A Case for the Coordination of the e-Science Ecosystem

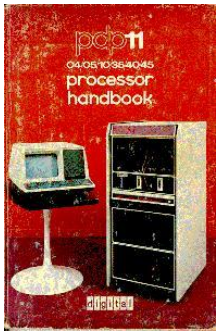
Francisco Brasileiro
fubica@dsc.ufcg.edu.br

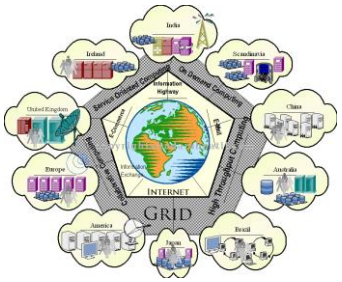
Universidade Federal de Campina Grande
Departamento de Sistemas e Computação
Laboratório de Sistemas Distribuídos

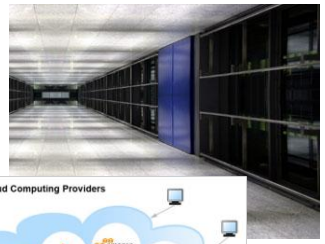
High Perfomance Scientific Computing in Clusters, Grid, and Cloud Computing Systems, Montivideo, November 8th, 2012

Universidade Federal de Campina Grande

e-Science = Science supported by ICT



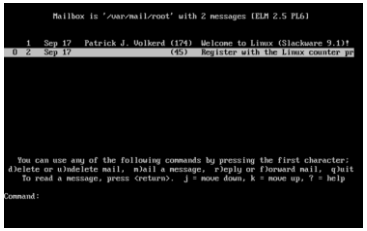


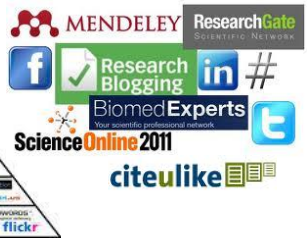


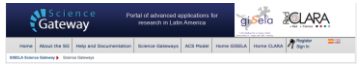
processing & storage

1940's-1970's 1960's-1980's 1980's-present 1990's-present Mid 2000's-present

communication & cooperation





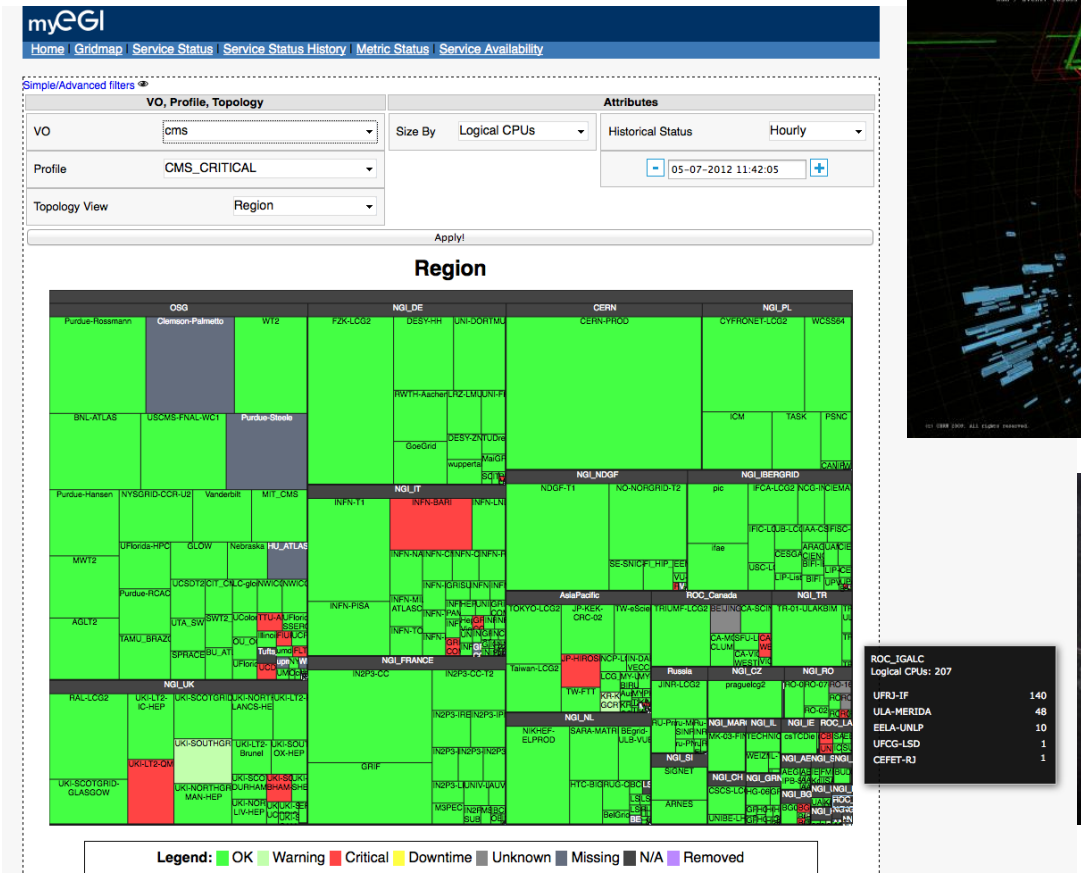


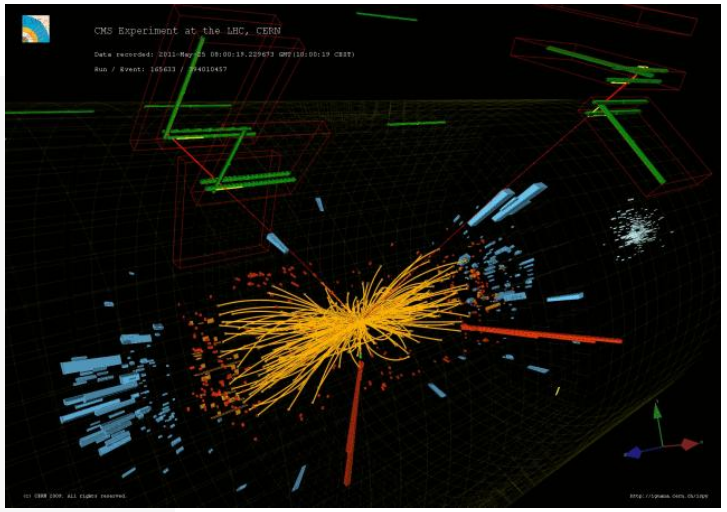
Big science

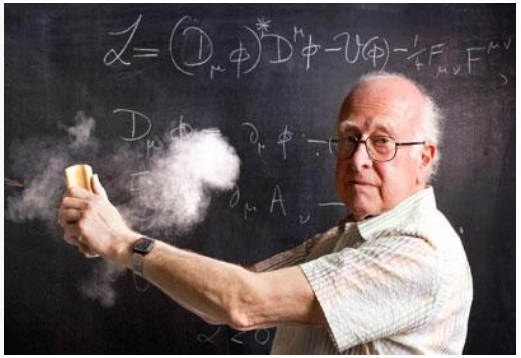
Source: Popular Science, http://www.popsci.com/

Success stories

- Higgs boson

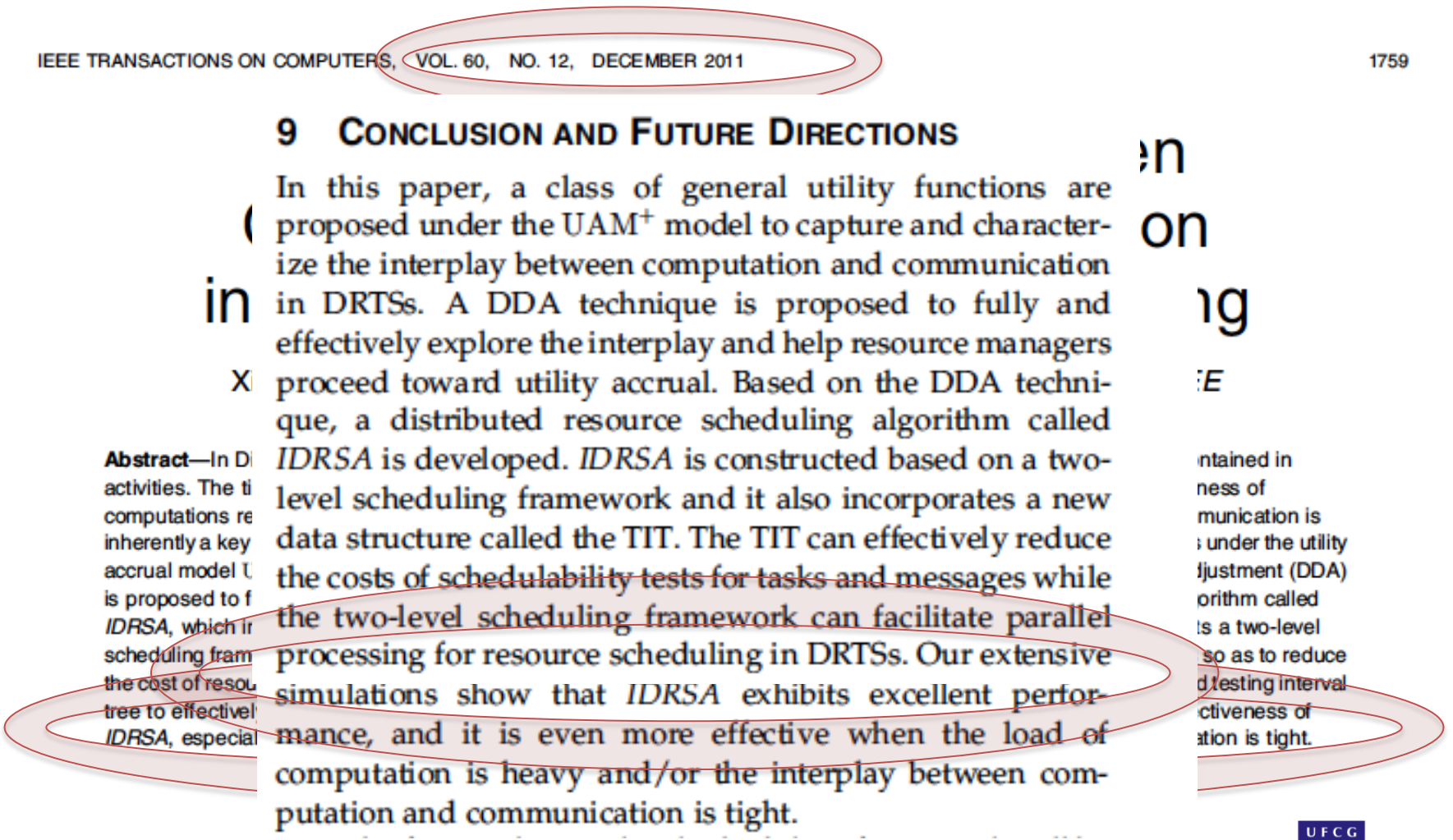






Universidade Federal de Campina Grande

What about “small” science?



IEEE TRANSACTIONS ON COMPUTERS, VOL. 60, NO. 12, DECEMBER 2011 1759

9 Conclusion and Future Directions

In this paper, a class of general utility functions are proposed under the UAM$^+$ model to capture and characterize the interplay between computation and communication in DRTSs. A DDA technique is proposed to fully and effectively explore the interplay and help resource managers proceed toward utility accrual. Based on the DDA technique, a distributed resource scheduling algorithm called *IDRSA* is developed. *IDRSA* is constructed based on a two-level scheduling framework and it also incorporates a new data structure called the TIT. The TIT can effectively reduce the costs of schedulability tests for tasks and messages while the two-level scheduling framework can facilitate parallel processing for resource scheduling in DRTSs. Our extensive simulations show that *IDRSA* exhibits excellent performance, and it is even more effective when the load of computation is heavy and/or the interplay between computation and communication is tight.

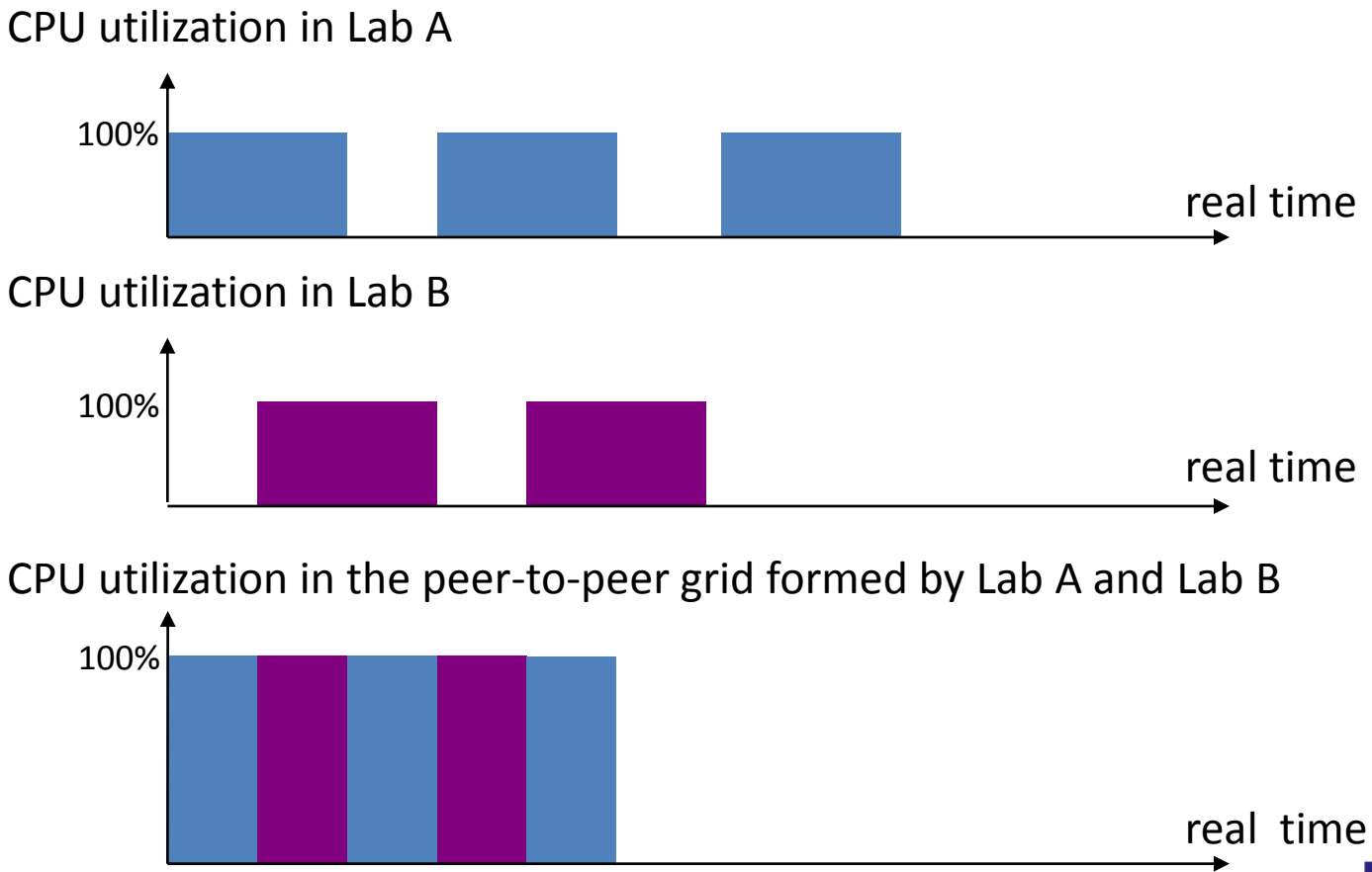
Big gap!

- How to “survive” without appropriate access to ICT?
- Not just lack of resources, but also lack of qualified scientists
  - e-Science requires a **change in mentality**
  - Access to computing resources needs to be widespread
    - if not to provide a fairer “battlefield”, **at least to support adequate training**

How can we reduce this gap?

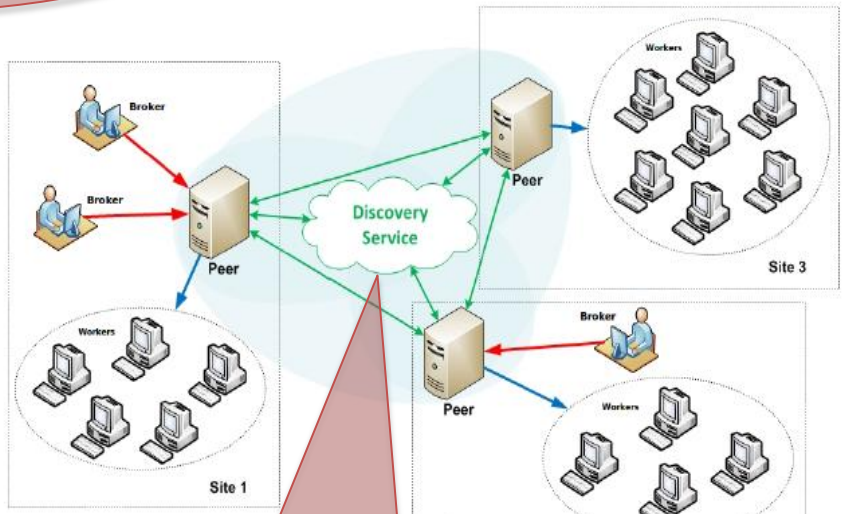
Leverage self-interest

- Peer-to-peer grids (and clouds!)



Leverage symbiosis

This is a typical cluster where BoT jobs are hated!

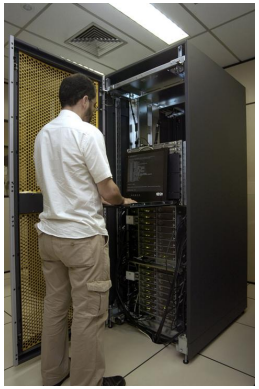


… and this is a typical opportunistic P2P grid where BoT jobs are loved!

“Render unto Caesar the things which are Caesar's, and unto God the things that are God’s” (Matthew, 22:21)

Table 1 Mean makespan for baseline scenario: Grids working independently

Grid type	Contention	Mean makespan (*s*)
Service	Low	67,882
Service	High	154,719
Opportunistic with NoF	Low	35,197
Opportunistic with NoF	High	66,559



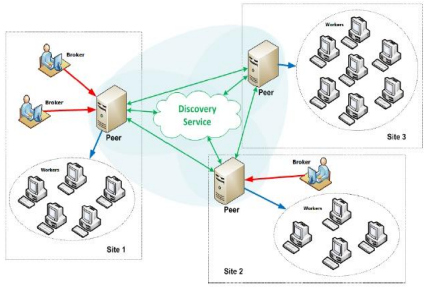
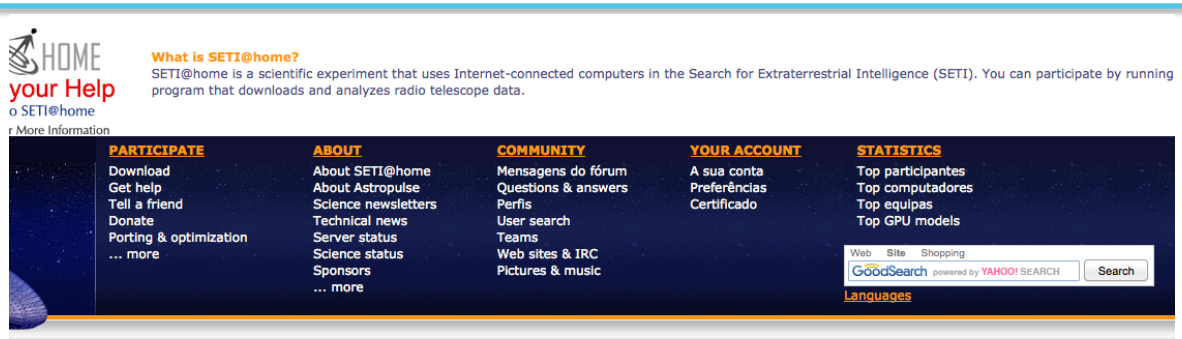


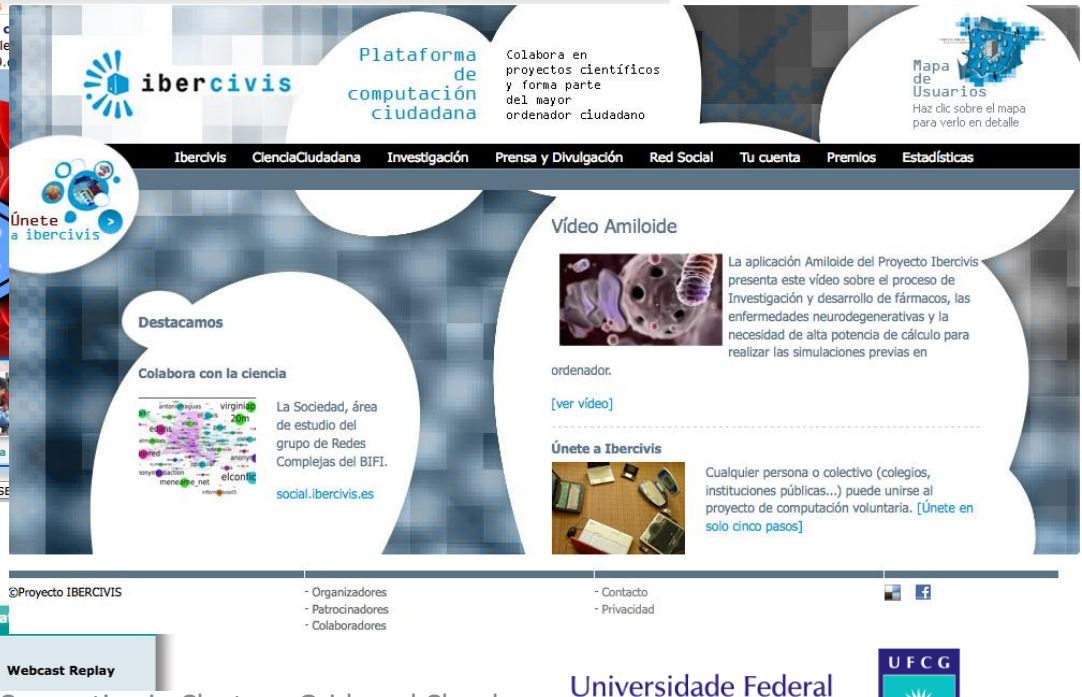
Table 5 Mean makespan for full interoperability with arbitration

Contention		Grid type			
		Service		Opportunistic	
Service	Opportunistic	Mean makespan (*s*)	Relative benefit	Mean makespan (*s*)	Relative benefit
Low	Low	43,234	1.57	36,332	0.97
Low	High	43,254	1.57	43,241	1.54
High	Low	44,268	3.50	36,955	0.95
High	High	44,595	3.47	45,145	1.47

Francisco Vilar Brasileiro, Matheus Gaudencio, Rafael Silva, Alexandre Duarte, Diego Carvalho, Diego Scardaci, Leandro Neumann Ciuffo, Rafael Mayo, Herbert Hoeger, Michael Stanton, Raul Ramos, Roberto Barbera, Bernard Marechal, Philippe Gavillet: **Using a Simple Prioritisation Mechanism to Effectively Interoperate Service and Opportunistic Grids in the EELA-2 e-Infrastructure**. Journal of Grid Computing. 9(2): 241-257 (2011).

Leverage altruism



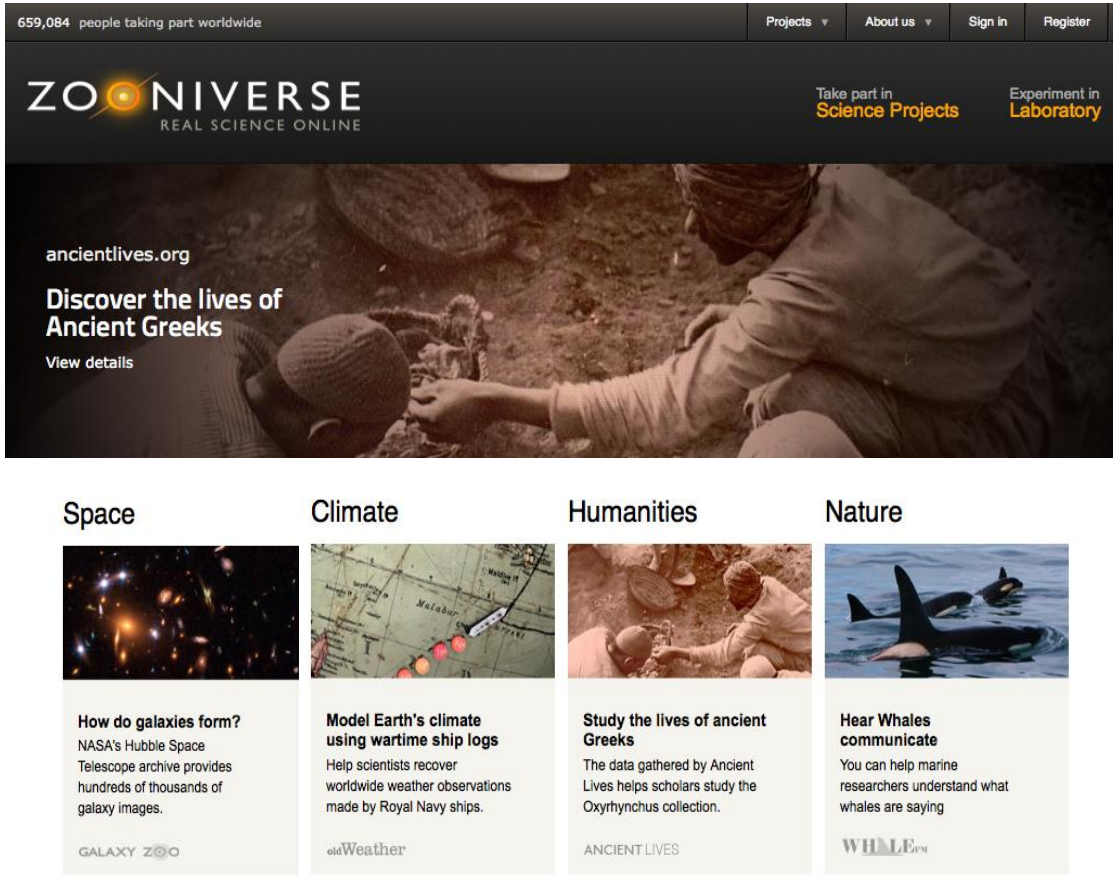


Scientific dissemination

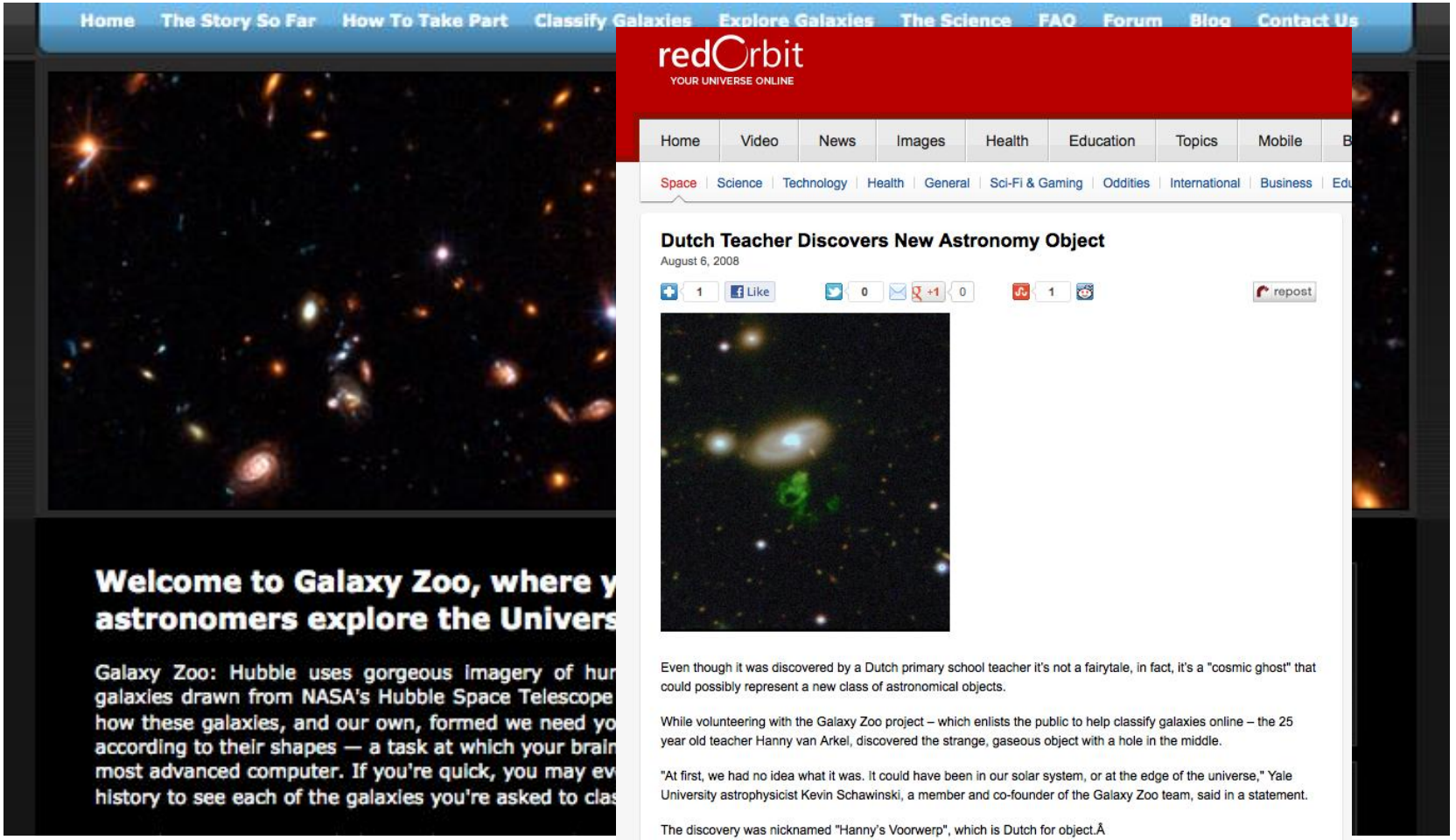
- This is a key requirement for the success of volunteer computing infrastructures
  - Gather more resources
- Create society awareness about the scientific activity
  - Increase the interest of youngsters for the scientific career
  - Get explicit scientific contribution from the citizens
    - **Citizen science**

Where's Wally?

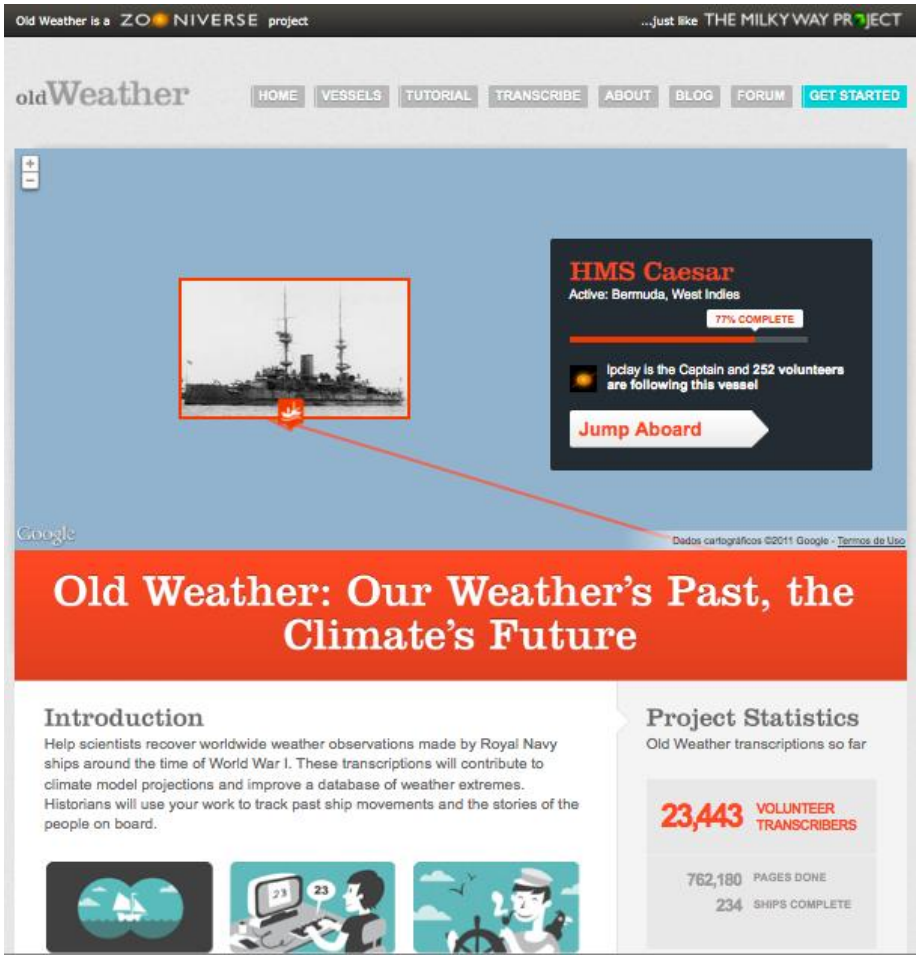
Citizen science projects



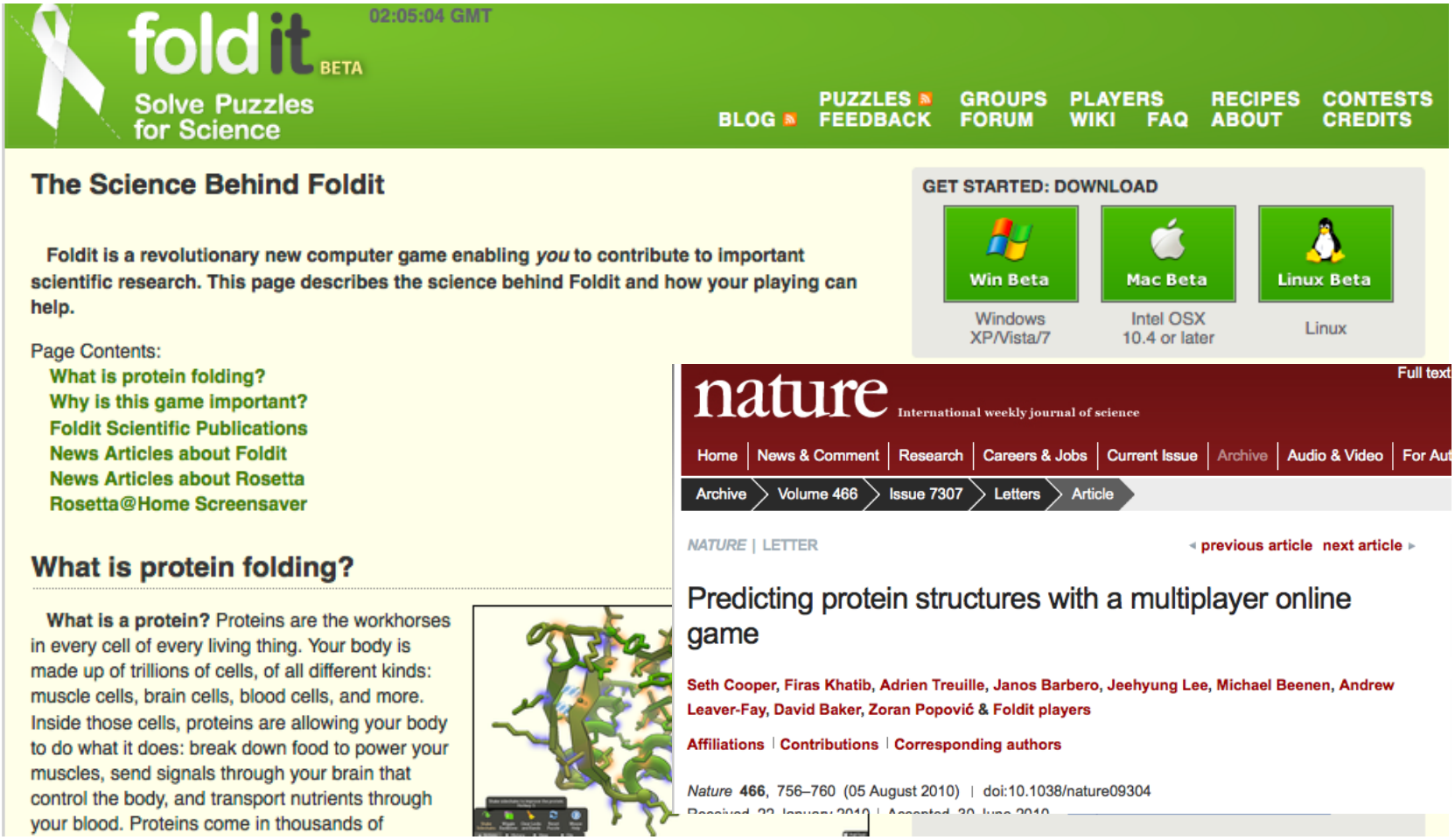
Classifying galaxies



Transcribing scientific data

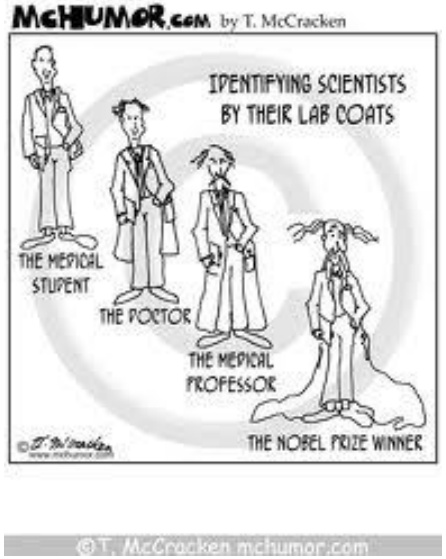


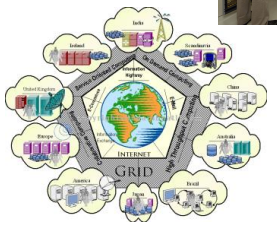
Or just having plain fun!

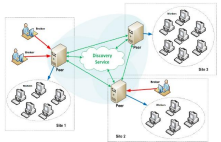


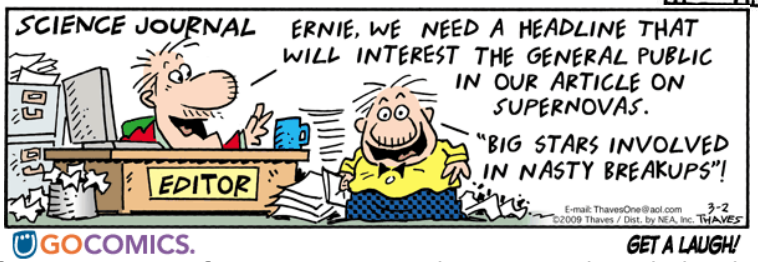
In conclusion …

- The e-Science ecosystem is rather broad and diverse









Thanks for your attention

- … and let's coordinate ourselves!