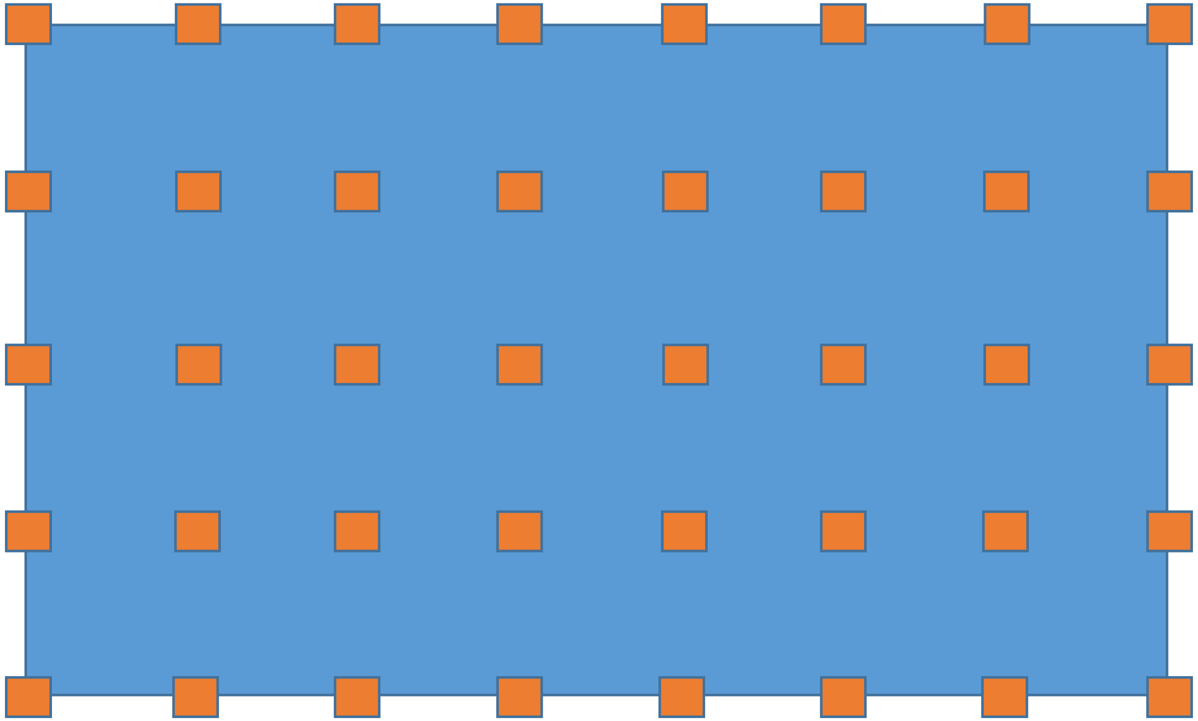
Parallel Applications - I

Lecture 18

March 20, 2024

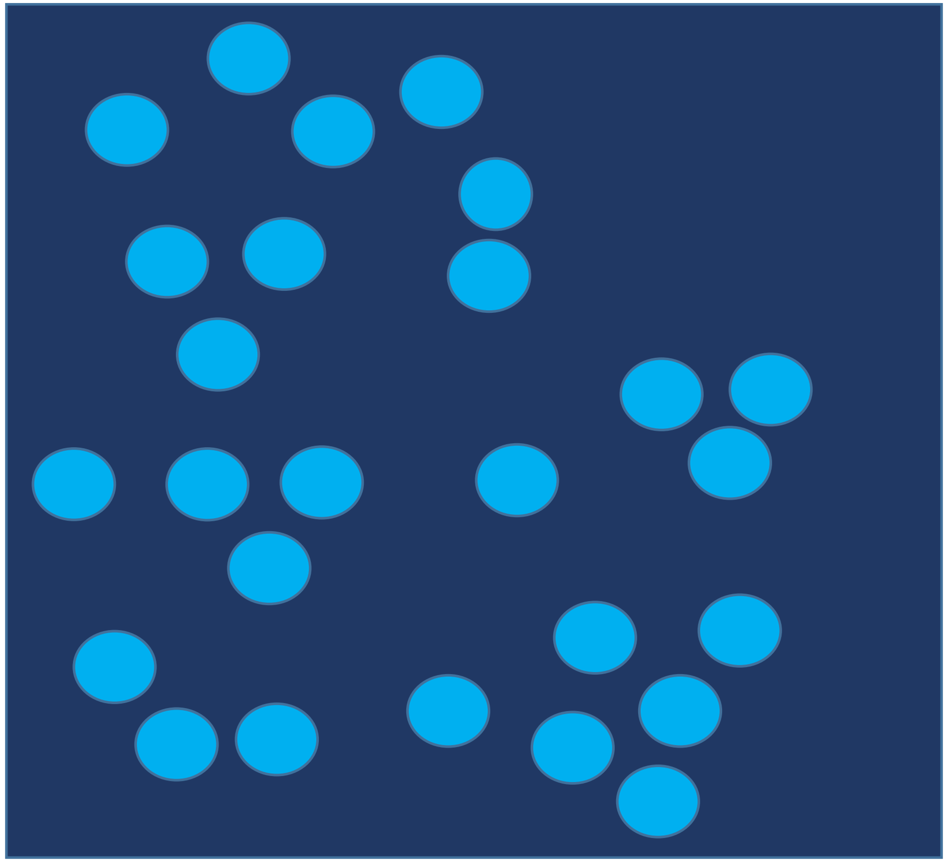
Mathematical Modeling



- Many real-world problems solved using ordinary and partial differential equations
- Discretization of domain
- Finer the grid, more accurate is the solution
- Domain decomposition (regular/uniform grid)

Parallelization of Irregular Application – Cosmology Application

N-body Simulation



Problem

- N bodies exert force on each other
- Model velocities and positions of particles over time

Applications

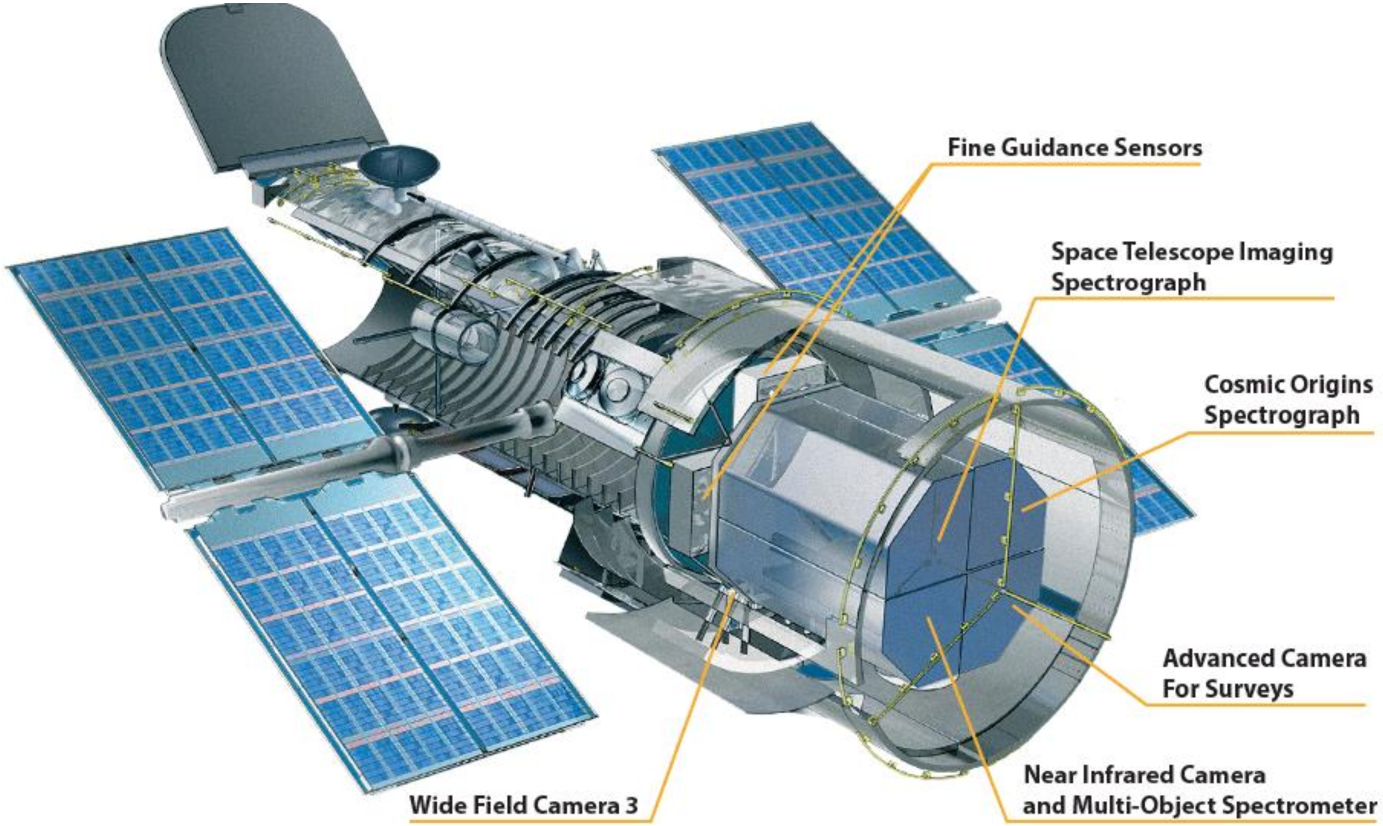
- Evolution of the universe
- Crack propagation in a material

Sky Observations

[Source: IDP]

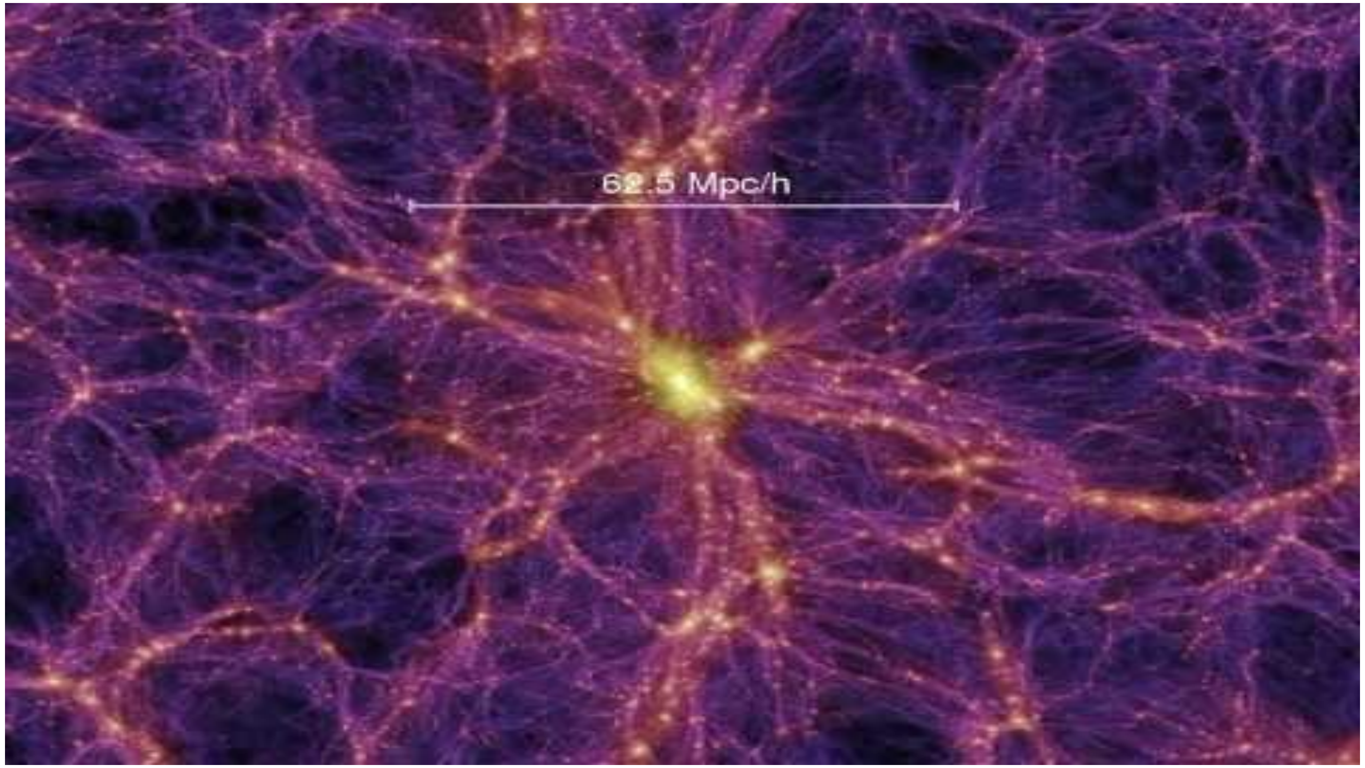
Data

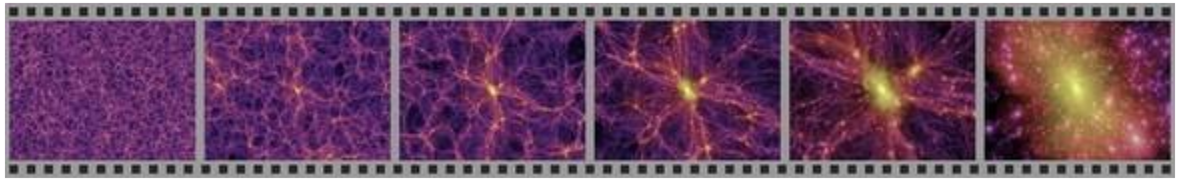
Rate at which science data is transmitted to the ground	1 megabit per second
Average amount of science data captured weekly	150 gigabits
Rate at which commands are sent to Hubble	32 kilobits per second
Amount of data that can be stored on the spacecraft	24 gigabits (or about 3 gigabytes)
Size of data archive	~340 terrabytes



[Source: NASA]

The Millennium Simulation (MPA)

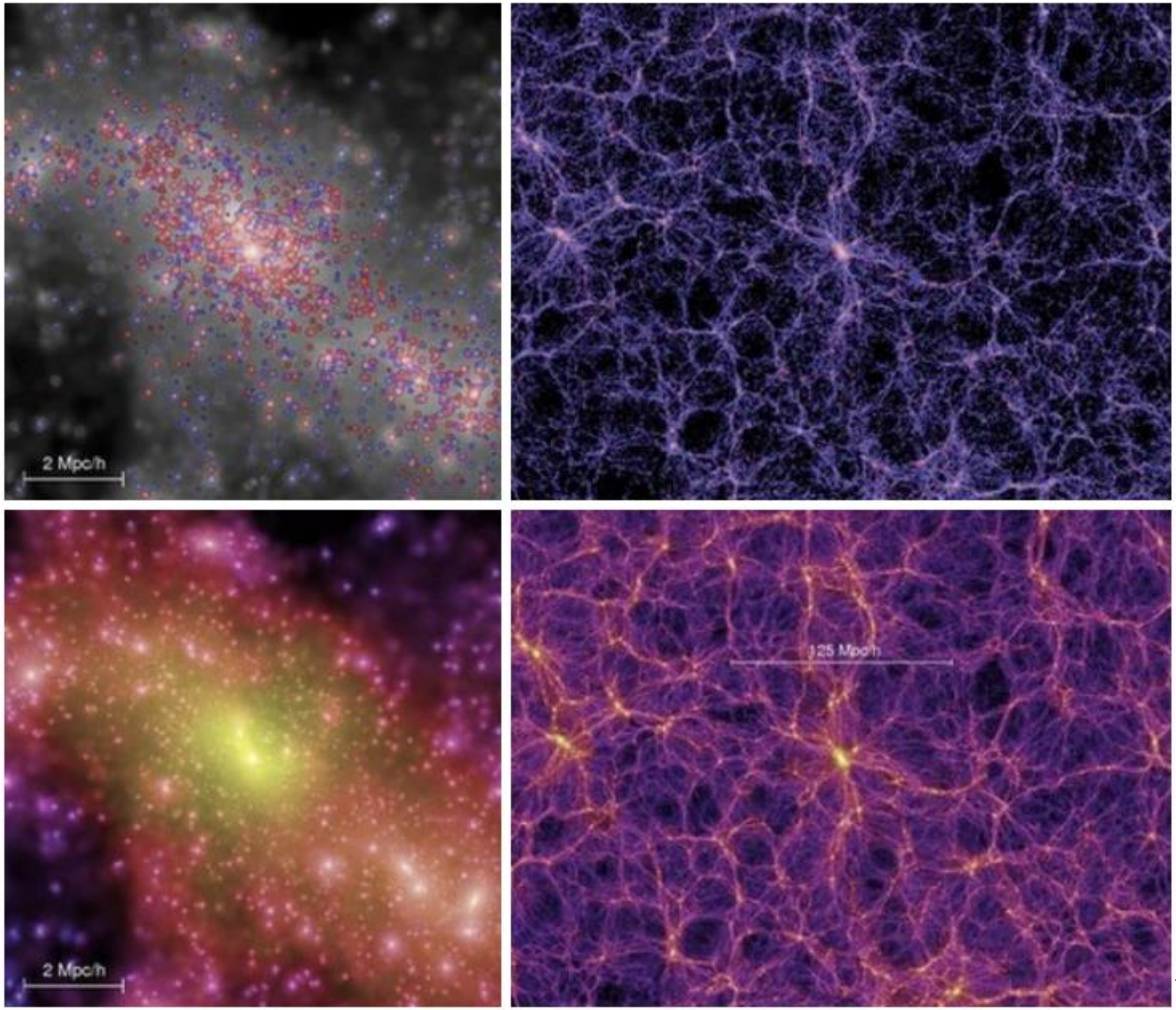




"The Millennium Run used more than 10 billion particles to trace the evolution of the matter distribution in a cubic region of the Universe over 2 billion light-years on a side. It kept busy the principal supercomputer at the Max Planck Society's Supercomputing Centre in Garching, Germany for more than a month."

https://wwwmpa.mpa-garching.mpg.de/galform/virgo/millennium/

Comparison

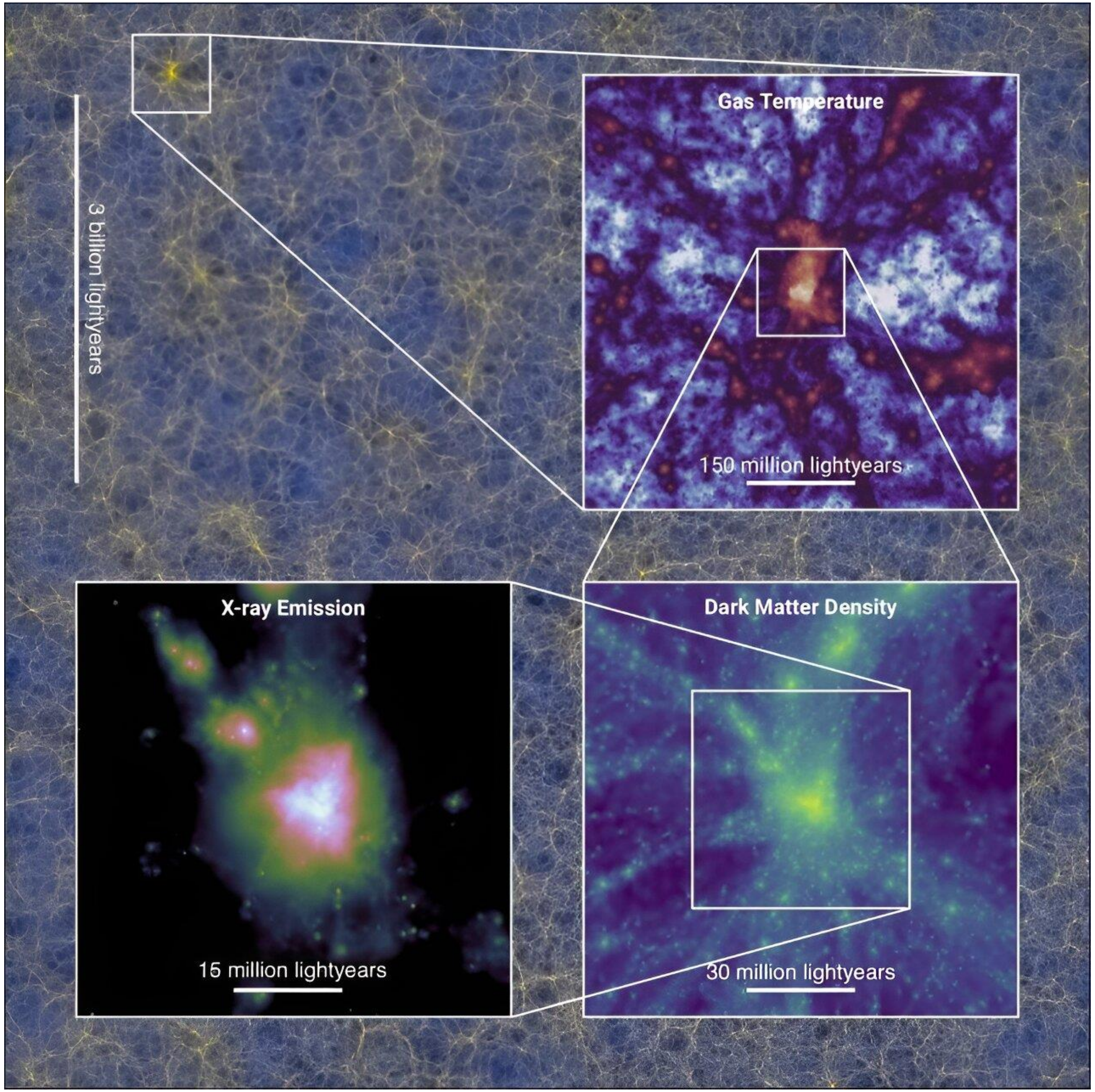


FLAMINGO

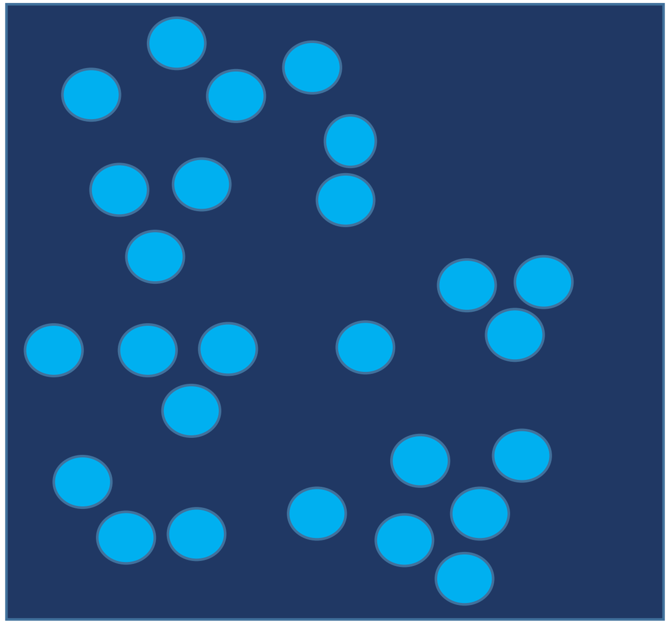
October 23, 2023

Astronomers carry out largest ever cosmological computer simulation

phys.org



Cosmological Simulations



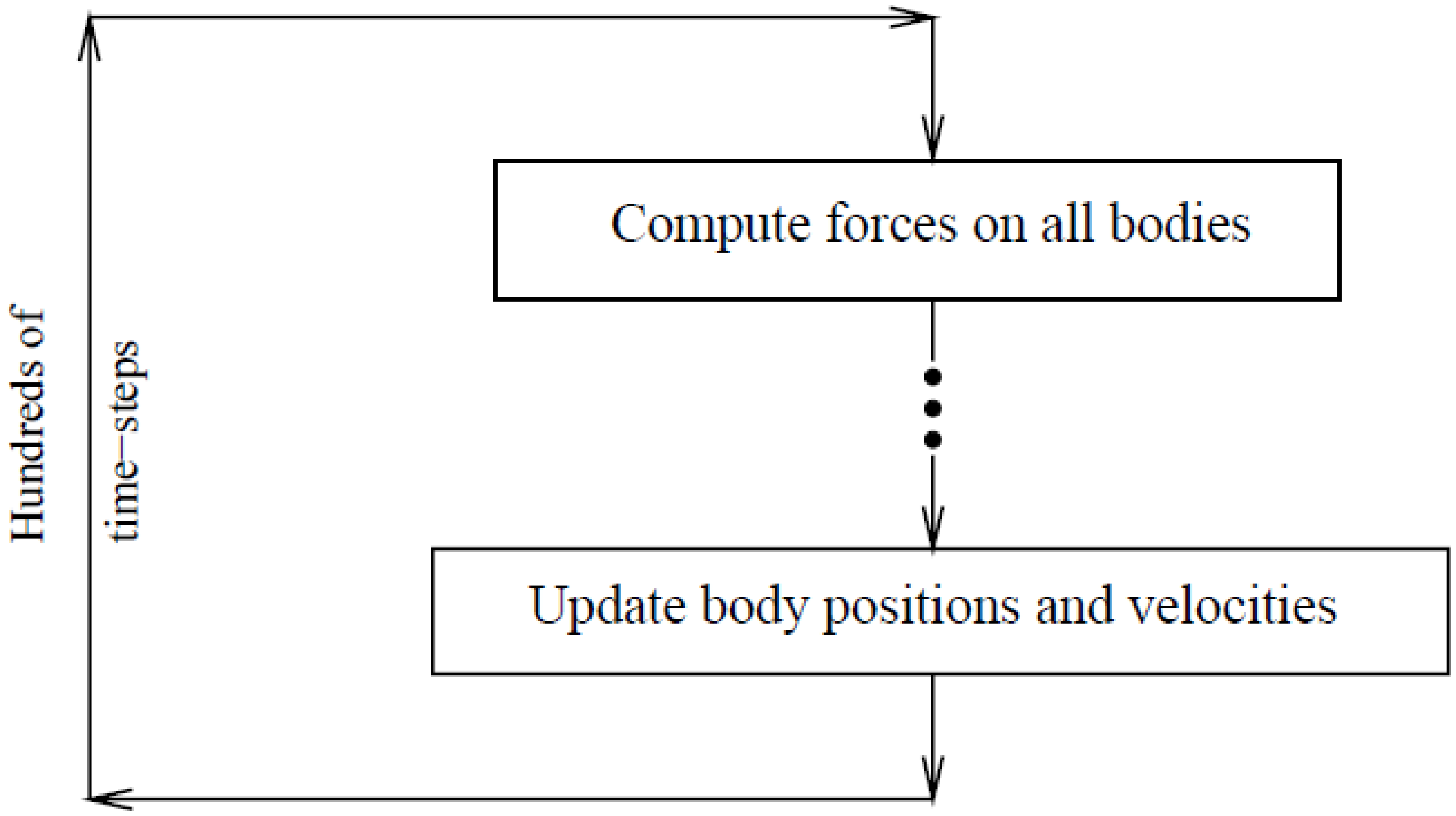
- Net force on every celestial object estimated
- Fields – three-dimensional position, velocity, acceleration, and mass
- Thousands of time steps
- Positions updated every time step, evolving system
- Spatial distribution is irregular

NEW YORK — All the stars, planets and galaxies that can be seen today make up just 4 percent of the universe. The other 96 percent is made of stuff astronomers can't see, detect or even comprehend.

These mysterious substances are called dark energy and dark matter. Astronomers infer their existence based on their gravitational influence on what little bits of the universe can be seen, but dark matter and energy themselves continue to elude all detection.

space.com

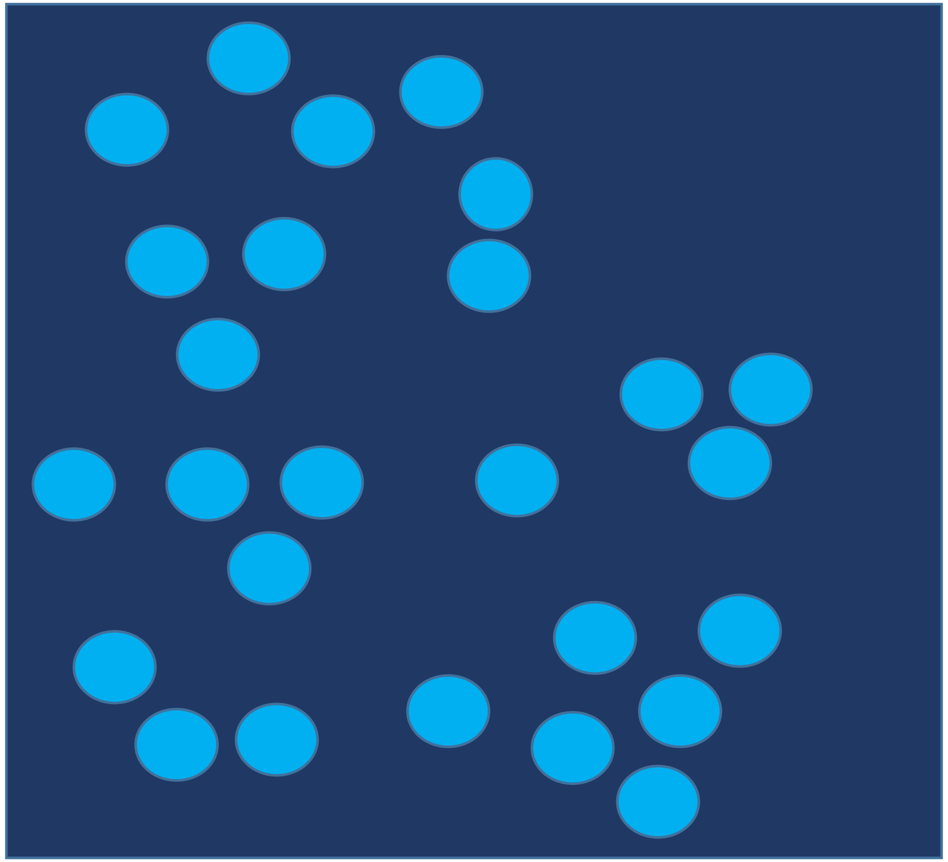
Iterative Computations



- Compute net force on every celestial object estimated by every other celestial object.
- Gravitational forces of each mass on others may be computed.
- $O(N^2)$ operations every time step.

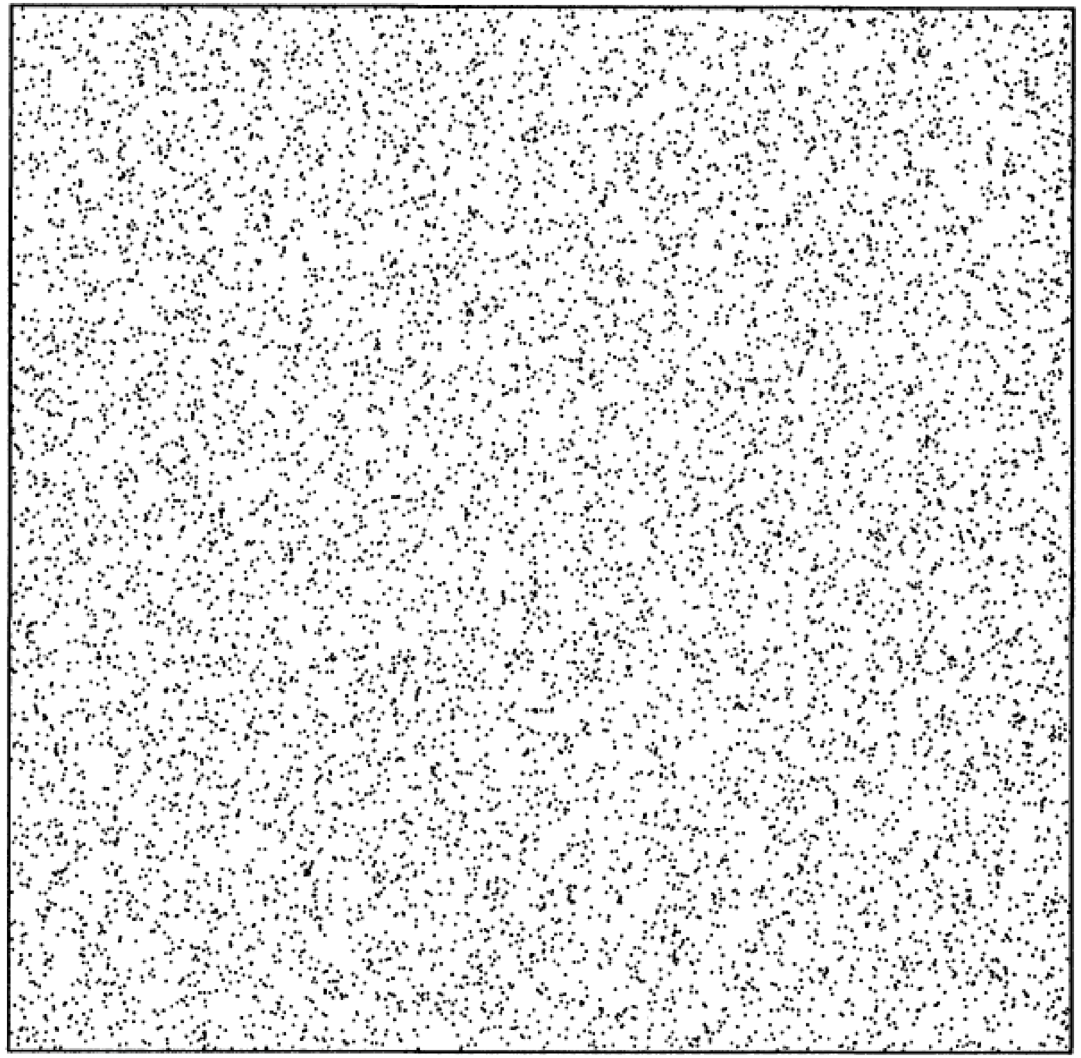
Load Balancing and Data Locality in Adaptive Hierarchical *N*-Body Methods: Barnes-Hut, Fast Multipole, and Radiosity
Singh J.P., Holt C., Totsuka T., Gupta A., Hennessy J.

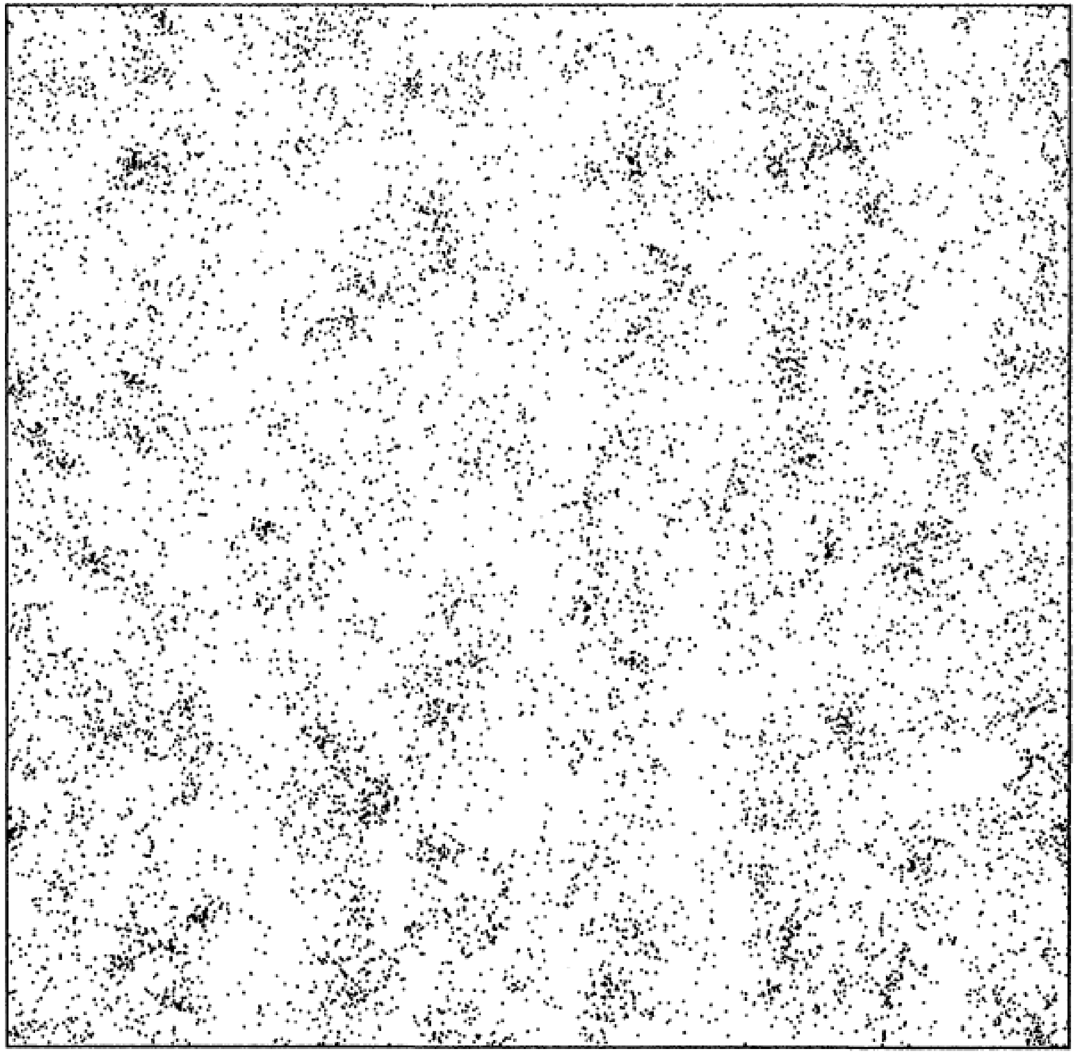
N-body Simulation



- *An Efficient Program for Many-body Simulation, Andrew W. Appel, SIAM Journal of SSC, 1985*
- O(N^2) -> O (N log N) operations every time step
- 8000 hours → 20 hours

A 10,000-galaxy Simulation

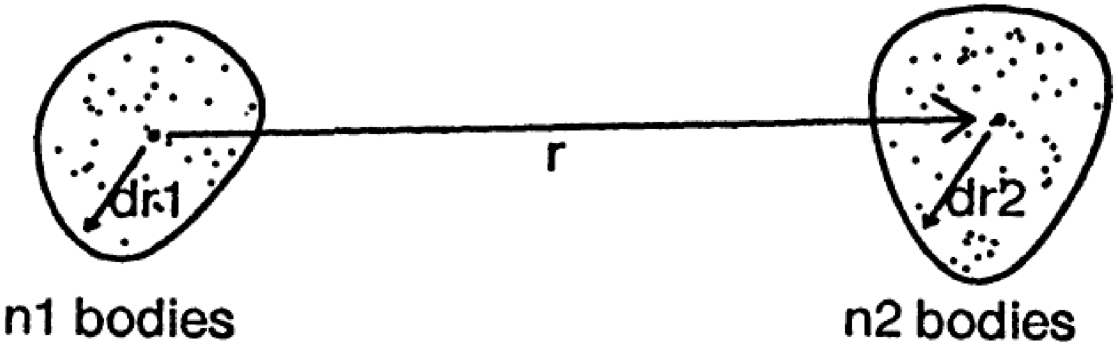




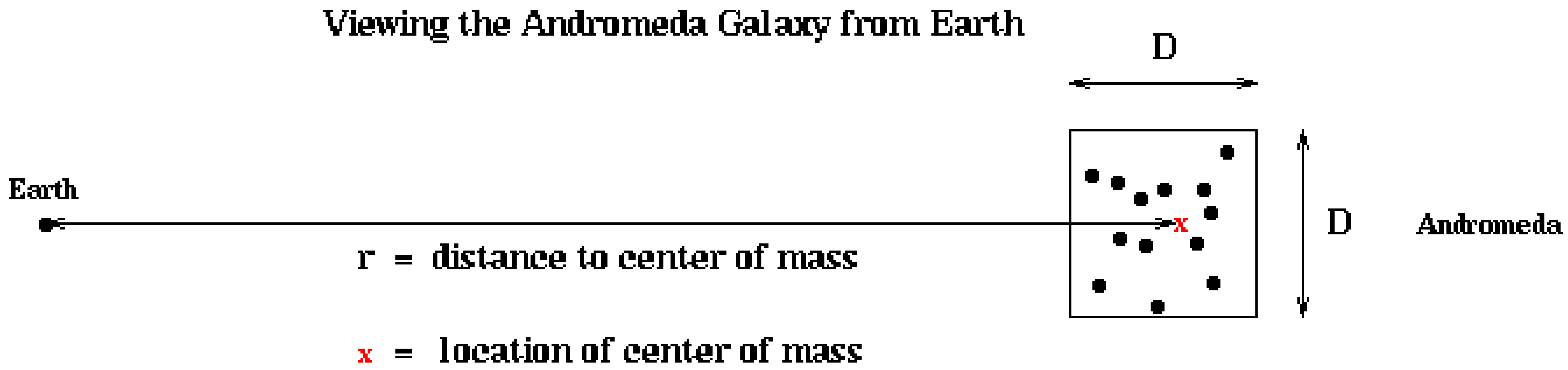
Barnes-Hut Method

A hierarchical O(N log N) force-calculation algorithm, Barnes and Hut, Nature 1986

- Effect of a large group of bodies may be approximated by a single equivalent body
- Center of mass approximation for distant bodies



Example of Approximation



Credit: James Demmel lecture notes

Load Balancing and Data Locality in Adaptive Hierarchical N-body Methods: Barnes-Hut, Fast Multipole, and Radiosity – Singh et al.

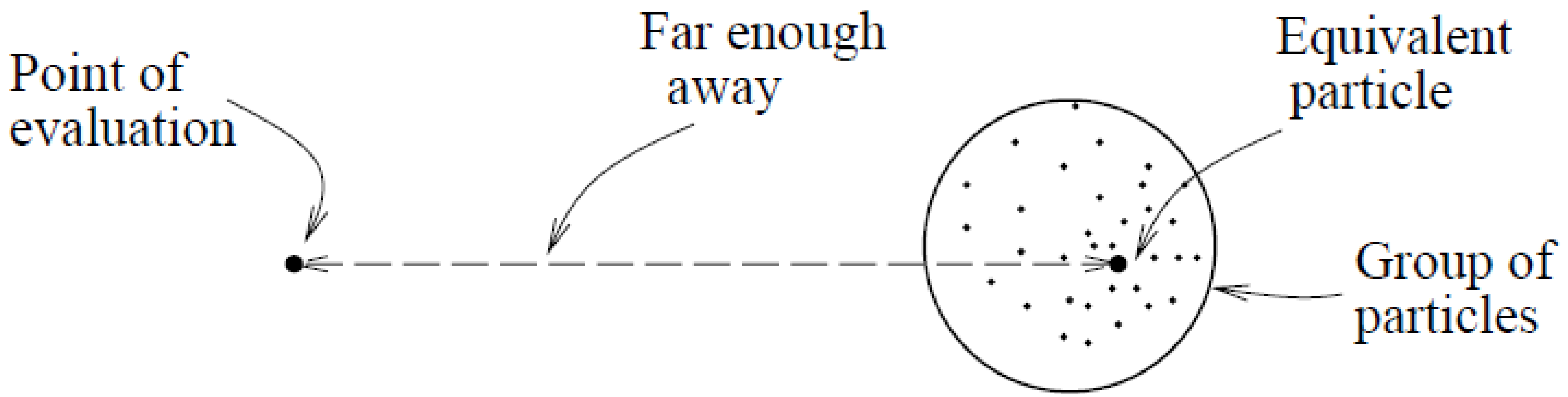


Figure 1: Approximation of a group of particles by a single equivalent particle.

Approximation

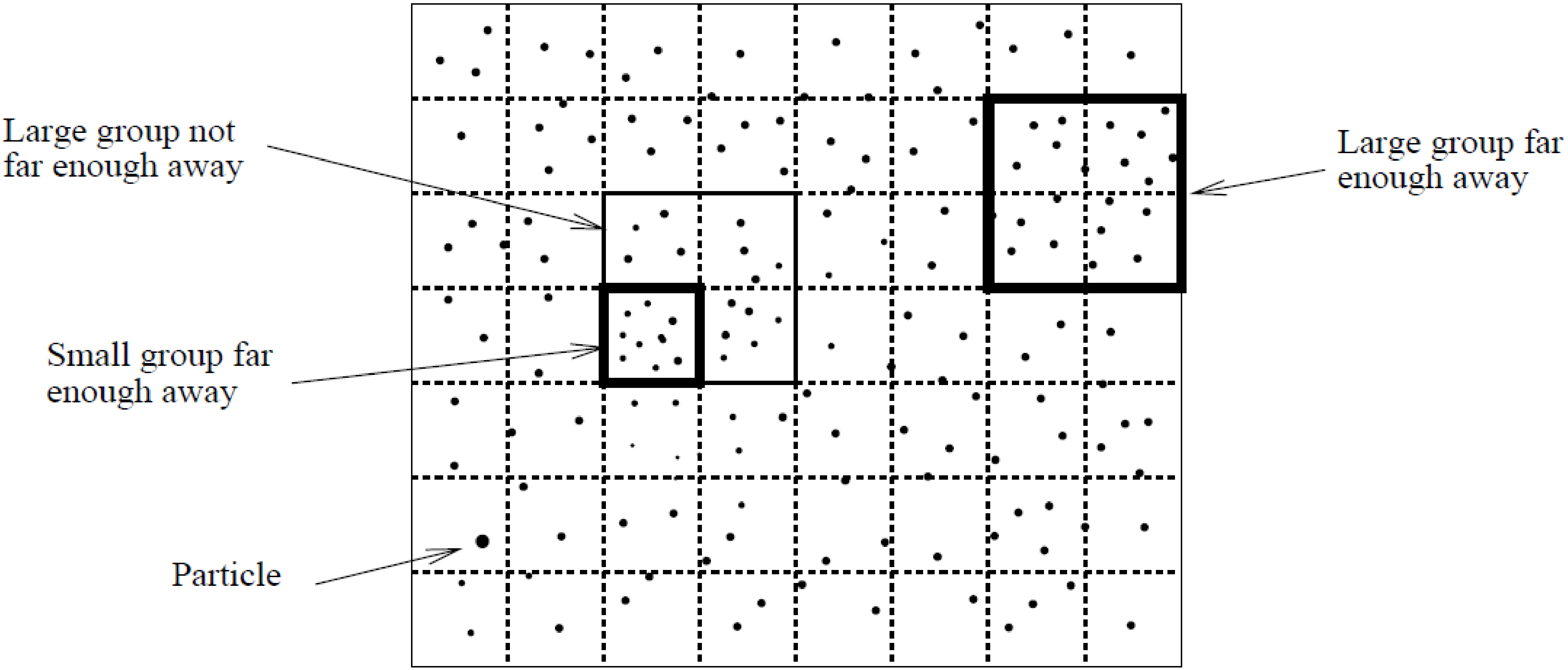
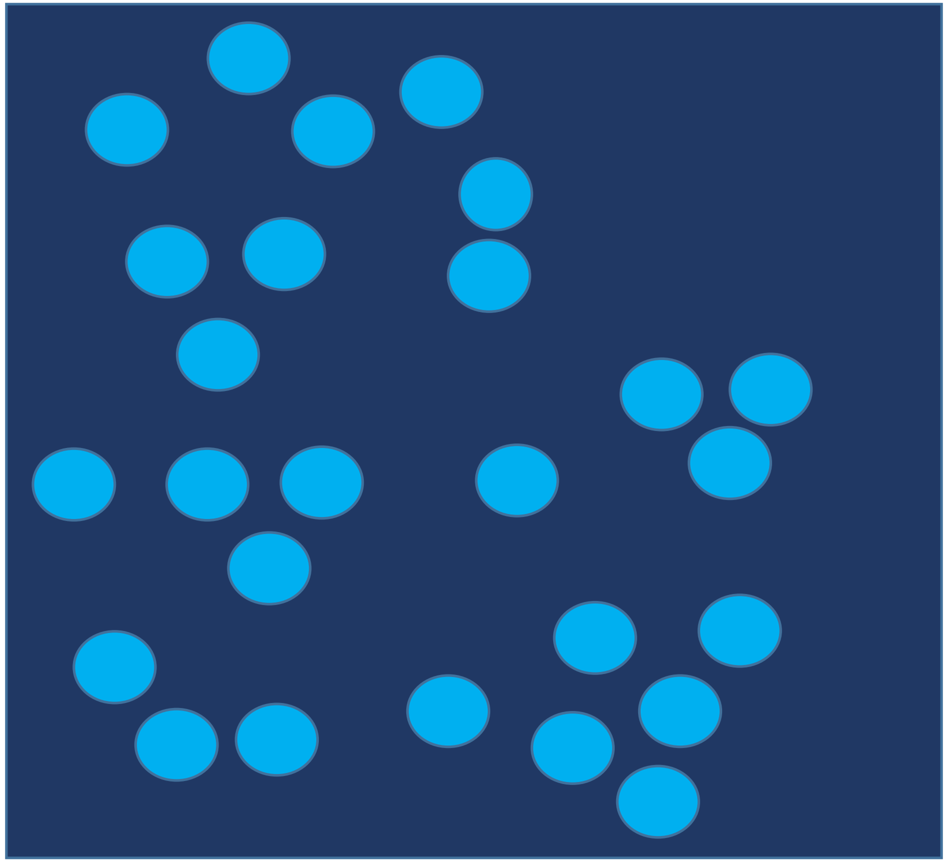


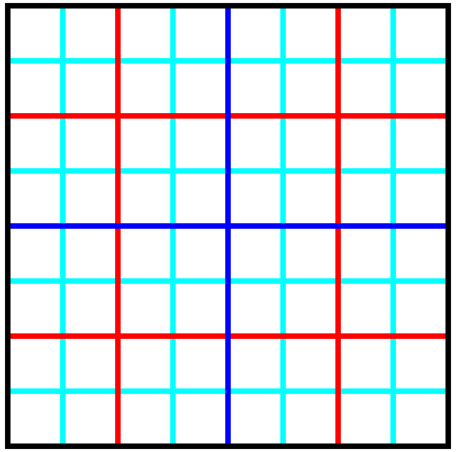
Figure 2: Hierarchical approximation of cells by equivalent particles.

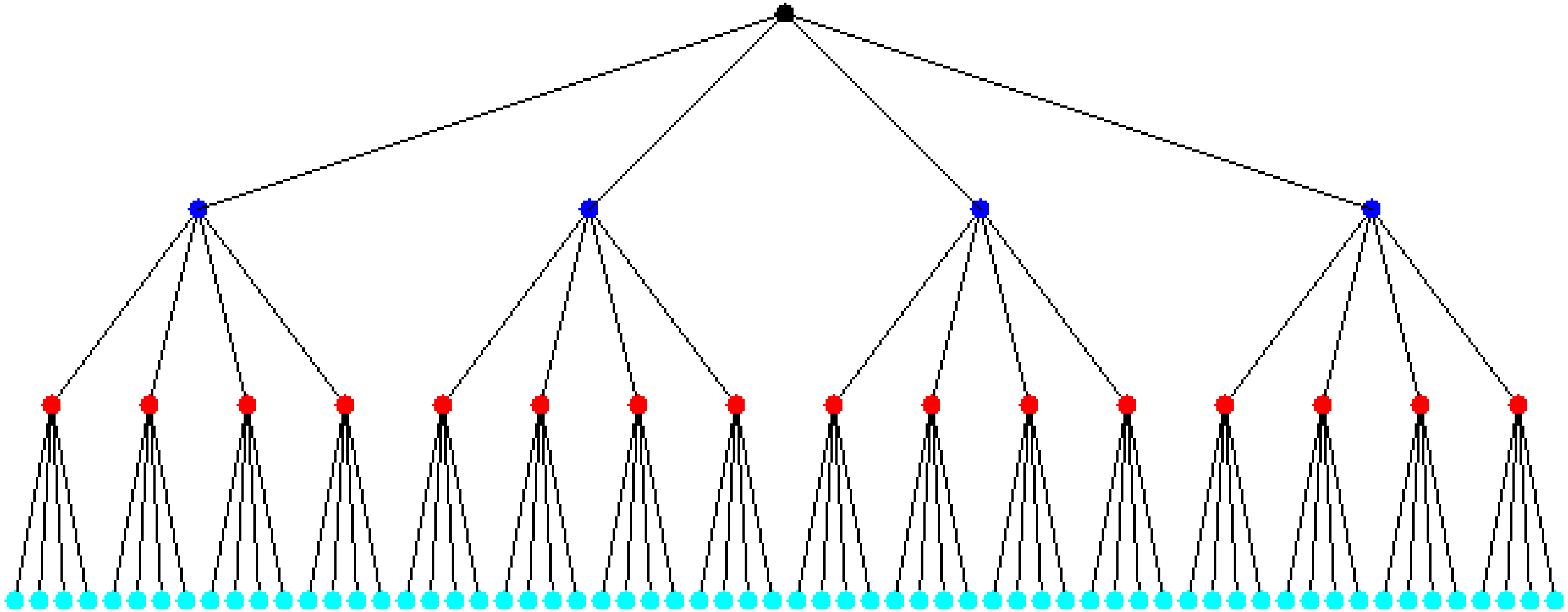
N-body Simulation Domain Decomposition



- Sub-divide the space into 8 sub-regions
- Sub-divide each region based on the number of bodies

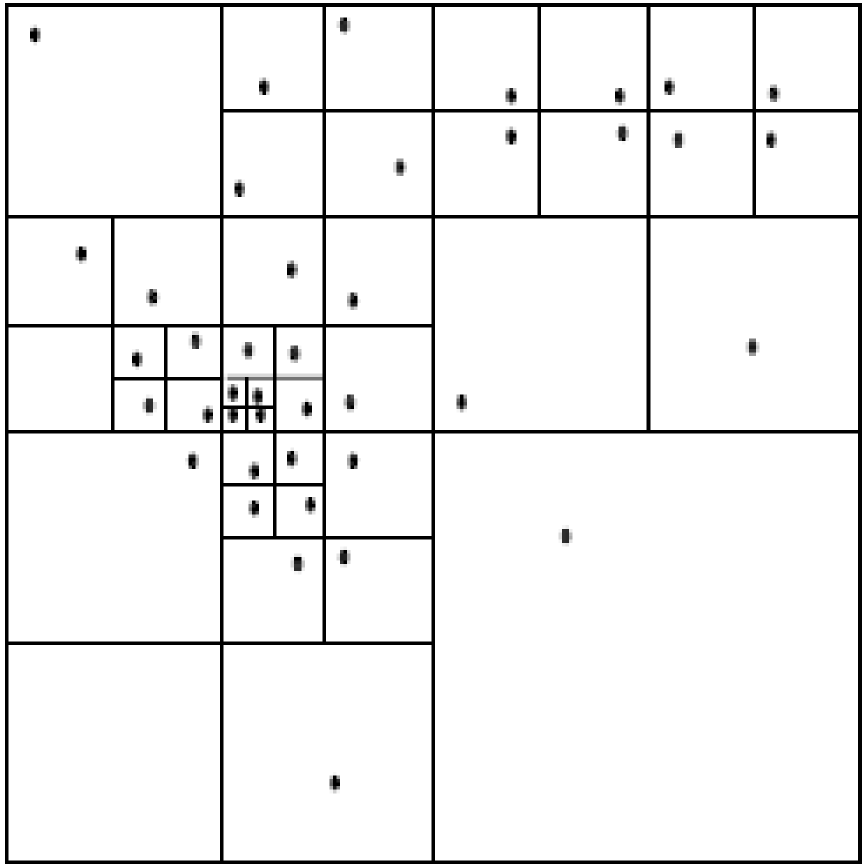
Quadtree

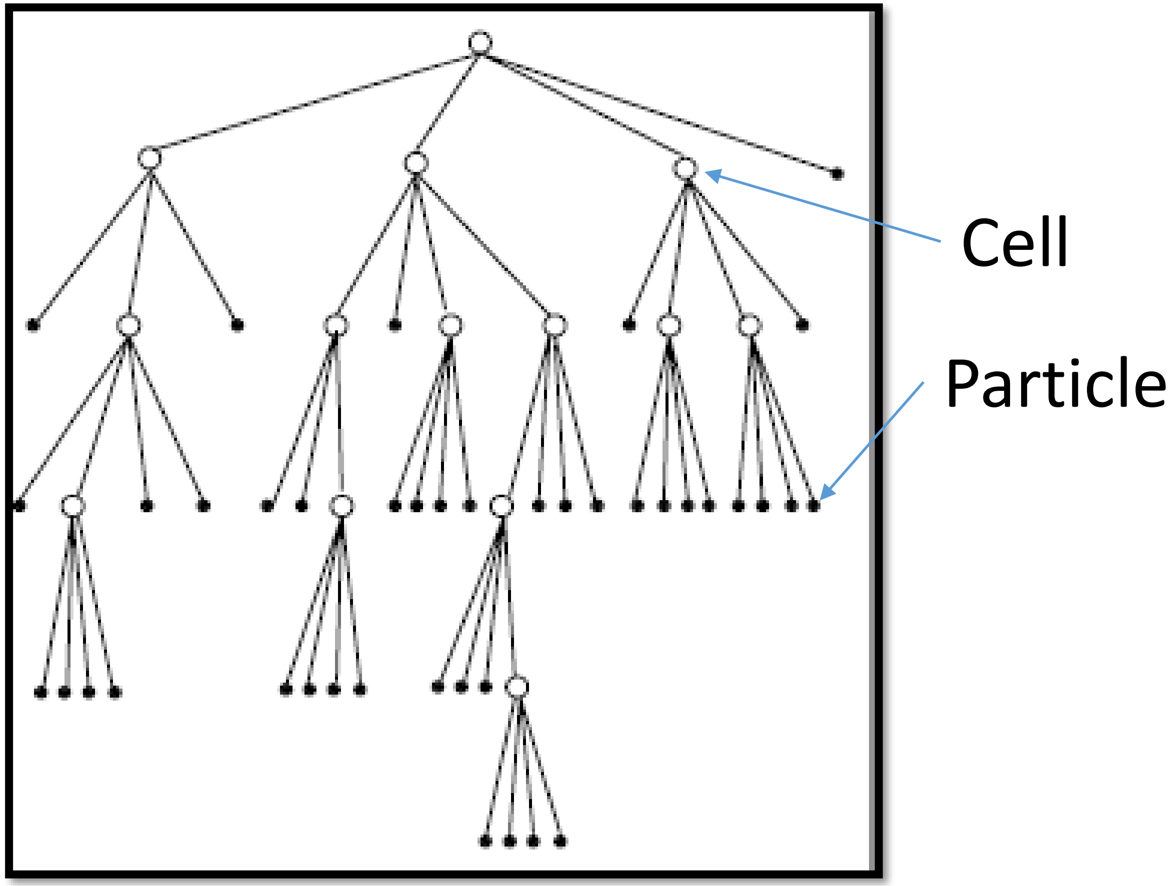




Credit: James Demmel lecture notes

Spatial Domain



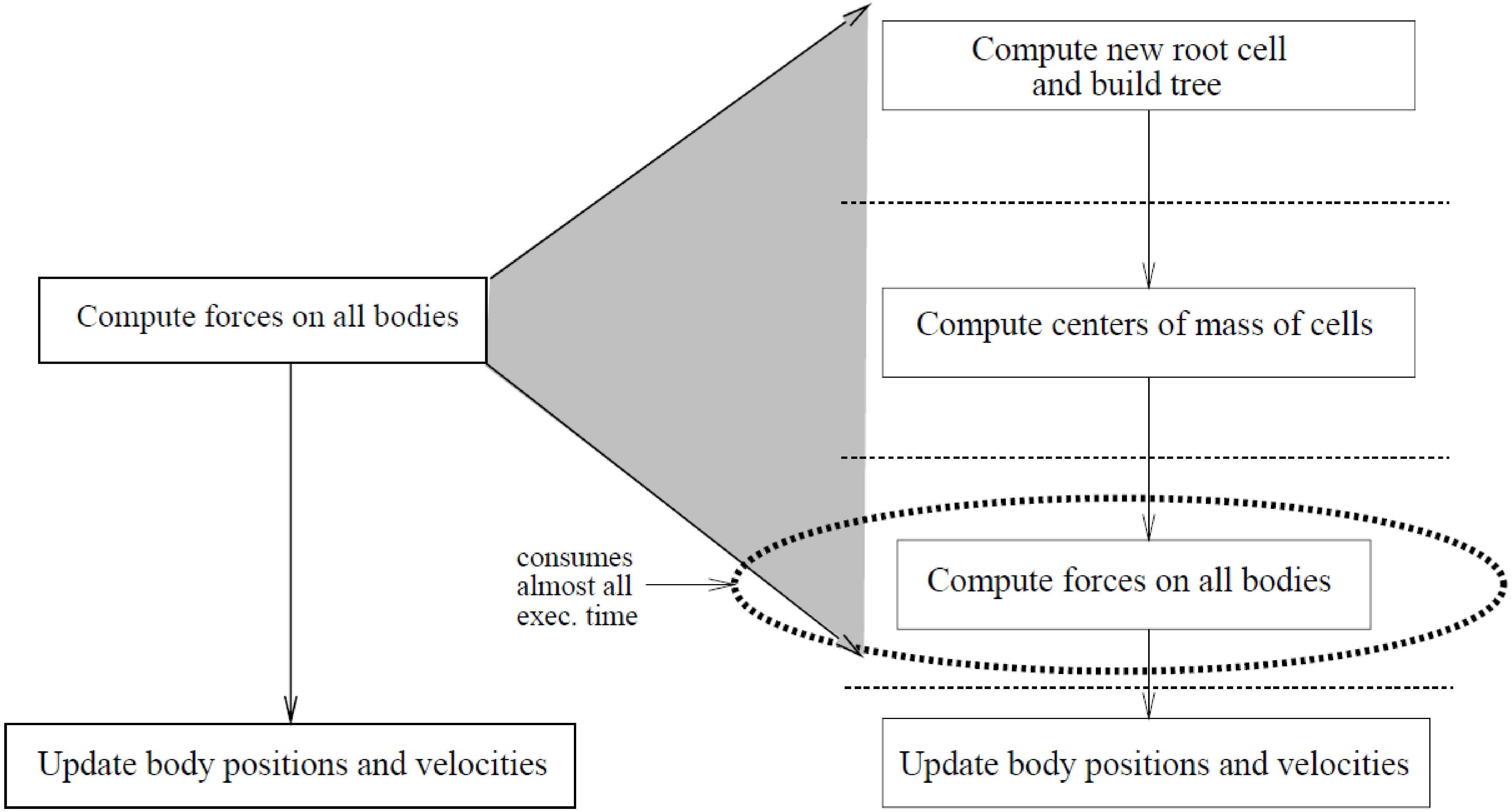


Source: Culler et al. and “Load Balancing and Data Locality in Adaptive Hierarchical *N*-Body Methods: Barnes-Hut, Fast Multipole, and Radiosity”, Singh J.P., Holt C., Totsuka T., Gupta A., Hennessy J.

Algorithm

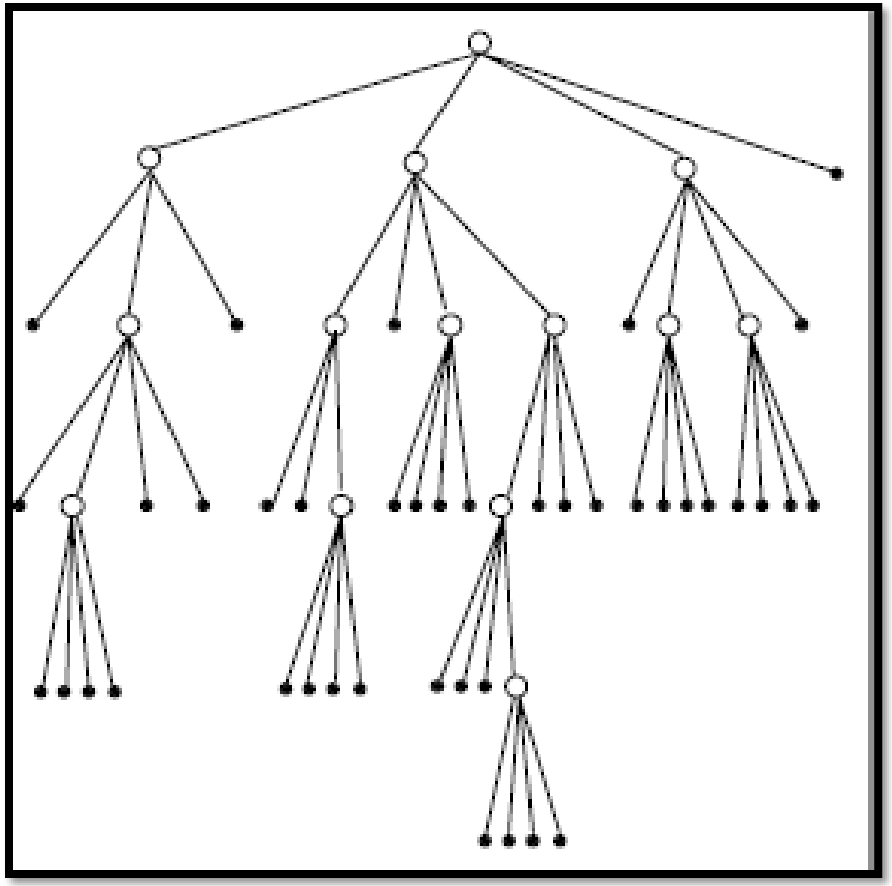
- Build tree
- Compute center of mass (CoM) and total mass
- Compute forces
- Update particle properties

A Time Step



Compute new root cell and build tree
Compute forces on all bodies
Compute centers of mass of cells
consumes almost all exec. time
Compute forces on all bodies
Update body positions and velocities
Update body positions and velocities

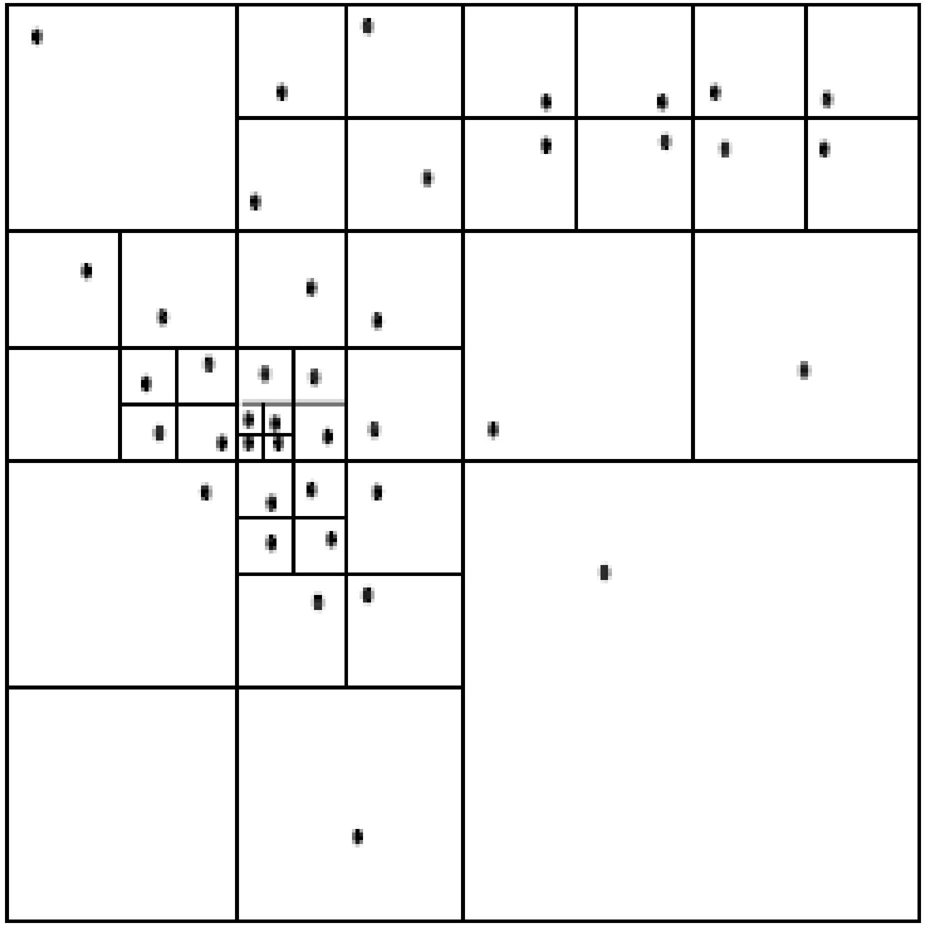
N-body Simulation – Force Computations



- Traverse the oct-tree for every mass
- If the CoM is far away, approximate by a single mass
  - Length of cell / distance < threshold
  - Sub-tree need not be expanded
- O(N log N) force computations
- Arrays per particle
  - 3D position, velocity, acceleration, mass

N-body Simulation

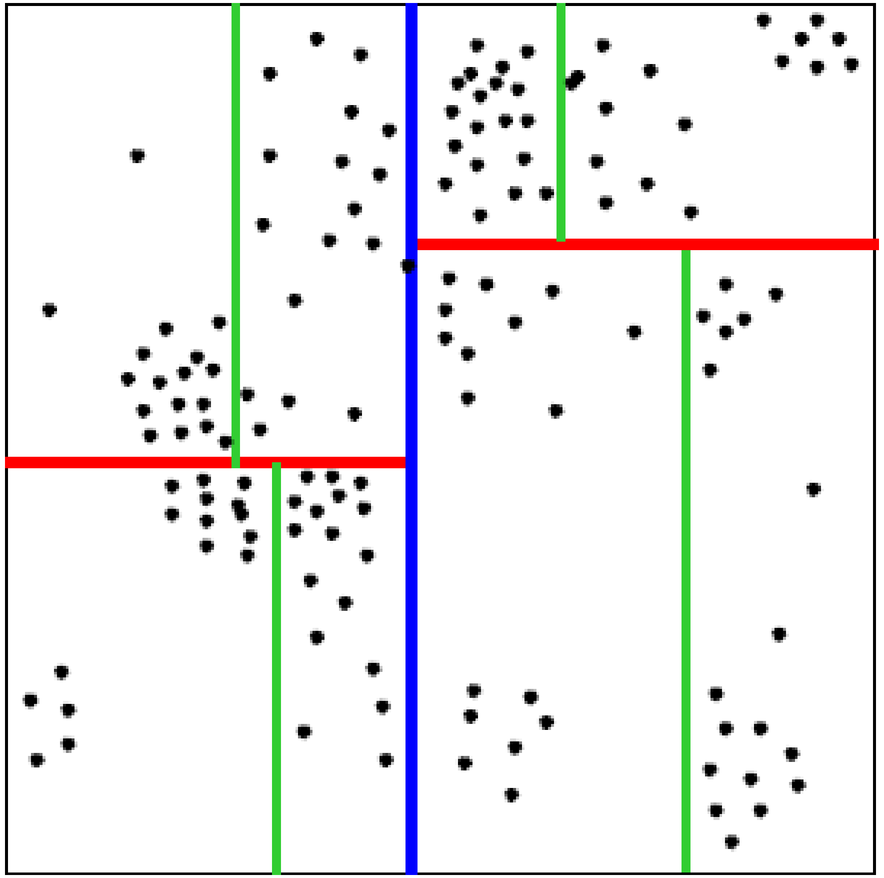
We achieved a good domain decomposition but …



Performance considerations

- Positions change across time steps
  - Dynamic decomposition
  - Repartitioning almost every time step
  - Rebuild the tree every time step
- Irregular communication pattern
  - Logical process neighbors may own spatially dis-contiguous chunks
- Synchronization between steps
- Partitions are not contiguous in space

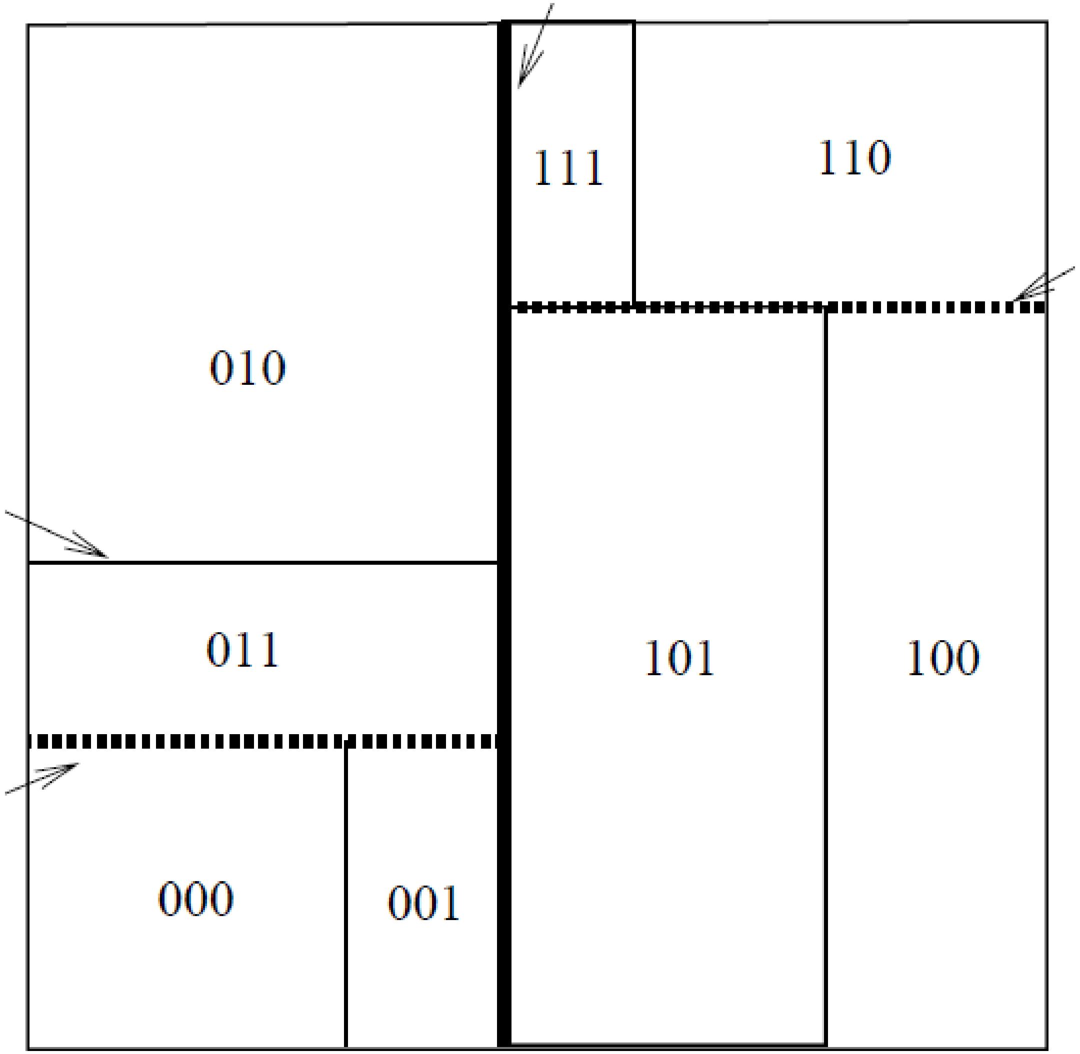
Orthogonal Recursive Bisection (ORB)



Source: Culler et al.

- Preserves physical locality
- Parallel median finder
- Alternate division along the dimensions
- Binary tree for P processes
  - Depth: log P

ORB



Scalability

"The scalability of a parallel system is a measure of its capacity to increase speedup in proportion to the number of processing elements." – Introduction to Parallel Computing

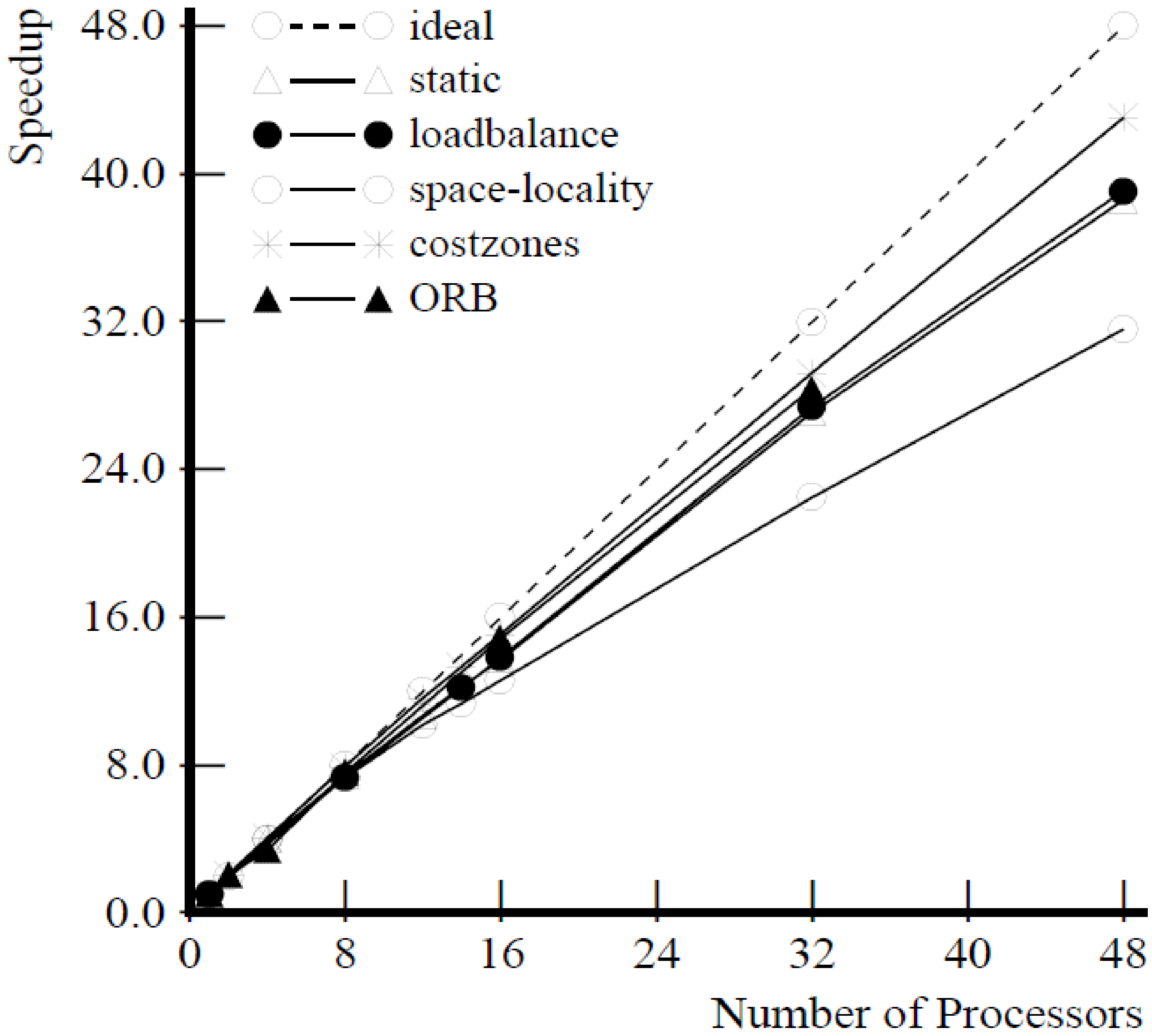
Strong scaling

- Fixed problem size
- Increase number of processes
- Efficiency decreases, in general – why?

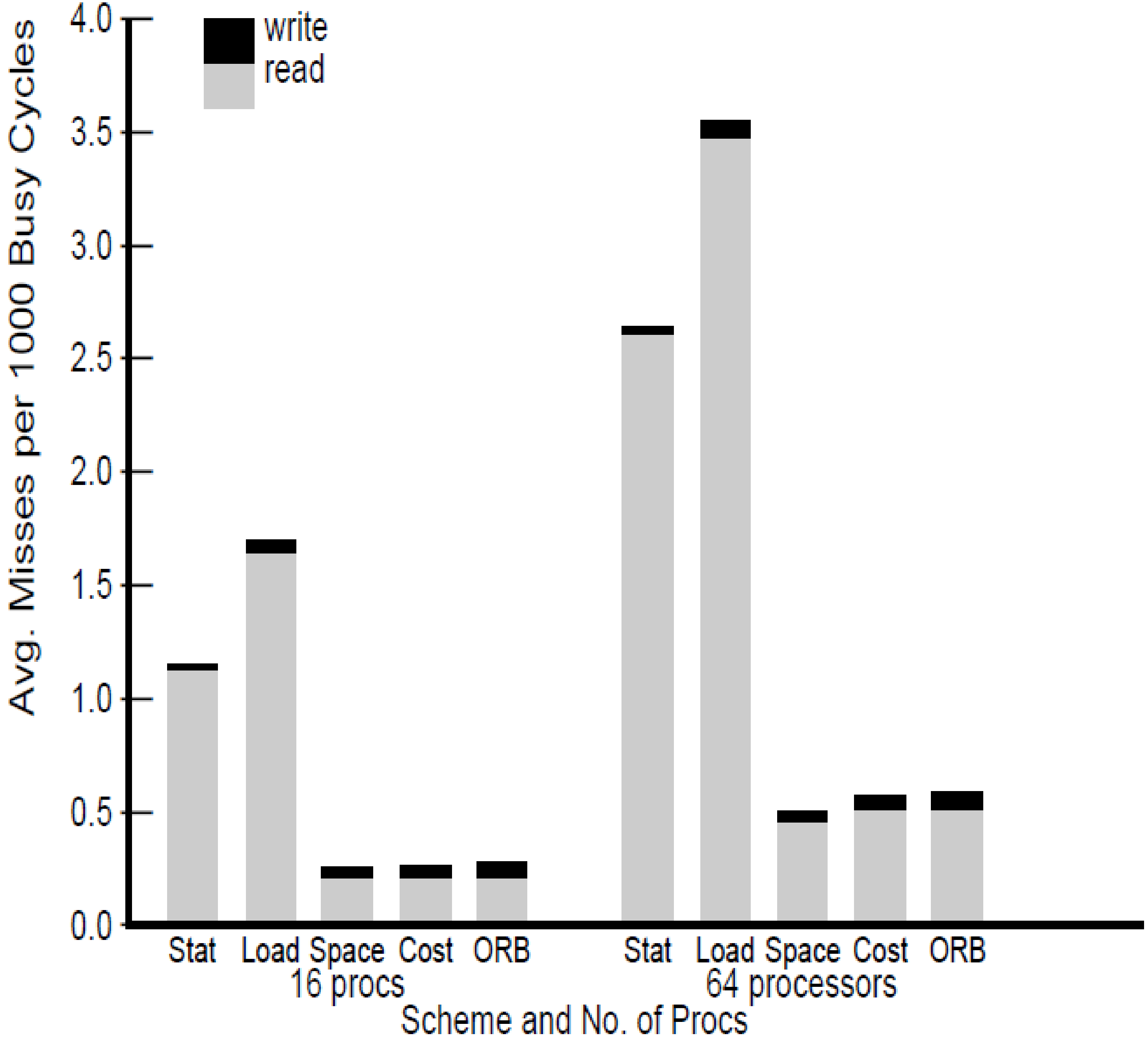
Weak scaling

- Fixed problem size per process
- Increase number of processes
- Increase problem size

Results

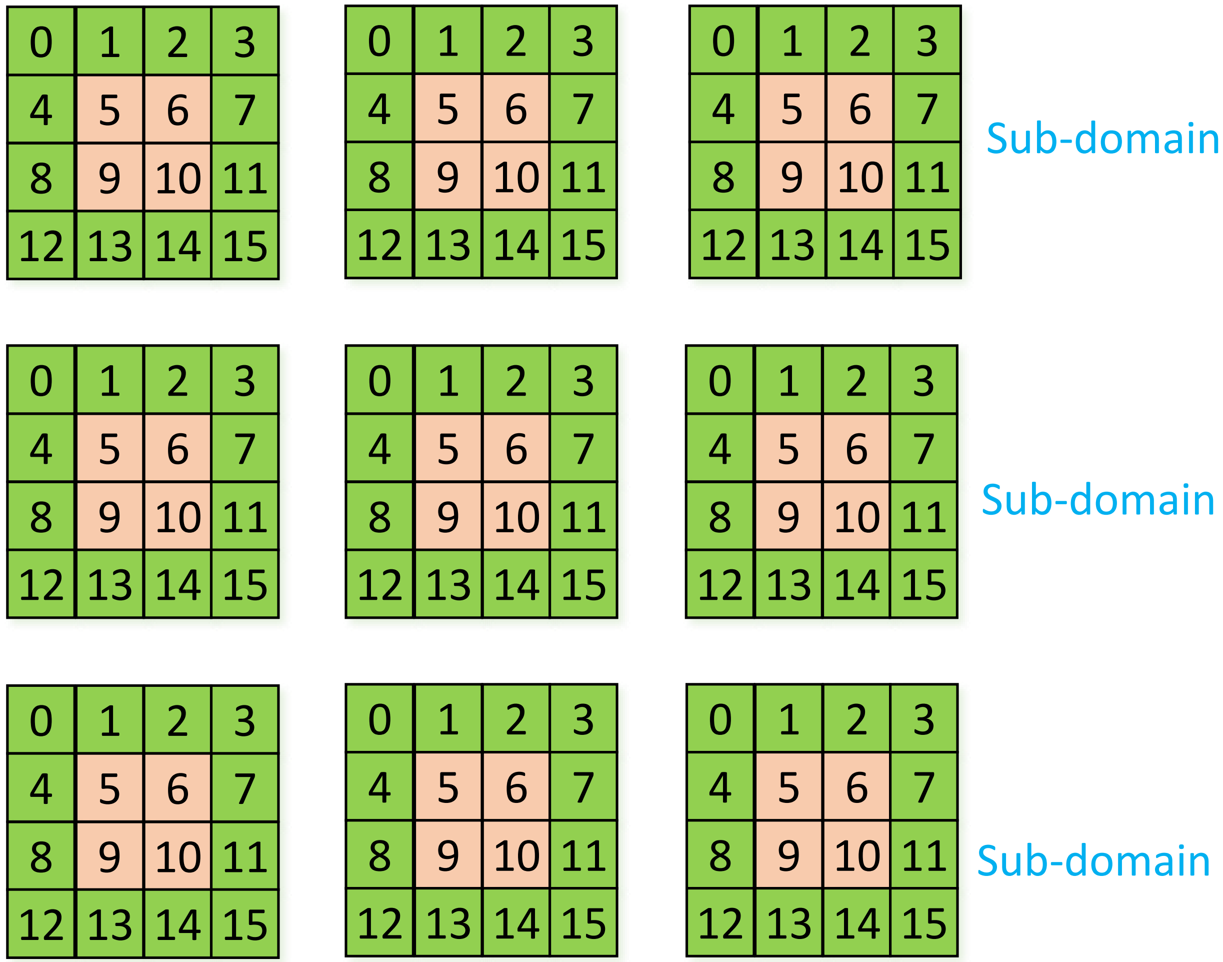

Speedup
48.0
40.0
32.0
24.0
16.0
8.0
0.0
ideal
static
loadbalance
space-locality
costzones
ORB
0
8
16
24
32
40
48
Number of Processors

Results



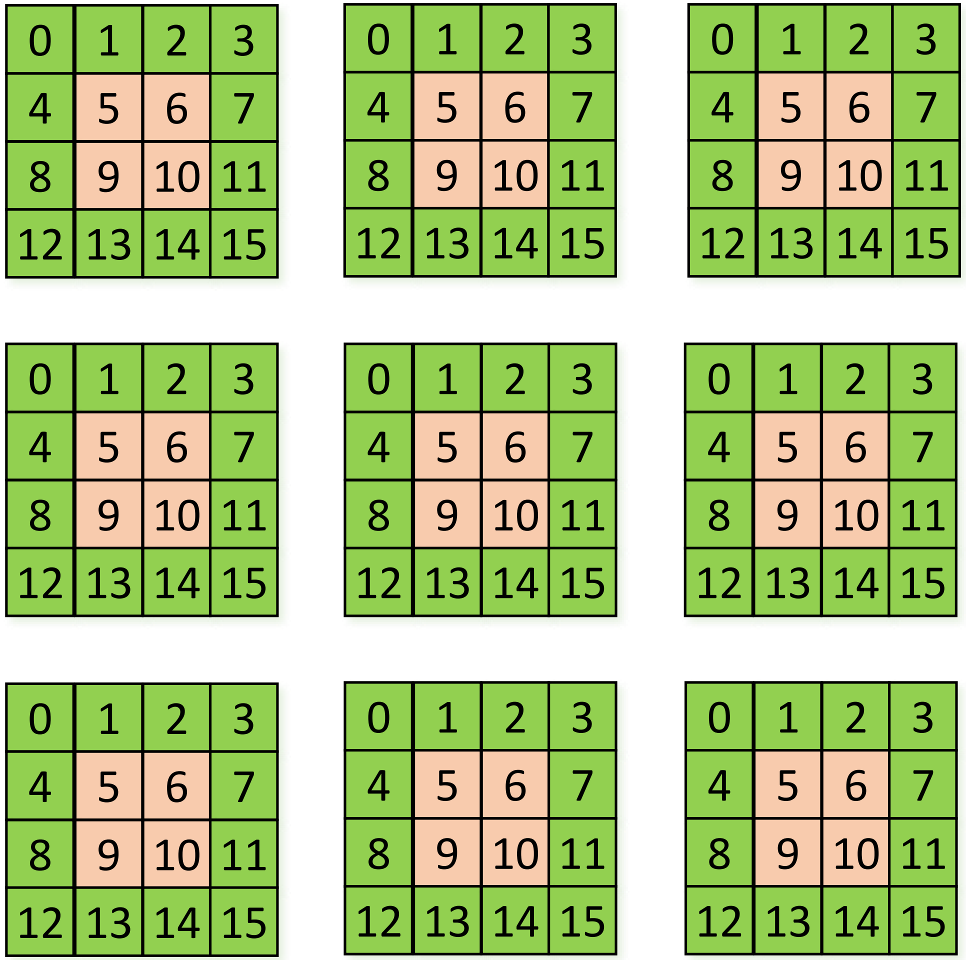
write
read
Avg. Misses per 1000 Busy Cycles
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
Stat
Load
Space
Cost
ORB
16 procs
Stat
Load
Space
Cost
ORB
64 processors
Scheme and No. of Procs

Halo Exchange



What are the internode communications for 3 nodes (9 processes)

Assignment 2 – Hierarchical Halo Exchange



- Communicate halo regions grouped by node
- Use intra-node sub-communicators
- Reduce the number of inter-node communications