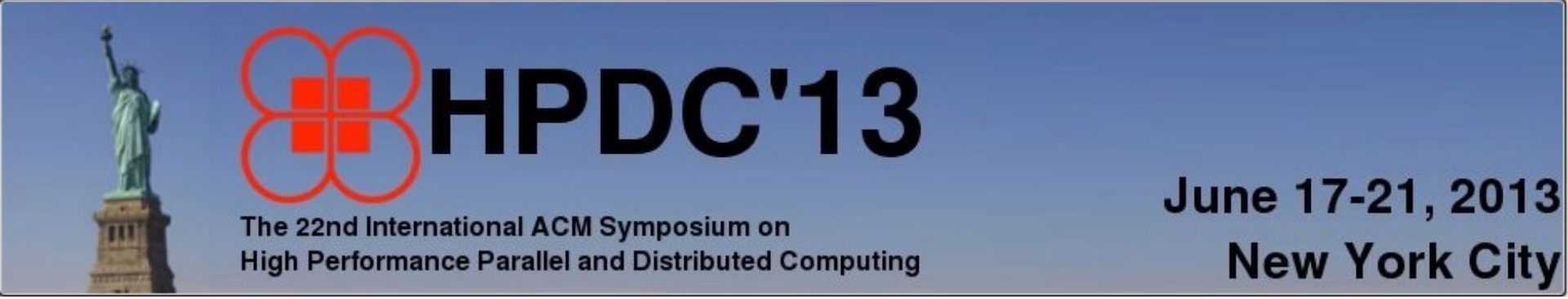


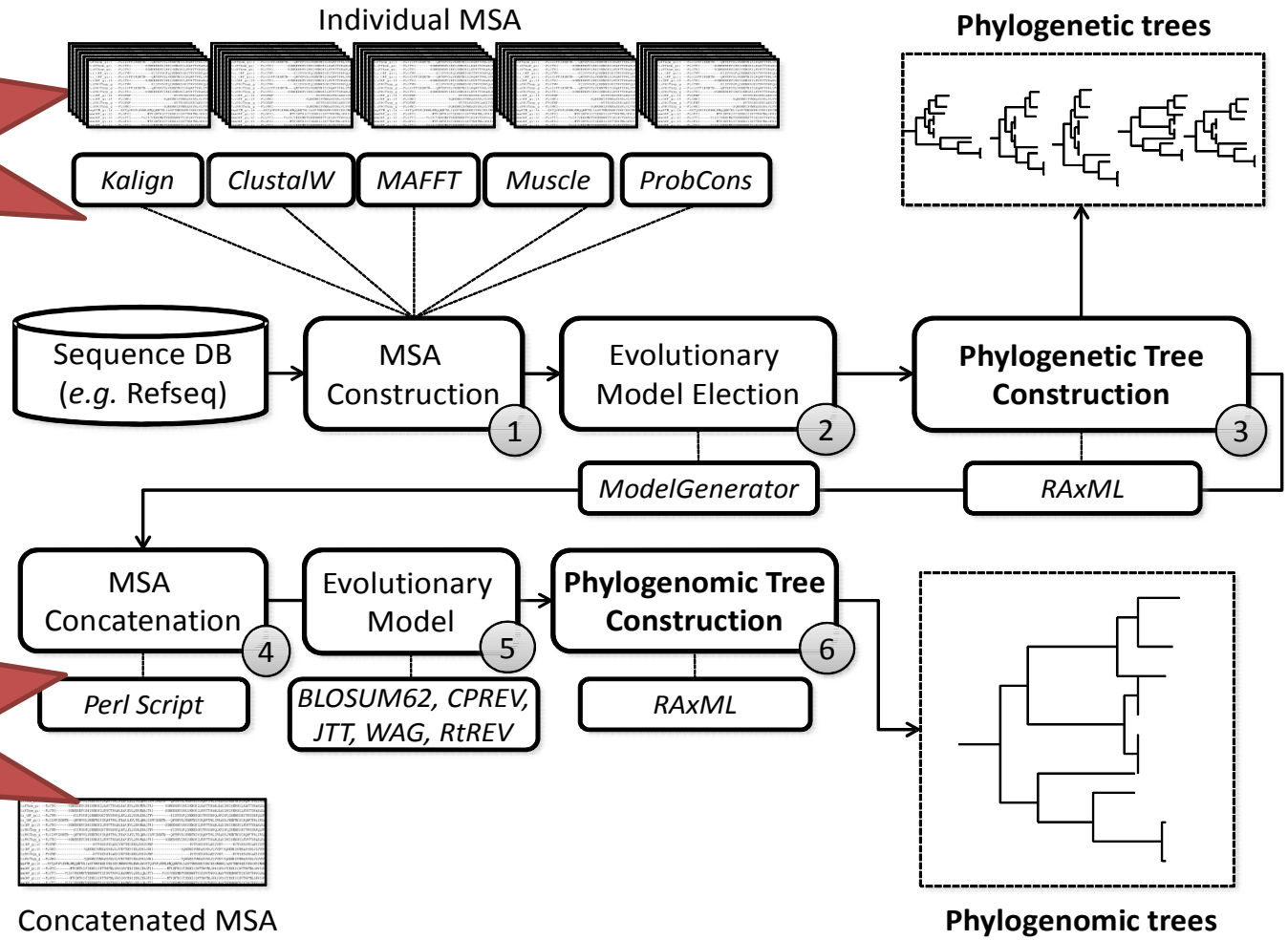
Dimensioning the Virtual Cluster for Parallel Scientific Workflows in Clouds

Daniel de Oliveira, Vitor Viana, Eduardo Ogasawara, Kary Ocaña and Marta Mattoso

Motivating Example: Phylogenetics

Data-intensive and CPU-intensive applications

Several scientists do not access clusters and grids



Scientific Workflows

- Scientific workflows are abstractions that represent the chain of activities within a scientific experiment
- They are managed by Scientific Workflow Management Systems (SWfMS)
- Provenance is a key issue

Scientific Workflows in Clouds

- Many scientists are migrating their experiments to clouds, including bioinformatics ones
- They are not required to assemble expensive computational infrastructure to execute their experiments
- Resources are instantiated under demand

Scientific Workflows in Clouds

- There are some SWfMS that manage the execution of workflows in clouds
  - Swift
  - Pegasus
  - Tavaxy
  - SciCumulus
- In most of these SWfMS a virtual cluster is created in the cloud to execute the workflow in parallel

How to dimension the cloud environment?

- However, clouds have to be configured before workflow execution, *i.e.* The virtual machines have to be instantiated.
- How?
  - Define the type of virtual machine to be used
  - Define the amount of virtual machines of each type to use
  - Brute force search?

How to dimension the cloud environment?

- Existing approaches are primarily focused on optimizing the scheduling plan and task priority
- Native cloud services can scale the amount of virtual machines horizontally and vertically , but they are general purpose
  - Services and do not take into account the specificities of scientific workflows.
  - We may find several different programs with different behaviors, each with a specific execution time

Goals

- Cloud dimension can increase or reduce the total execution time of the workflow
- If a near optimal initial solution is generated it can be further refined by adaptive solutions
- We have used a previously proposed cost function for scheduling workflows in clouds in a fitness function using genetic algorithms as metaheuristics.
- This optimization is implemented in a component named SciDim for the initial dimensioning of the cloud environment before workflow execution

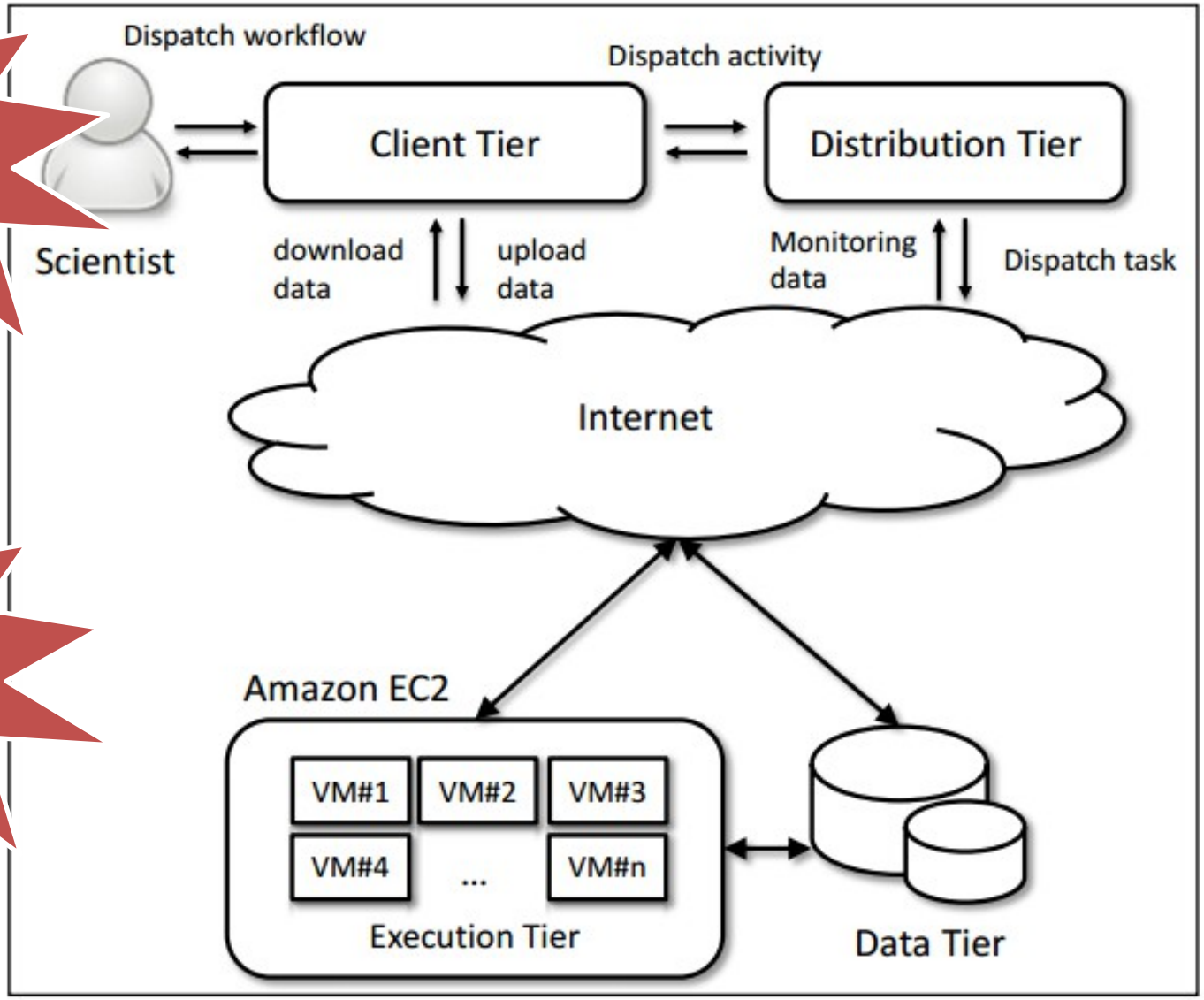
SciCumulus Workflow Engine

- Workflow engine that manages the parallel execution of workflow activities in cloud environments
- Parameter sweep
- Data parallelism
- Static and adaptive modes

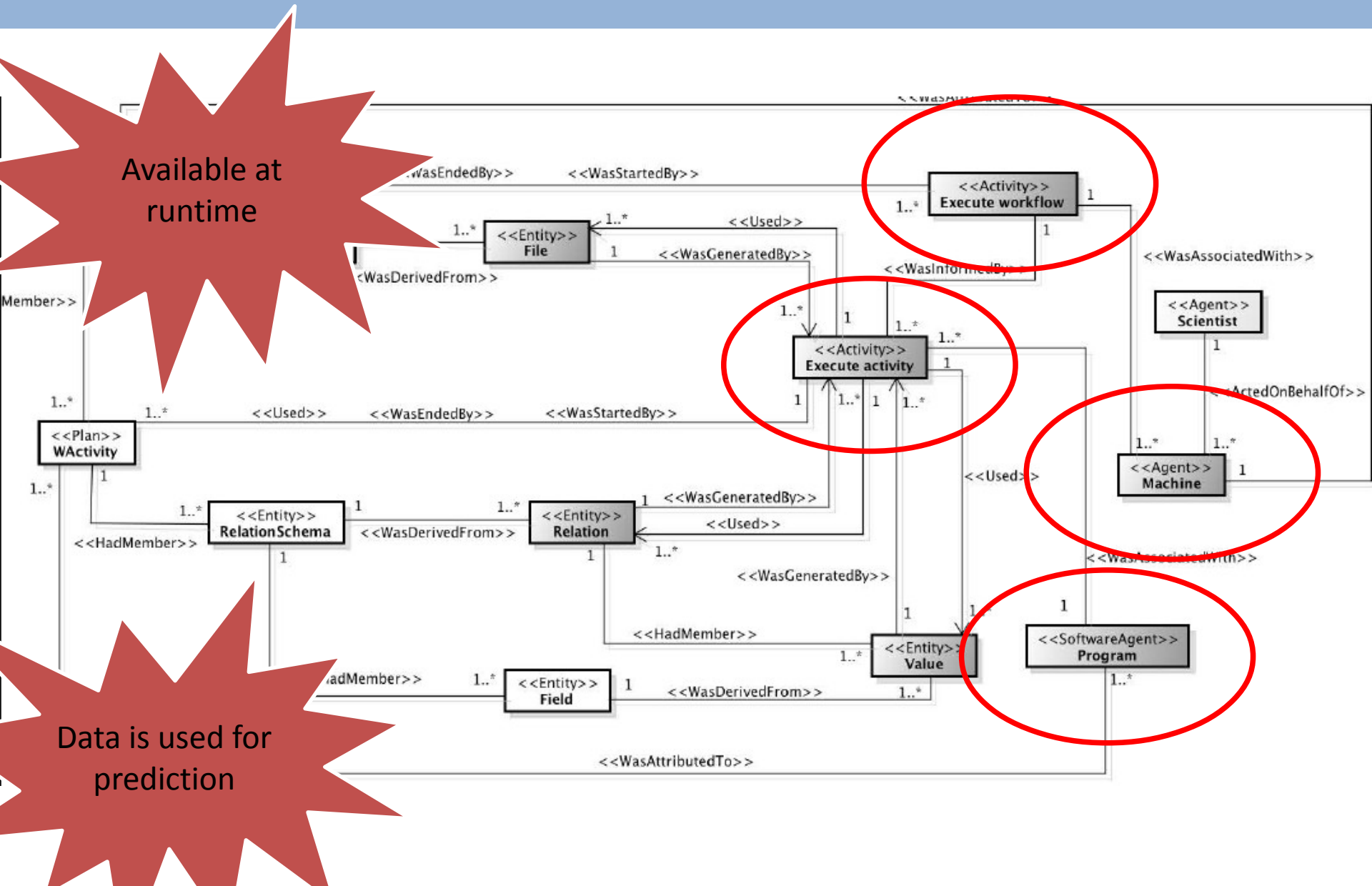
SciCumulus Architecture

Static and adaptive Mode

May be coupled to VisTrails



SciCumulus Provenance Model



SciDim

- SciDim is an optimization component coupled to SciCumulus
- Responsible for evaluating several configurations of the environment to choose the most suitable for a specific scientific workflow
- The main objective is to inform SciCumulus what types of virtual machines and how many virtual machines of each type must be instantiated

SciDim

- It is based on genetic algorithms metaheuristic
- GA are relatively simple and fast algorithms
- Indicated for cases where the search space is large and the problem does not require the optimal solution
- Modeled using JGAP framework

Chromossome Structure

- Each chromossome is composed by four genes
- Each gene is associated to a specific type of virtual machine
- We used Amazon EC2 model:
  - Micro
  - Large
  - Extra-large
  - Extra-large high capacity

3 Objective Fitness Function

- The cost function used in SciDim is the 3-Objective weighted cost function proposed by Oliveira *et al.*
- The cost function is used to simulate the entire execution of the workflow to dimension the cloud.

$$f(ca_i, VM_j) = \alpha_1 \times P(ca_i, VM_j) + \alpha_2 \times R(ca_i, VM_j) + \alpha_2 \times M(\varphi)$$

$$g(x) = 1 - \left\lceil \frac{\left(\sum_{i=1}^{n} f(cai, vmj)\right)}{|CA|} \right\rceil$$

3 Objective Fitness Function

- For each value of g(x) we have a total execution time and financial cost associated
- These values have to be compared with the constraints set by the scientist for deadline and budget.

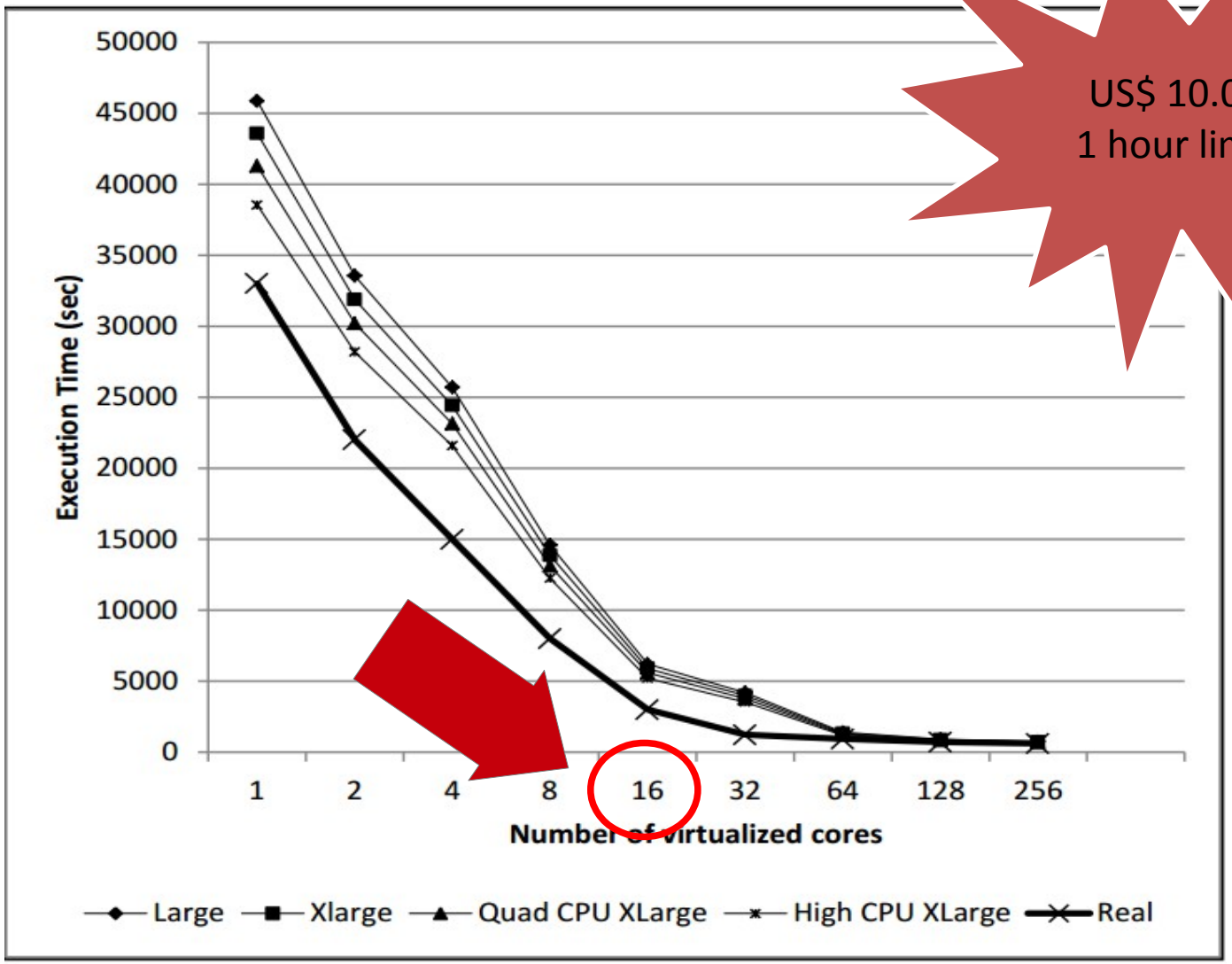
GA Execution

- We used standard values widely used in similar problems in the literature.
- The initial population of chromosomes : 100 indiv.
- 200 generations since we cannot consume a long time in the environment configuration
- Crossing rate: 35%
- Rate of mutation of individuals: 8%
- Default values for JGAP

SciDim Calibration

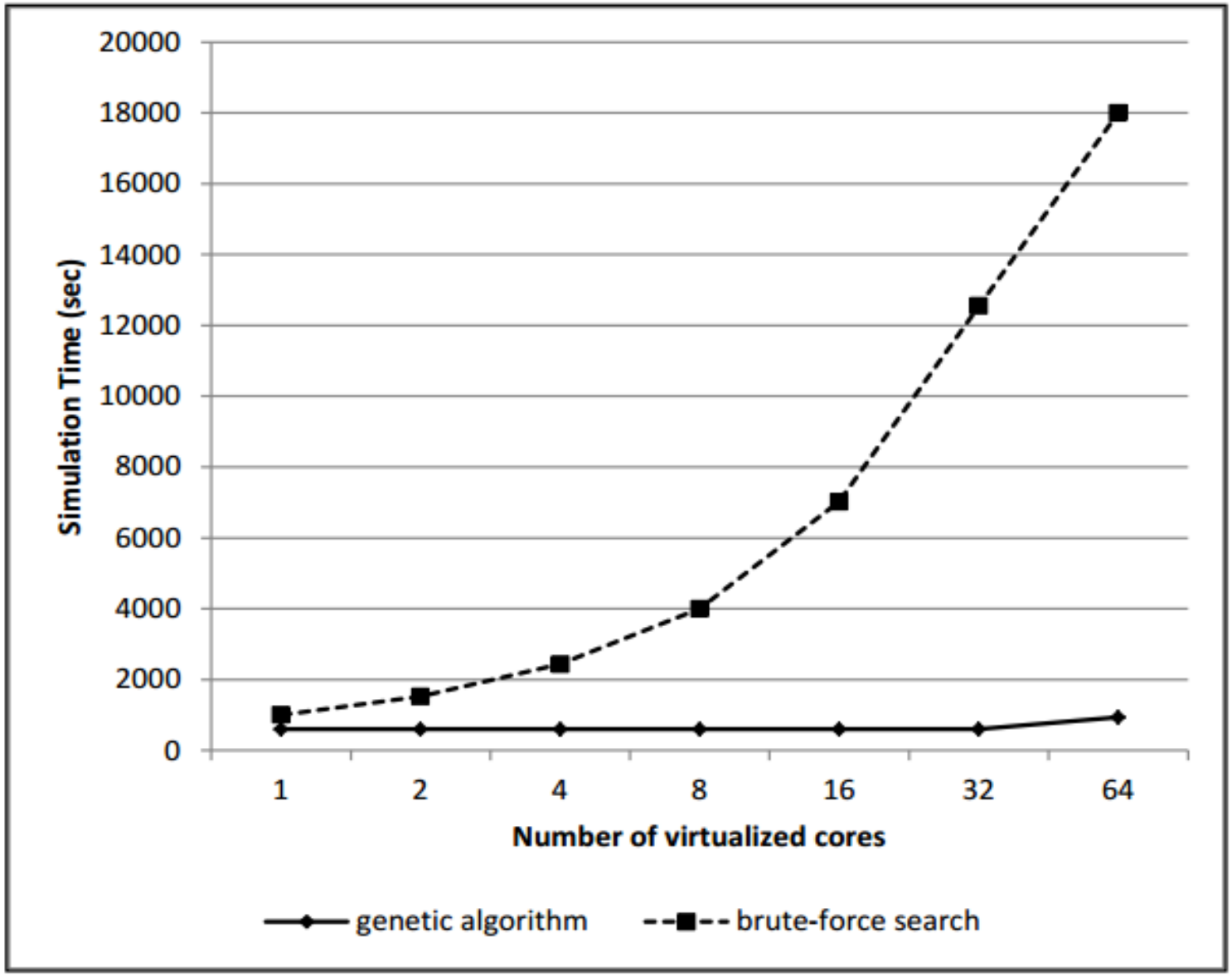
- We used X-ray Cristalography scientific workflow (proposed for provenance challenge series)
- In this calibration, the simulations were performed using only one virtual machine type
- Each execution processed 2,000 input images separated into groups of 3 images, generating a set of 667 activities to be processed in parallel.
- We used 356 previous executions for calibration
  - Provenance in practice!

SciDim Calibration

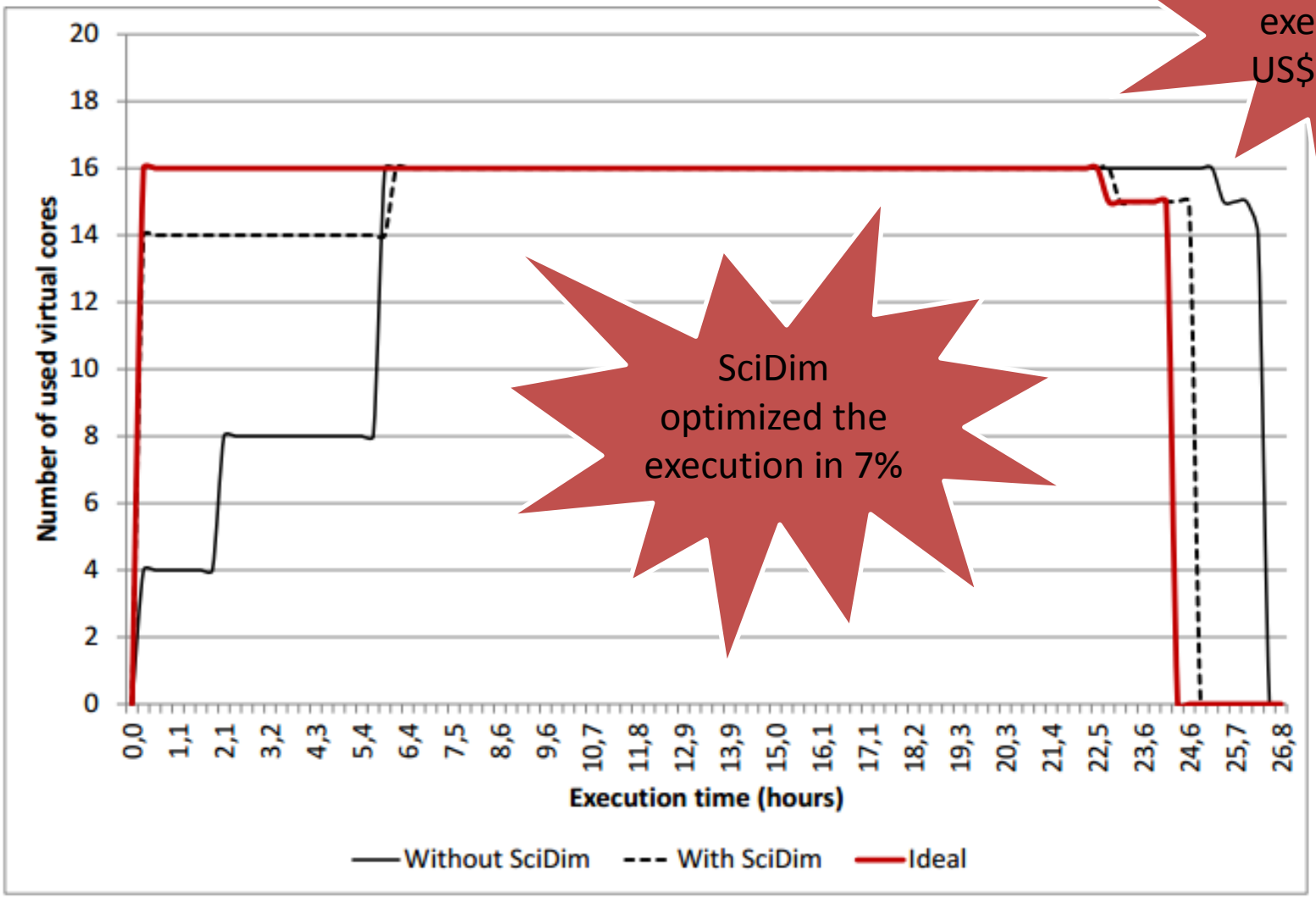


US$ 10.00
1 hour limit

Brute-force Comparison



Using Adaptive Mode in SciCumulus



SciPhy workflow executed wit US$ 75.00 lim

SciDim optimized the execution in 7%

Conclusions

- Dimensioning the cloud environment can reduce total execution time and financial costs involved
- SciDim is used for dimensioning the amount of resources to allocate for a parallel workflow
- Using genetic algorithms SciDim was able to reduce the total execution time of a real workflow in SciCumulus in 2 hours
- It needed 5 minutes to generate the initial dimensioning of the virtual cluster

Conclusions (2)

- Future work includes the implementation of new metaheuristics in SciDim such as .

Acknowledgments

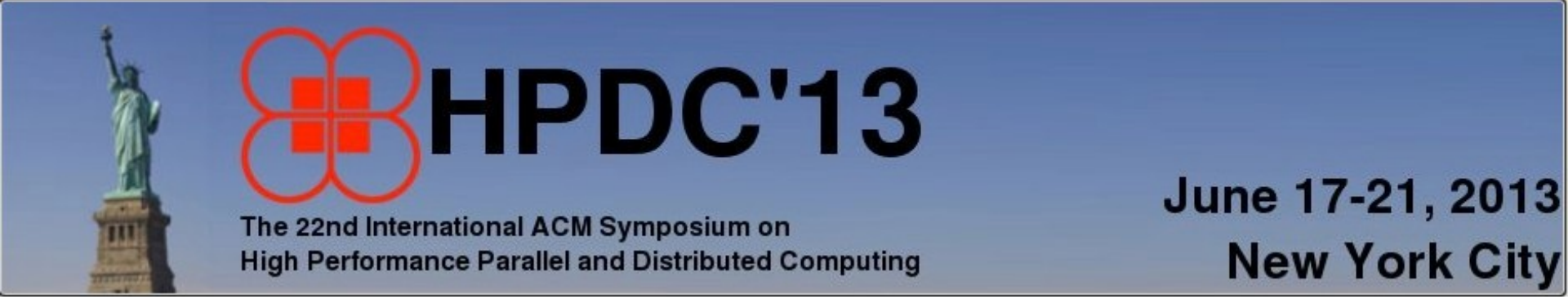
CNPq
Conselho Nacional de Desenvolvimento Científico e Tecnológico

COPPE UFRJ
Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia

CAPES

Instituto de Computação UFF

FAPERJ
Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro



Thank you!

Daniel de Oliveira, Kary Ocaña, Eduardo Ogasawara, Jonas Dias
João Gonçalves and Marta Mattoso

danielcmo@ic.uff.br

Download: http://sourceforge.net/projects/scicumulus/