HPC-LEAP
EUROPEAN JOINT DOCTORATES

Heterogeneous implementation of the D2Q37 Lattice Boltzmann Method

Alessandro Gabbana

Università degli studi di Ferrara
Bergische Universität Wuppertal

September 27, 2016

Outline

Outline

Lattice Boltzmann Method

- ► Lattice Boltzmann Method: Computational fluid dynamics method for solving complex fluid flows.
- ► Second order approximation of the **Navier**-**Stokes** equations.
- ► A set of virtual particles called **populations** arranged at the edges of a discrete regular mesh.
- ► Particles only have a finite number of velocity directions:

$$\vec{v} \rightarrow \{\vec{e_i}, i = 1, \ldots, m\}$$

Lattice Boltzmann Method

- Lattice Boltzmann Method: Computational fluid dynamics method for solving complex fluid flows.
- Second order approximation of the **Navier-Stokes** equations.
- A set of virtual particles called **populations** arranged at the edges of a discrete regular mesh.
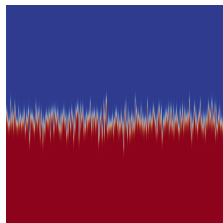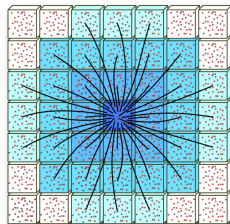- Particles only have a finite number of velocity directions:

$$\vec{v} \rightarrow \{\vec{e_i}, i = 1, \ldots, m\}$$

Lattice Boltzmann Method

- Lattice Boltzmann Method: Computational fluid dynamics method for solving complex fluid flows.

- Second order approximation of the **Navier-Stokes** equations.

- A set of virtual particles called **populations** arranged at the edges of a discrete regular mesh.

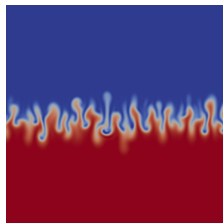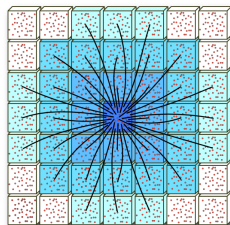- Particles only have a finite number of velocity directions:

$$\vec{v} \to \{\vec{e}_i, i = 1, \ldots, m\}$$

Lattice Boltzmann Method

- Lattice Boltzmann Method: Computational fluid dynamics method for solving complex fluid flows.

- Second order approximation of the **Navier-Stokes** equations.

- A set of virtual particles called **populations** arranged at the edges of a discrete regular mesh.

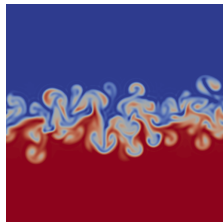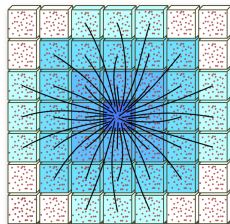- Particles only have a finite number of velocity directions:

$$\vec{v} \to \{\vec{e_i}, i = 1, \ldots, m\}$$

Lattice Boltzmann Method

- Lattice Boltzmann Method: Computational fluid dynamics method for solving complex fluid flows.

- Second order approximation of the **Navier-Stokes** equations.

- A set of virtual particles called **populations** arranged at the edges of a discrete regular mesh.

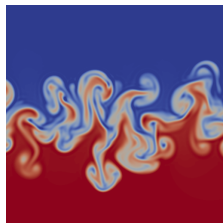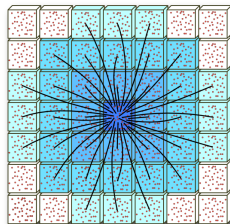- Particles only have a finite number of velocity directions:

$$\vec{v} \rightarrow \{\vec{e_i}, i = 1, \ldots, m\}$$

Lattice Boltzmann Method

- Lattice Boltzmann Method: Computational fluid dynamics method for solving complex fluid flows.

- Second order approximation of the **Navier-Stokes** equations.

- A set of virtual particles called **populations** arranged at the edges of a discrete regular mesh.

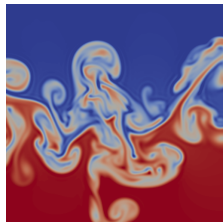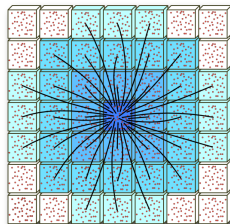- Particles only have a finite number of velocity directions:
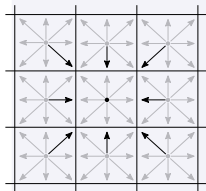
$$\vec{v} \to \{\vec{e_i}, i = 1, \ldots, m\}$$

Lattice Boltzmann Equation

$$f_i(x + e_i \Delta t, t + \Delta t) = f_i(x, t) + \frac{\Delta t}{\tau} \left(f_i^{eq}(x, t) - f_i(x, t) \right), \ i = 1 \dots m$$

Lattice Boltzmann Equation

$$f_i(x + e_i\Delta t, t + \Delta t) = f_i(x,t) + \frac{\Delta t}{\tau}\left(f_i^{eq}(x,t) - f_i(x,t)\right), \ i = 1 \ldots m$$

$$\tilde{f}_i(x,t) = f_i(x - e_i\Delta t, t), \quad i = 1 \ldots m$$

Lattice Boltzmann Equation

$$f_i(x + e_i \Delta t, t + \Delta t) = f_i(x, t) + \frac{\Delta t}{\tau} \left(f_i^{eq}(x, t) - f_i(x, t) \right), \; i = 1 \ldots m$$
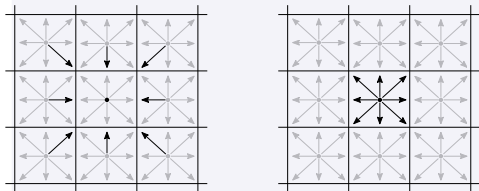
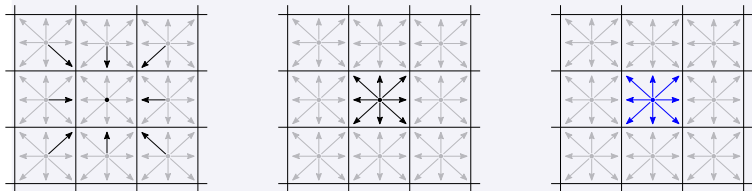$$\tilde{f}_i(x, t) = f_i(x - e_i \Delta t, t), \quad i = 1 \ldots m$$

Lattice Boltzmann Equation

$$f_i(x + e_i \Delta t, t + \Delta t) = f_i(x, t) + \frac{\Delta t}{\tau} \left(f_i^{eq}(x, t) - f_i(x, t) \right), \ i = 1 \dots m$$

$$\tilde{f}_i(x, t) = f_i(x - e_i \Delta t, t), \quad i = 1 \dots m$$

$$f_i(x, t + \Delta t) = \tilde{f}_i(x, t) + \frac{\Delta t}{\tau} \left(f_i^{eq}(x, t) - \tilde{f}_i(x, t) \right), \ i = 1 \dots m$$

Computational Scheme

1: **for all** time step **do**
2:   $<$ Set boundary conditions $>$
3:   **for all** lattice site **do** {in parallel}
4:     $<$ Propagate $>$
5:     $<$ Collide $>$
6:   **end for**
7: **end for**

- ► In principle *propagate* and *collide* could be fused.
- ► Convenient for benchmarking to keep the memory bound section separated from the compute intensive one.

Outline

1 Lattice Boltzmann Method

2 Programming heterogeneous architectures
- Directive based programming models
- Strategies for accelerator-based implementations

3 Data Layout Optimization

4 Load Balancing

5 Performances & Results

Many architectures... many implementations!

Multiphase lattice Boltzmann on the Cell Broadband Engine

F. Belletti[1], L. Biferale[2], F. Mantovani[1], S. F. Schifano[3], F. Toschi[4][5] and R. Tripiccione[1]

[1] Dipartimento di Fisica and INFN, Università di Ferrara - Ferrara, Italy
[2] Dipartimento di Fisica and INFN, Università di Tor Vergata - Rome, Italy
[3] Dipartimento di Matematica and INFN, Università di Ferrara - Ferrara, Italy
[4] Istituto per le Applicazioni del Calcolo, CNR - I-00161 Rome, Italy
   INFN, Sezione di Ferrara - I-44100 Ferrara, Italy
[5] Department of Physics and Department of Mathematics and Computer Science
   Eindhoven University of Technology - 5600 MB Eindhoven, The Netherlands
   and International Collaboration for Turbulence Research

Summary. — Computational experiments are one of the most used and flexible investigation tools in fluid dynamics. The Lattice Boltzmann Equation is a well established computational method particularly promising for multi-phase flows at micro and macro scales. Here we present preliminary results on performances of the

Many architectures... many implementations!

Many architectures... many implementations!

An Optimized Lattice Boltzmann Code for BlueGene/Q

Marcello Pivanti[1], Filippo Mantovani[2], Sebastiano Fabio Schifano[1]✉, Raffaele Tripiccione[1], and Luca Zenesini[1]

[1] Università di Ferrara and INFN, Ferrara, Italy
schifano@fe.infn.it
[2] Facultät für Physik, Univesität Regensburg, Regensburg, Germany

Abstract. In this paper we describe an optimized implementation of a Lattice Boltzmann (LB) code on the BlueGene/Q system, the latest generation massively parallel system of the BlueGene family. We consider a state-of-art LB code, that accurately reproduces the thermo-hydrodynamics of a 2D-fluid obeying the equations of state of a perfect gas. The regular structure of LB algorithms offers several levels of algorithmic parallelism that can be matched by a massively parallel computer architecture. However the complex memory access patterns associated to our LB model make it not trivial to efficiently exploit all available parallelism. We describe our implementation ...

Many architectures... many implementations!

Many architectures... many implementations!

Many architectures... many implementations!

Many architectures... many implementations!

SPECIAL ISSUE PAPER

Performance and portability of accelerated lattice Boltzmann applications with OpenACC

Enrico Calore[1], Alessandro Gabbana[1], Jiri Kraus[2], Sebastiano Fabio Schifano[3,*,†]
and Raffaele Tripiccione[1]

[1]*Dip. di Fisica e Scienze della Terra, University of Ferrara, and INFN, Ferrara, Italy*
[2]*NVIDIA GmbH, Würselen (Germany)*
[3]*Dip. di Matematica e Informatica, University of Ferrara, and INFN, Ferrara, Italy*

SUMMARY

An increasingly large number of HPC systems rely on heterogeneous architectures combining traditional multi-core CPUs with power efficient accelerators. Designing efficient applications for these systems have been troublesome in the past as accelerators could usually be programmed using specific programming languages threatening maintainability, portability, and correctness. Several new programming environments try to tackle this problem. Among them, OpenACC offers a high-level approach based on compiler directives to mark regions of existing C, C++, or Fortran codes to run on accelerators. This approach directly addresses code portability, leaving to compilers the support of each different accelerator, but one has to carefully assess the relative costs of portable approaches versus computing efficiency. In this paper, we address precisely this issue, using as a test-bench a massively parallel lattice Boltzmann algorithm. We first describe our multi-node implementation and optimization of the algorithm, using OpenACC and MPI. We then benchmark

Solutions for performance portability in HPC

Goal

Can we have a single performance-portable code capable of running efficiently on recent heterogeneous architectures?

Tested solutions

- OpenCL
  - Low level approach
  - Future support for GPUs uncertain
- Directive based programming models
  - High level programming approach.
  - Portability becomes a duty of the compiler.
  - Several standards (OpenMP4.x, OpenACC).

Accelerator-based programming model

- ▶ Host-centric model
- ▶ Abstraction supporting both many-core (GPUs, MIC) and multi-core architectures.

Strategies for accelerator-based implementations

Two possible approaches

1. Map compute intensive sections onto the device
2. Heterogeneous implementation

Implementation aspects common to both approaches:
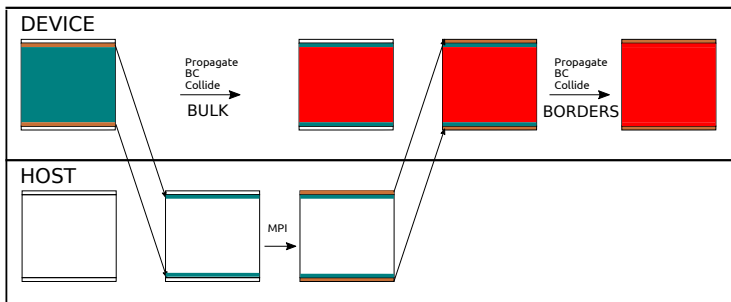
- ▶ Full-matrix lattice representation with equidistant Cartesian coordinates
- ▶ Two copies of the lattice stored in memory in order to avoid data dependencies
- ▶ External halo-layers used to implement boundary conditions
- ▶ Control on data movements

Strategies for accelerator-based implementations

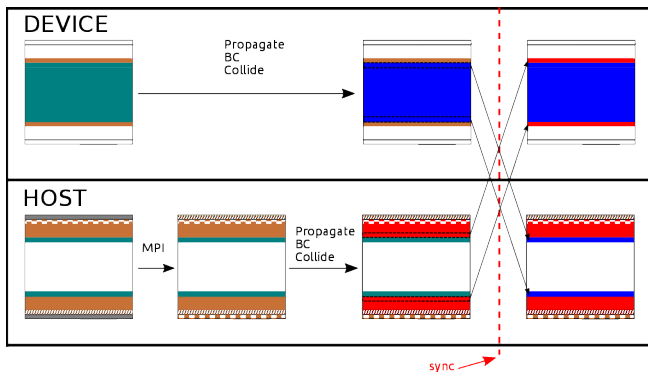## Offload computational intensive sections onto the device

- ▶ Pros (I) : Overlap communications with computation.
- ▶ Pros (II): Needs to optimize only the code targeting the accelerator.
- ▶ Cons: The host sits idle for most of the simulation time.

Strategies for accelerator-based implementations

Heterogeneous implementation

- ▶ Pros: Fully exploits compute capabilities of the cluster.
- ▶ Cons (I) : Need for a data-layout optimizing performances on both host and accelerator.
- ▶ Cons (II): Requires load-balancing.

Outline

Array of Structures (AoS)

```
// AoS data-type definition

#define N (LX*LY)
typedef struct {
  data_t p0; // population 0
  data_t p1; // population 1
  ...
  data_t p8; // population 8
} pop_t;

aos_t lattice[N];
```
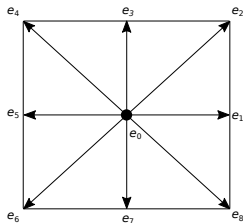
```
// snippet of collide code kernel
// computing density rho

idx = (ix*NY)+iy;

rho = 0.0;

for(p = 0; p < NPOP; p++)
  rho = rho + lattice[ NPOP*idx + p ];
```
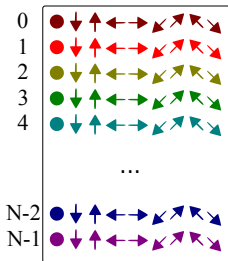

AoS

Array of Structures (AoS): Performances

		Data Structure			
Architecture	AoS				
Haswell	10.17				
Broadwell	18.91				
Xeon Phi	21.58				
Tesla K80	16.06				
AMD Hawaii	6.07				

- ▶ Performance figures in term of MLUPS (Million Lattice Update Per Second).
- ▶ Simulations performed on a 2160x8192 lattice.

Structure of Arrays (SoA)

```
// SoA data-type definition

#define N (LX*LY)
typedef struct {
  data_t p0[N]; // population 0
  data_t p1[N]; // population 1
  ...
  data_t p8[N]; // population 8
} pop_t;

soa_t lattice;
```
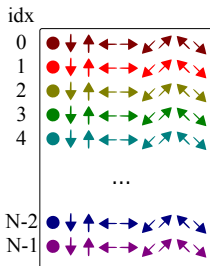
```
// snippet of collide code kernel
// computing density rho

idx = (ix*NY)+iy;

rho = 0.0;

for(p = 0; p < NPOP; p++)
  rho = rho + lattice[ (p*NX*NY) + idx ];
```
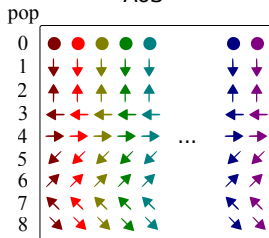

AoS

SoA

Structure of Arrays (SoA): Performances

		Data Structure		
Architecture	AoS	**SoA**		
Haswell	10.17	**9.53**		
Broadwell	18.91	**16.60**		
Xeon Phi	21.58	**9.68**		
Tesla K80	16.06	**77.97**		
AMD Hawaii	6.07	**16.14**		

- ► Performance figures in term of MLUPS (Million Lattice Update Per Second).
- ► Simulations performed on a 2160x8192 lattice.

Clusterized Structure of Arrays (CSoA)

```c
// cluster definition
typedef struct {
  data_t c[VL];
} vdata_t;

// CSoA data-type definition
typedef struct {
  vdata_t p[NPOP][NX*(NY / VL)];
} csoa_t;

csoa_t lattice;
```
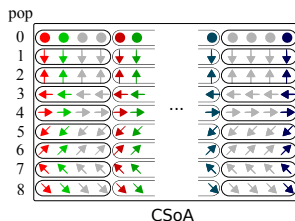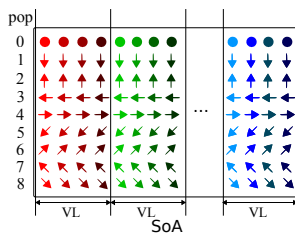
```c
// snippet of collide code kernel
// computing density rho
idx = ix*(NY / VL) + iy;

#pragma omp simd
for(k = 0; k < VL; k++){
  rho.c[k] = 0.0;
}

#pragma unroll novector
for (p = 0; p < NPOP; p++){
  #pragma omp simd
  for(k = 0; k < VL; k++)
    rho.c[k] += lattice->p[p][idx].c[k];
}
```


SoA

CSoA

Clusterized Structure of Arrays (CSoA): Performances

	Data Structure			
Architecture	AoS	SoA	**CSoA**	
Haswell	10.17	9.53	**16.22**	
Broadwell	18.91	16.60	**26.69**	
Xeon Phi	21.58	9.68	**30.28**	
Tesla K80	16.06	77.97	**80.34**	
AMD Hawaii	6.07	16.14	**32.74**	

- ▶ Performance figures in term of MLUPS (Million Lattice Update Per Second).
- ▶ Simulations performed on a 2160x8192 lattice.

Clusterized Array of Structure of Arrays (CAoSoA)

```
// CSoA data-type definition
typedef struct {
  data_t c[VL];
} vdata_t;

typedef struct {
  vdata_t p[NPOP];
} caosoa_t;

caosoa_t lattice[NX*(NY / VL)];
```
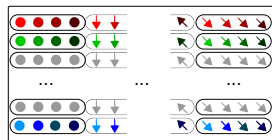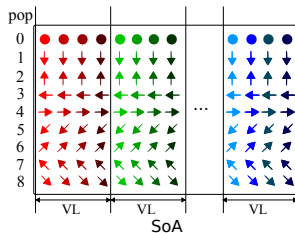
```
// snippet of collide code kernel
// computing density rho
idx = ix*(NY / VL) + iy;

#pragma omp simd
for(k = 0; k < VL; k++){
  rho.c[k] = 0.0;
}

#pragma unroll novector
for (p = 0; p < NPOP; p++){
  #pragma omp simd
  for(k = 0; k < VL; k++)
    rho.c[k] += lattice[idx].p[p].c[k];
}
```


SoA

CAoSoA

Clusterized Array of Structure of Arrays (CAoSoA): Performances

	Data Structure			
Architecture	AoS	SoA	CSoA	**CAoSoA**
Haswell	10.17	9.53	16.22	**17.68**
Broadwell	18.91	16.60	26.69	**31.10**
Xeon Phi	21.58	9.68	30.28	**40.41**
Tesla K80	16.06	77.97	80.34	**80.19**
AMD Hawaii	6.07	16.14	32.74	**36.65**

▶ Performance figures in term of MLUPS (Million Lattice Update Per Second).

▶ Simulations performed on a 2160x8192 lattice.

Outline

Load Balancing

$$T_{\texttt{exe}} = \max\{T_{\texttt{acc}}, T_{\texttt{host}} + T_{\texttt{mpi}}\} + T_{\texttt{swap}}$$

$T_{\texttt{acc}} \propto (LX - 2M)LY \cdot \tau_d$

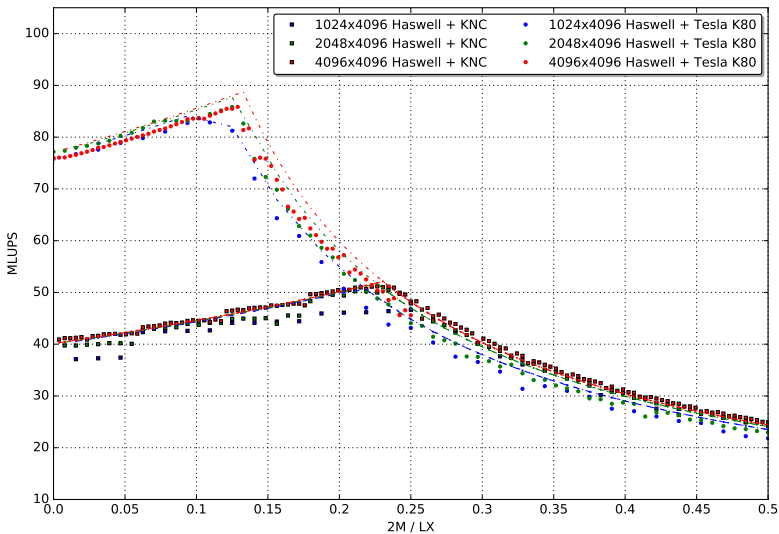$T_{\texttt{host}} \propto (2M)LY \cdot \tau_h$

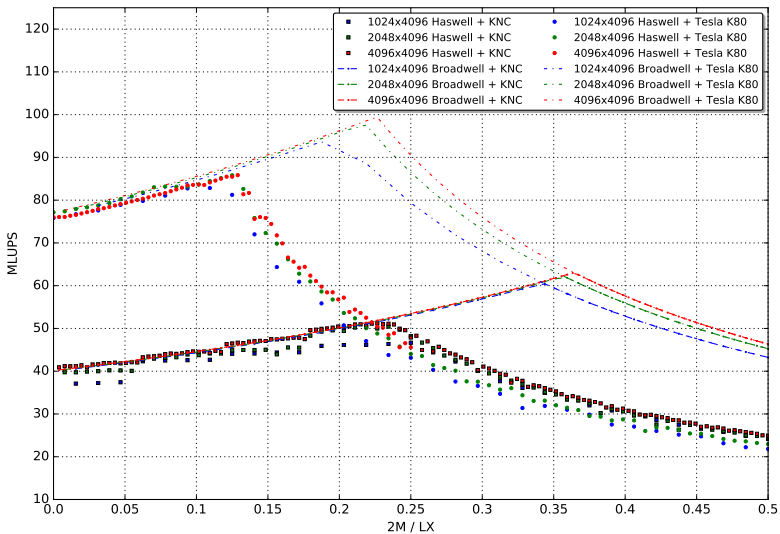$T_{\texttt{mpi}} \propto \tau_c$ (constant since we are using a 1D partitioning)

Autotuning:

1. Get an estimate for τ_d, τ_h, τ_c by mini-benchmarks
2. Compute M such that T_{exe} by solving

$$T_{acc}(M) = T_{host}(M) + T_{mpi}(M)$$

Load Balancing: Testing the Model
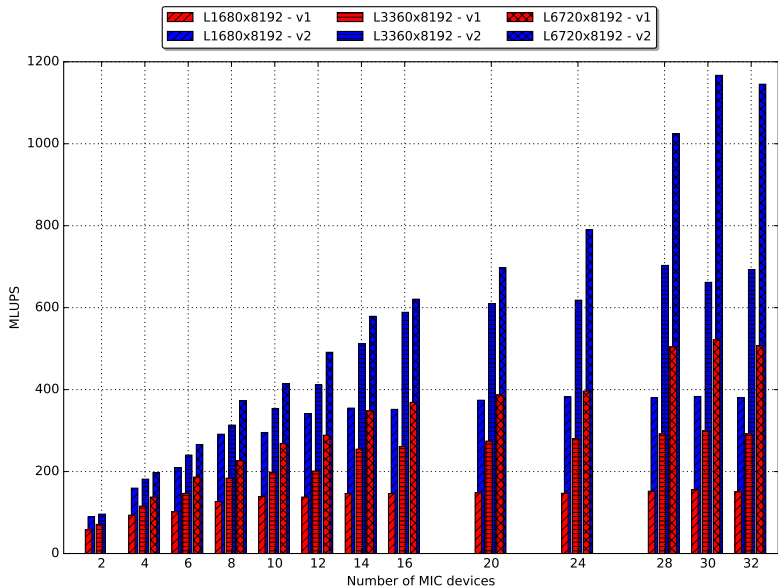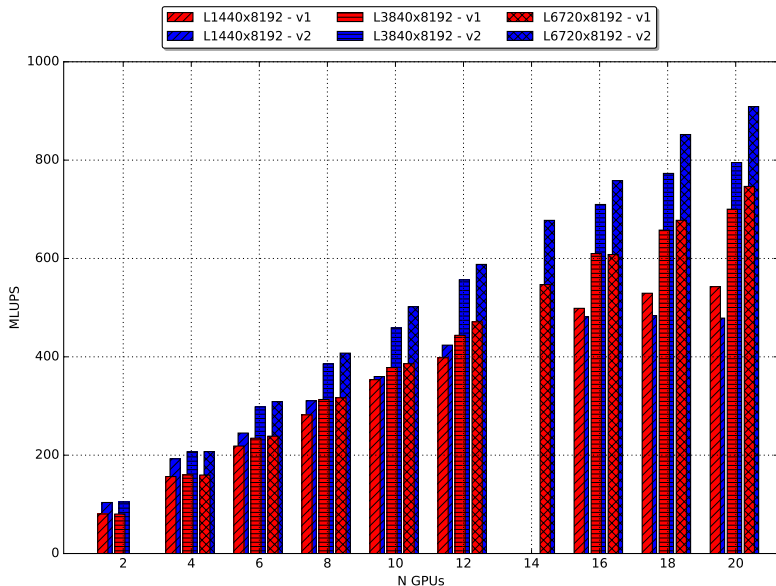
Load Balancing: Performance Forecasting

Outline

Tuning of cluster dimension

Scalability performances for Haswell + KNC

Scalability performances for Haswell + Tesla K80

Outline

Coming Next..

Heterogeneous implementation of the D2Q37 Lattice Boltzmann Method

Alessandro Gabbana

Università degli studi di Ferrara
Bergische Universität Wuppertal

September 27, 2016