MIDLAB Middleware Laboratory

Università di Roma "La Sapienza"
Dipartimento di Ingegneria Informatica,
Automatica e Gestionale

Resource location and scheduling issues in cloud/grid computing

Adriano Cerocchi

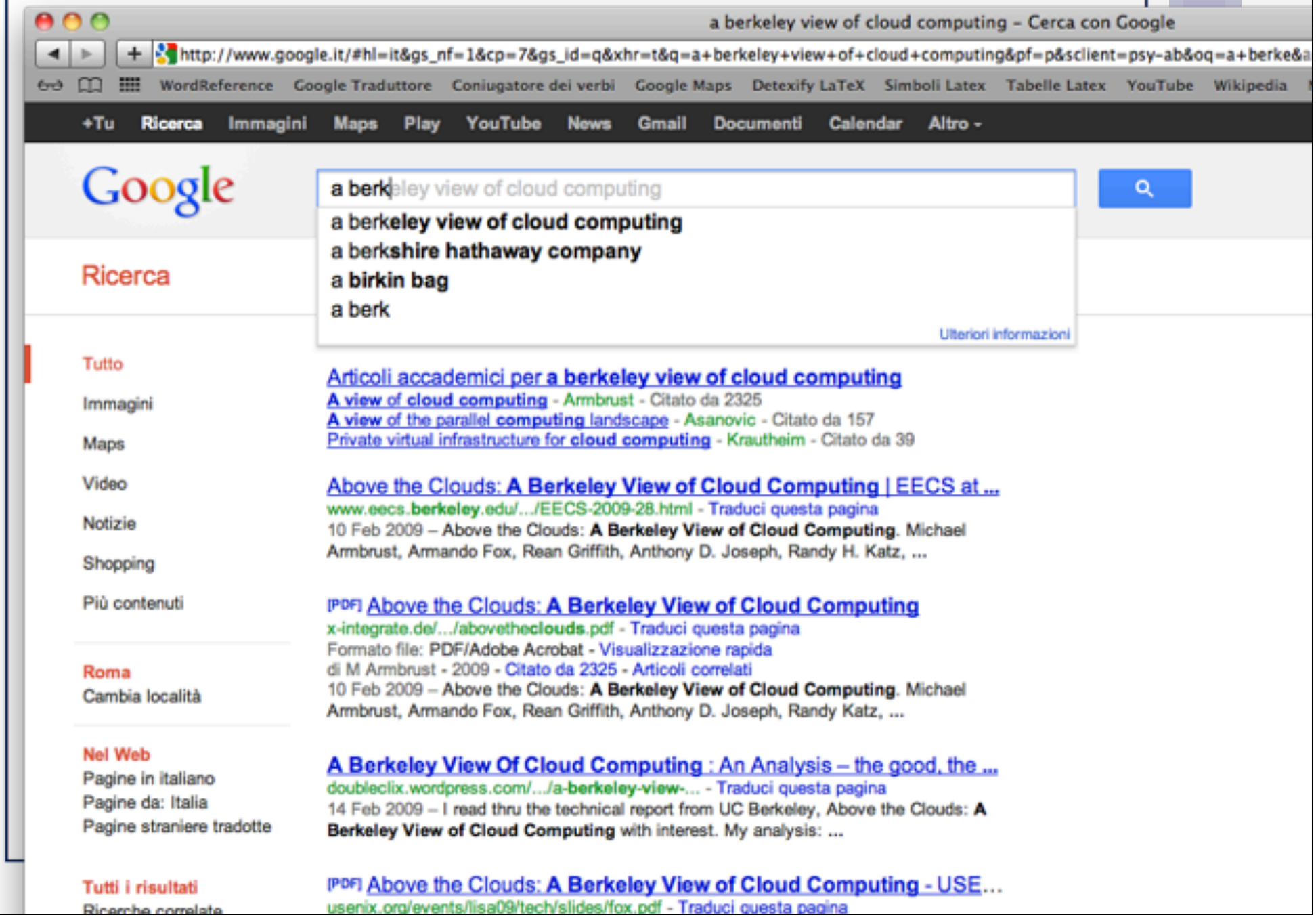
cerocchi@dis.uniroma1.it

www.dis.uniroma1.it/~cerocchi

Sapienza University of Rome

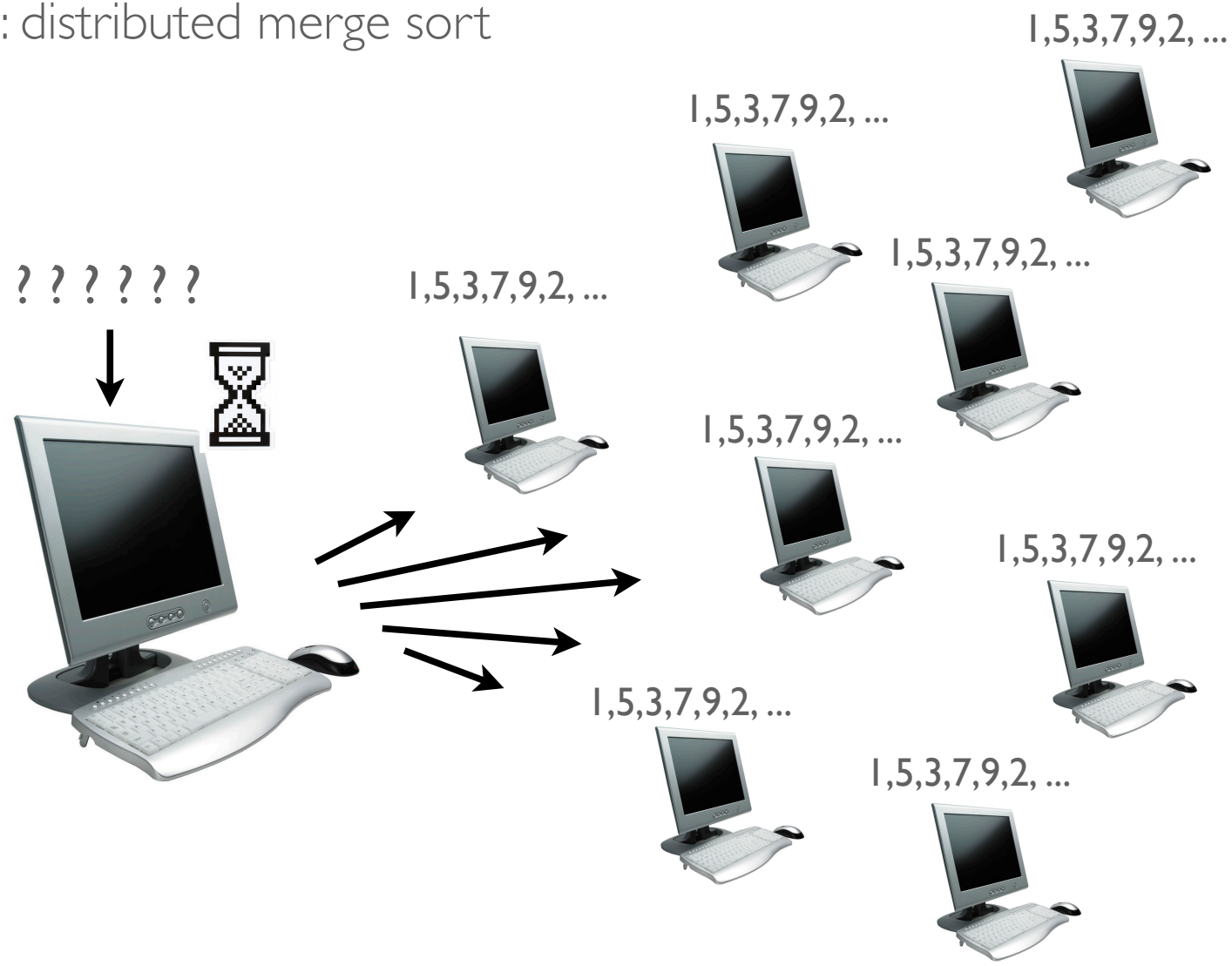
- Seminar Schedule:
  - Lecture 1 (G. di Luna) - privacy, security and cloud
  - Lecture 2 (A. Cerocchi) - resource location and scheduling issues in cloud computing
    - components of a grid computing systems
    - schedulers
    - the resource location component
    - the importance of the locality awareness - the delayed scheduler example
    - from centralised to distributed solution
    - preliminary results on distributed volunteer computing
    - conclusions
  - Lecture 3 (L. Aniello) - storage management in cloud computing

- Forewords...

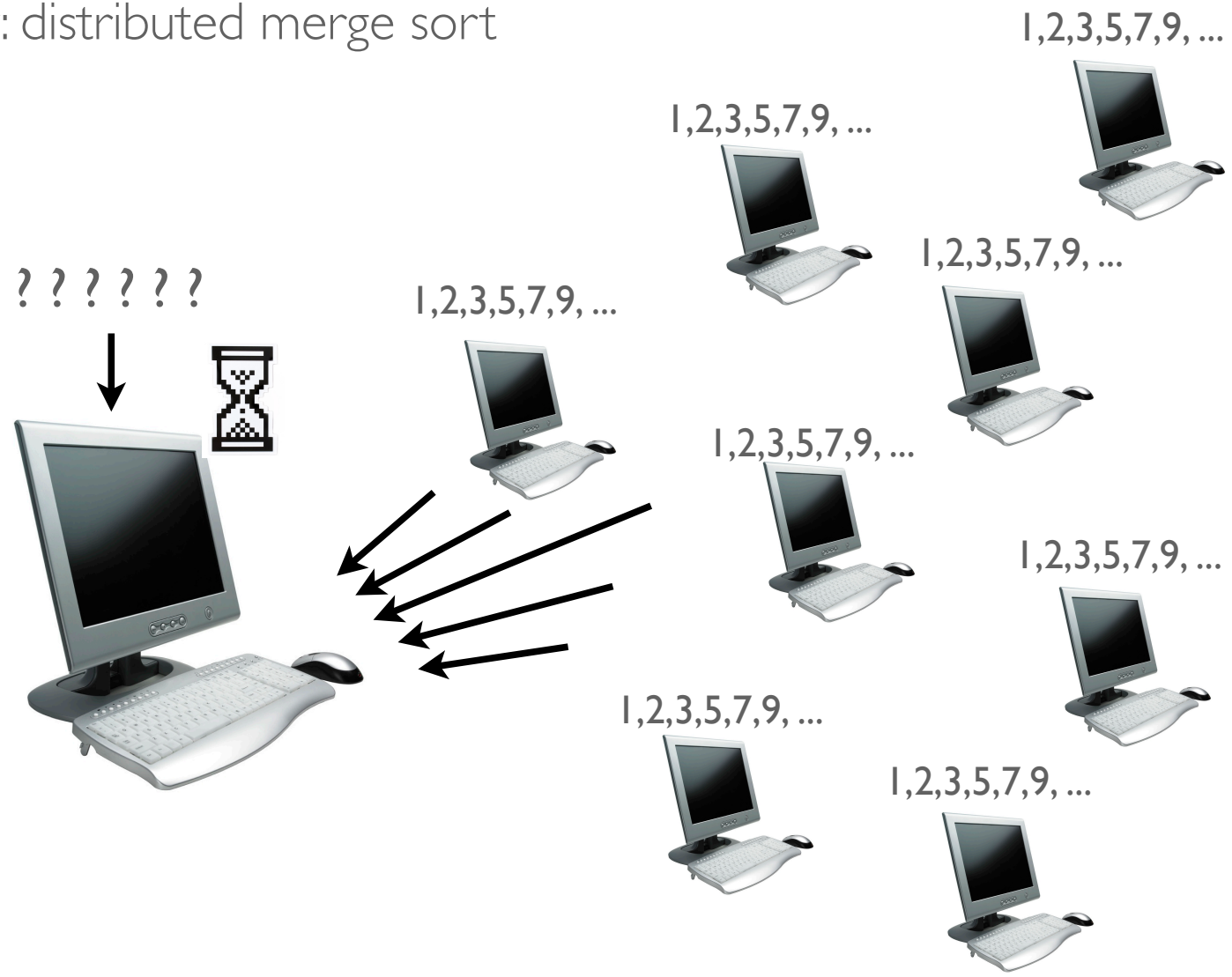

a berkeley view of cloud computing - Cerca con Google
http://www.google.it/#hl=it&gs_nf=1&cp=7&gs_id=q&xhr=t&q=a+berkeley+view+of+cloud+computing&pf=p&sclient=psy-ab&oq=a+berke&a
WordReference Google Traduttore Coniugatore dei verbi Google Maps Detexify LaTeX Simboli Latex Tabelle Latex YouTube Wikipedia
+Tu Ricerca Immagini Maps Play YouTube News Gmail Documenti Calendar Altro
Google
a berkeley view of cloud computing
a berkeley view of cloud computing
a berkshire hathaway company
a birkin bag
a berk
Ulteriori informazioni
Ricerca
Tutto
Immagini
Maps
Video
Notizie
Shopping
Più contenuti
Roma
Cambia località
Nel Web
Pagine in italiano
Pagine da: Italia
Pagine straniere tradotte
Tutti i risultati
Articoli accademici per a berkeley view of cloud computing
A view of cloud computing - Armbrust - Citato da 2325
A view of the parallel computing landscape - Asanovic - Citato da 157
Private virtual infrastructure for cloud computing - Krautheim - Citato da 39
Above the Clouds: A Berkeley View of Cloud Computing | EECS at ...
www.eecs.berkeley.edu/.../EECS-2009-28.html - Traduci questa pagina
10 Feb 2009 – Above the Clouds: A Berkeley View of Cloud Computing. Michael Armbrust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy H. Katz, ...
[PDF] Above the Clouds: A Berkeley View of Cloud Computing
x-integrate.de/.../abovetheclouds.pdf - Traduci questa pagina
Formato file: PDF/Adobe Acrobat - Visualizzazione rapida
di M Armbrust - 2009 - Citato da 2325 - Articoli correlati
10 Feb 2009 – Above the Clouds: A Berkeley View of Cloud Computing. Michael Armbrust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy Katz, ...
A Berkeley View Of Cloud Computing : An Analysis – the good, the ...
doubleclix.wordpress.com/.../a-berkeley-view-... - Traduci questa pagina
14 Feb 2009 – I read thru the technical report from UC Berkeley, Above the Clouds: A Berkeley View of Cloud Computing with interest. My analysis: ...
[PDF] Above the Clouds: A Berkeley View of Cloud Computing - USE...
usenix.org/events/lisa09/tech/slides/fox.pdf - Traduci questa pagina

- Components of a grid computing system - a driver example
- mission: Terabyte array ordering
- how: distributed merge sort

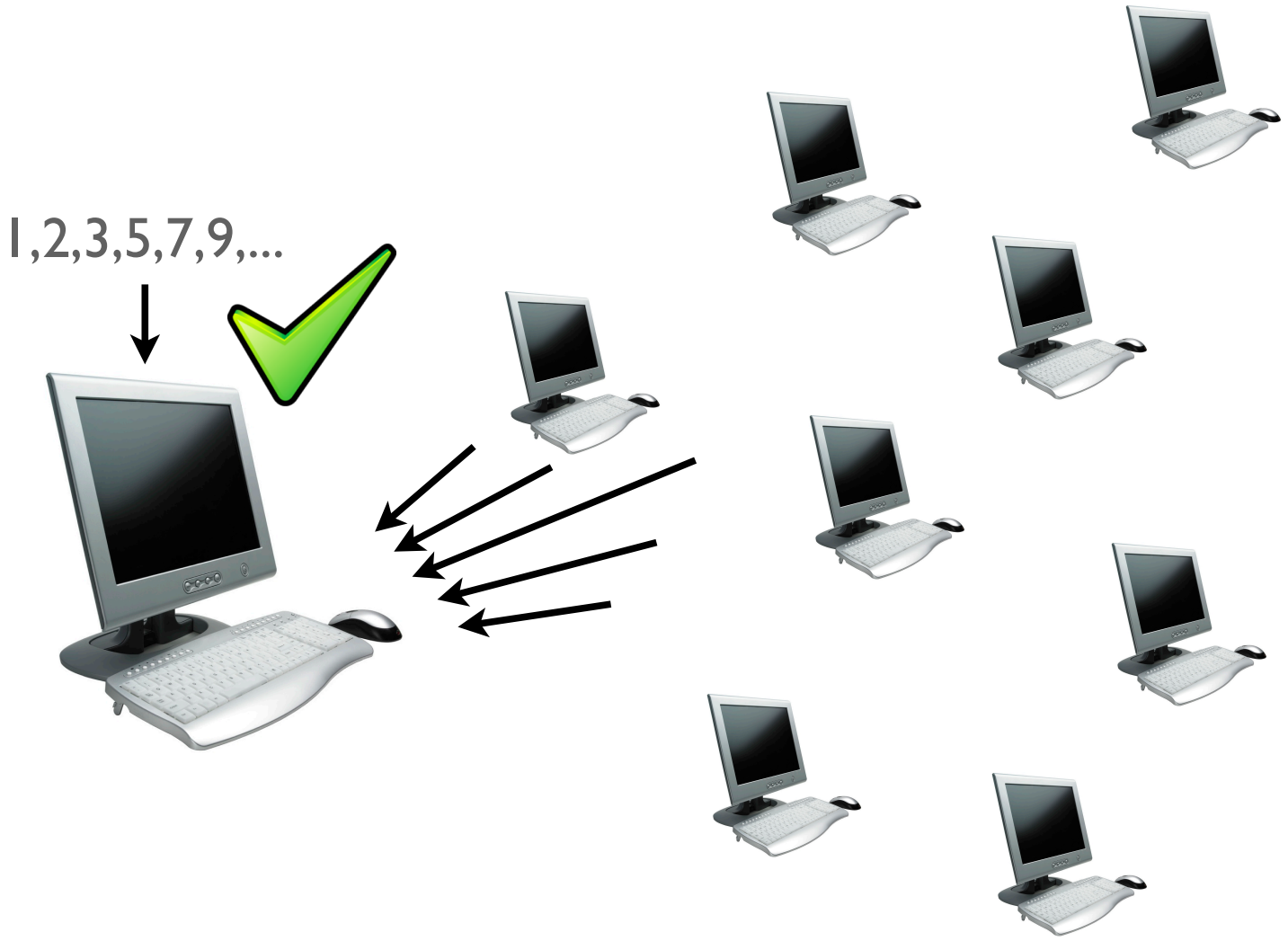
- Components of a grid computing system - a driver example
- mission: Terabyte array ordering
- how: distributed merge sort



- Components of a grid computing system - a driver example
- mission: Terabyte array ordering
- how: distributed merge sort

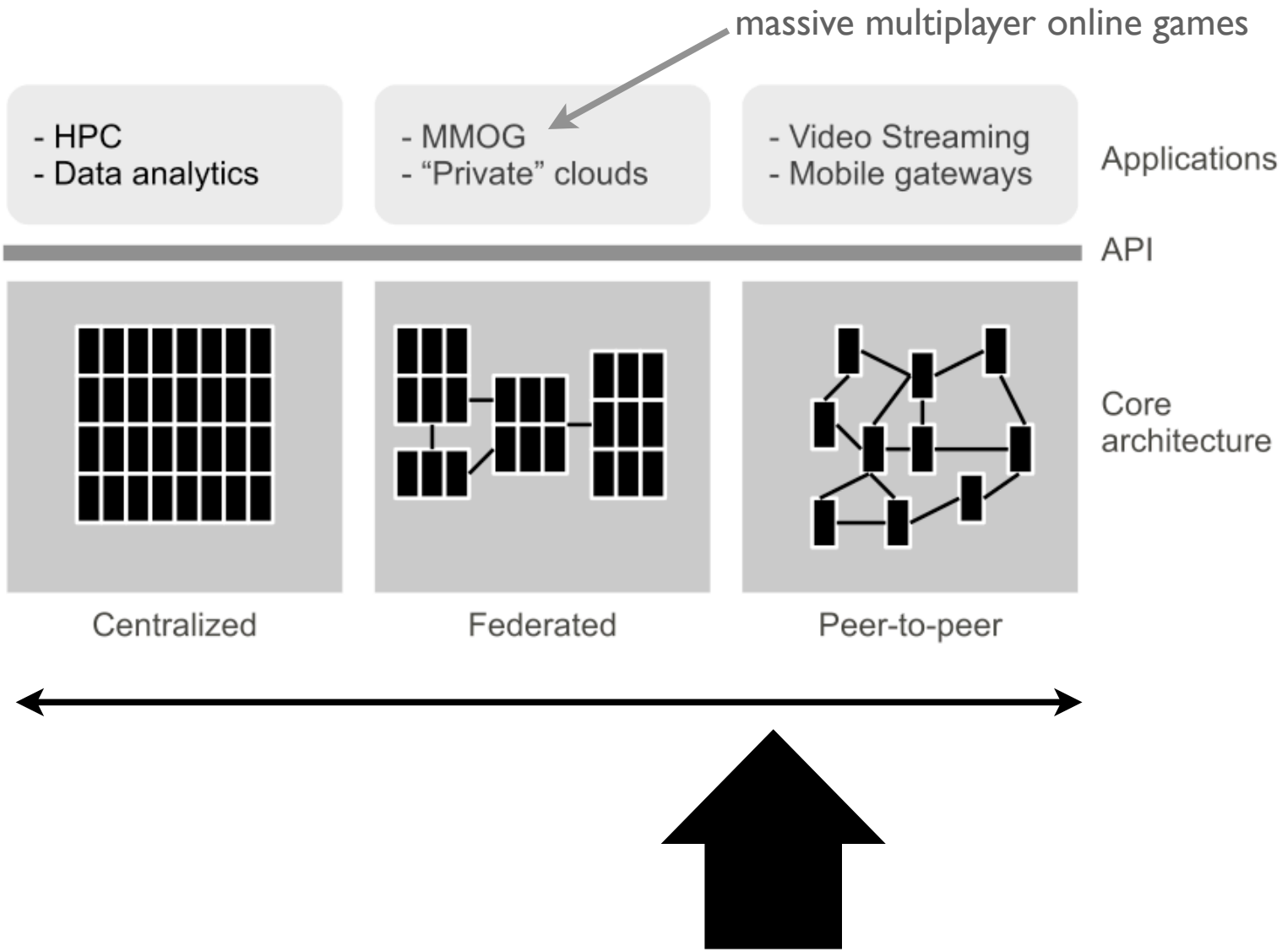


- Components of a grid computing system - a driver example
  - mission: Terabyte array ordering
  - how: distributed merge sort



- Components of a grid computing system - a driver example
- mission: Terabyte array ordering
- how: distributed merge sort
- ***discussion 1: which are the weaknesses of the proposed approach?***
- ***discussion 2: what-if the requested operation was "word count" ?***

Components of a grid computing system - overview

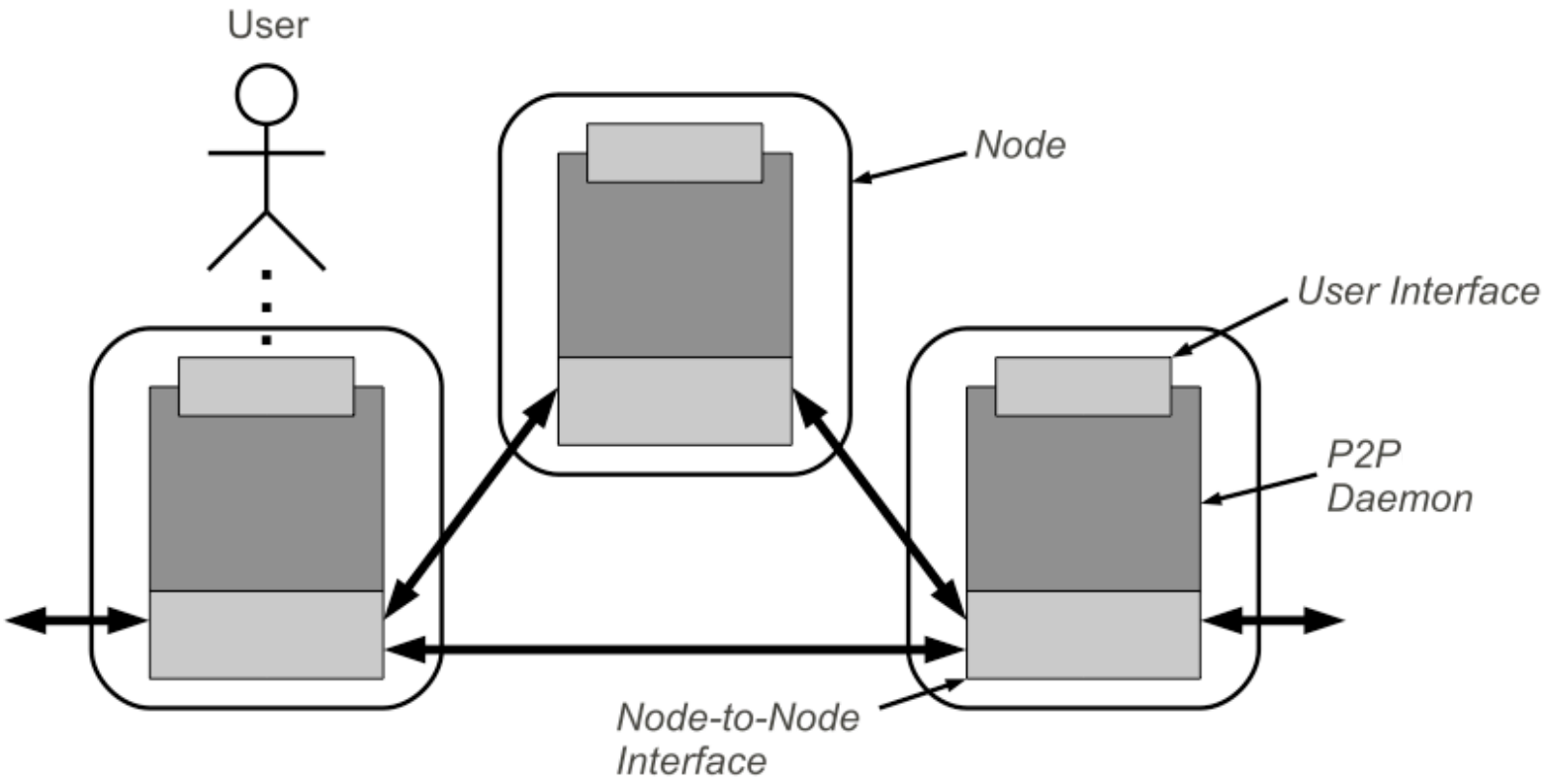


Components of a grid computing system - overview

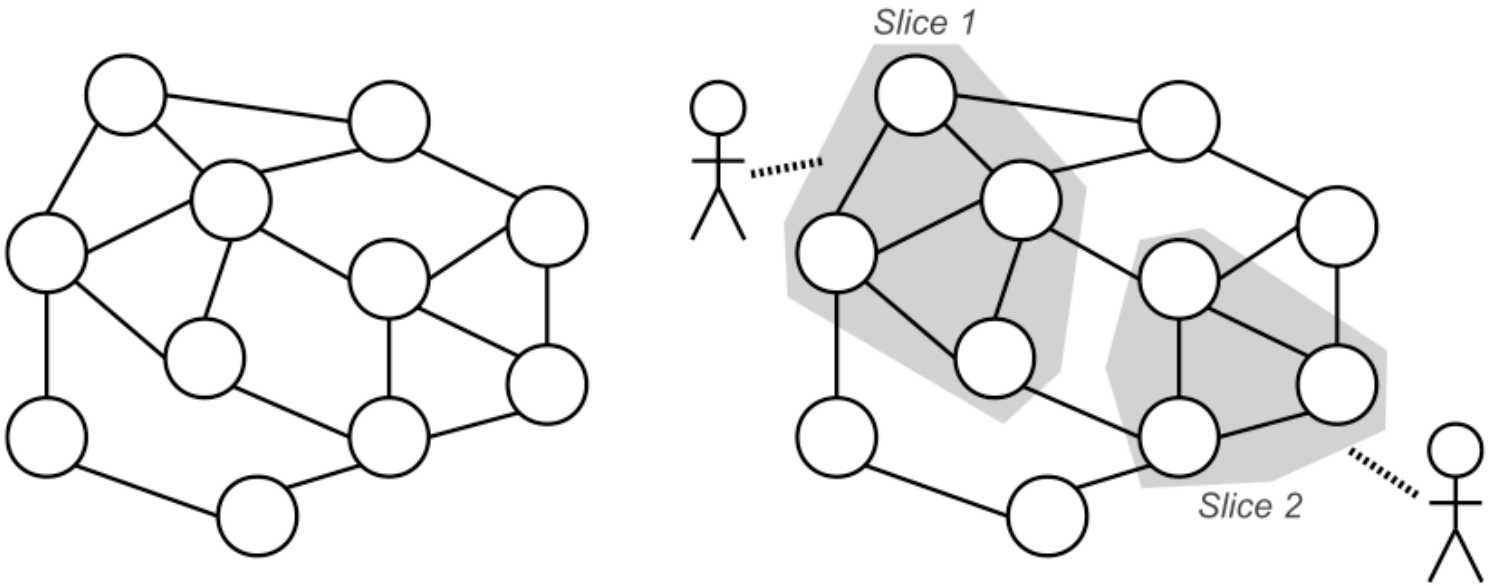
Infrastructure as a service

IaaS Cloud	**Volunteer Computing**	**P2P Cloud**
Single resource provider	Multiple resource providers	Multiple resource providers
Virtualized environment	Runs specific applications	Virtualized environment
High reliability	Unpredictable reliability	Unpredictable reliability
Local or Geographic scale	Geographic scale	Local or Geographic scale
Public, private or hybrid	Public	Public, private or hybrid

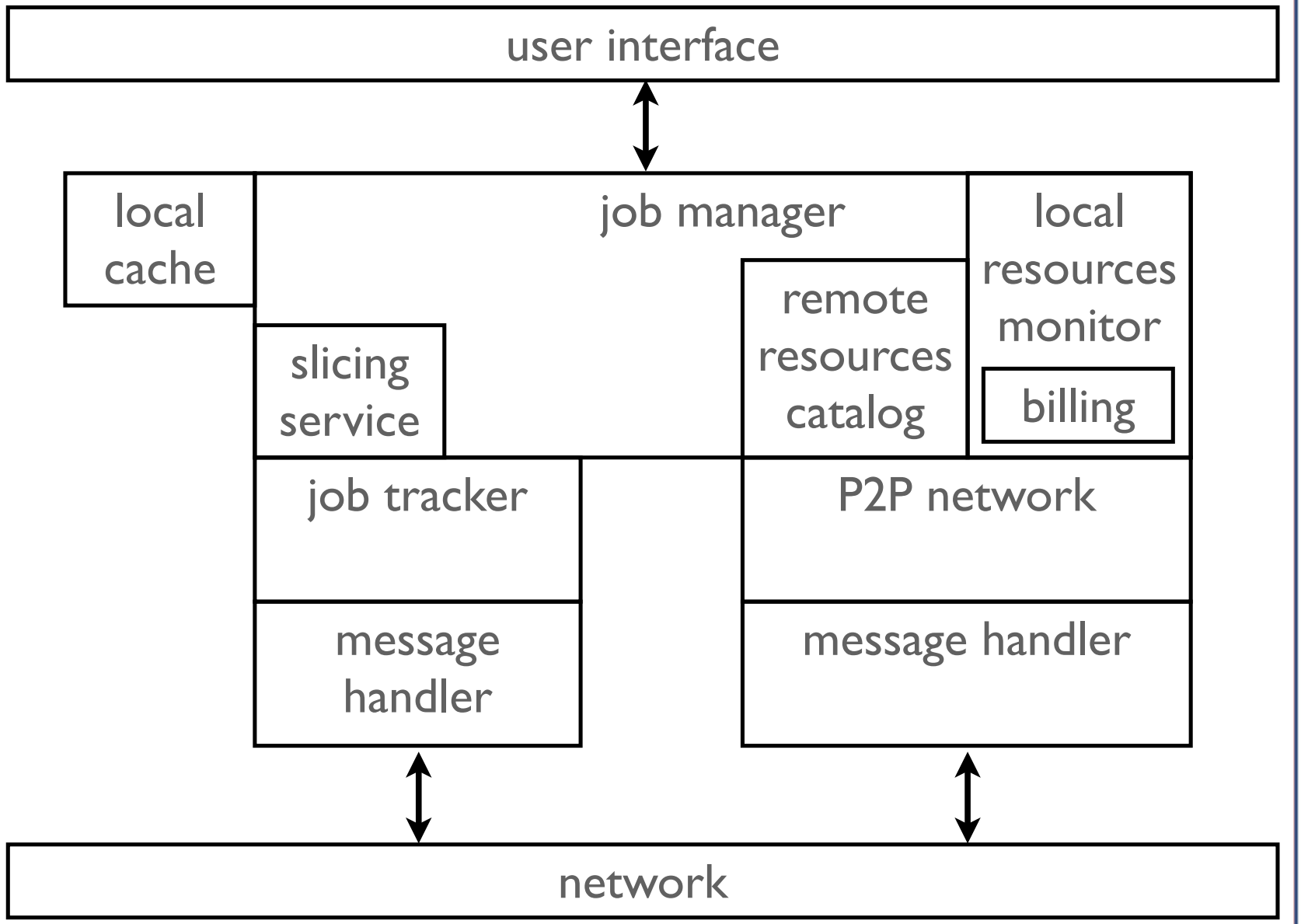
- Components of a grid computing system - overview



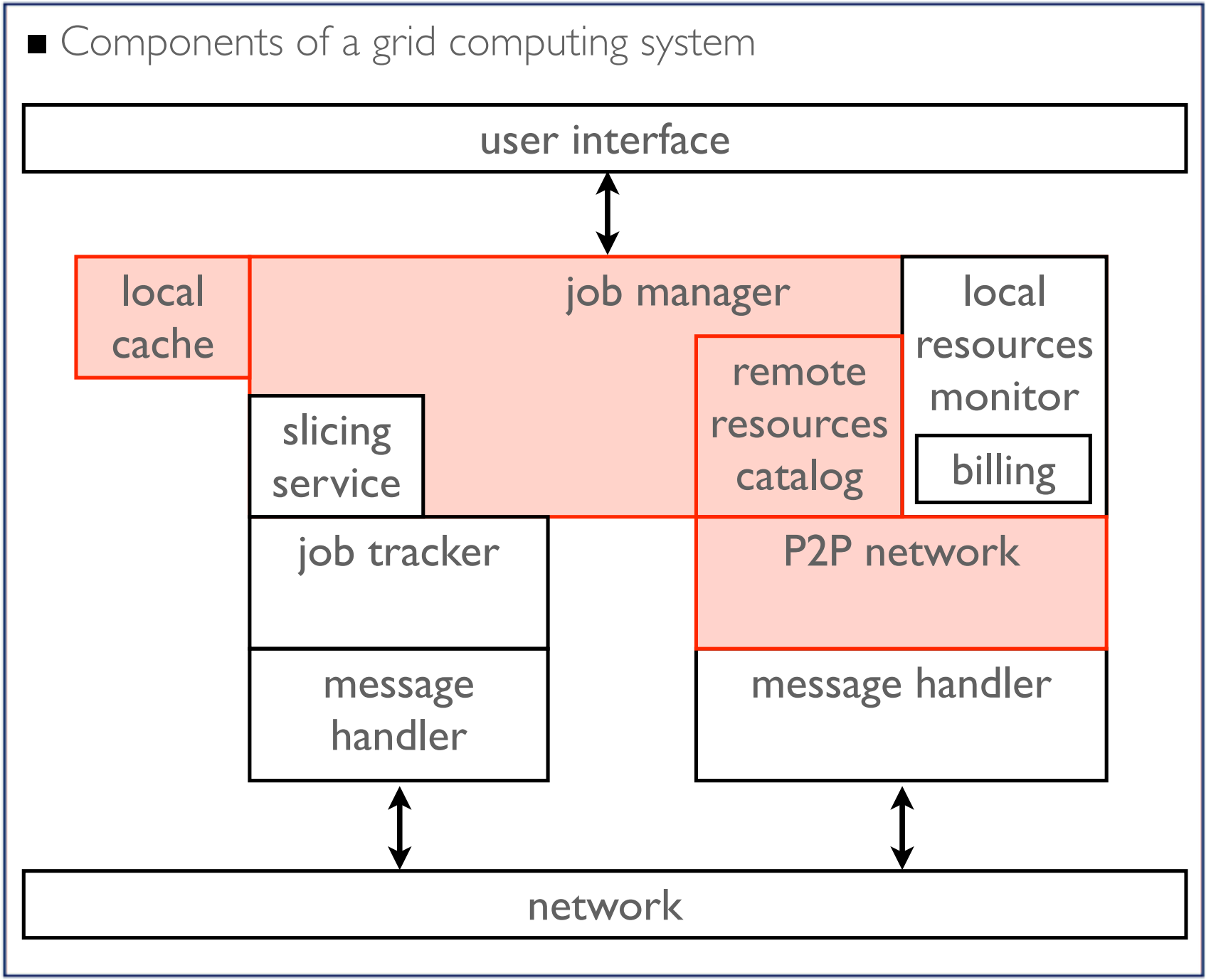
- Components of a grid computing system - overview



- Components of a grid computing system



■ Components of a grid computing system



- Schedulers
- it is a well-known concept
- like in home automation the grid (cloud) computing schedulers are a little bit more complex, they are "constrained schedulers"
- the mission of a constrained scheduler is the same of an OS-scheduler: maximising the throughput
- constraints:
  - constrained resources (i.e. CPU, memory, storage of a single machine are not independent concept, if a machine is running a CPU-bounded task (without using the storage), we cannot assume that the storage can offer the same performance of a free-of-task machine)
  - job inter-dependency: the schedule must be aware about the inter-dependency of the task and the data

- Schedulers
- other issues:
  - resource model: cpu-time? memory-size? machine-time? It depends from the architecture and the scope of the grid
  - impossibility (at the state of the art) of inter-scheduler communication: the main scheduler cannot deal with the local one

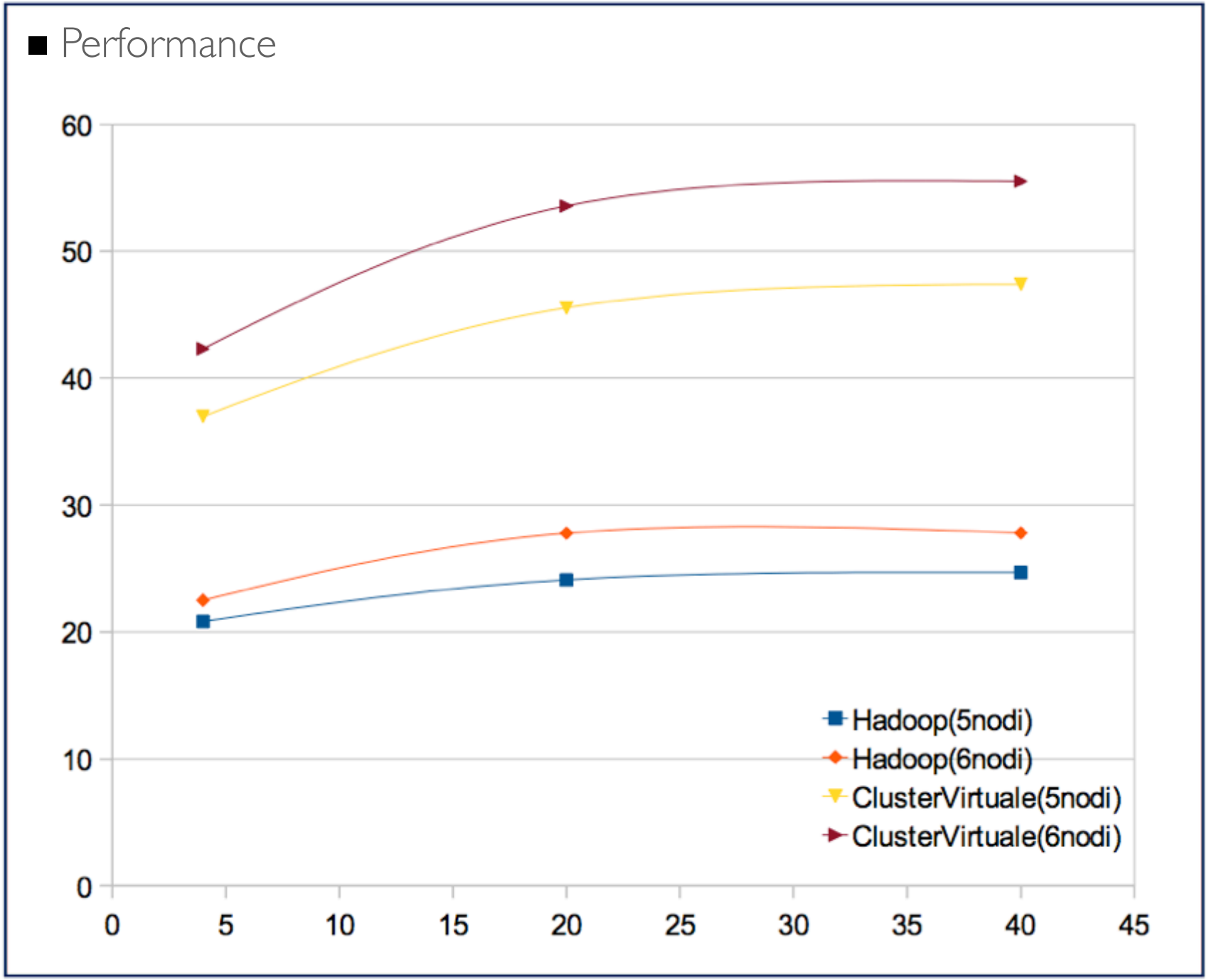
- The resource location component
- one of the main issue of a grid-computing scheduler is the knowledge of the available resources
- the scheduler and the resources are usually divided by the network then it means update delays
- the "alloc", "free" and "realloc" operation of the resources is performed during the task assigning/conclusion, i.e. there is no-change of assignment during the execution of a task
  - what-if the scheduler is not centralised?

- Stoica's presentation

- The resource location component
- if the scheduler is not centralised it is so difficult to have a clear view of the available resources
- if the number of the scheduler is fixed and well-known it is possible to apply a "replication schema" for updating the resource catalog
- if the number of the scheduler is unknown it is mandatory the usage of a mechanism of queries for aggregate and distribute the resource catalog
- the querying system must be reliable, the scheduler need to retrieve the resources that "perfectly" fit its own requirements

- Virtual Node presentation

Performance



Concluding Remarks and Projects

Grid/Cloud Computing over P2P

The huge success of the cloud computing is moving the research interest toward federation of cloud services. The extreme envision of this trend is represented by a cloud/grid computing service developed in a p2p environment. In the past, projects like Seti@home focus the attention on the possibility to exploit the wasted computation power of the computers connected to the internet (a standard user consumes less than 50% of the computational power of its machine). These solutions have mainly two limits: (i) they have a centralized control and (ii) they are suitable for a single application. Scope of the project is to investigate the state of the art in order to list the weaknesses/strengths of each approach envisioning (optional) a possible architecture for a real p2p computing.

Riferimento 1: http://www.adrianocerocchi.it/public/download/1.Volunteer%20Computing%20and%20Desktop%20Cloud%20The%20Cloud@Home%20Paradigm.pdf
Riferimento 2: http://www.adrianocerocchi.it/public/download/3.design%20and%20implementation%20of%20a%20p2p%20cloud%20system.pdf
Riferimento 3: http://www.adrianocerocchi.it/public/download/6.%20managing%20clouds,%20a%20case%20for%20a%20fresh%20look%20at%20large%20unreliable%20dynamic%20networks.pdf

Concluding Remarks and Projects

The cost of a cloud

The data centers used to create cloud services represent a significant investment in capital outlay and ongoing costs. the student is invited to examine the costs of cloud service data centers today. The cost breakdown reveals the importance of optimizing work completed per dollar/euro invested. Unfortunately, the resources inside the data centers often operate at low utilization due to resource stranding and fragmentation, moreover, a relevant variety of organizations such university, department, etc... invests a large amount of money for very small cluster (i.e. up to 16 machines) that are dramatically under-used because the limited experience of the development team and the limited needs of it. The student is invited to evaluate the cost of a grid computing platform proposing (optional) a study about the feasibility of a grid-cluster made of very low-cost machines (i.e. eee PCs or embedded devices).

Riferimento 1: http://www.adrianocerocchi.it/public/download/12.The%20Cost%20of%20a%20Cloud,%20Research%20Problems%20in%20Data%20Center%20Networks.pdf
Riferimento 2: http://www.pcengines.ch/alix3d3.htm
Riferimento 3: http://www.raspberrypi.org/faqs