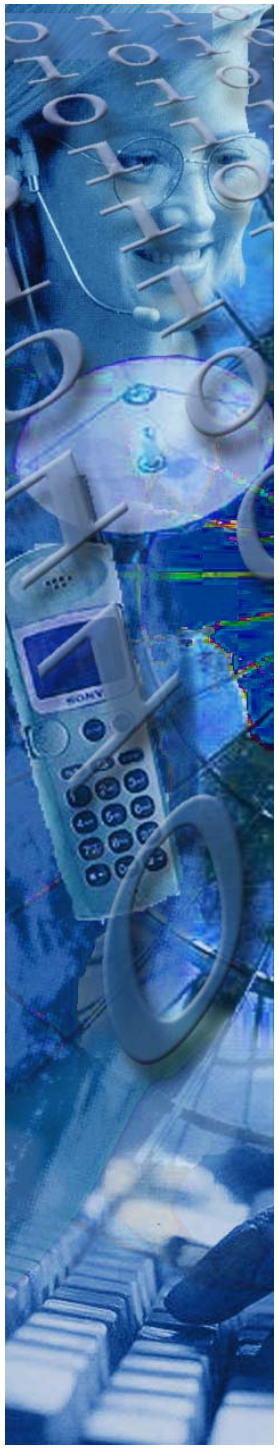


International Conference on **G**rid computing, high-perform**A**nce and **D**istributed **A**pplications **GADA'07**
November 29-30, 2007 - Vilamoura, Algarve, Portugal

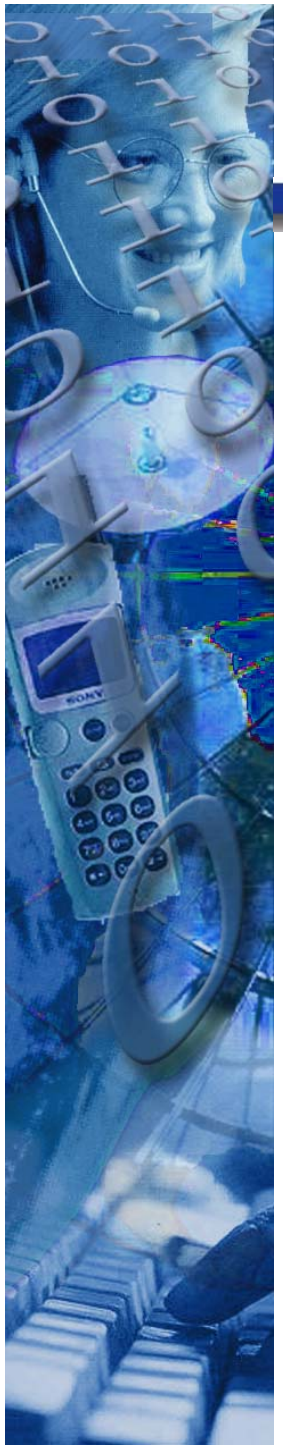
SNMP-based Monitoring Agents and Heuristic Scheduling for Large-scale Grids

Edgar Magaña[1-2], Laurent Lefèvre[3], Masum Hasan[1], and Joan Serrat[2]

[1] Cisco Systems, Inc.
emagana, masum@cisco.com

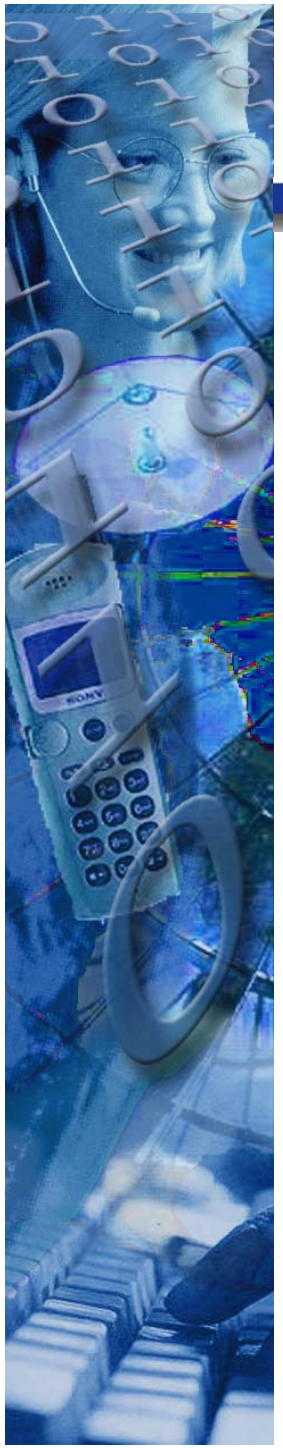
[3] INRIA RESO / Université de Lyon
LIP UMR 5668 (CNRS, ENS Lyon, INRIA, UCB), France
laurent.lefevre@inria.fr

[2] Universitat Politècnica de Catalunya
serrat@tsc.upc.edu



Outline

- Introduction
- Grid Resource Management (GRM)
- Policy-based Grid Management Architecture
- SBLOMARS Monitoring Agents Approach
- BLOMERS Heuristic Scheduler Approach
- Grid5000 Experiments
- Conclusions and Future Work



Introduction (I)

- **Grid Computing Definition:**

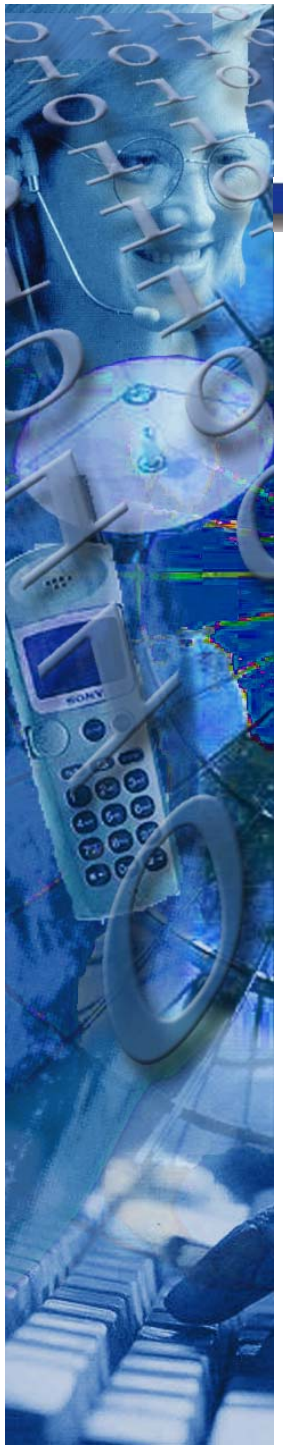
"Grid is an interconnected collections of geographically distributed and heterogeneous hardware and software infrastructure that provides dependable, consistent, pervasive, and inexpensive access to high-end computational capabilities". [Foster_Grid2]

– Grid Systems has to fulfill following three points:

» **Coordinates Resources** that are not under **centralized management**.

» Uses **standard, open, generic protocols and interfaces**.

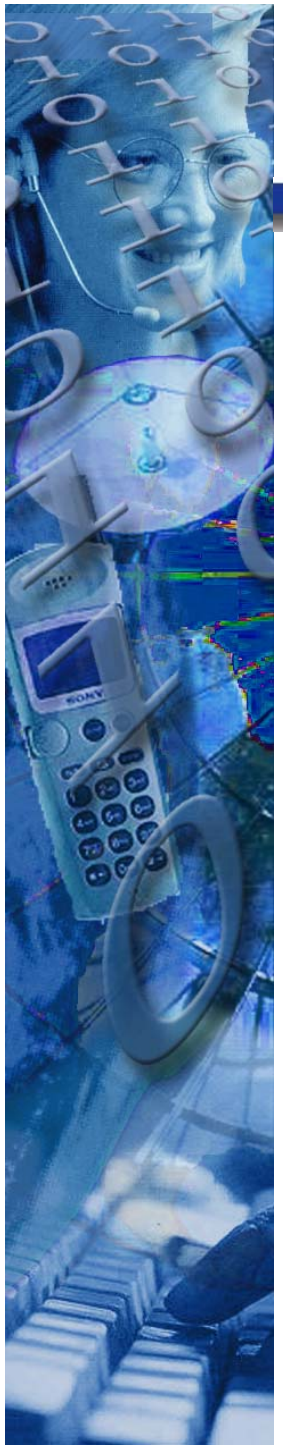
» Provides non-trivial **Quality of Services (QoS).**



Introduction (II)

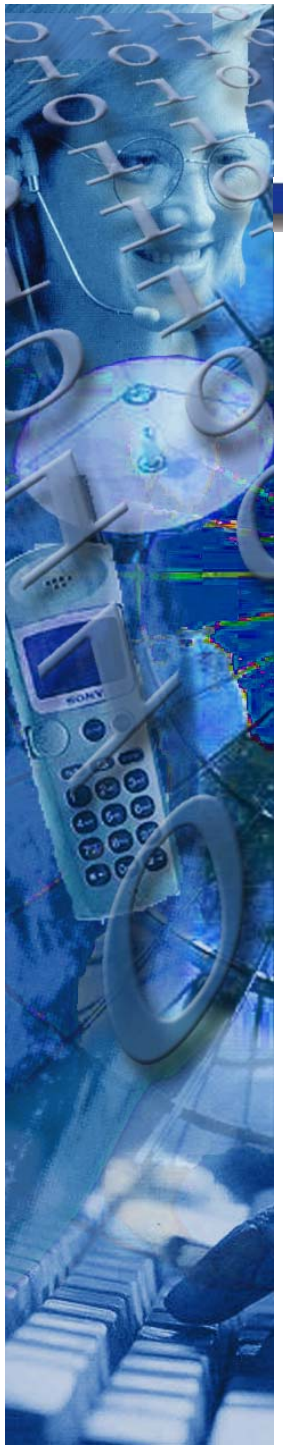
Grid Computing Tendencies:

- Grid Services should be provided to users **regardless** of **network technology, administrative domain** or **operative platform (Heterogeneity).**
- Effective access to **large amount** of computing, network and storage resources, reducing procurement, deployment, maintenance and operational cost.
- Network Performance: Fault-tolerance, **Scalability** and **Flexibility**.



Outline

- Introduction
- **Grid Resource Management (GRM)**
- Policy-based Grid Management Architecture
- SBLOMARS Monitoring Agents Approach
- **BLOMERS Heuristic Scheduler Approach**
- Grid5000 Experiments
- Conclusions and Future Work

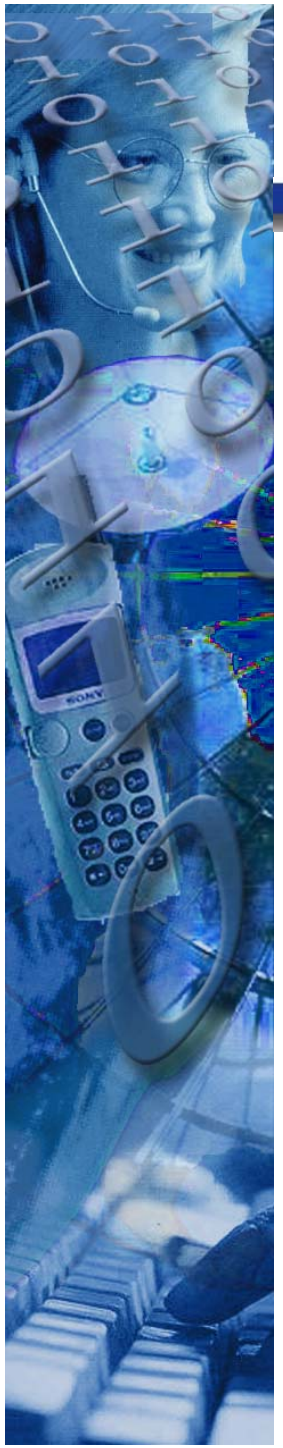


Grid Resource Management (I)

Definition

It is the ability to discover, allocate, negotiate, monitor, and manage the use of network-accessible capabilities in order to achieve various end-to-end or global qualities of service*.

*J. Nabrzyski, J. M. Schopf and J. Weglarz, "Grid Resource Management State of the Art and Future Trends“ Kluwer Academic Publishers. Boston, USA October 2003. ISBN 1-4020-7575-8.



Grid Resource Management (II)

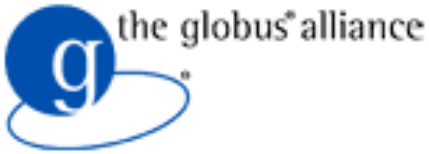
Main Activities:

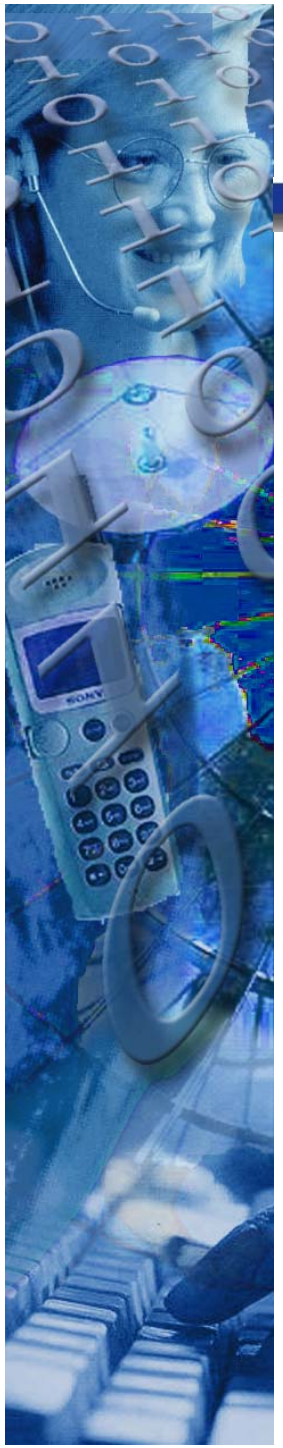
- GRM is regarded as a vital component of the Grid infrastructure.
- It coordinates and shares multiple kinds of resources efficiently.
- GRM must fulfill strict functional requirements from heterogeneous, and sometimes conflicting, domains (e.g., the users', applications and networks domains).
- It must adhere to non-functional requirements that are also rigid, such as reliability and efficiency in terms of time consumption and load on the host nodes.

Grid Resource Management (III)

Therefore…

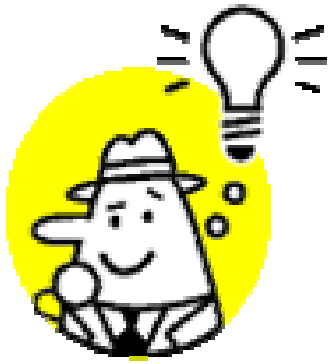
- GRM is a very challenging issue. Where, several project are proposing their approaches:



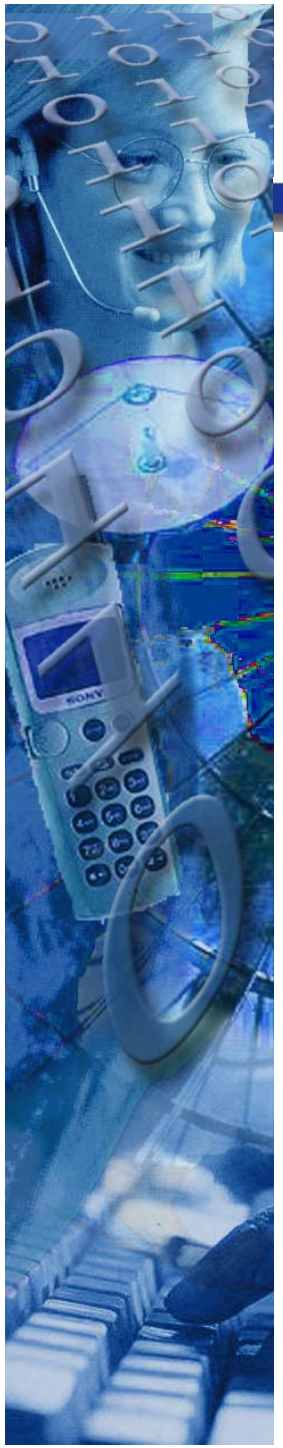


Grid Resource Management (IV)

- **Remaining Problems:**
  - Swift and dynamic allocation and reservation of computational resources.
  - **Algorithms** to find available resource.
  - Reduce resources **analysis** and **scheduling** times. **(Makespan)**
  - **Dynamic** resources discovery and analysis.
  - **Makespan** is normally reduced but L**oad Balancing** is not taking into account.
  - Allocation of network resources per service.



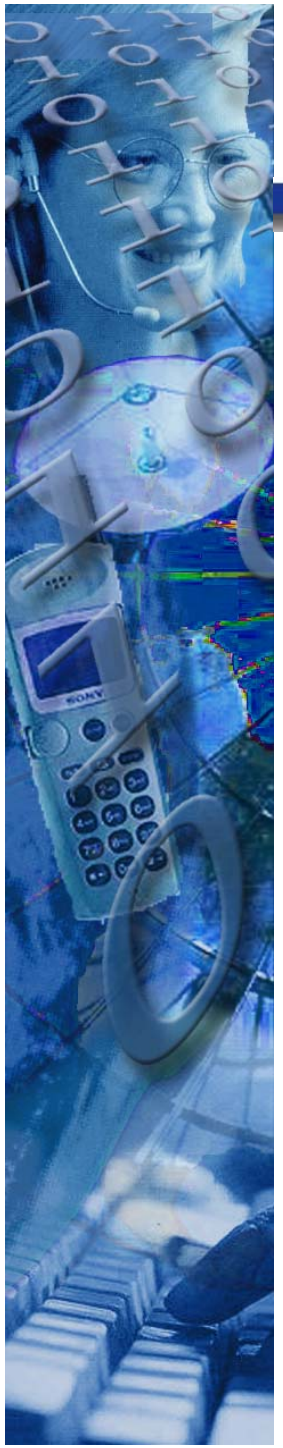
Divide and Conquer



Grid Resource Management (V)

Three Grid Management Phases:

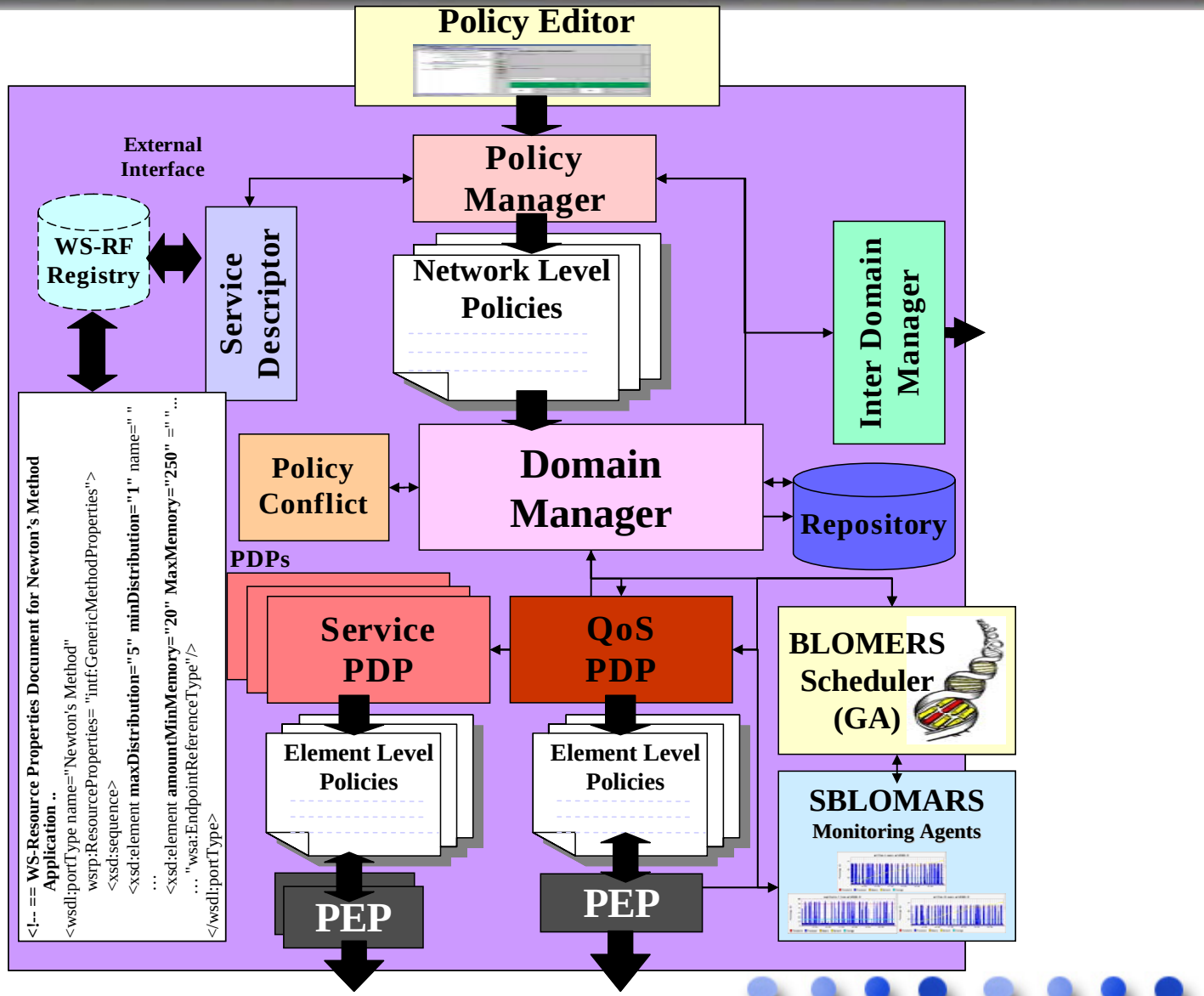
- **Resource Discovery and Monitoring (SBLOMARS)**
  - Which resources are available to a given user
  - Selecting resource source to search in more detail
  - Filtering out resources that do not meet the minimal job requirements
- **Resource Scheduling (BLOMERS)**
  - Given a group of possible resources must be selected on which to schedule the job
  - Algorithms for scheduling should be applied
  - Load Balanced techniques should be take into account
  - Setting up times also have to be considered
- **Job Allocation and Execution (PbGRMA)**
  - Job submission through simple commands or policies
  - Monitoring progress and evaluation of the performance
  - Notification to users when jobs have done
  - Cleaning up any temporal information on the server side

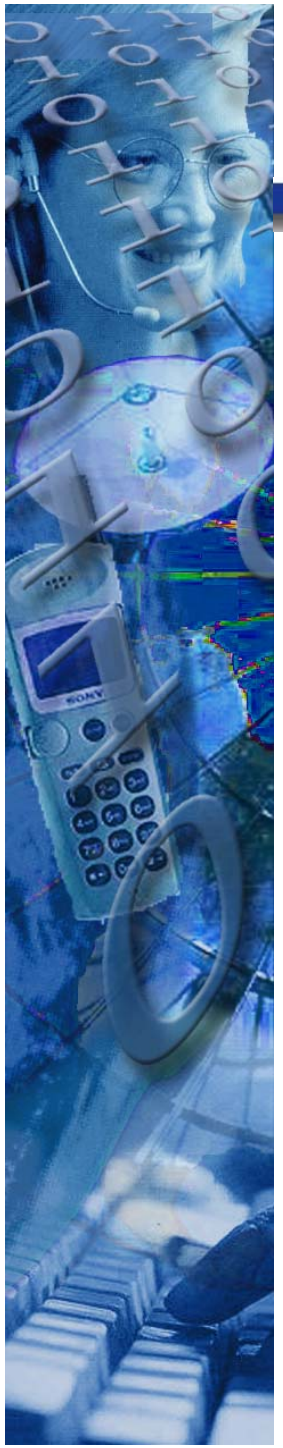


Outline

- Introduction
- Grid Resource Management
- **Policy-based Grid Management Architecture**
- SBLOMARS Monitoring Agents Approach
- **BLOMERS Heuristic Scheduler Approach**
- Grid5000 Experiments
- Conclusions
- Ongoing and Future Work

Policy-based Grid Management Architecture





Policy-based Grid Management Architecture*

Features:

– We obtain a gorgeous synergy by coupling Policy-based Technology and SNMP-based Agents and Genetic Algorithms.

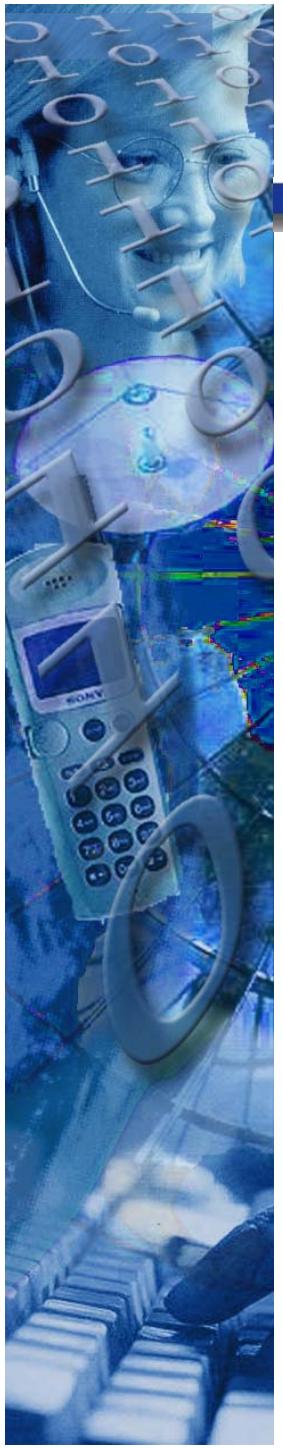
– Simplifies Grid Services deployment and management

– Support for dynamic, reconfigurable on demand, secure and highly customizable computing storage and networking environments

– Dynamic extensibility and flexibility of the architecture

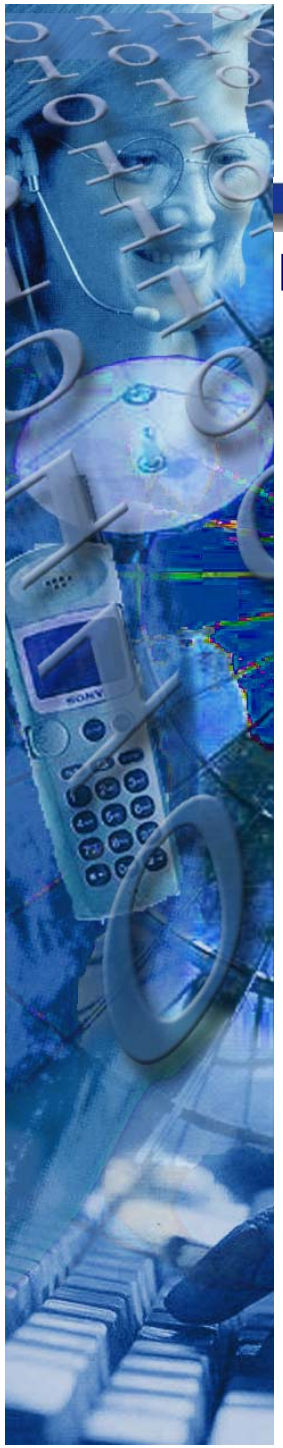
– Deployment and Activation of Grid services in all planes

*E. Magaña, L. Lefevre and J. Serrat. “Autonomic Management Architecture for Flexible Grid Services Deployment Based on Policies”. ARCS'07, Zurich, Switzerland. 2007.



Outline

- Introduction
- Grid Resource Management (GRM)
- Policy-based Grid Management Architecture
- **SBLOMARS Monitoring Agents Approach**
- **BLOMERS Heuristic Scheduler Approach**
- Grid5000 Experiments
- Conclusions and Future Work



SBLOMARS Monitoring Agents Approach (I)

Definition and Features:

– **SNMP-based** Balanced Load Monitoring Agents for Resource Scheduling

– It is a pure **decentralized monitoring system** in charge of permanently capturing computational resource performance based on **autonomous distributed agents**.

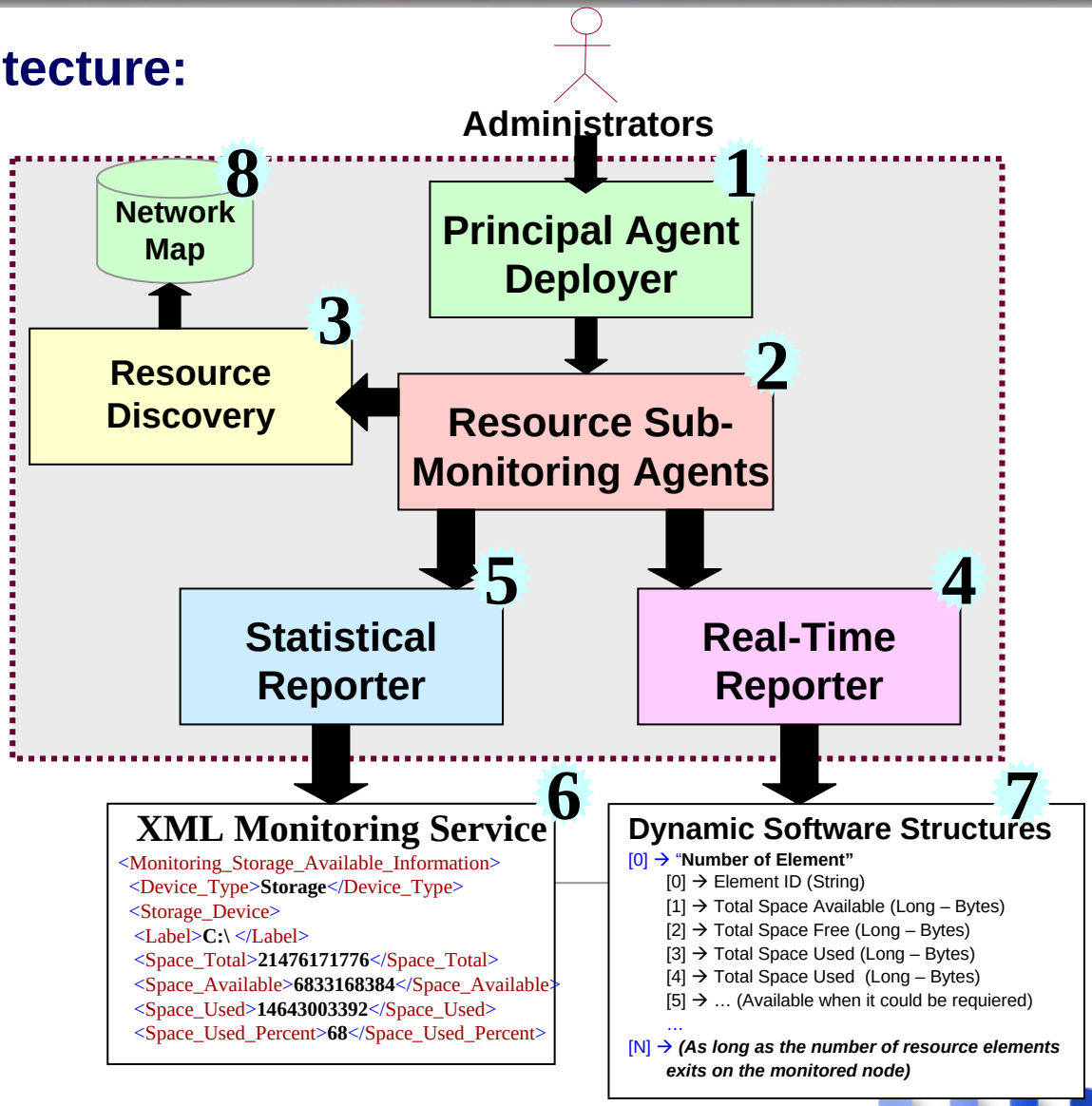
– it integrates SNMP technology and thus, offers an alternative solution to handle **heterogeneous resources**.

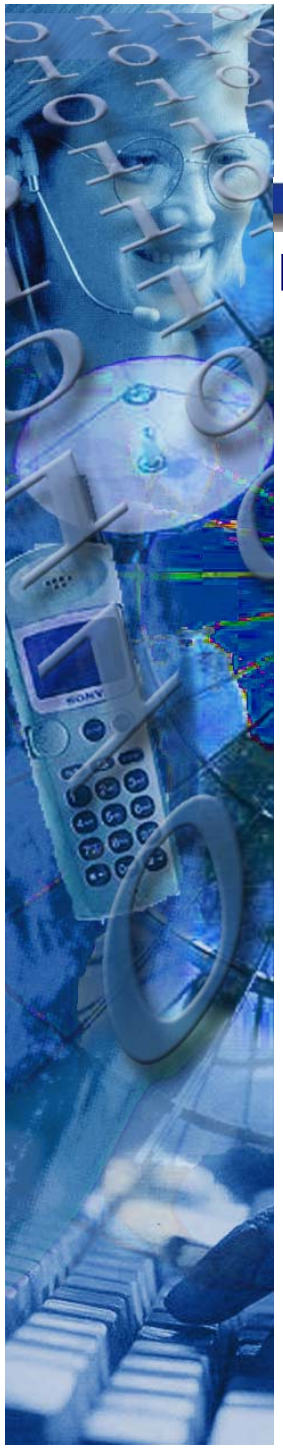
– it implements complex **dynamic software structures**, which are used to monitor from simple personal computers to robust multiprocessor systems or clusters with even multiple hard disks and storage partitions.

– It distributes the monitoring activities into a set of **sub-monitoring instances** which are specific per each kind of computational resource to monitor (processor, memory, software, network and storage)

SBLOMARS Monitoring Agents Approach (I)

Architecture:





SBLOMARS Monitoring Agents Approach (I)

Data Structures:

XML Reports "Storage"

```
<?xml version="1.0" encoding="UTF-8" ?>
<!-- Edited with Agent BLOMERSXML v1.0 …
<!-- Monitoring Resources Service xmlns:xsi= …
<Monitoring_Storage_Available_Information>
  <Device_Type>Storage</Device_Type>
  <Number_of_Elements>3</Number_of_Elements>
   <Storage_Device>
    <Label>C:\ Label: Serial Number f010b634</Label>
    <Space_Total>21476171776</Space_Total>
    <Space_Available>6833168384</Space_Available>
    <Space_Used>14643003392</Space_Used>
    <Space_Used_Percent>68</Space_Used_Percent>
    </Storage_Device>
  <Storage_Device>
    <Label>G:\ Label:Disco local Serial Number 302e</Label>
    <Space_Total>10733957120</Space_Total>
    <Space_Available>3095842816</Space_Available>
    <Space_Used>7638114304</Space_Used>
    <Space_Used_Percent>71</Space_Used_Percent>
    </Storage_Device>
  <Storage_Device>
    <Label>H:\ Label:SHARED Serial Number 48f893</Label>
    <Space_Total>34290843648</Space_Total>
    <Space_Available>13172244480</Space_Available>
    <Space_Used>21118599168</Space_Used>
    <Space_Used_Percent>61</Space_Used_Percent>
  </Storage_Device>
</Monitoring_Storage_Available_Information>
```

Dynamic Software Structures - "Storage"

[0] → "**Number of Element**"
- [0] → Element ID (String)
- [1] → Total Space Available (Long – Bytes)
- [2] → Total Space Free (Long – Bytes)
- [3] → Total Space Used (Long – Bytes)
- [4] → Total Space Used Percentage (Long – Bytes)
- [5] → … (Available when it could be requiered)
- …
- …

[1] → "**Number of Element**"
- [0] → Element ID (String)
- [1] → Total Space Available (Long – Bytes)
- [2] → Total Space Free (Long – Bytes)
- [3] → Total Space Used (Long – Bytes)
- [4] → Total Space Used Percentage (Long – Bytes)
- [5] → … (Available when it could be requiered)
- …
- …

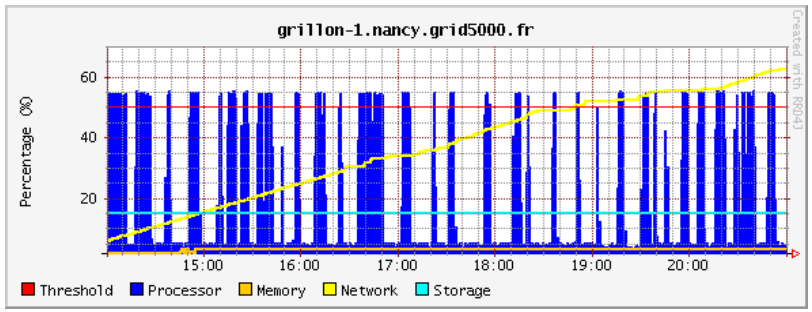
…

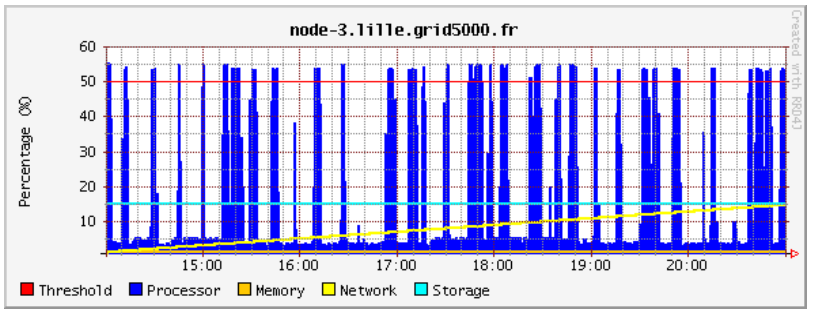
[N] → ***(As long as the number of resource elements exits on the monitored node)***

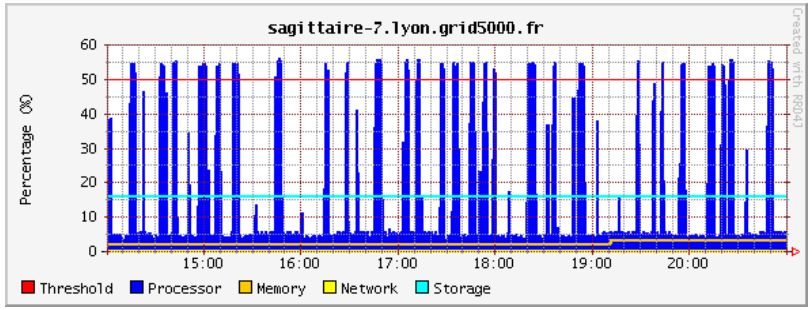
SBLOMARS Monitoring Agents Approach (II)

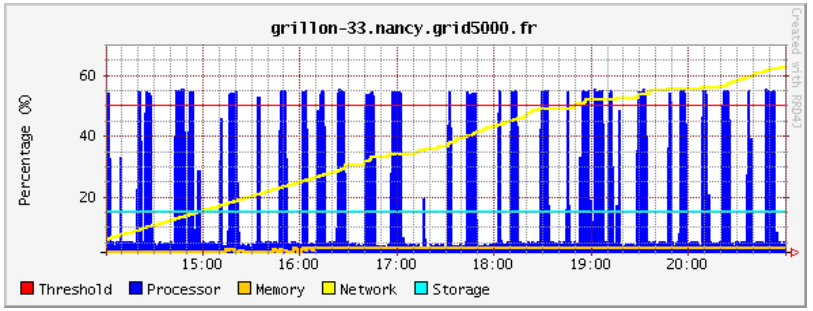
Graphical User Interface:

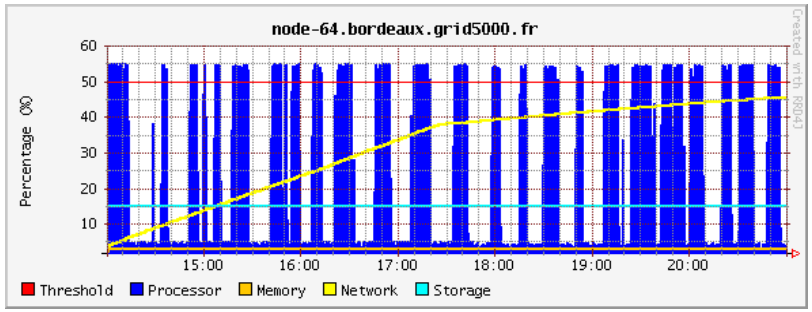
Grid 5000 CPU Resource Performance Monitoring by SBLOMARS

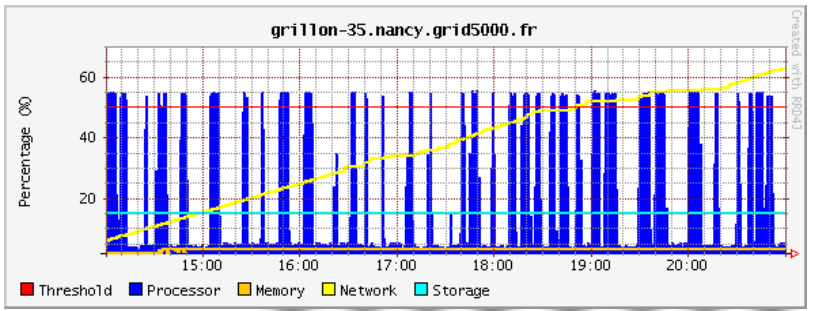


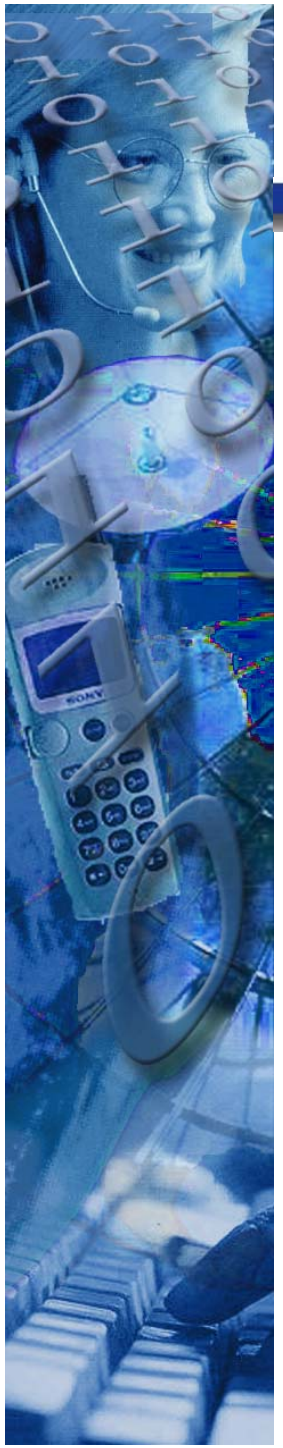






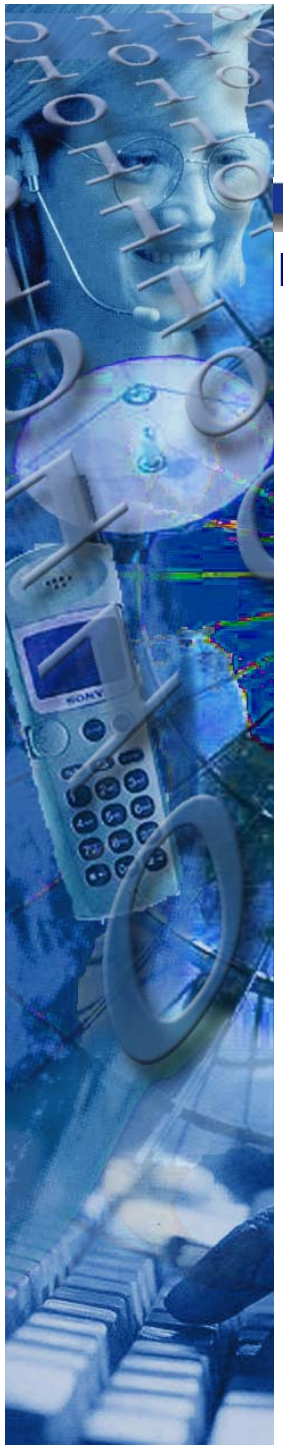






Outline

- Introduction
- Grid Resource Management (GRM)
- Policy-based Grid Management Architecture
- SBLOMARS Monitoring Agents Approach
- **BLOMERS Heuristic Scheduler Approach**
- Grid5000 Experiments
- Conclusions and Future Work



BLOMERS Heuristic Scheduler Approach (III)

Definition and Activities:

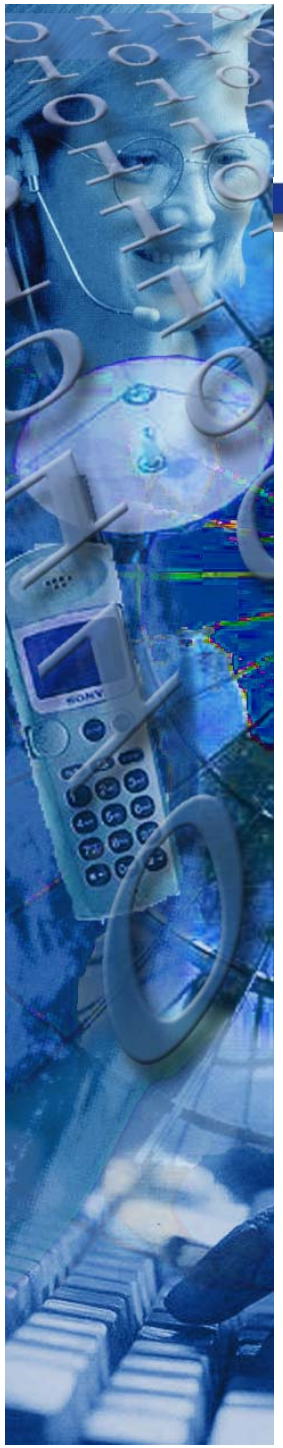
– Balanced Load Multi-Constrain Resource Scheduler (BLOMERS)

– This scheduler makes use of the statistical resource availability information generated by SBLOMARS agents.

– This procedure examines the set of available resources, generates a number of candidates and evaluates the candidate resources to select a final subset to be allocated and communicates the results.

– Every resource (not node) is assigned an ID which is the reference to generate new populations in our approach.

– The reference ID is taken in its binary representation to perform "**Mutation** and **Crossover**" operations.



BLOMERS Heuristic Scheduler Approach (IV)

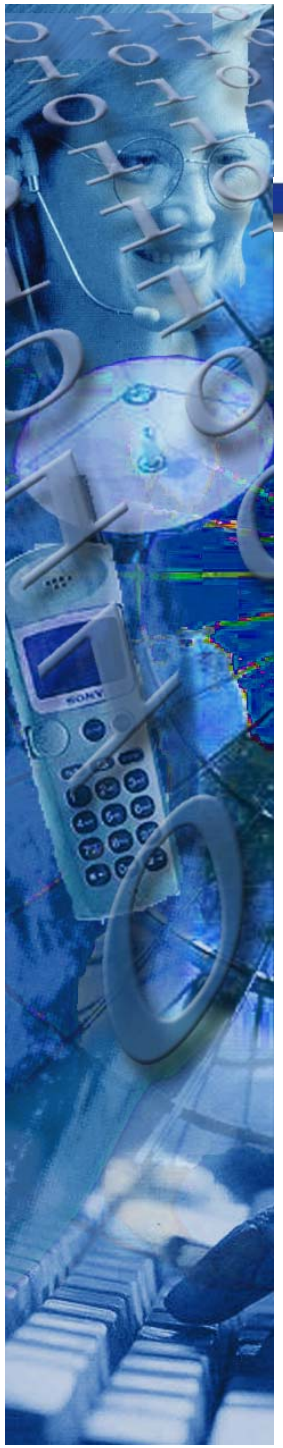
Pseudo Code:

```
CleaningBuffer (Pk)
Initialize (k, Pk);
Evaluate (Pk);
Do
 {
 Select_Resource_Candidates (Pk);
 Crossover (Pk);
    IF Evaluate(Pk+1) == Minimal Constraints;
      Ends Do-While;
    ELSE
      Mutation(Pk+1);
    IF Evaluate(Pk+2) == Minimal Constraints;
      Ends Do-While;
  }
Deliver (k_solution);
```

Where:

(Pk) *Is the selected Population (set of resources)*

(k) *Is one kind of resource (memory, storage, etc.)*



BLOMERS Heuristic Scheduler Approach (V)

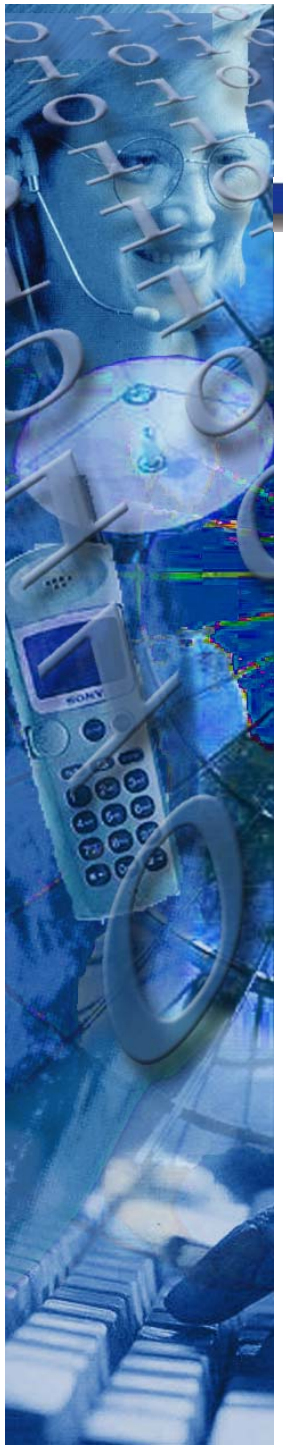
Crossover:

			CHROMOSOME
Storage_0 → ID_24: 147.83.106.199:6400	Parent 1	=	0000 0110 00
Storage_1 → ID_45: 147.83.106.167:6401	Parent 2	=	0000 1100 01

			CHROMOSOME
Storage_1 → ID_25: 147.83.106.199:6401	Child 1	=	0000 0110 **01**
Storage_0 → ID_44: 147.83.106.167:6400	Child 2	=	0000 1100 **00**
…	…		…

Mutation:

			CHROMOSOME
Storage_1 → ID_25: 147.83.106.199:6401	Parent 1	=	0000 0110 01
Storage_N → ID_57: 147.83.**206**.199:6401	Child 1	=	0000 1110 01

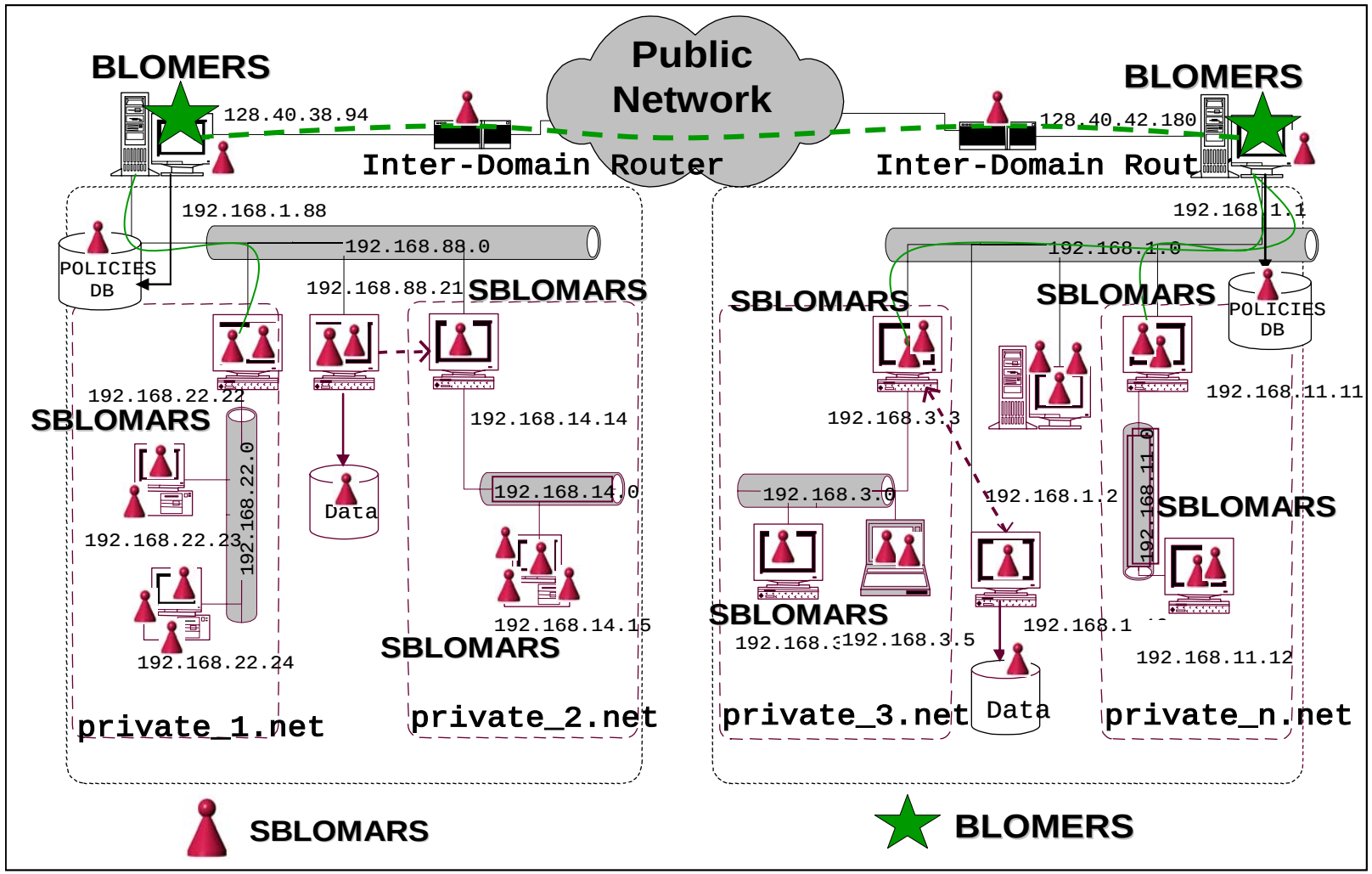


Outline

- Introduction
- Grid Resource Management (GRM)
- Policy-based Grid Management Architecture
- SBLOMARS Monitoring Agents Approach
- BLOMERS Heuristic Scheduler Approach
- **Grid5000 Experiments**
- Conclusions and Future Work

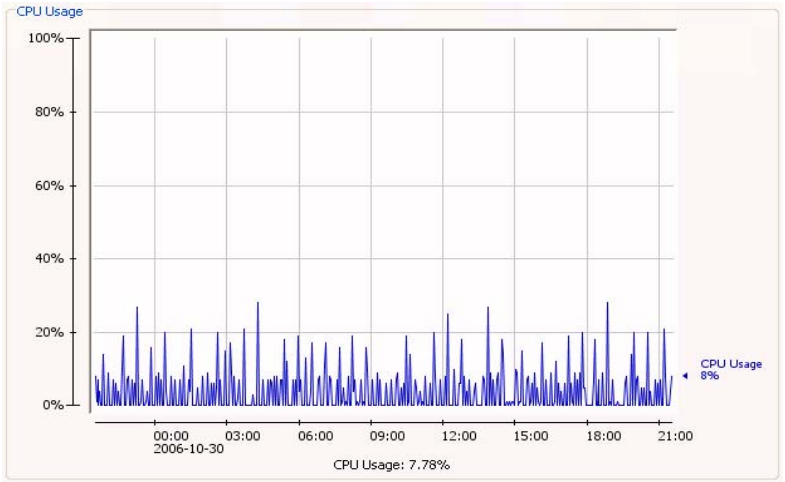
SBLOMARS and BLOMERS Approach (I)

General Scenario:

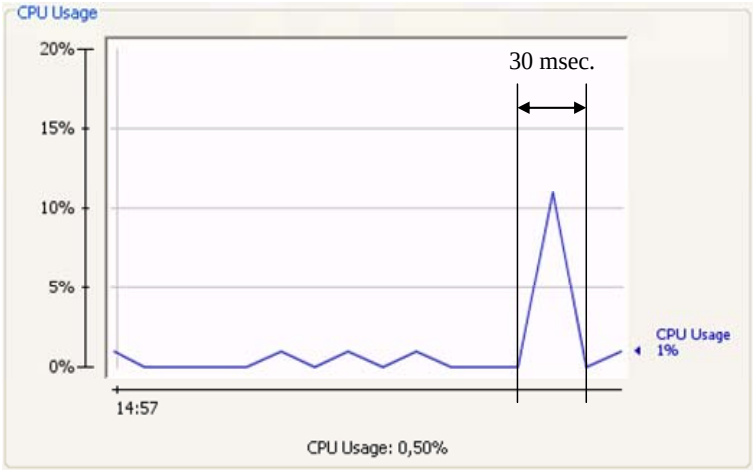


Grid5000 Experiments (I)

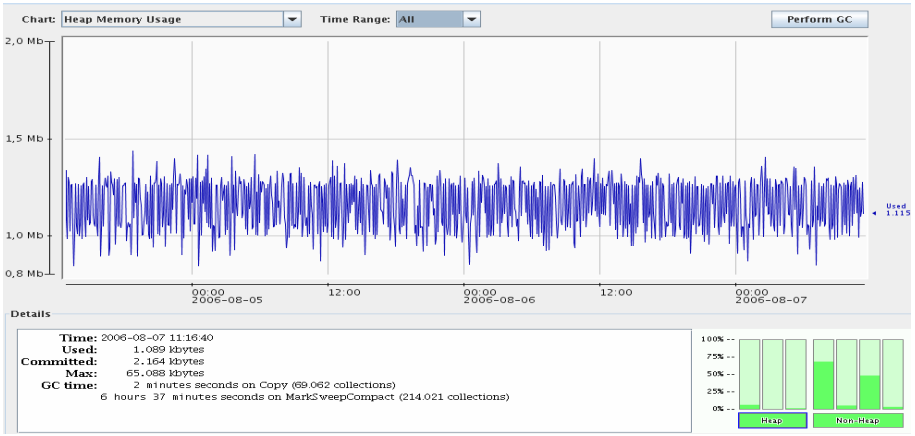
SBLOMARS PERFORMANCE:



Twenty-four Hours CPU Usage



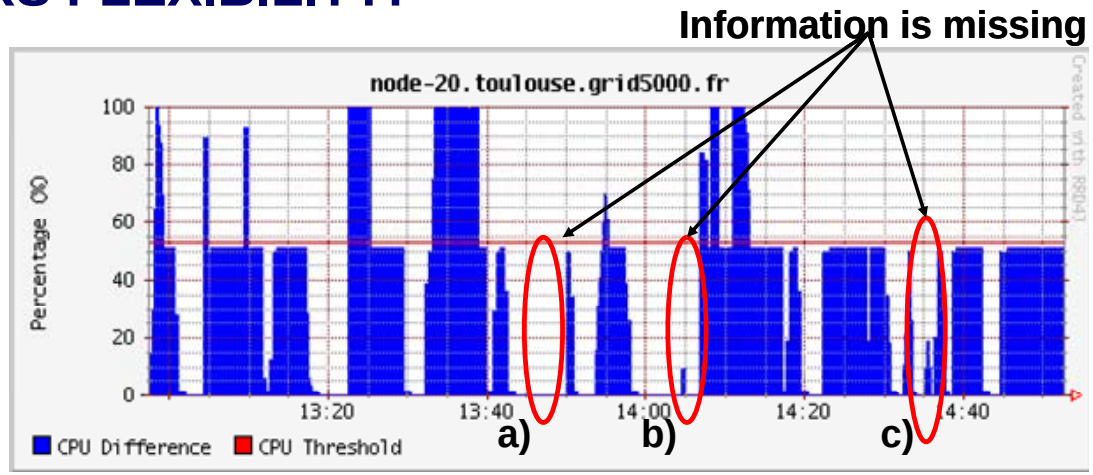
Sixty Seconds CPU Usage



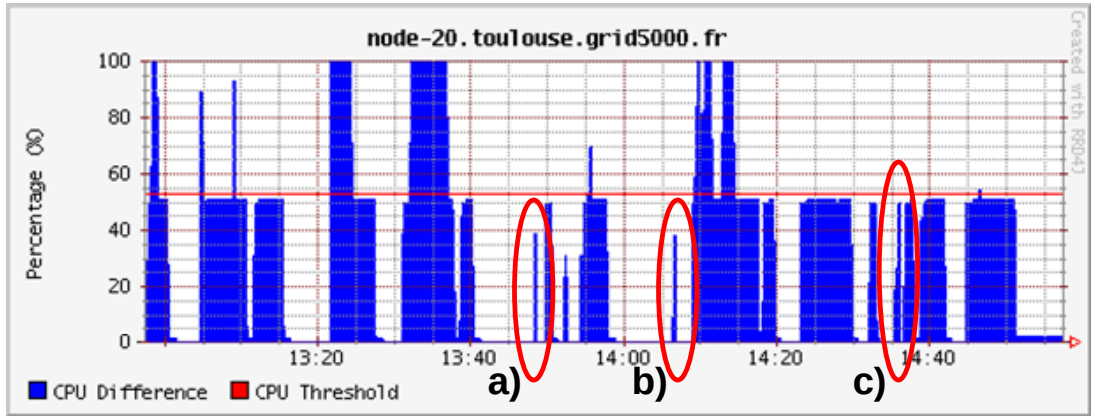
Forty-eight Hours Memory Usage

Grid5000 Experiments (II)

SBLOMARS FLEXIBILITY:



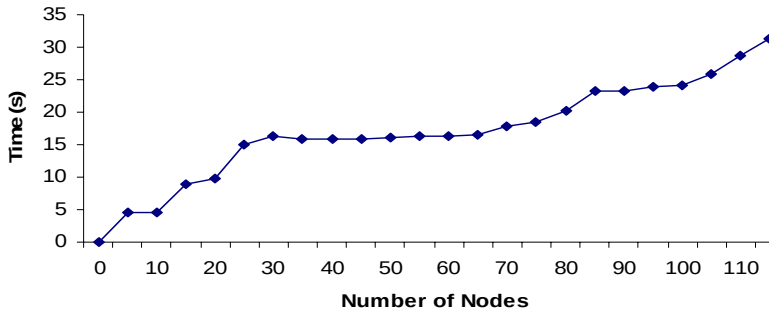
Fix Timing between SNMP Traps



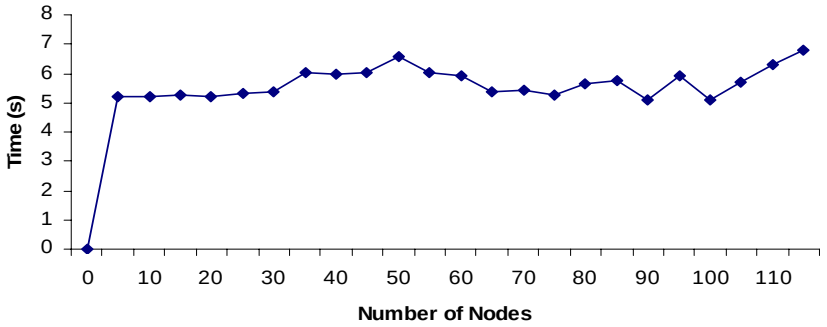
Auto-configuration between SNMP Traps

Grid5000 Experiments (III)

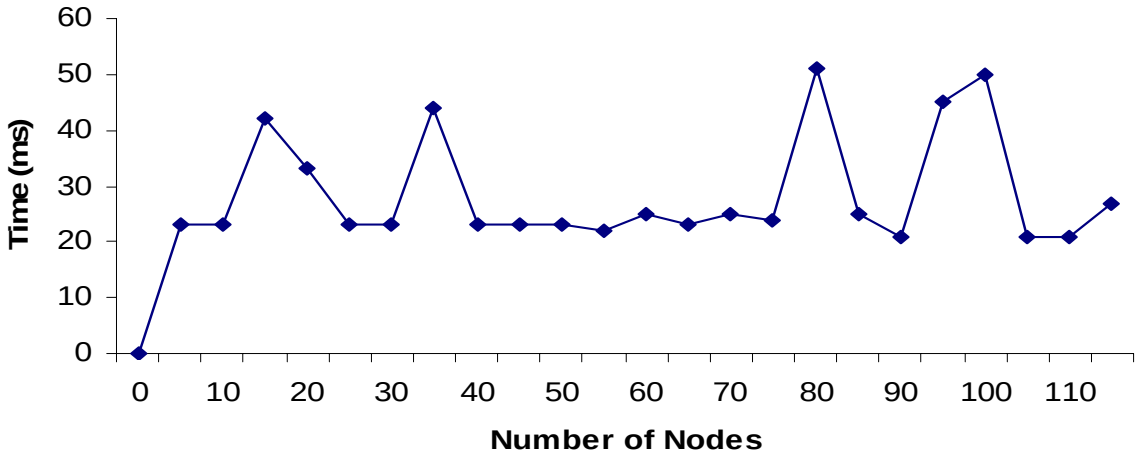
SBLOMARS SCALABILITY:



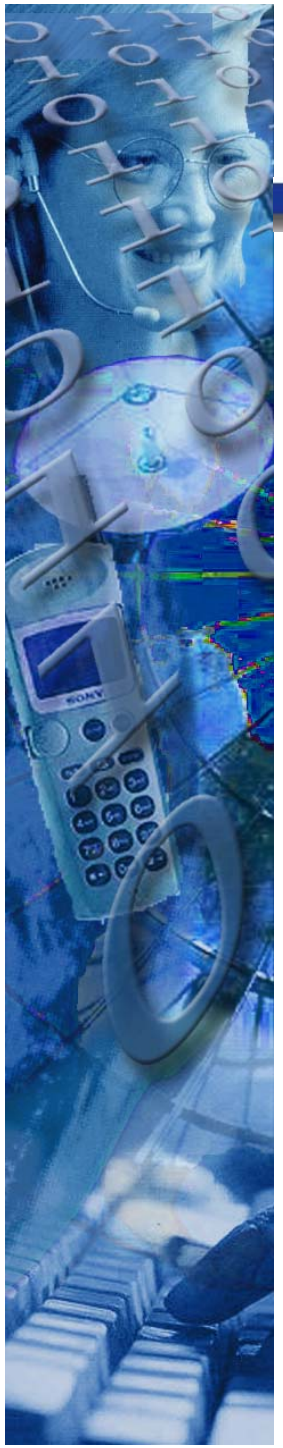
SBLOMARS Configuration Time



SBLOMARS Starting Time



SBLOMARS Responding Time



Grid5000 Experiments (IV)

- Grid5000 : 4000 CPUs/cores on 10 sites around 10Gbit/s networks. Heterogeneous architectures. http://www.grid5000.fr

- The total amount of nodes performing this experiment were 115.

- Every node was running a processor generator application to simulate processor load.

- Each scheduling algorithm was working along 120 minutes (2hrs) receiving 30 jobs every 60 seconds. On every node were also running a background processes generator. It was running randomly for the whole experiment (6hrs)

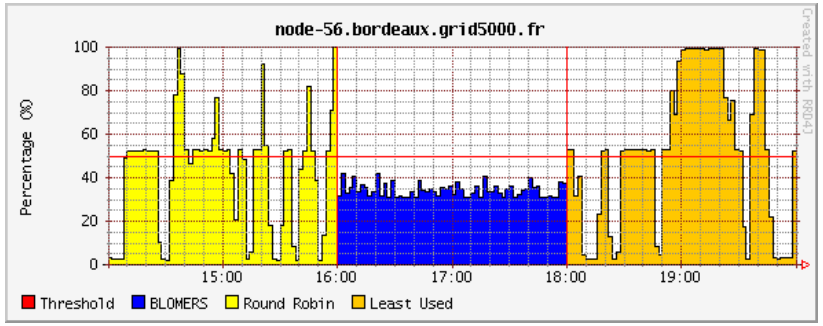
- **Round-robin:** This algorithm schedules every job received to the next available node from a list of nodes available.

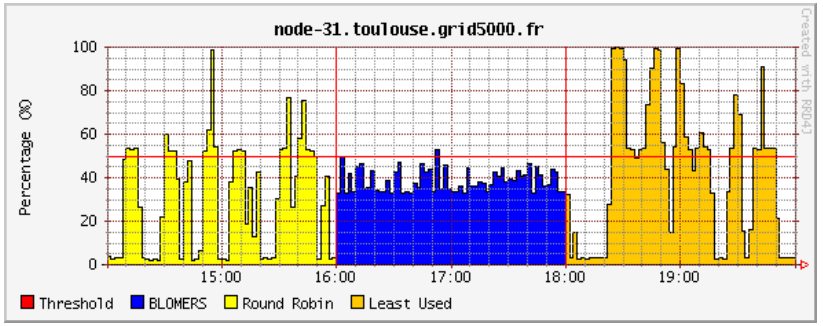
- **Least Used:** This algorithm schedules based on the average of the least used node.

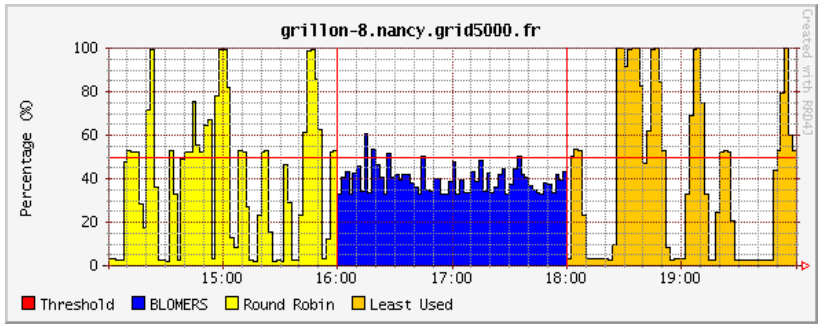
http://nmg.upc.es/~emagana/sblomars/grid5000.html

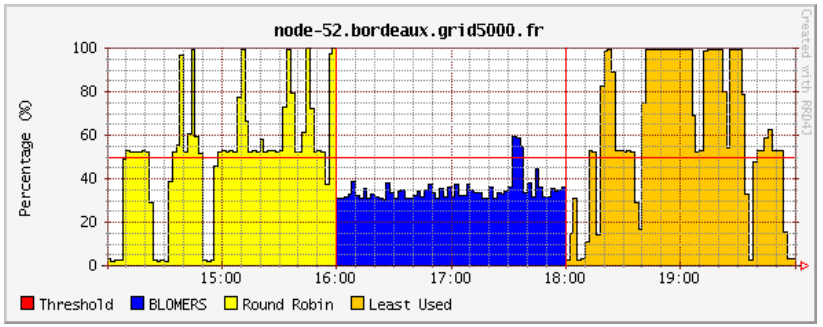
Grid5000 Experiments (v)

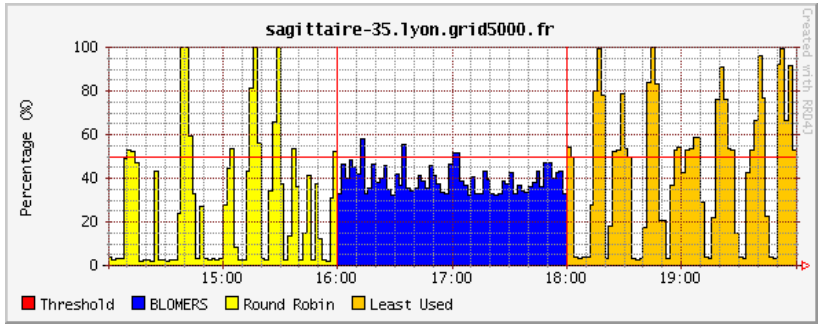
Grid 5000 CPU Resource Performance Monitoring by SBLOMARS

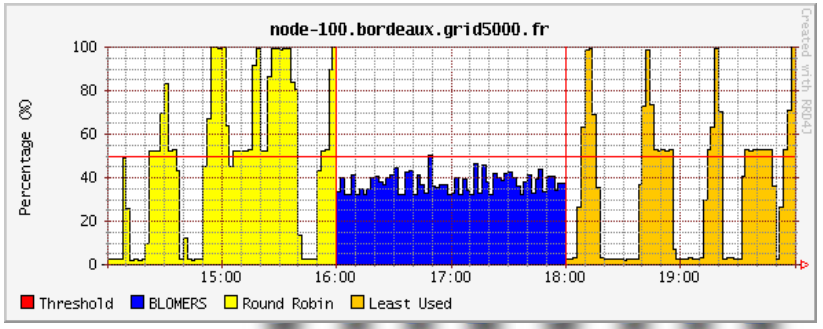


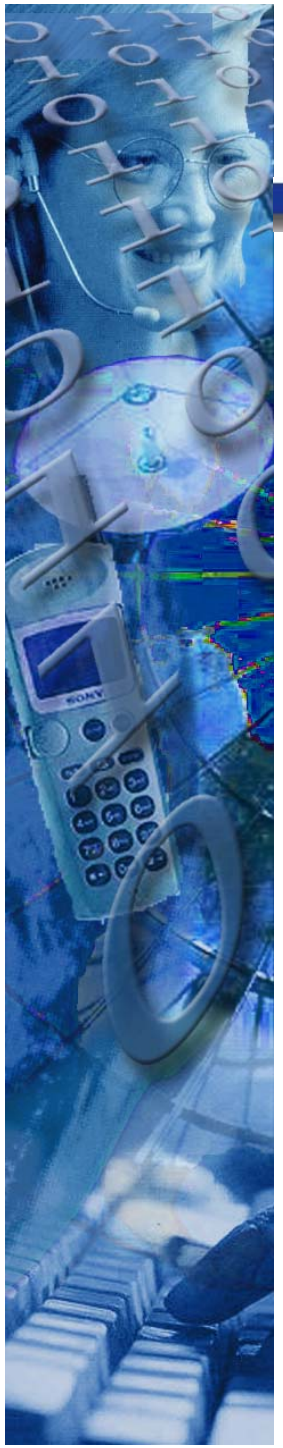






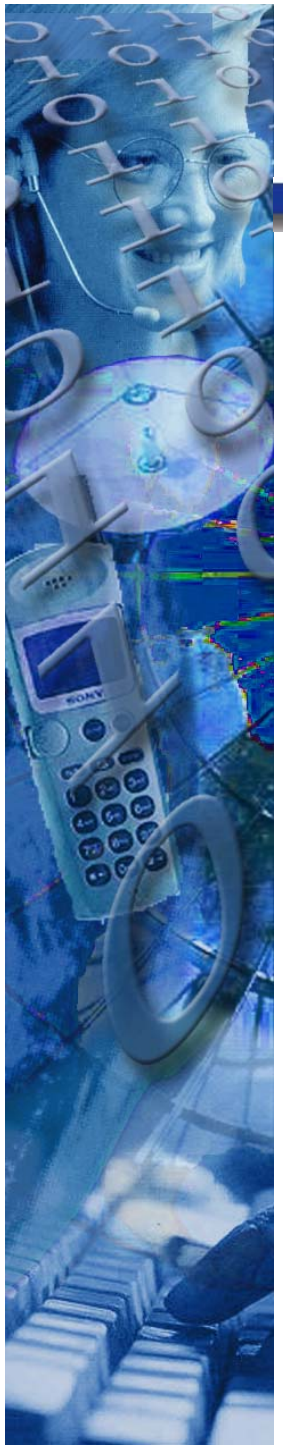






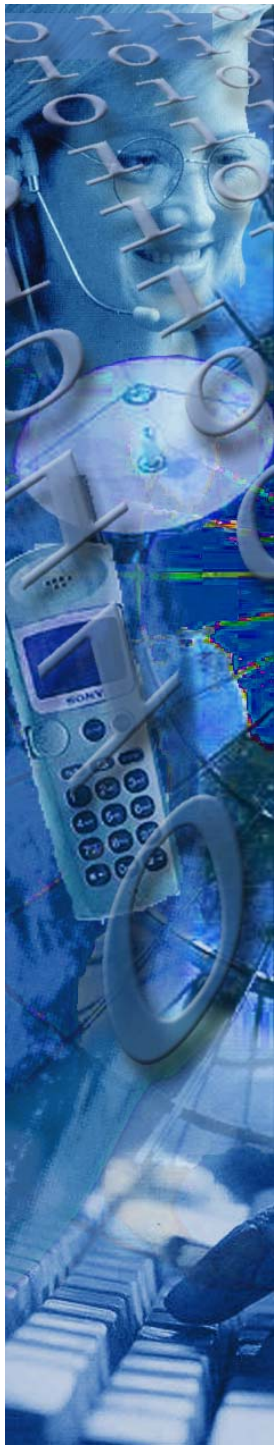
Outline

- Introduction
- Grid Resource Management (GRM)
- Policy-based Grid Management Architecture
- SBLOMARS Monitoring Agents Approach
- BLOMERS Heuristic Scheduler Approach
- Grid5000 Experiments
- **Conclusions and Future Work**



Conclusions and Future Work

- BLOMERS implements a Genetic Algorithm, which it offers a parallelism to multi-constraint service requests avoiding to be enclosed into a local minima.
- It improves resource load-balancing and reduces the makespan in any scheduling.
- BLOMERS is a scalable system based on a distributed agents design and its flexibility allows it to handle heterogeneous devices.
- It is able to schedule large numbers of services in real scenarios, such as Grid5000 testbed.
- Current evaluation results do not include yet, the effect of network latency and other communication impairments.
- We are including network performance between end-to-end edge routers as an entry parameter for our genetic algorithm. Resources with high latencies or jitter will be taken into account.



Questions?

Laurent Lefèvre
http://perso.ens-lyon.fr/laurent.lefevre
laurent.lefevre@inria.fr

Edgar Magaña

http://nmg.upc.es/~emagana/
emagana@cisco.com
emagana@nmg.upc.edu

http://nmg.upc.es/~emagana/sblomars/grid5000.html