CS-E4690 – Programming parallel supercomputers D

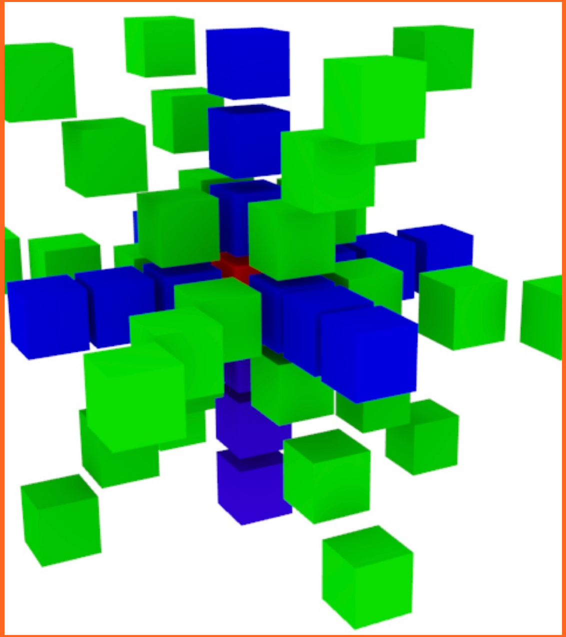
2nd Lecture

Maarit Korpi-Lagg

maarit.korpi-lagg@aalto.fi

31.10.2023

Aalto University
School of Science



Lecture 2

Basic definitions and the importance of the network

- **Course practicalities repetition: 5 min**
- **Unbroken phone: key concepts recap (30 min)**
- **Exercise 2 as group work:**
  - 4 topics (from sheet2 in GitLab)
  - Row-wise group work (10 mins work, 2 mins for discussing answers)
  - Synthesis (5 mins max)
- **Signing up for points and feedback (2 mins)**

Break-down of learning objectives

Lecture1

Introduction to the current HPC landscape

Understanding how this course fits into that

Establishing understanding of the learning outcomes, specifically answering the question: “What are programming during this course?”

Lecture2

Learning basic definitions and taxonomies

Understanding the importance of the “network”

Learning basic performance models

Understanding the concept of a well-performing software in large-scale computing.

Lecture3

Becoming knowledgeable of the modern landscape of distributed memory programming

Understanding why in this course we will concentrate on low-level programming models

Getting acquainted with MPI: basics and synchronous and asynchronous point-to-point communication

Break-down of learning objectives

Lecture4

Learning more about MPI:

One-sided point-to-point communications

Collective communications

Lecture5

Programming MP hybrid architectures

Becoming knowledgeable of the spectrum of options

Understanding efficiency issues

Lecture6

Programming hybrid architectures with accelerators

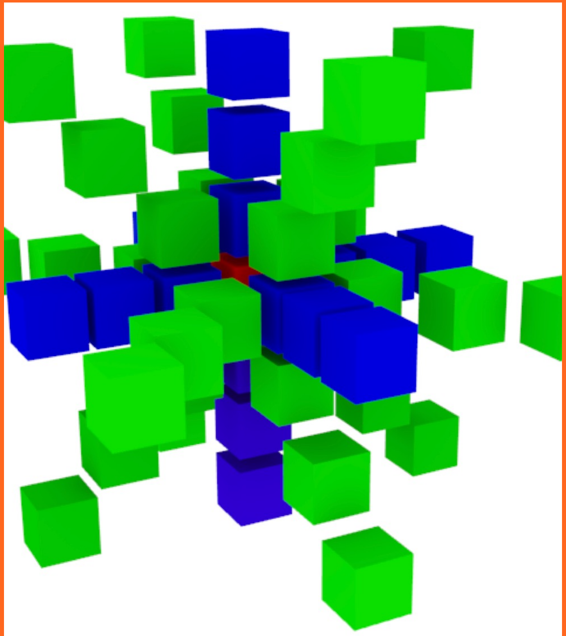
Acquiring knowledge of CUDA-MPI programming model

Lecture 2

Basic definitions and the importance of the network

- **Course practicalities repetition: 5 min**
- **Unbroken phone: key concepts recap (30 min)**
- **Exercise 2 as group work:**
  - 4 topics (from sheet2 in GitLab)
  - Row-wise group work (10 mins work, 2 mins for discussing answers)
  - Synthesis (5 mins max)
- **Signing up for points and feedback (2 mins)**

Unbroken phone: Key concepts



Aalto University
School of Science

Flynn’s taxonomy

What is it in short?

Green(s) explain ~ 1 min to the group

Red/yellow explains it to the class ~ 1 min

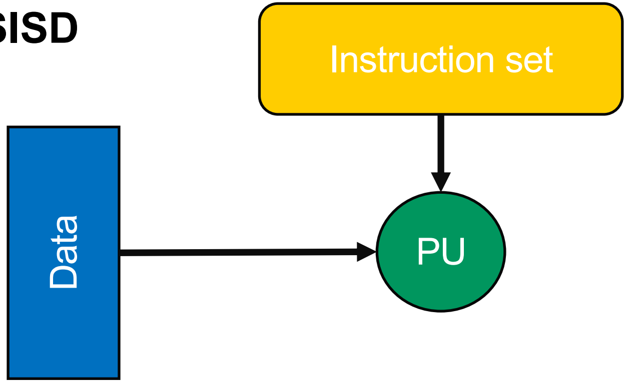
Model answer

Flynn’s taxonomy

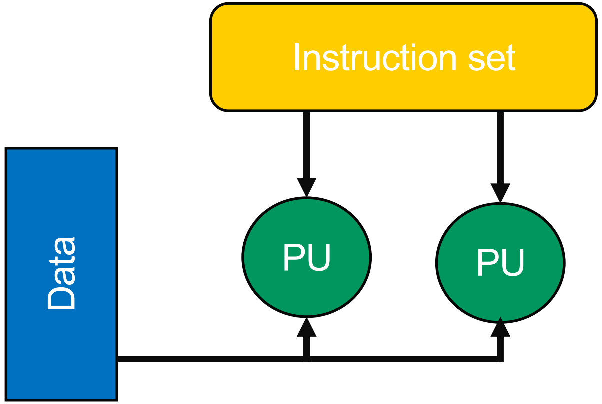
It is a coarse-grained classification of computer architectures using only three abstraction levels: processing unit, data streams and instruction streams. In its original form it contains four different classes.

Flynn’s taxonomy

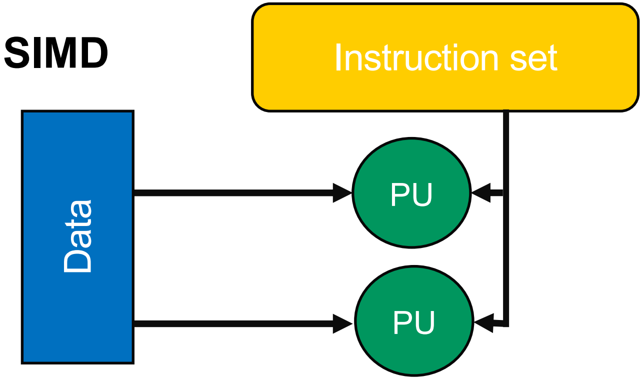
SISD



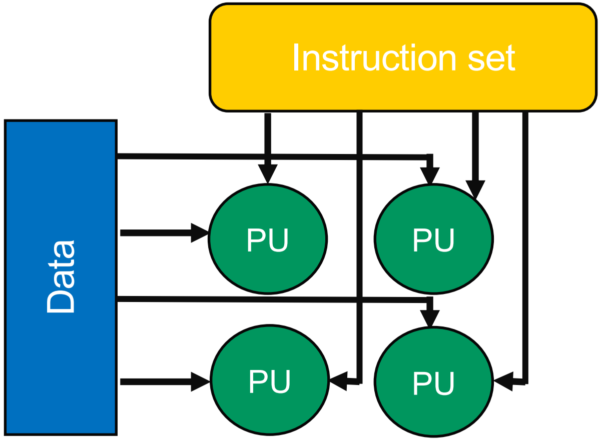
MISD



SIMD



MIMD



Iterative stencil loop

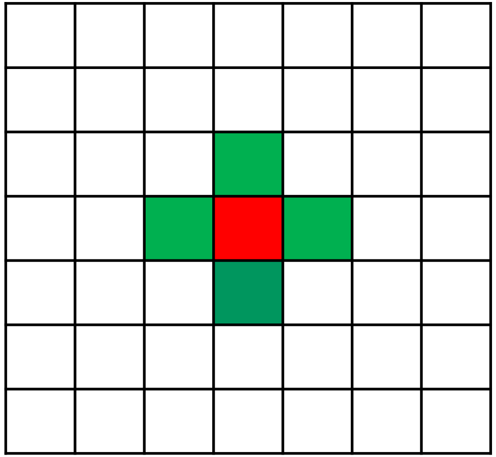
What is it in short?

Green(s) explain ~ 1 min to the group

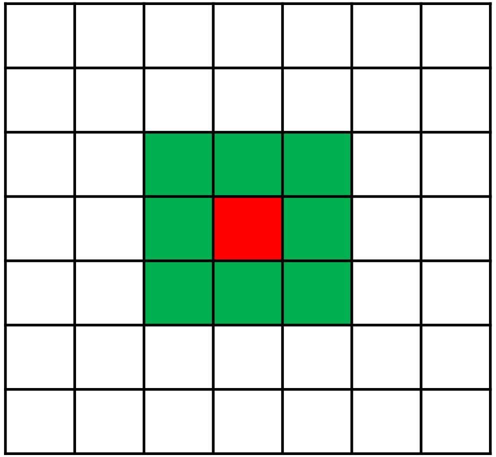
Red/yellow explains it to the class ~ 1 min

Model answer

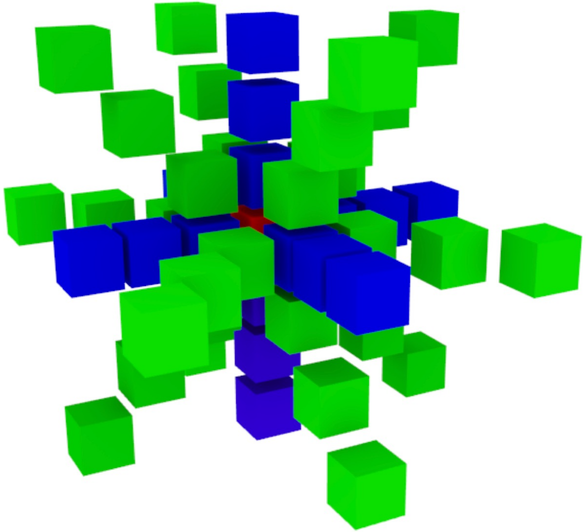
Iterative stencil loop



2D von Neumann



2D Moore



3D 55-point stencil

Recurring update pattern of array elements based on their neighbors.

Amdahl’s law

What is it in short?

Green(s) explain ~ 1 min to the group

Red/yellow explains it to the class ~ 1 min

Model answer

Amdahl's law

A fixed computational problem is given to increasing number of processing elements. The serial parts of the code strongly dominate the scale-up.

Gustafsson’s law

What is it in short?

Green(s) explain ~ 1 min to the group

Red/yellow explains it to the class ~ 1 min

Model answer

Gustafsson’s law

When increasing the number of processing elements, the problem size is allowed to grow. The serial part of the code no longer constrain the speed up.

Interconnect topology: bisection or diameter

Pick one: what is it in short?

Green(s) explain ~ 1 min to the group

Red/yellow explains it to the class ~ 1 min

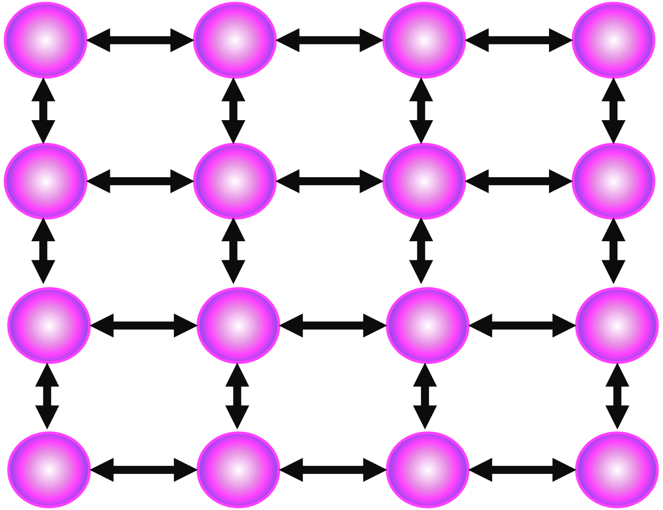
Model answer

Interconnect topology: bisection and diameter

Bisection: minimum number of links that divide the network into two equal halves (estimates worst case bandwidth)

Diameter: maximum number of links between nodes **over a path with minimal distance** (worst case routing distance)

2D mesh

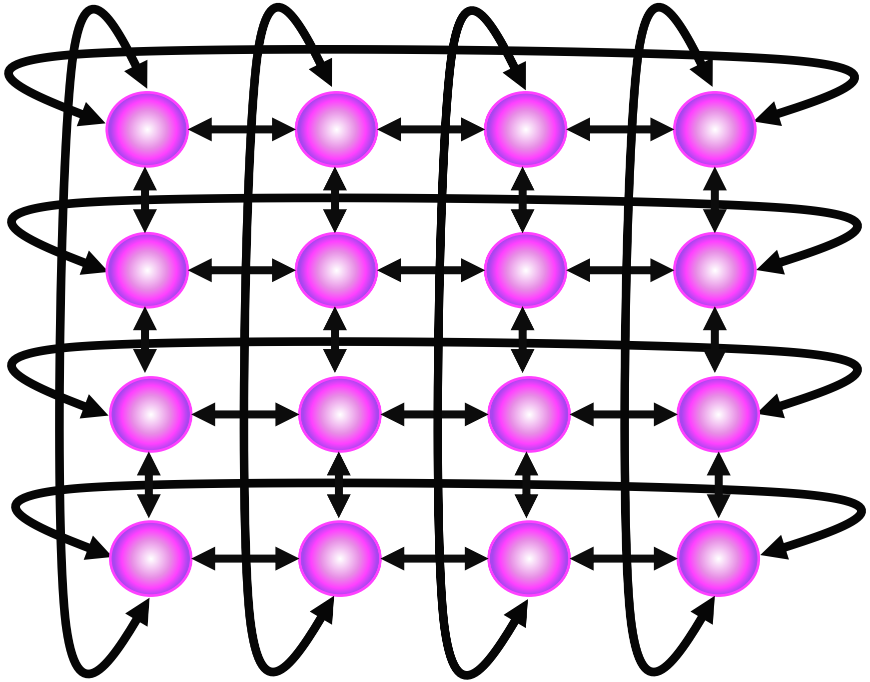


a. Two-dimensional 4x4 mesh.

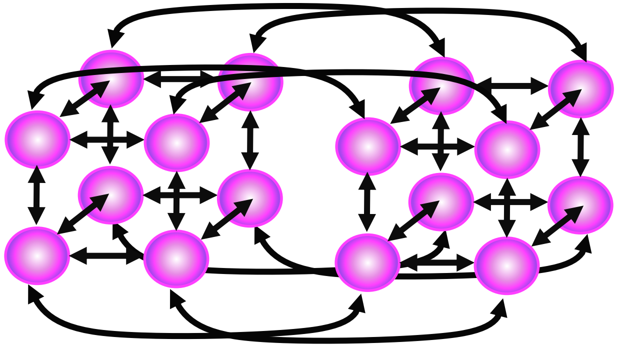
Bisection: 4, Diameter: 6, generalization $\sqrt{N}$, $2(\sqrt{N}-1)$

2D torus

b. Two-dimensional 4x4 torus.
Bisection: 8, Diameter: 4, generalization $2\sqrt{N}$, $\sqrt{N}$



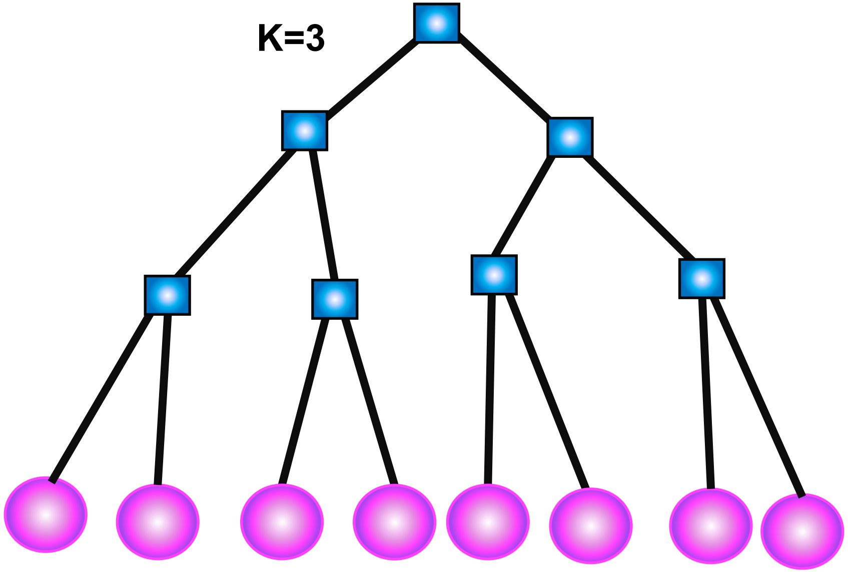
“8+8” Hypercube



c. 8+8 hypercube.
Bisection: 8, Diameter: 4, generalization $N/2$, $log_2 N$

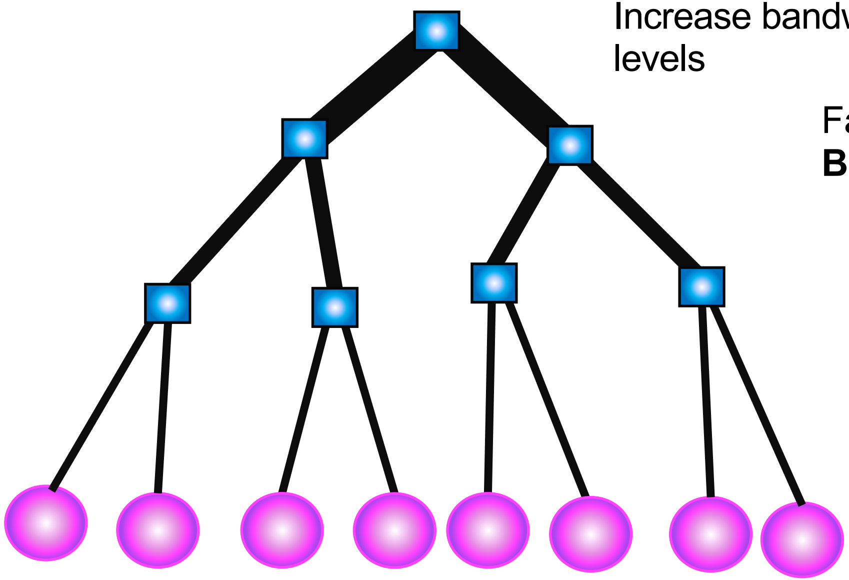
K-Binary tree

3-binary tree.
Bisection: 1 (constant) Diameter: 6, generalization 1 , $2log_2N$



Topology: modern HPC

Fat tree



Increase bandwidth at high levels

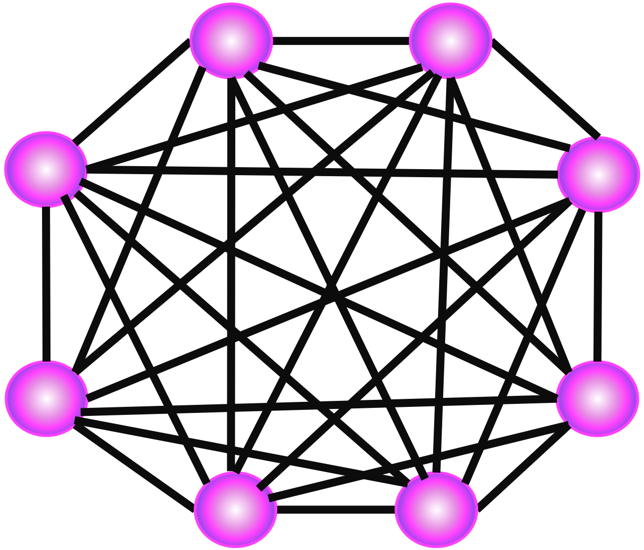
Fat 3-binary tree.
Bisection: 4, Diameter: 6

Topology: modern HPC

Dragonfly

Fully connected graph

Minimizes diameter and
maximises bisection



Diameter 1, Bisection 16

Aalto University
School of Science

Lecture 2

Basic definitions and the importance of the network

- **Course practicalities repetition: 5 min**
- **Unbroken phone: key concepts recap (30 min)**
- **Exercise 2 as group work:**
  - 4 topics (from sheet2 in GitLab)
  - Row-wise group work (10 mins work, 2 mins for discussing answers)
  - Synthesis (5 mins max)
- **Signing up for points and feedback (2 mins)**

Which one are you now?

I am still a bit lost

I understood more, but still would have some questions

I understood much more


Great work!
SCAN ME