# HPC Cloud @ SURFsara

## Design a Parallel Application

Natalie Danezi <anatoli.danezi@surfsara.nl>

UNESCO-IHE Institute for Water Education

SURF SARA

# Outline

- Parallel application:

  - What is the need? … Laptop is not enough

  - How to make it parallel? … Scalability

- Running a parallel application on the HPC cloud

- Hands-on Part B: demo

UNESCO-IHE
Institute for Water Education

SURF SARA

> **SURFsara** @SURFsara_NL · Jun 22
>
> Why do I Need SURFsara Cloud Facility in My Research? #abl #cfd slideshare.net /aliabbasicivil … via @SlideShare

Highly demanding applications for **compute** and **storage** resources to:

- Analyse larger computational domains

- Analyse larger volume of data

- Achieve higher accuracy

- Retrieve the results faster

# Scalability: system & software
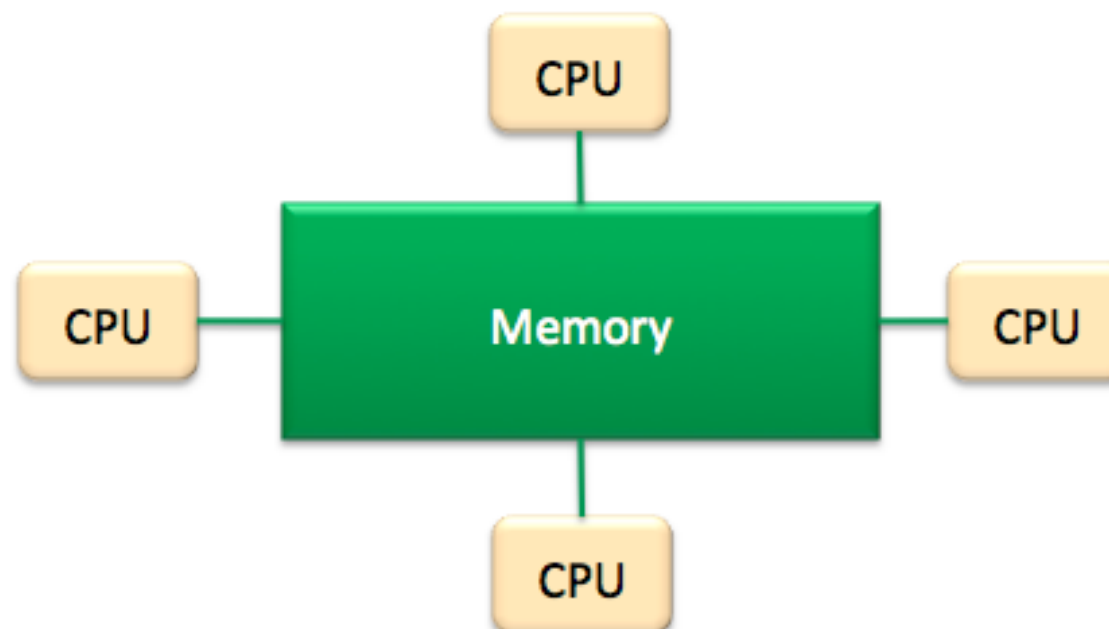
- System architecture:

    - Distributed memory systems

    - Shared memory systems

- Software optimisation:

    - Dig in your code, e.g. MPI, OpenMP

    - Out of the box tools, e.g. XBeach, Delft3D, D-Flow FM

UNESCO-IHE
Institute for Water Education

SURF SARA

# System architecture
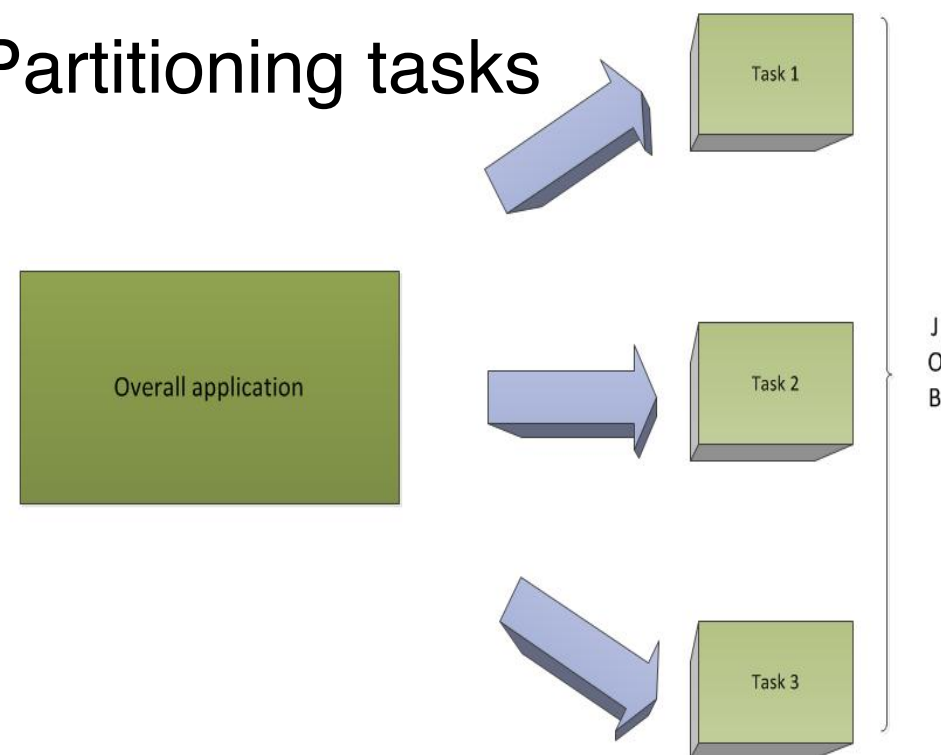
- Distributed memory systems
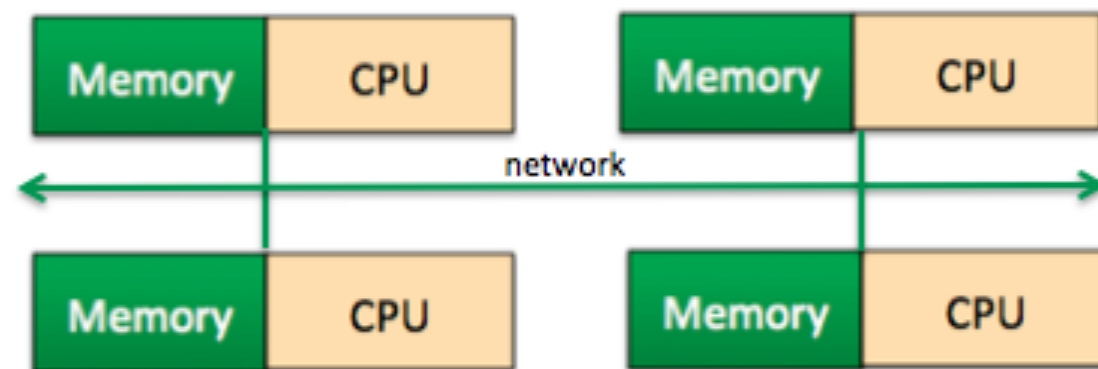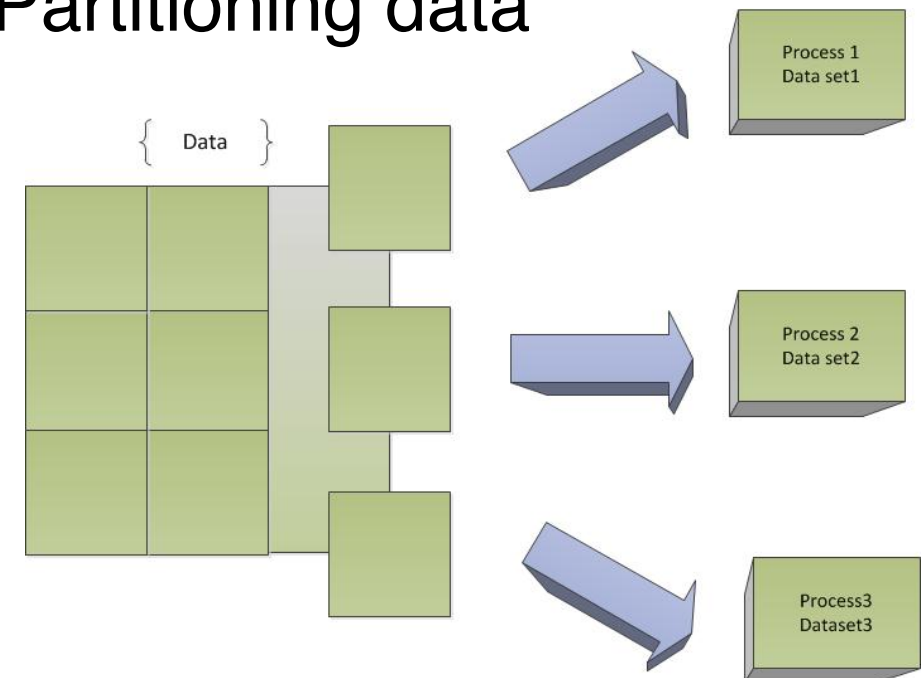


- Shared memory systems

# Software optimisation 1/2

- Dig in your code

  - MPI, OpenMP - advanced programming skills

  - Optimisation techniques - partitioning tasks / data
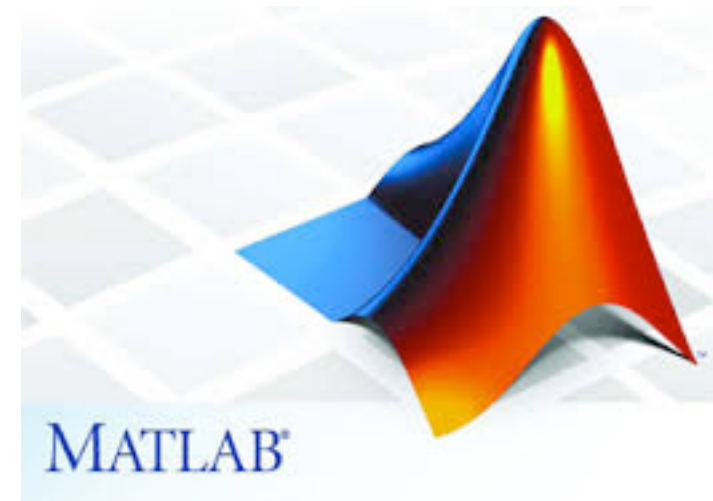
Partitioning tasks



Partitioning data
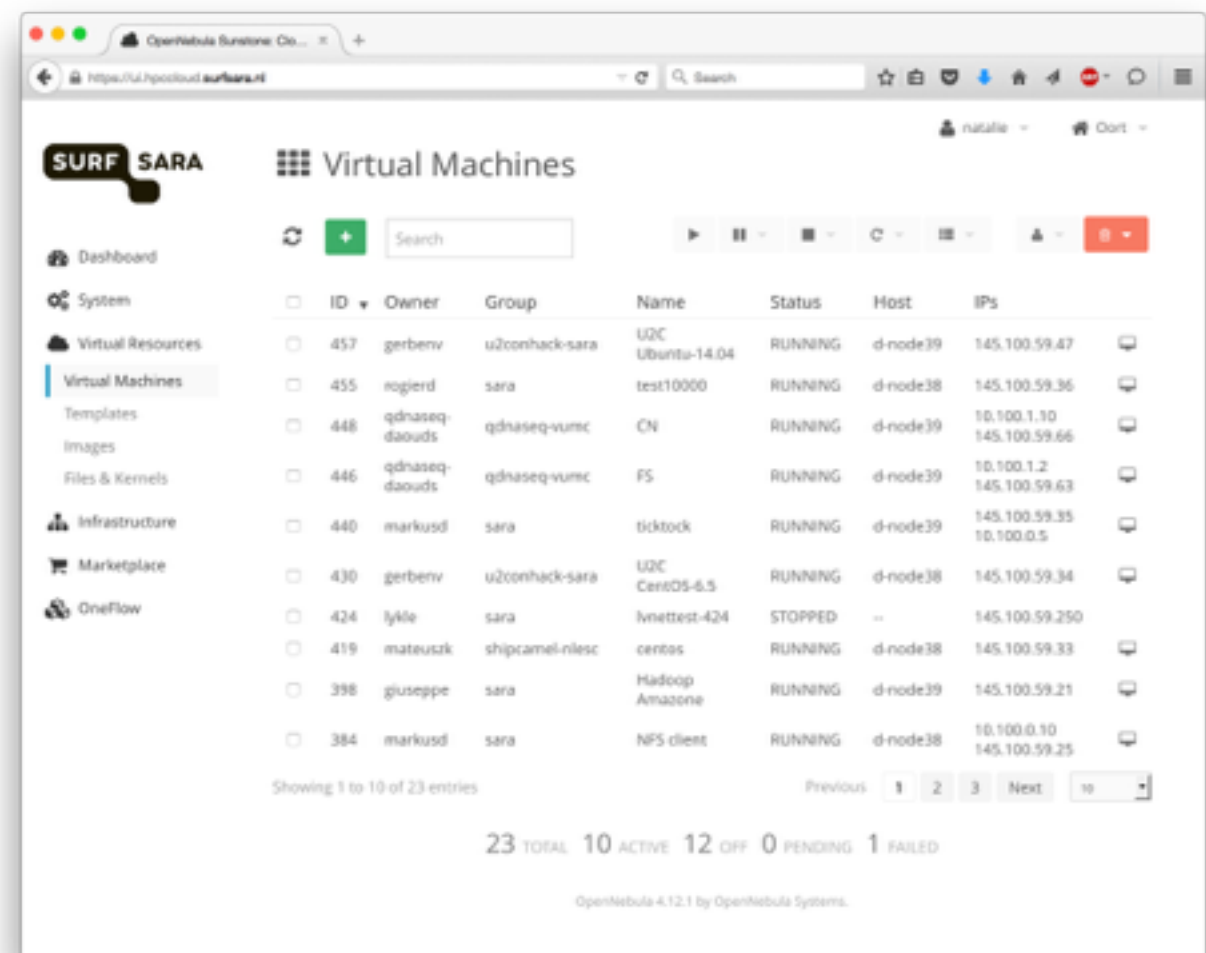
# Software optimisation 2/2
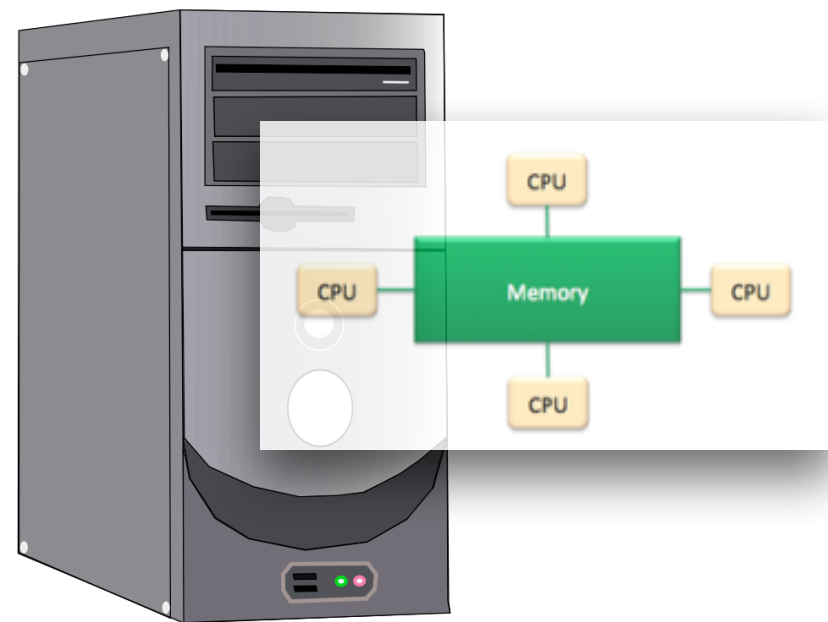
- Out of the box tools

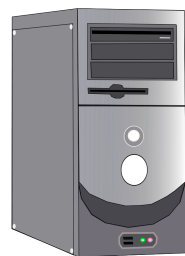# Running a parallel application on the HPC cloud

- Access the UI

- Prepare OS Image, e.g. XBeach

- Create the Template

- Instantiate VM(s)

  - single-VM, single-core

  - single-VM, multi-core

  - multi-VM, multi-core

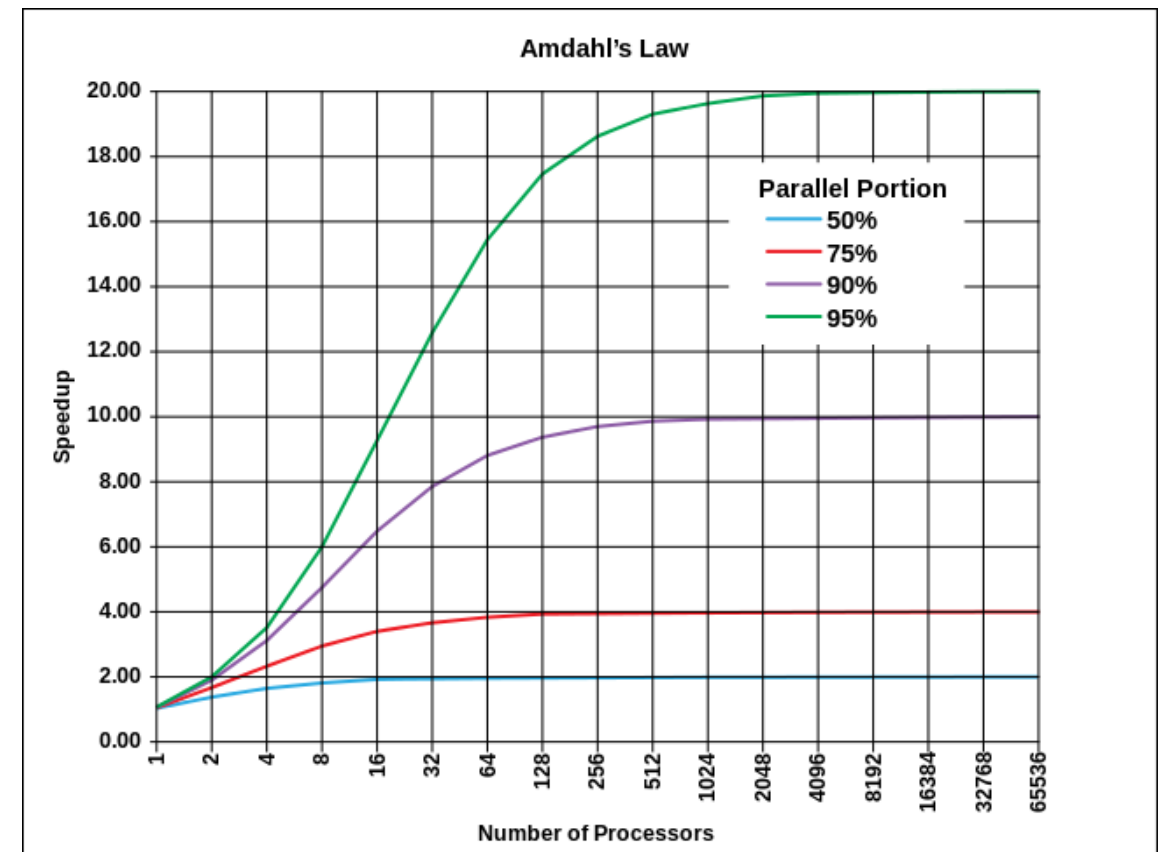    (master/workers cluster)

# Scaling up vs. Scaling out



Scale up

Scale out

# What makes HPC difficult?

- Scaling up
  - ◉ Amdahl's law
- Scaling out
  - ◉ More machines - more problems
  - ◉ Machines fail
  - ◉ Networks fail
  - ◉ Heterogeneous hardware
  - ◉ Latencies
  - ◉ Data locality
- Key questions…

# Key questions

Define your specific requirements:

- Number of cores, Memory, Wall-clock time, data volumes

- Pipelines: what is shared in tasks, what needs to be communicated

- Data locality (private or shared)

- Wall-clock vs. CPU time

- Operating system, software (licensing programs) and databases

- Network interfaces (private or public) - virtual cluster

- Balance effort: system / software scalability

# Hands-on **Part B**

- Continue from:

    **http://doc.hpccloud.surfsara.nl/wshop-uihe-part-B-2015-Dec-11**

- Work in *pairs*

    - each with your *own credentials* on your own laptop

- Follow the instructions at your *own pace*

- Call *Natalie and Ander*

    - *for a) doubts; b) when feeling stuck; c) "food for brain" hints*

UNESCO-IHE
Institute for Water Education

SURF SARA