensing
- Deep-survey lightcone generated ($z<3.5$)
- On-the-fly modelling of baryonic effects? (by adding a displacement field about halo centers prior to analysis)

Thank-you. Questions? (also see WP5 on redmine for info on available data)