

Challenges facing HPC

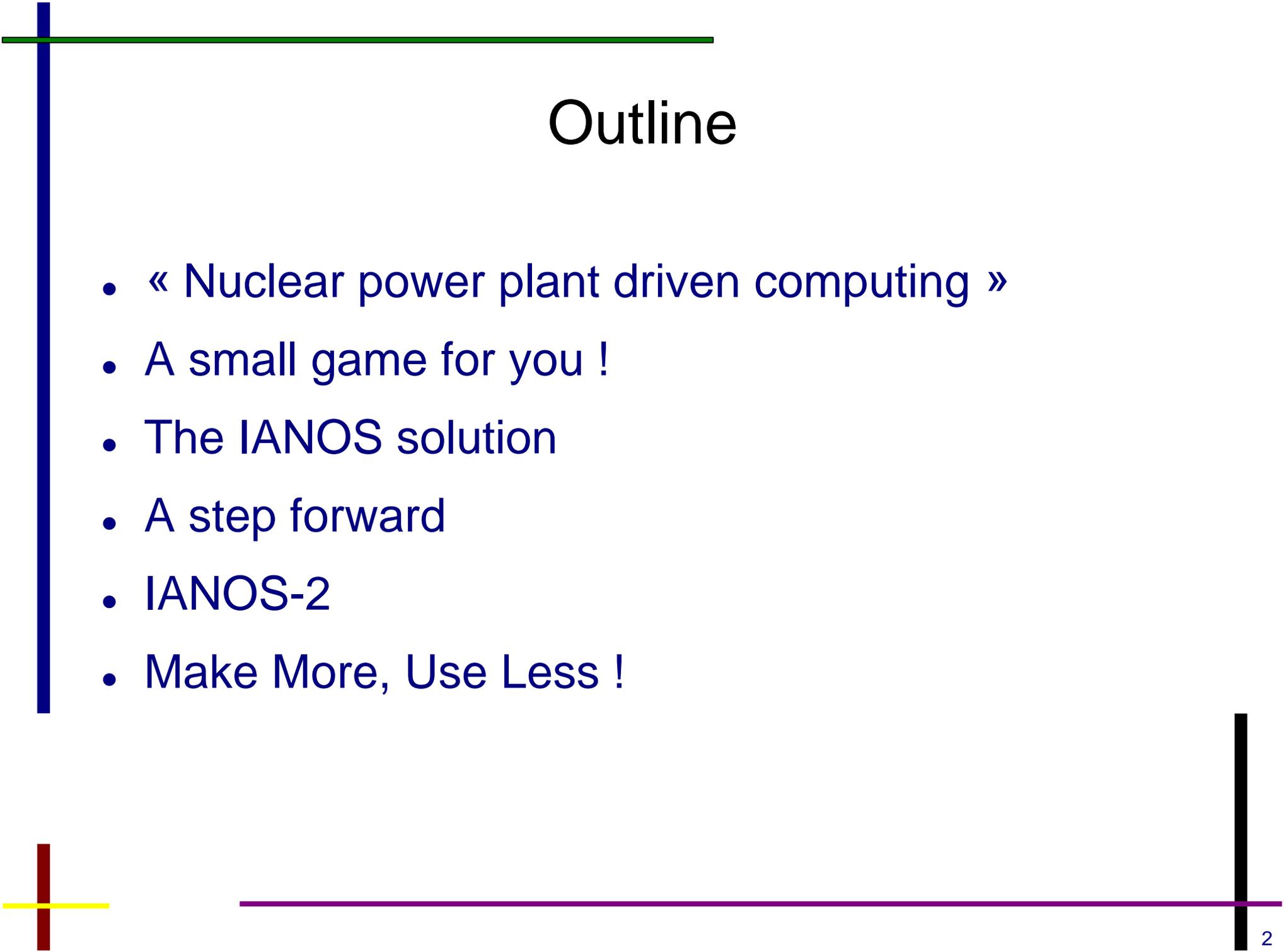
is “Green-IT” only a fad ?

Vincent Keller

Vincent.Keller@scai-extern.fraunhofer.de

University of Bonn, Germany

*SOS 13 Workshop, Hilton Head, SC, USA
March 10th, 2009*



Outline

- « Nuclear power plant driven computing »
- A small game for you !
- The IANOS solution
- A step forward
- IANOS-2
- Make More, Use Less !

Known stuff (facts) ...

- According to Koomey et al. (LBNL), data centers electric consumption worldwide :
 - equals to 16 2GW-Nuclear Power Plants in 2005 : 280 TWh
 - Half of it used for cooling
 - doubles every 5 years
- HPC machines operate 24/7/365
- Applications rarely use the resources in the best way

Observations (known)

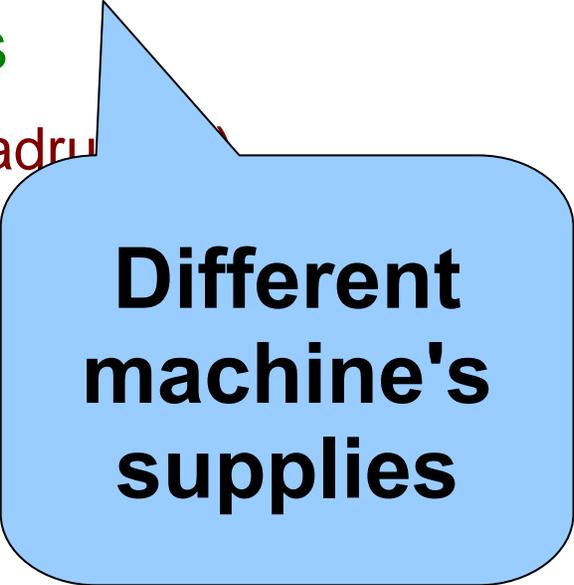
- Different types of HPC applications
 - Serial
 - Memory bound
 - CPU performance bound
 - Parallel
 - Embarrassingly parallel
 - p-2-p dominated
 - Multi-cast dominated
 - etc...



**Different
application's
needs**

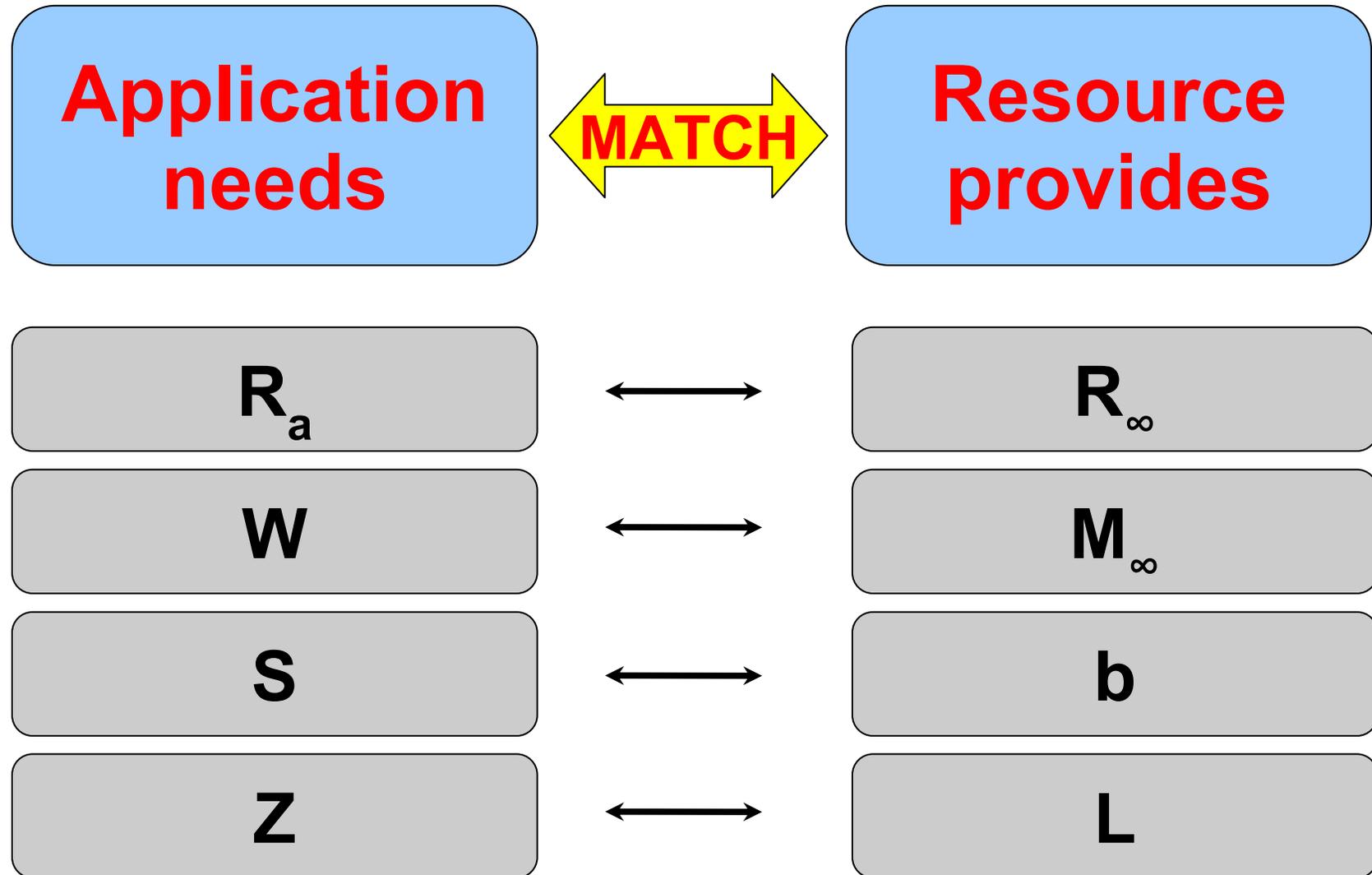
Observations (known)

- Different types of HPC resources
 - Commodity clusters
 - Computational nodes characteristics
 - Interconnection network
 - Massively parallel machines
 - BG/P, Cell BE based (Roadrunner)
 - New architectures
 - GPU, etc...
 - Etc ...



**Different
machine's
supplies**

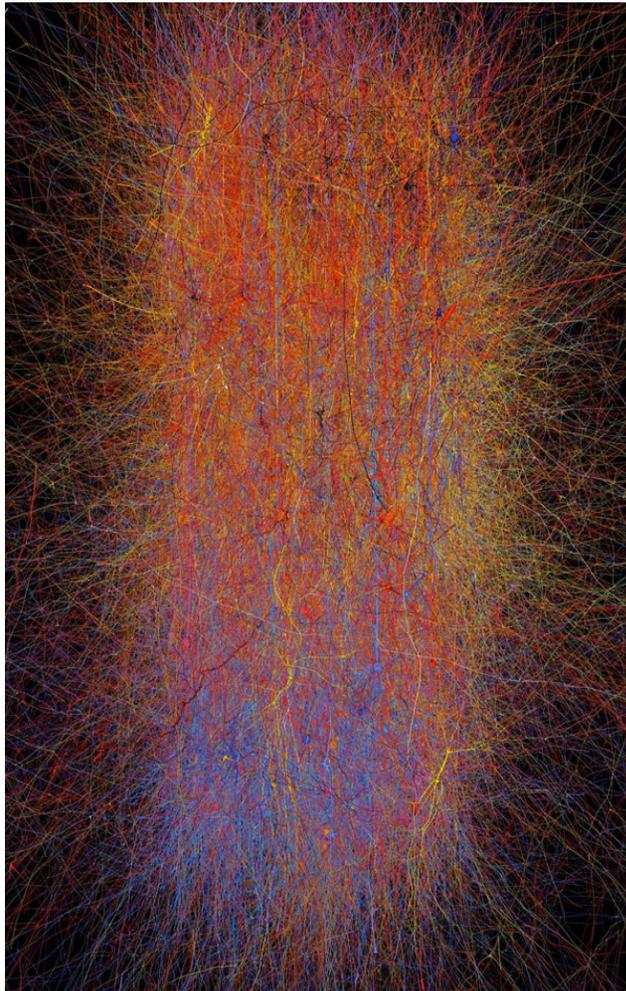
Matching needs-supplies



(...)
Small game

Which machine for ?

Which machine for ?



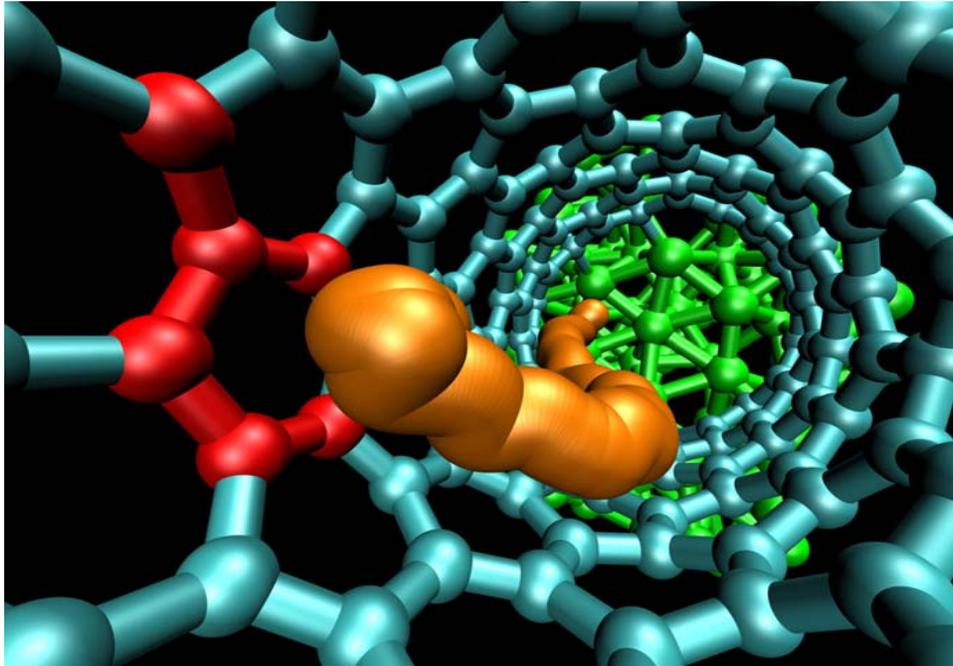
Bluebrain (EPFL)

- Network intensive

→ **An IBM BG/P**



Which machine for ?



Large CPMD simulation

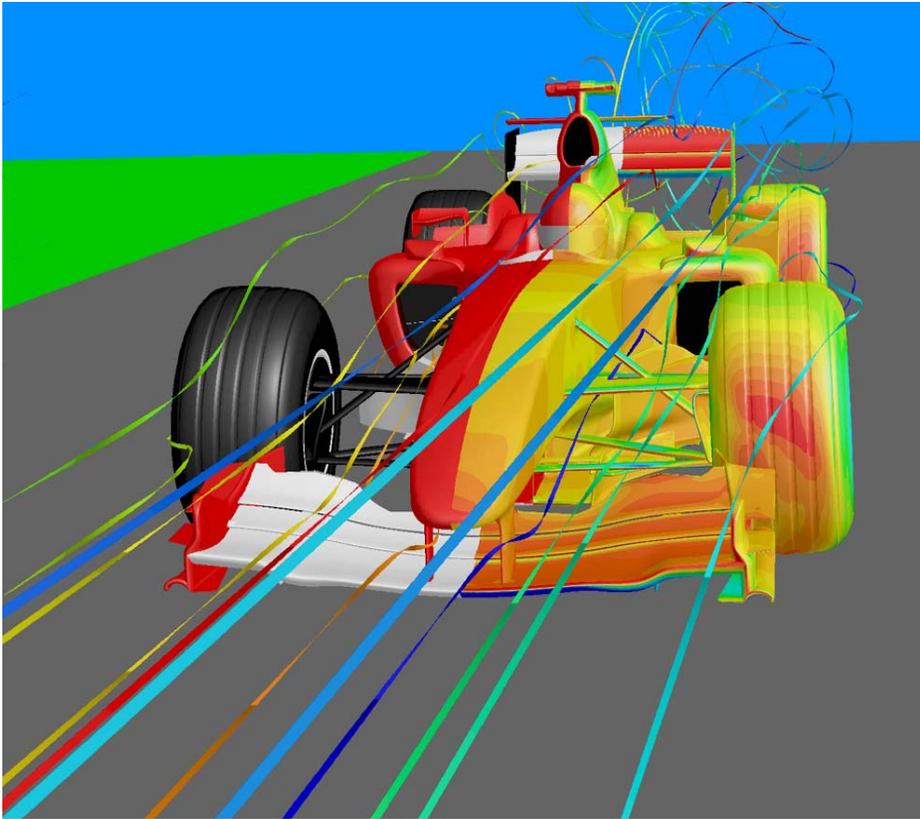
- CPU intensive
- Network intensive

→ A Cray



Which machine for ?

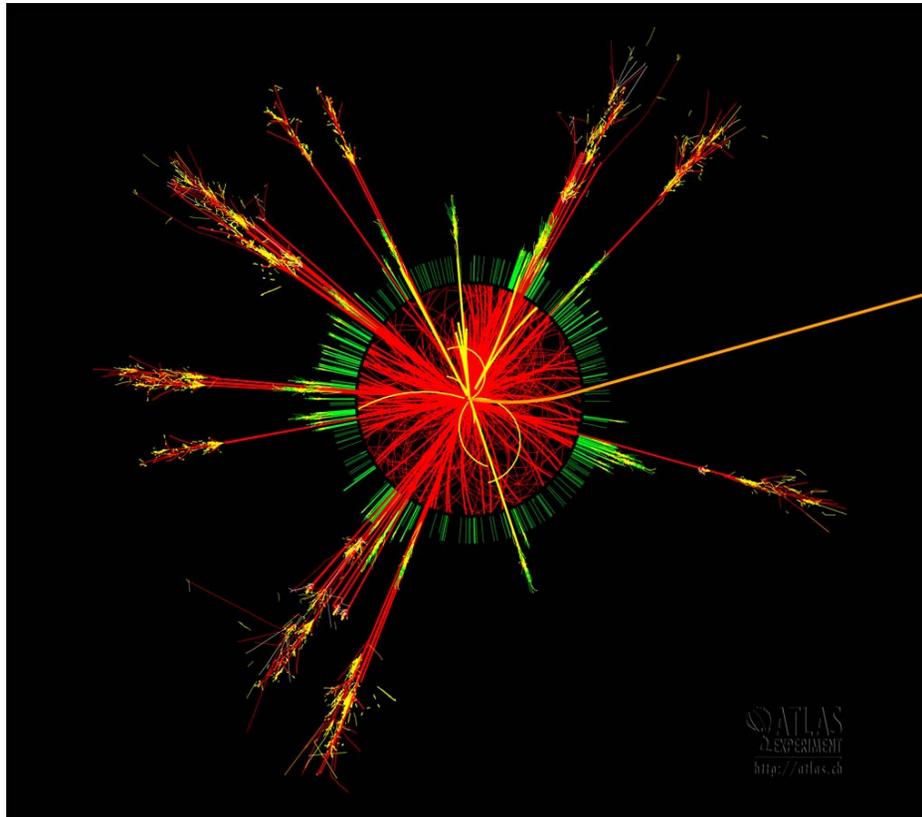
- CPU intensive
 - loosely coupled
- **A commodity cluster**



CFD



Which machine for ?



HEP @ATLAS (CERN)

- CPU intensive
- Embarrassingly parallel

→ A large quantity
of laptops



Which machine for ?



→ Roadrunner ?



Energy reduction through :

- Reduce electric consumption on each core
- Better manage a node
- Better use a node
- Use the right node
- Use the right resource to the application
- Stop outperformed resources
- Send the application to the best resource
- Detect poorly implemented apps and give hints to improve them



The challenge

- **FIRST :**
- Recognize the needs of the submitted app
- Knowledge of what a resource provide
- Find a metric to match both
- **THEN :**
- Predict the execution time on the resources
- Predict when the app will end on the resource
- Take into account user QoS ... **SUBMIT !**

supply-and-demand brokering under user's QoS

- Characterize
 - the applications (demander)
 - the resources in the “Grid” (supplier)
- User's QoS (time, money constraints)
- Find the best match

**Resource
Brokering**





An implementation exists :

- IANOS = Intelligent ApplicationN-Oriented System
- First β -version working across Switzerland and Germany
- Collaborative work between multiple specialists around Europe
 - Users
 - Developers
 - Computing Centers

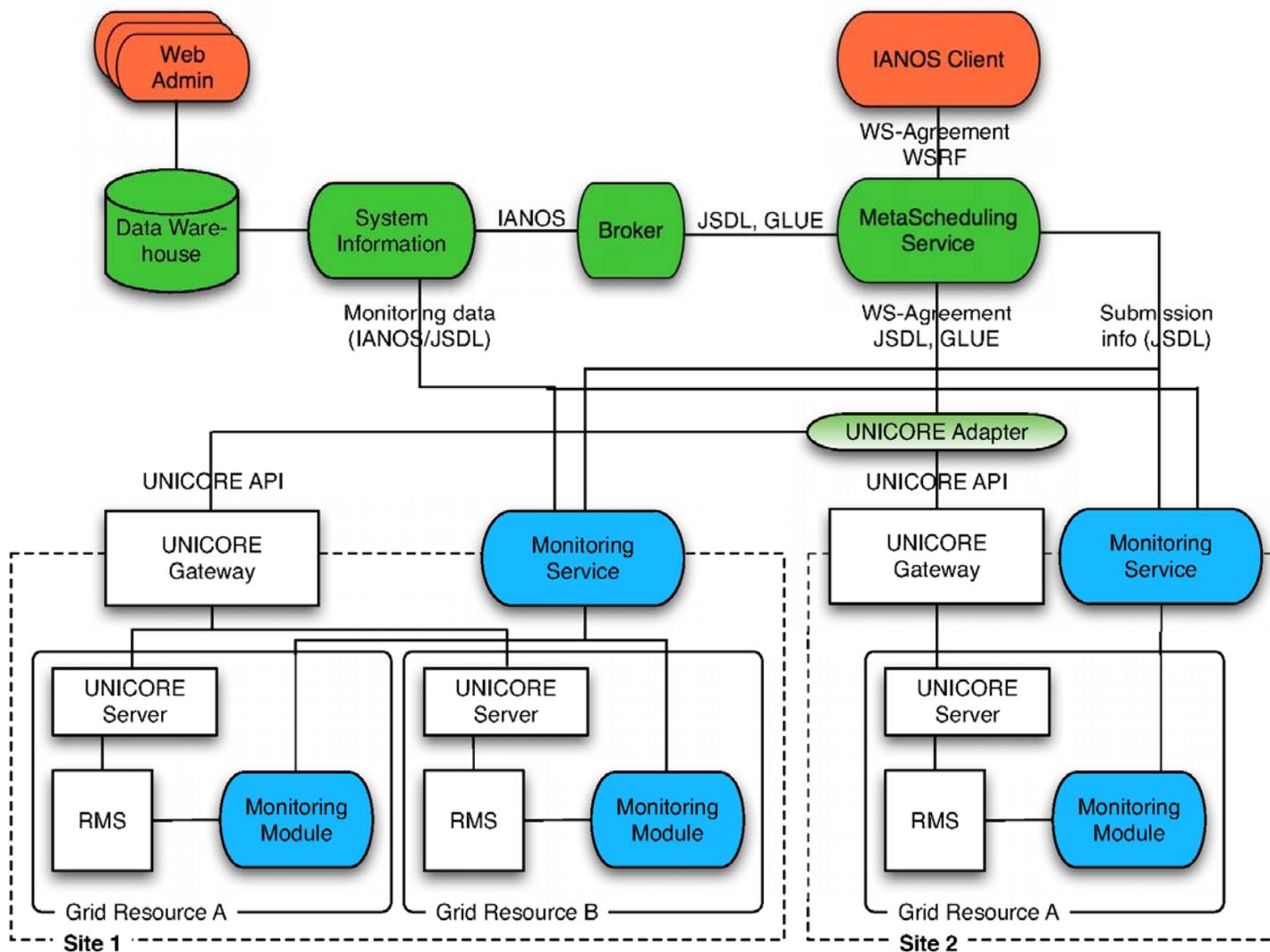
IANOS: the double-faced God

- Takes into account:
 - Applications
 - Resources
- Fulfill:
 - User's QoS
 - CC's reqs



Ἰανός (Janus in Latin) is the double faced God of the bridges and doors. That statue is on display in Roma (Vatican City)

IANOS architecture



IANOS inside

- Low-level monitoring of the running applications
- Decision algo based on historical data
- Minimization of a cost function (exec cost, licenses cost, energy cost, etc..)
- Execution time prediction based on performance prediction
- Availabilities of the resources at submission time

IANOS in action

IANOS Client

Request applications: **MXM:1.0-2**

N1: 10
N2: 1000
N3: 1
N4: 1

minimum cost minimum time optimal expert

alpha: beta:
gamma: delta:

Submit application **MXM:1.0-8** Job monitoring Terminate job

```
[PLEIADES2] Start Time: [7/11/08 10:53 AM] End Time: [7/11/08 10:54 AM] Ex_Time: [15 Se
[PLEIADES1] Start Time: [7/11/08 10:54 AM] End Time: [7/11/08 10:54 AM] Ex_Time: [17 Se
[PLEIADES2+] Start Time: [7/11/08 10:53 AM] End Time: [7/11/08 10:53 AM] Ex_Time: [5 Se
[Fhl_SCAI_PACK] Start Time: [7/11/08 10:54 AM] End Time: [7/11/08 10:54 AM] Ex_Time: [14
application [MXM:1.0] has been submitted successfully on the Grid site [PLEIADES2]
```

User GUI

Admin console

IANOS Scheduling Middleware System Information Web Console

Host: Execution

Edit/Delete Host CFM Parameters

PLEIADES2	
C/	
2.0	
7.0	
0.7	
0.0	
2006-01-01 20:00:00	[1995-12-31 23:59:59]
8750 C	
C/	
C.1	
25000 D	
27000 D	
85000 D	
C.0	
45000 D	
C.0	
24000 D	

Terminé

Produce more, Use the same energy
(and optimize the set of machines)!

- Point out machines efficiencies :
 - Too costly, never chosen : DECOMMISSION IT !
 - Not available, always chosen : BUY MORE RESOURCES OF THAT TYPE !
 - Predict better suited machines
- Optimizes the overall usage by scheduling applications on best fitted machines.

Research directions : IANOS-2

- Large scale HPC Grid
- Increase the number of applications
- Fine tuning of the brokering model
- Tests with other Grid middlewares
- More functionalities (checkpoint-restart, application deployment process, etc..)
- Simulator to use archived data

IANOS-2 : Proposal in EU-FP7

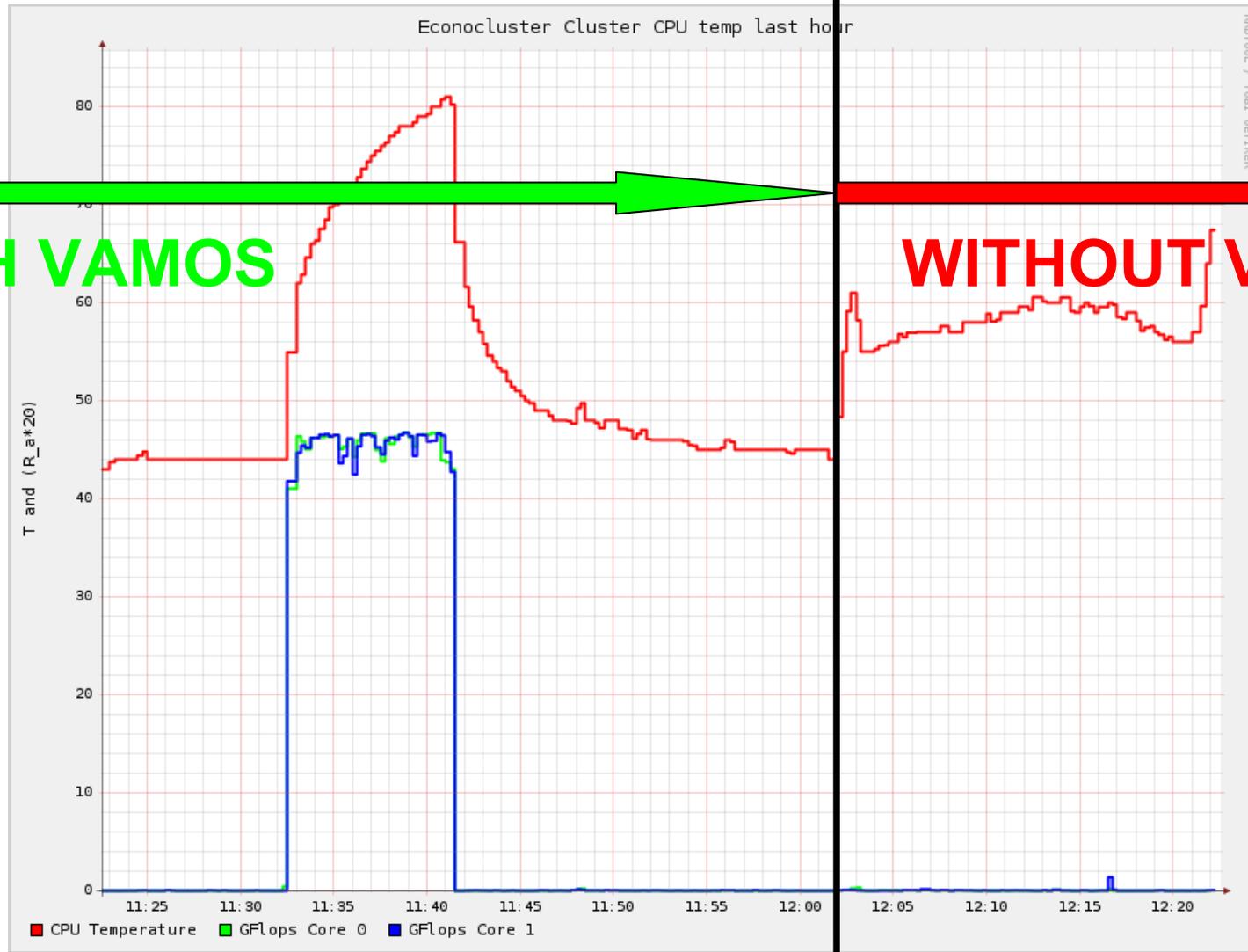
Is it possible to go a step further ?

- Use DVS-capable processors !
- Tune the machine to the needs of the applications (not the opposite)
- Which metric(s) is(are) relevant ?
- First tests with VAMOS (Veritable Application Monitoring System)
- “automatic” help to improve bad implemented applications

Make (compute) More, Use Less !

WITH VAMOS

WITHOUT VAMOS



Is “Green-IT” only a fad ?

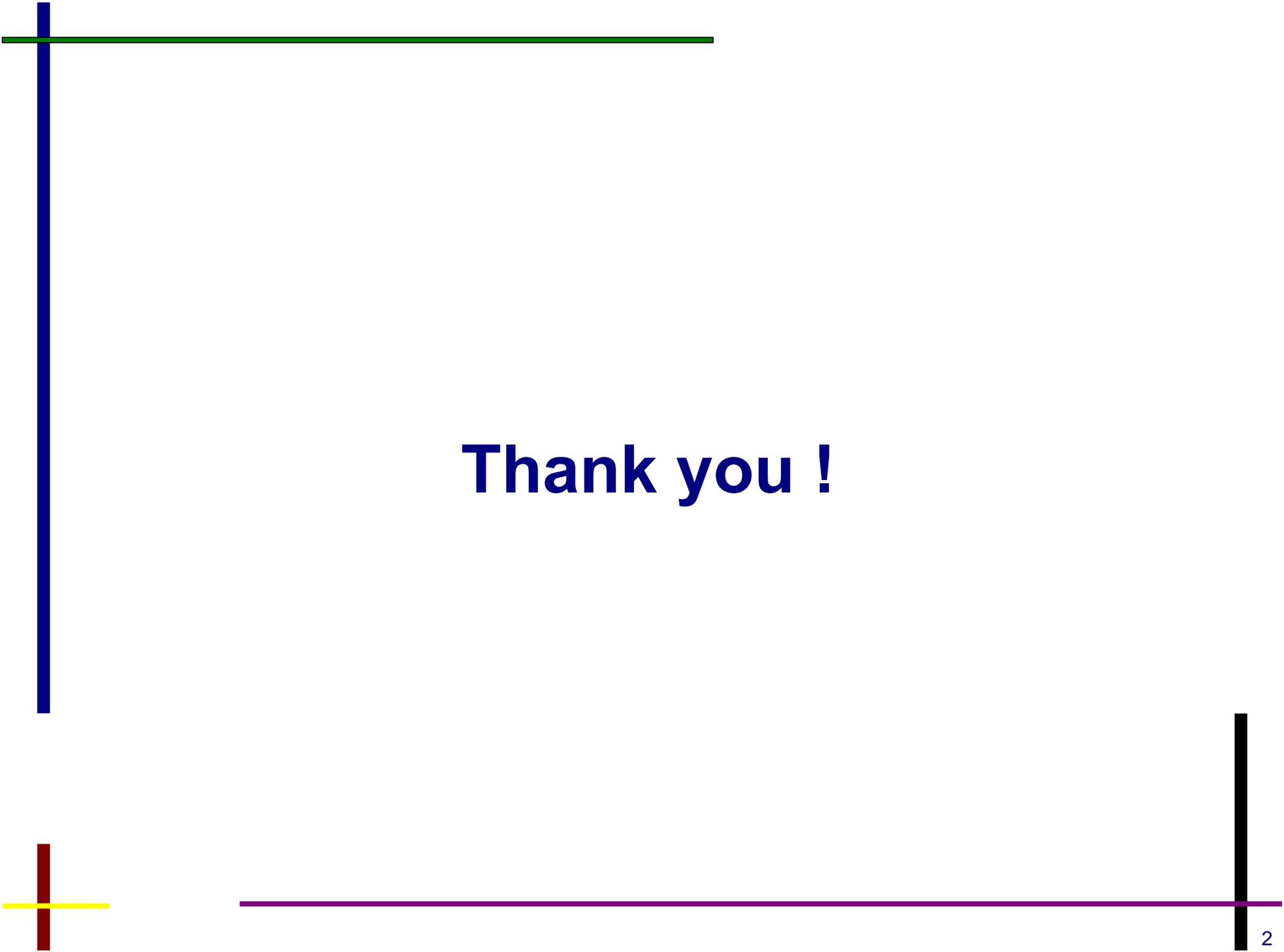
NO !

If Koomey's predictions are correct, in 2010, IT-related energy worldwide will approach 10 % of the 2008 power consumption of USA .

Green-IT does not mean less performance !

Green-IT is one of the keys for success in the HPC field

<http://www.energy.eu>



Thank you !