

Clusters, grids,  
clouds

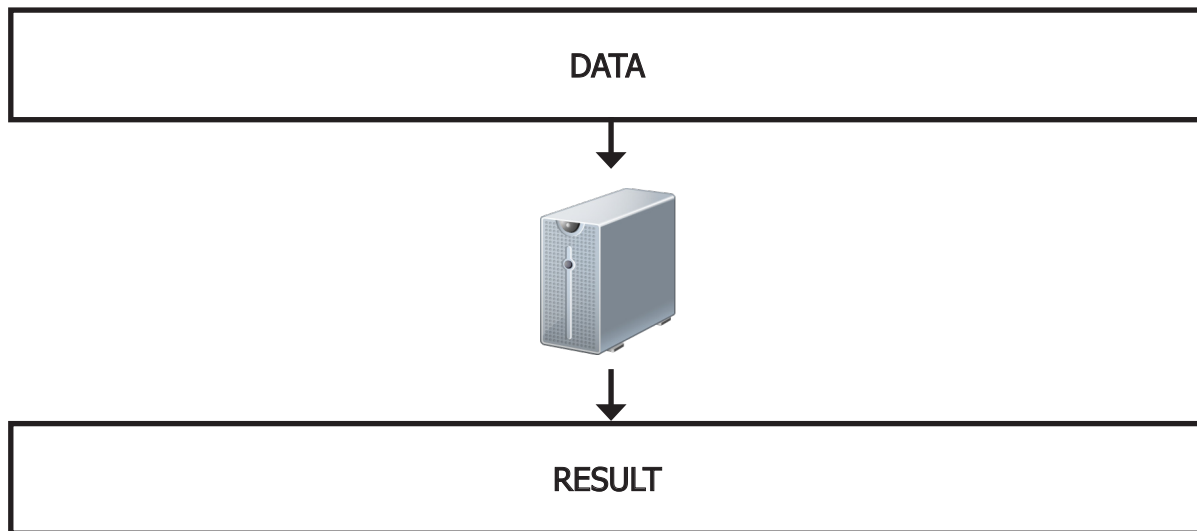
# Parallelization



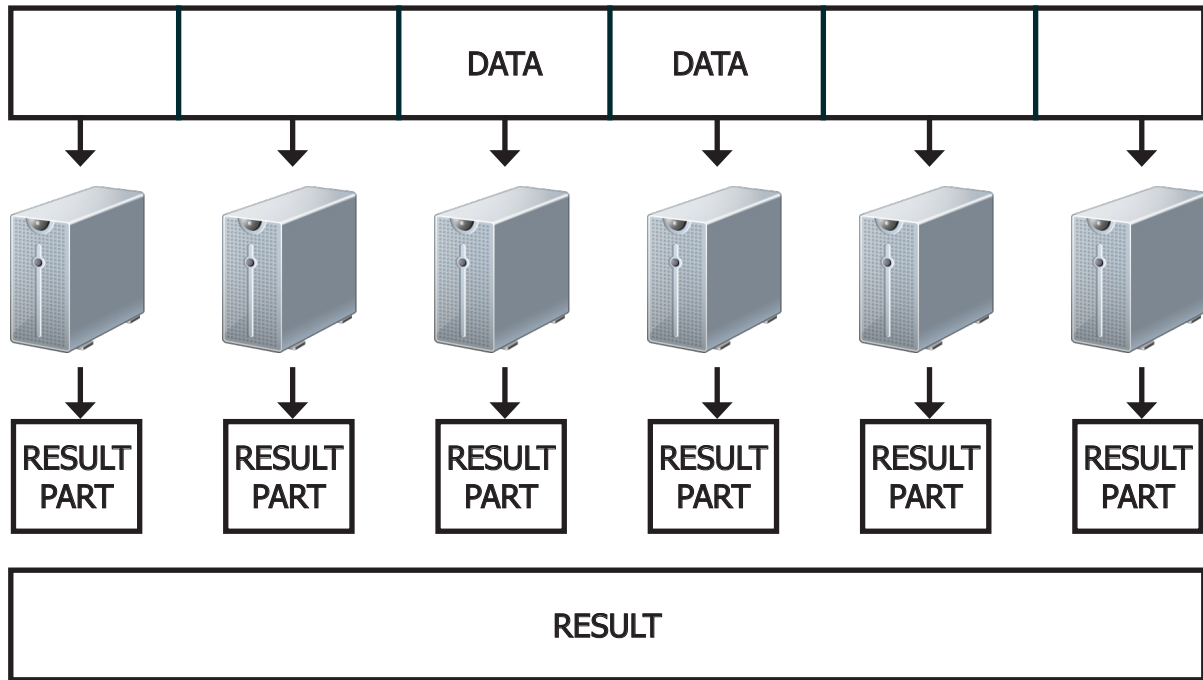
# Parallelization



Parallelization: this takes 1 hour



Parallelization: this takes 10 minutes



## Parallelization:



- this is possible only for “embarrassingly parallel” problems
- e.g. split the reads aligned to the genome and call variants separately

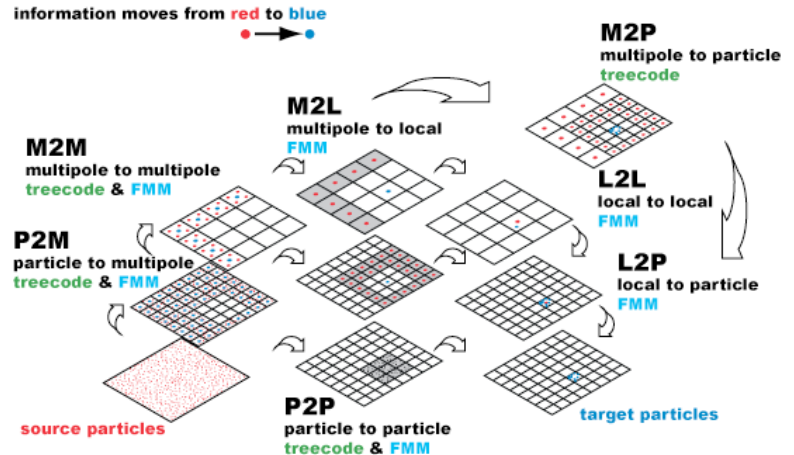
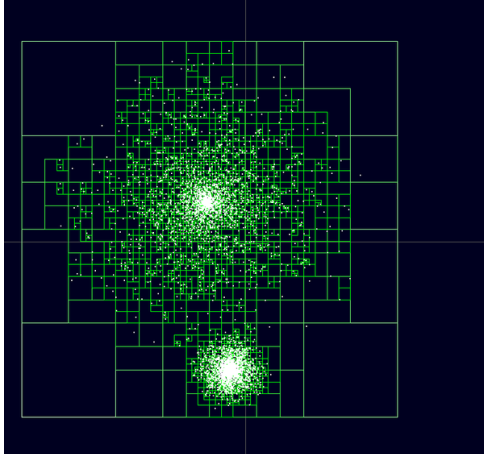
## Parallelization:

- not all problems are like this
- e.g. simulating galaxy collisions



# Parallelization:

- not all problems are like this
- e.g. simulating galaxy collisions





Parallelization: sometimes it does not help



What is the difference  
between cluster, grid and cloud?

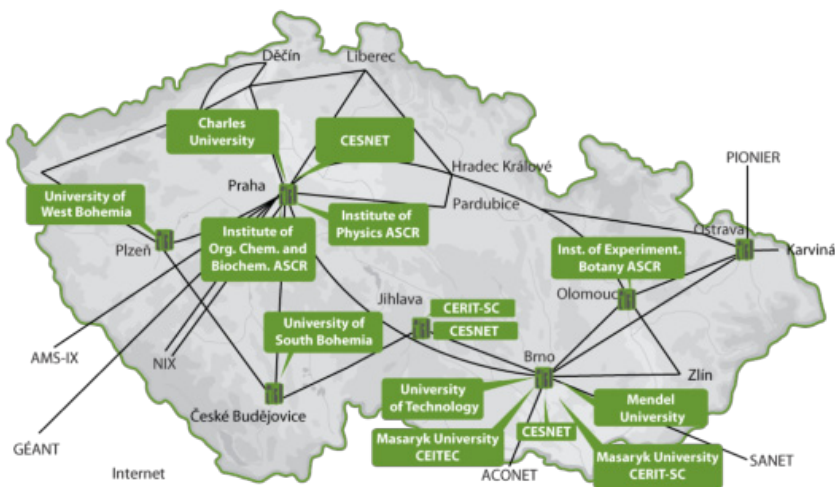


# Grid is a network of clusters





# MetaCentrum is a grid.



## MetaCentrum:

- machines (~10,000 CPU)
- software licenses (lotta good stuff;-)
- shared storage (1 PB + 13 PB hierarchical)
- job scheduler (necessity)
- support (fast and good)

# Cloud is:

- a service – “VirtualBox online”

The screenshot displays the OpenNebula Sunstone web interface. On the left is a sidebar with navigation links: Dashboard, Virtual Resources (expanded), Virtual Machines, Templates, Images, Files & Kernels, Infrastructure (expanded), Datastores, Virtual Networks, and Marketplace. The main content area is titled 'Virtual Machines' and shows summary statistics: 3 TOTAL, 1 ACTIVE, 1 OFF, 0 PENDING, and 0 FAILED. The user 'liborm' is logged in. Below the statistics is a table of virtual machines.

ID	Owner	Group	Name	Status	Host	IPs	VNC
6338	fridrich	metacloud	METACLOUD-Scilinux-6-x86_64-PV-6338	RUNNING	dukan3.ics.muni.cz	147.251.9.233	
6335	liborm	metacloud	NGS kurz	STOPPED	--	147.251.252.223 10.18.1.152	
6334	liborm	metacloud	liborm-test	UNDEPLOYED	--	147.251.252.222 10.18.1.151	

At the bottom of the table, there is a pagination control showing 'Showing 1 to 3 of 3 entries' and a page indicator '1'.

Amazon EC2 is a cloud.

MetaCentrum provides a cloud in test mode.





Cloud is:

- a buzzword
- SaaS, PaaS ... ?!



And the supercomputers?

Supercomputers: clusters with high speed links



# Supercomputers: made in Czech Republic (Anselm)



# Supercomputers: made in Czech Republic (Anselm)

