Transient simulations of thermodynamics in Data Centers

Using the Lattice Boltzmann Method on Graphics Processing Units for Predicting Thermal Flow in Data Centers

Johannes Sjölund

7th December 2018

Table of Contents

Data Centers

- A facility used to host computer server and networking equipment.
- Servers are mounted in racks.
- Computer Room Air Conditioners (CRACs) cool equipment using heat exchangers.

Figure 1: Server racks in data center module POD 2 at RISE SICS North.

Data Centers

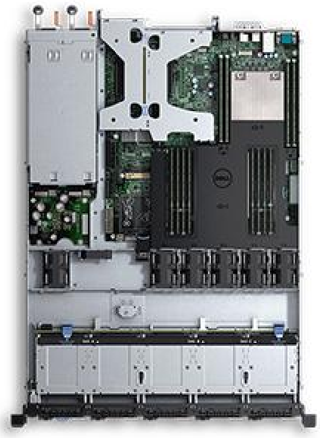


Figure 2: Dell R430 blade server with six internal fans.

Figure 3: Computer Room Air Conditioner, SEE Cooler HDZ-3.

Goals

- Use The Lattice Boltzmann Method (LBM) to simulate thermal conditions in a data center.
- Validate different LB methods and CFD models with sensor data from data center.
- Ability to use LBM as a training/testing system for air conditioning control systems.

The Lattice Boltzmann Method (LBM)

The LBM models the behavior of a group of particles as distribution functions in a uniform grid.

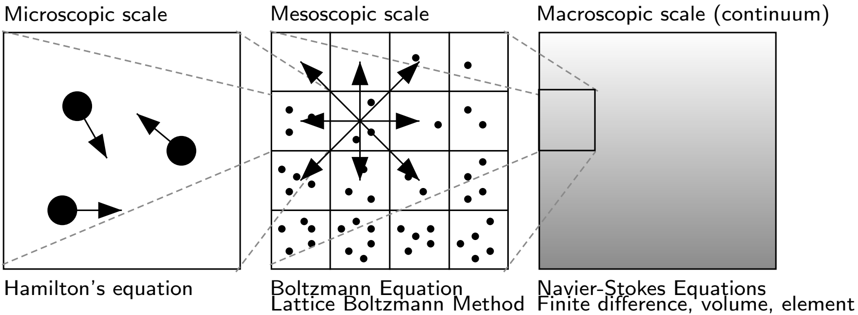


Figure 4: Techniques of simulations for different scales of fluid representations.

LBM: The Discrete Lattice Boltzmann Equation

The evolution of the system is described by the Discrete Lattice Boltzmann Equation

$$f_i(\vec{x} + \vec{e_i}\Delta t, t + \Delta t) = f_i(\vec{x}, t) + \Gamma(f_i(\vec{x}, t)). \tag{1}$$

Each distribution function f_i is associated with a direction. Γ is called a collision operator and can be implemented in different ways. Current implementation uses the Bhatnagar–Gross–Krook (BGK) method.

LBM: Streaming Step

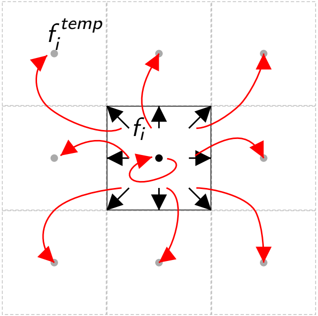


Figure 5: Lattice streaming step, representing advection in a fluid. All functions f_i are copied to the neighboring f_i^{temp} in parallel.

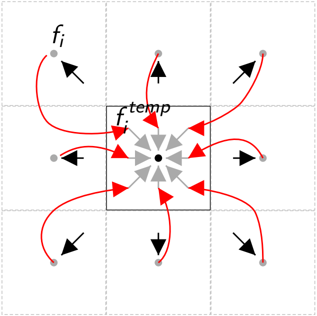


Figure 6: Also in the streaming step, the current site is filled with new distributions from the neighboring sites.

LBM: Collision Step

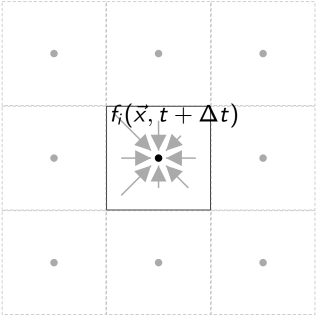


Figure 7: Collision step, representing diffusion in a fluid. Particles from adjacent sites collide locally in the current site (see BGK).

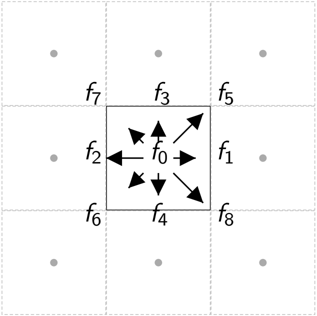


Figure 8: During the collide step the particle populations are redistributed. Both mass and momentum is conserved.

RAFSINE

- Originally written by Nicolas Delbosc during his Ph.D study in the School of Mechanical Engineering at the University of Leeds, England.
- Implements LBM (BGK model) in C++ with streaming-, collision- and boundary-steps accelerated by Nvidia CUDA.
- Simulates fluid behavior in real time or faster depending on domain size.
- OpenGL visualization of system evolution.

Figure 9: Nvidia GTX 1080 Ti.

RAFSINE

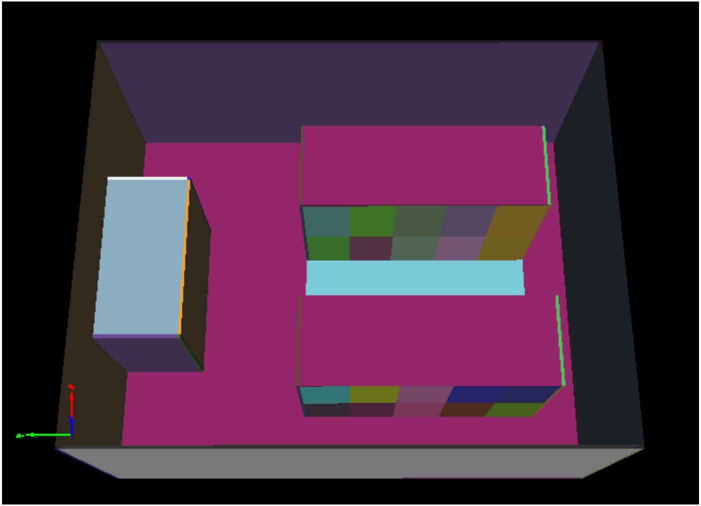


Figure 10: Geometry of an example data center.

RAFSINE

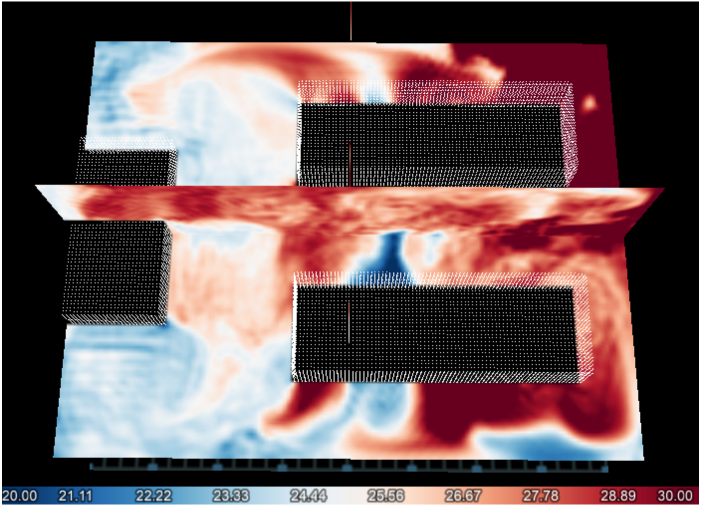


Figure 11: Heat flow in an example data center.

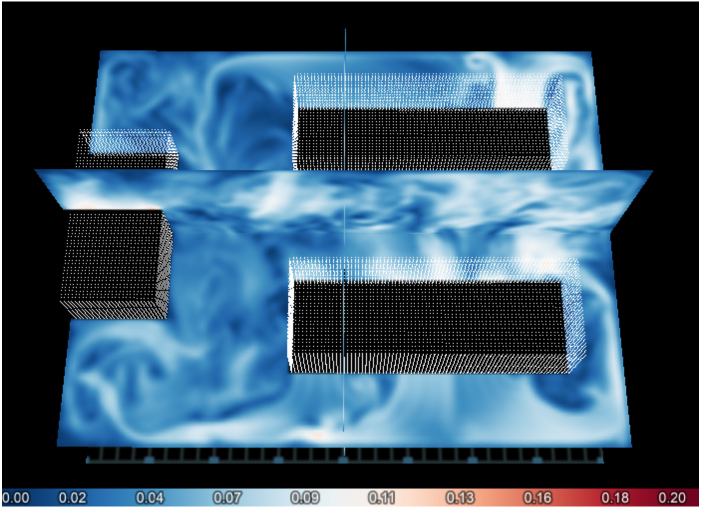


Figure 12: Air velocity in an example data center.

RAFSINE: Multiple GPUs

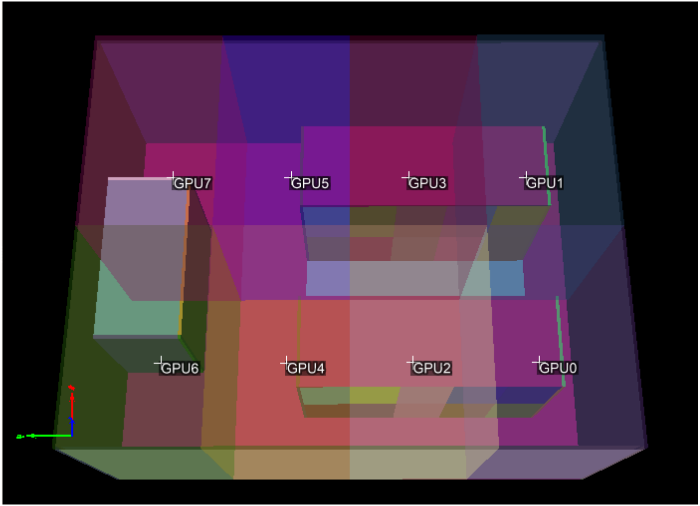


Figure 13: Domain decomposition of the data center.

RAFSINE: Multiple GPUs

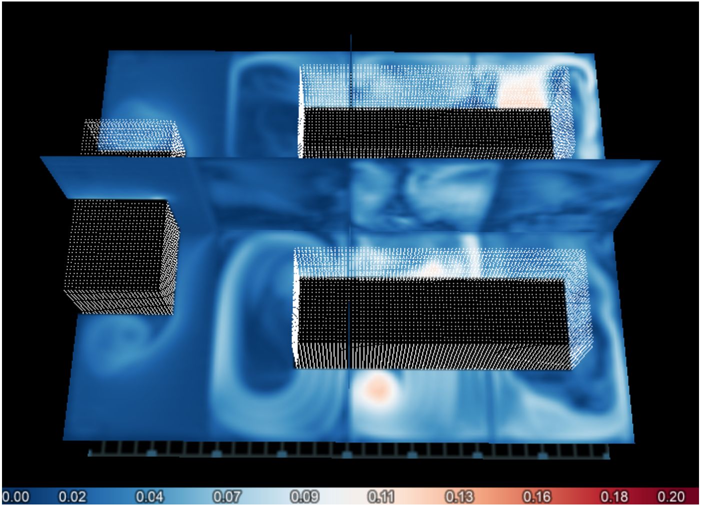


Figure 14: Air velocity when decompositions are simulated separately.

LBM: Halo Exchange

Each GPU calculates the LBM evolution of a group of grid cells (lattice sites). After each evolution step, the state of the border sites (halos) are copied to buffers on neighbouring GPUs. These states are read in the next evolution.

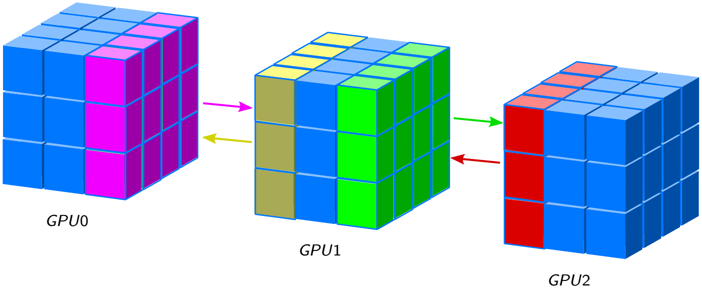


Figure 15: D3Q19 lattice site.

Number of halos to copy depends on the type of lattice site.

LBM: Halo Exchange

LBM discretizes the domain into a uniform grid of cells called lattice sites.

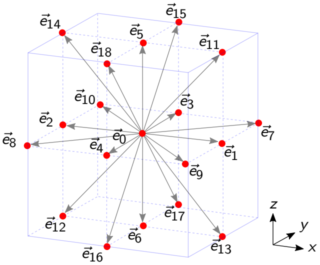
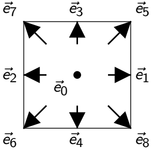


Figure 16: D3Q19 lattice site.



$\vec{e}_0 = (\ 0,\ \ 0)$	$w_0 = 4/9$
$\vec{e}_1 = (\ 1,\ \ 0)$	$w_1 = 1/9$
$\vec{e}_2 = (-1,\ \ 0)$	$w_2 = 1/9$
$\vec{e}_3 = (\ 0,\ \ 1)$	$w_3 = 1/9$
$\vec{e}_4 = (\ 0, -1)$	$w_4 = 1/9$
$\vec{e}_5 = (\ 1,\ \ 1)$	$w_5 = 1/36$
$\vec{e}_6 = (-1, -1)$	$w_6 = 1/36$
$\vec{e}_7 = (-1,\ \ 1)$	$w_7 = 1/36$
$\vec{e}_8 = (\ 1, -1)$	$w_8 = 1/36$

Figure 17: D2Q9 lattice site. Direction vectors $\vec{e}_i$ are lattice velocities, with their corresponding weight w_i.

LBM: Halo Exchange

- D3Q7 must copy **sides** to neighbours.
- D3Q19 must copy **sides** and **edges**.
- D3Q27 must copy **sides**, **edges** and **corners**.

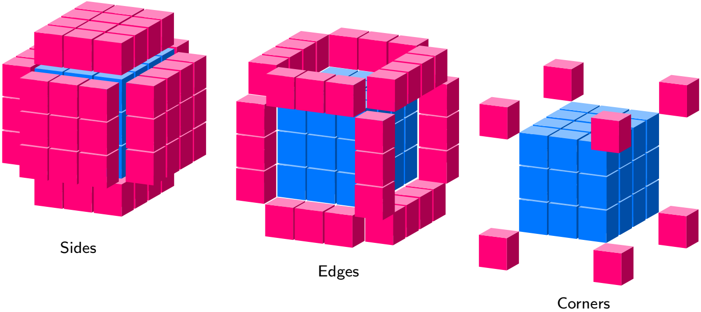


Figure 18: Illustration of groups of lattice sites which must be copied to up to 26 neighbouring GPUs.

GPU to GPU communication is performed in a peer-to-peer fashion (NVIDIA GPUDirect). Data is transfered on the PCI-Express bus without passing through a CPU.

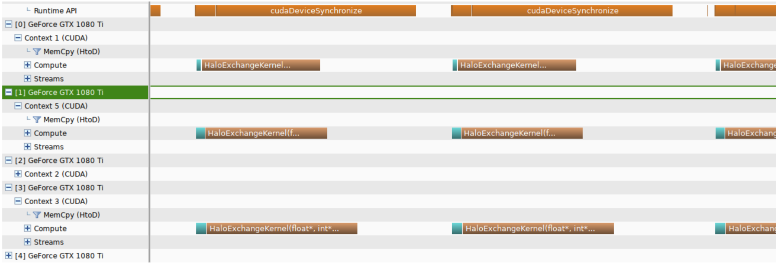


Figure 19: NVIDIA Visual Profiler.

This simple implementation works, but is limited by memory bandwidth. The amount of data can be reduced by only copying the distribution functions required by each neighbour.

Future work

- Optimize halo exchange CUDA kernel and reduce amount of data copied.
- Validate data center models of different sizes with measured data.
- Develop Cascaded LBM and MRT into RAFSINE.
- Develop a platform for training airflow control systems.

Thanks for listening!