Resource Management in Multi-Cluster Grid Configurations Adopting an Ontology-Fuzzy Approach

D.J. Ferreira and M.A.R. Dantas
Department of Informatics and Statistics (INE)
Federal University of Santa Catarina (UFSC)

Jinhui Qin and Michael A. Bauer
Department of Computer Science
University of Western Ontario

The 23rd Annual International Symposium on High Performance Computing Systems and Applications (HPCS)
June 14 – 17, 2009

UFSC

Western

Agenda

- Motivation
- Introduction
- Related Work
- Proposed Approach
- Experimental Environment and Results
- Conclusions and Future Work

Agenda

- Motivation
- Introduction
- Related Work
- Proposed Approach
- Experimental Environment and Results
- Conclusions and Future Work

Motivation

- *Resource management in multi-cluster grid configurations is a challenging task which requires an efficient scheduling of distributed and parallel applications;*

- *Within a grid environment, it is also desirable to have transparency on how to match available resources and communication links to applications;*

Motivation

- In this research work, we present an approach that considers the management of both *computer resources* and *communication links* through an *ontology-fuzzy approach*;

- The *ontology paradigm* is employed as a means to provide *a standard interface* in which users can express their requirements for desired resources;

- *Fuzzy logic algorithms* are used *to gather parameters* for matching processor usage, and communication based on dynamically monitored values;

UFSC

Western

Agenda

- Motivation
- Introduction
- Related Work
- Proposed Approach
- Experimental Environment and Results
- Conclusions and Future Work

Introduction

- *Multi-clusters*, if well orchestrated as in a grid environment, can represent significant computational power for the execution of several classes of applications, including parallel jobs;

- This *meta-computing environment* can be employed as a HPC facility inside a specific organization, or among different organizations, creating in both cases the concept of virtual organization (VO);

- *Inside a cluster configuration*, a local resource management system (RMS) will naturally provide *high priority to local applications*;

Introduction

- *Some* RMSs also support *advanced reservation* (AR), such as, MAUI/MOAB, PBS Professional, LSF and Loadlever;

- These RMS typically utilize *a private interface format* for the execution of advanced reservations;

- A *meta-scheduler* with an efficient algorithm for *co-reservation* and *co-allocation* is a key element in grid environments, especially those formed by *multi-cluster configurations*.

UFSC

Western

Agenda

- Motivation
- Introduction
- Related Work
- Proposed Approach
- Experimental Environment and Results
- Conclusions and Future Work

UFSC

Western

Related Works

- A number of *meta-schedulers projects* for grid environments have already been proposed with *different characteristics*;

- Examples of grid *meta-schedulers* are *Calana, Deco, Gridway, GridARS and Viola*;

- Another import aspect of *meta-schedulers* is the *challenge* of dealing with *different heterogeneous* local resource management systems in multi-cluster configurations;

- This problem occurs because of the utilization of *private heterogeneous interfaces* from AR modules;

-Only a few projects, such as *Community Scheduler Framework (CSF)* , promise to have a standard interface for information exchange between RMSs;

- However, these efforts *are far from the final goal.*

Agenda

- Motivation
- Introduction
- Related Work
- Proposed Approach
- Experimental Environment and Results
- Conclusions and Future Work

Proposed Approach

- The *ontology paradigm* has been considered an interesting solution to formal description of grid resources. Furthermore, it offers a base for the semantic matching;

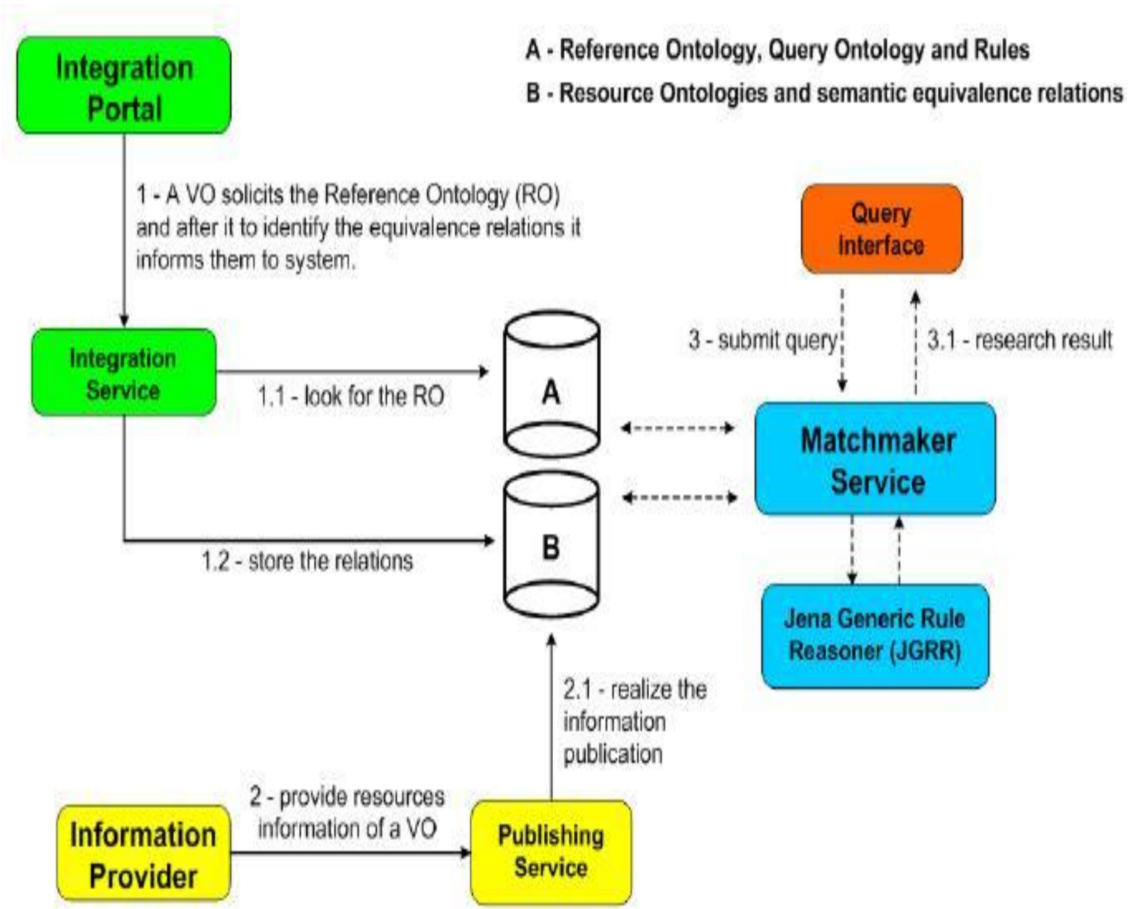


Fig. 1: The resource matching based on ontology

Proposed Approach

- Our approach modifies the original resource matching shown in Figure 1, by sending the submitted query to a *new Allocator* module that interfaces with the *fuzzy analysis module.*

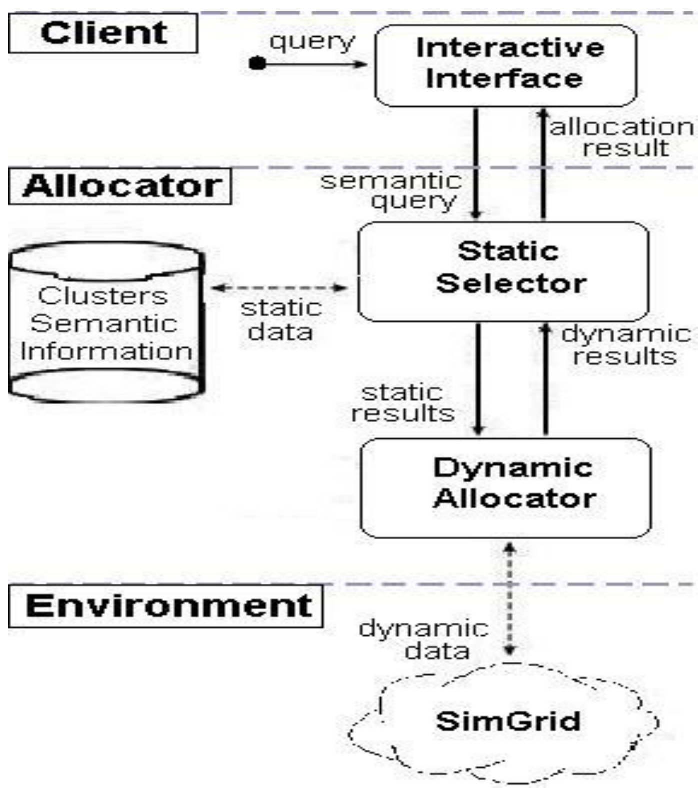


Fig. 2 : Main components of the approach

Proposed Approach

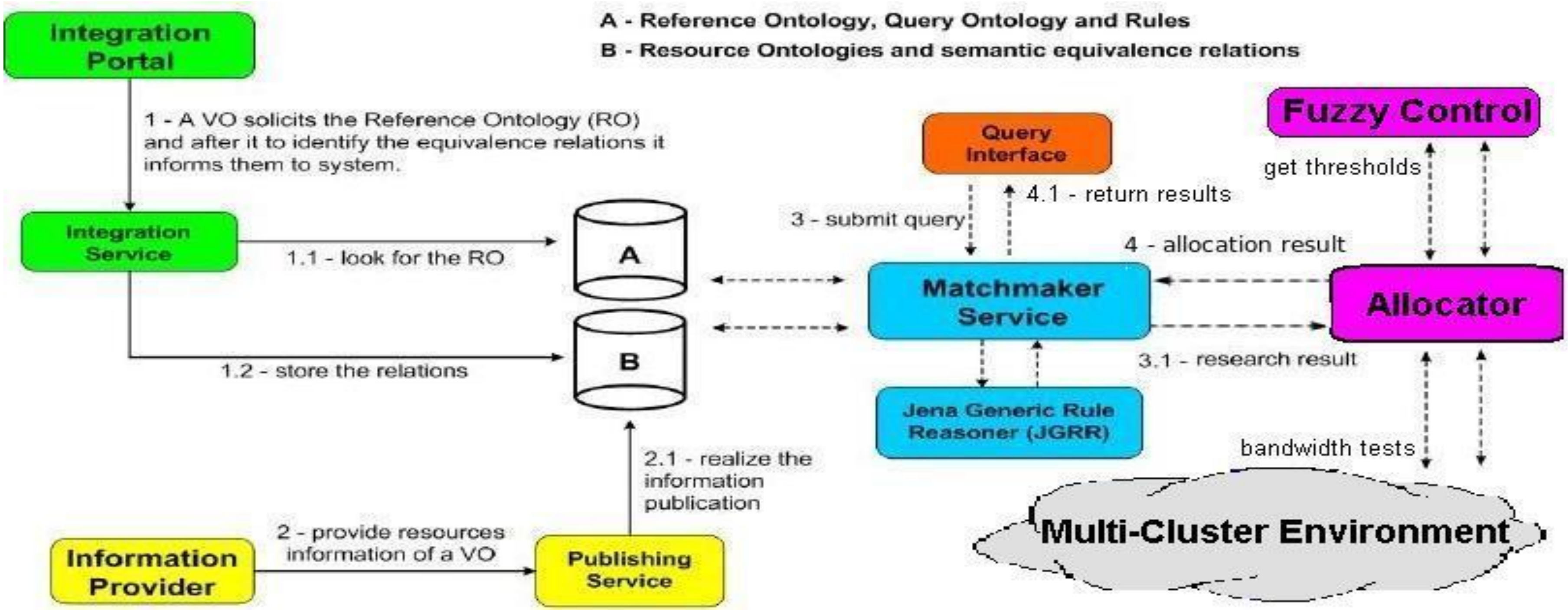
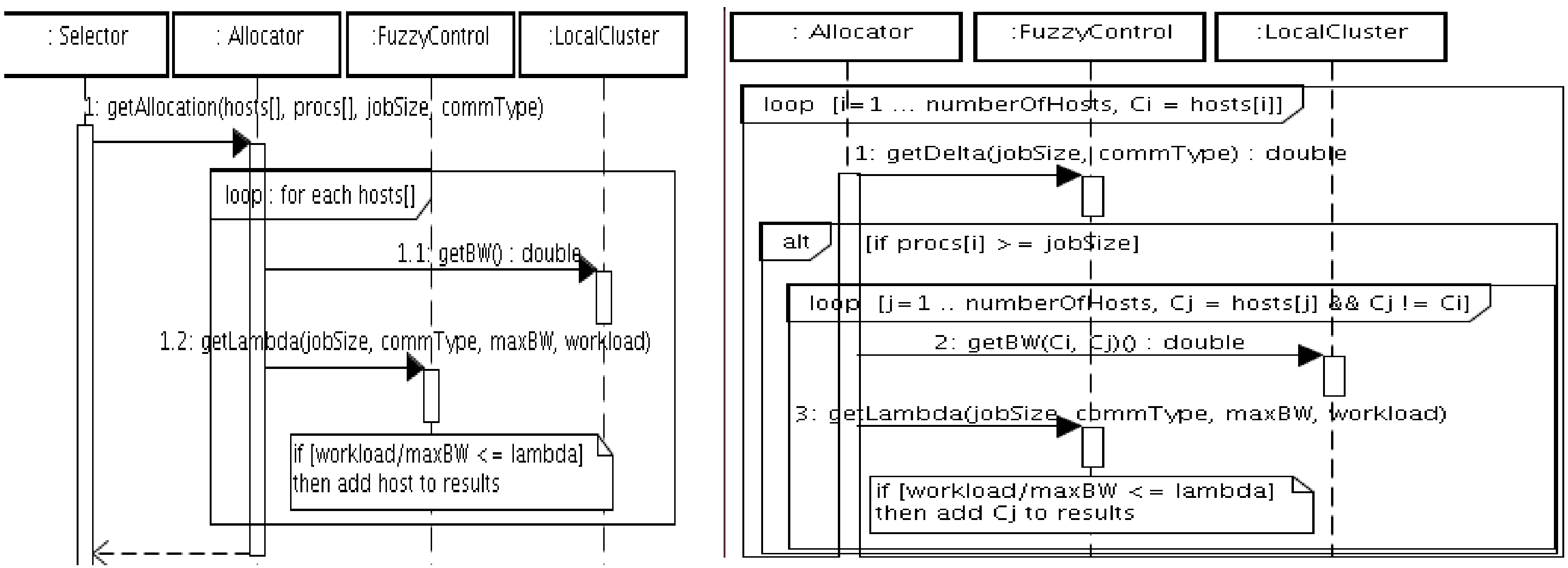


Fig. 3: The new proposed architecture

Proposed Approach



(a) Allocation through a unique cluster (b) Allocation among clusters

Fig. 4: Allocator sequence diagram

Agenda

- Motivation
- Introduction
- Related Work
- Proposed Approach
- Experimental Environment and Results
- Conclusions and Future Work

Experimental Environment and Results

- The *ontology components* were built utilizing *Protégé 3.3.1* and the multi-cluster environment was simulated using the *SimGrid software package*;

-*The goal* of the SimGrid project is to facilitate research in the area of distributed and parallel application scheduling on distributed computing platforms, ranging from a simple network of workstations to *large computational grids*;

-The SimGrid toolkit provides *core functionalities* for the simulation of distributed applications in *heterogeneous distributed environments*;

-This characteristic of the package is interesting since it enables simulation of *multi-cluster configurations;*

VO-A	# processors	VO-B	# processors
Cluster_01	7	Cluster_1	4
Cluster_02	5	Cluster_2	4
Cluster_03	3	-	-

Table 1: Multi-cluster configuration

Experimental Environment and Results

- After creating the multi-cluster configuration, we identified *a number of basic concepts* to be used in *both virtual organizations* for requesting resources;

- Table 2 presents how *these request concepts* are understood in virtual organizations A and B. Virtual organization A represents a Canadian environment, whereas virtual B is a Brazilian configuration.

Request's Concepts	VO-B's Concepts	VO-A's Concepts
min_number_of_cpus	[numero_cpus]	[number_of_cpus]
min_virtual_memory_available	[swap_livre_MB]	[free_virtual_memory]
min_clock_speed	[velocidade_cpu]	[max_clock_speed]
owner	[nome_conta]	[login]
number_of_cpus	[numero_cpus]	[number_of_cpus]
min_physical_memory_available	[memoria_livre_MB]	[free_physical_memory]
min_free_disk_space	[espaco_disco_livre_GB]	[available_space]
os_type	[SO, SistemaOperacional]	[os_type]
max_load_cpu_15min		[percentage_load_15min]
max_load_cpu_5min		[percentage_load_5min]
max_load_cpu_1min		[percentage_load_1min]

Table 2: The ontology set utilized for the experiments

Experimental Environment and Results

- The first test was a query where a user required a certain minimum amount physical memory, operating system type (UNIX) and *number of processors equal to 2*.

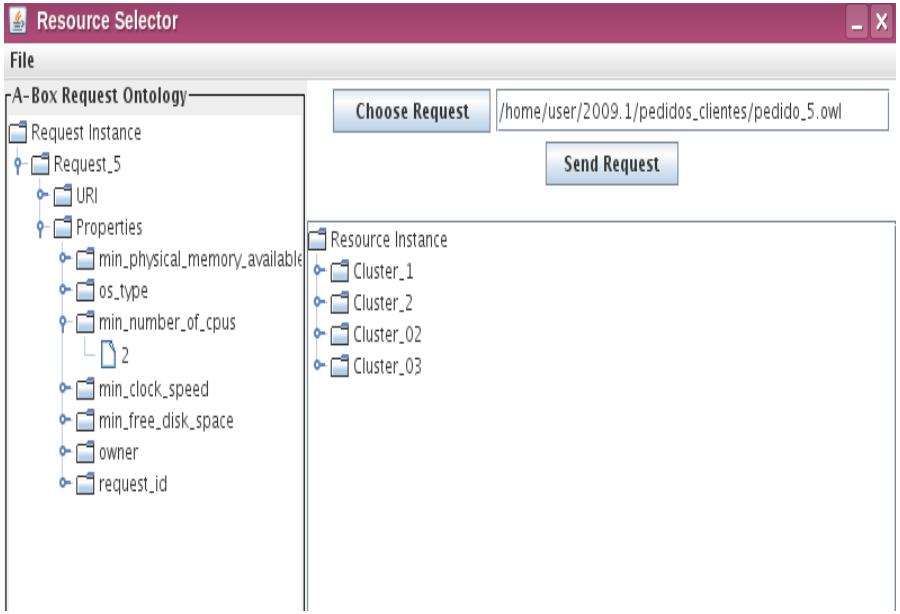


Fig. 5: The first test without dynamic information

Experimental Environment and Results

- Figure 6 shows the second test, where the *same query* was executed under the new system;
- In addition to the available environment, this figure *presents* numbers related to the *number of processors* in each node and *link workload to server*.

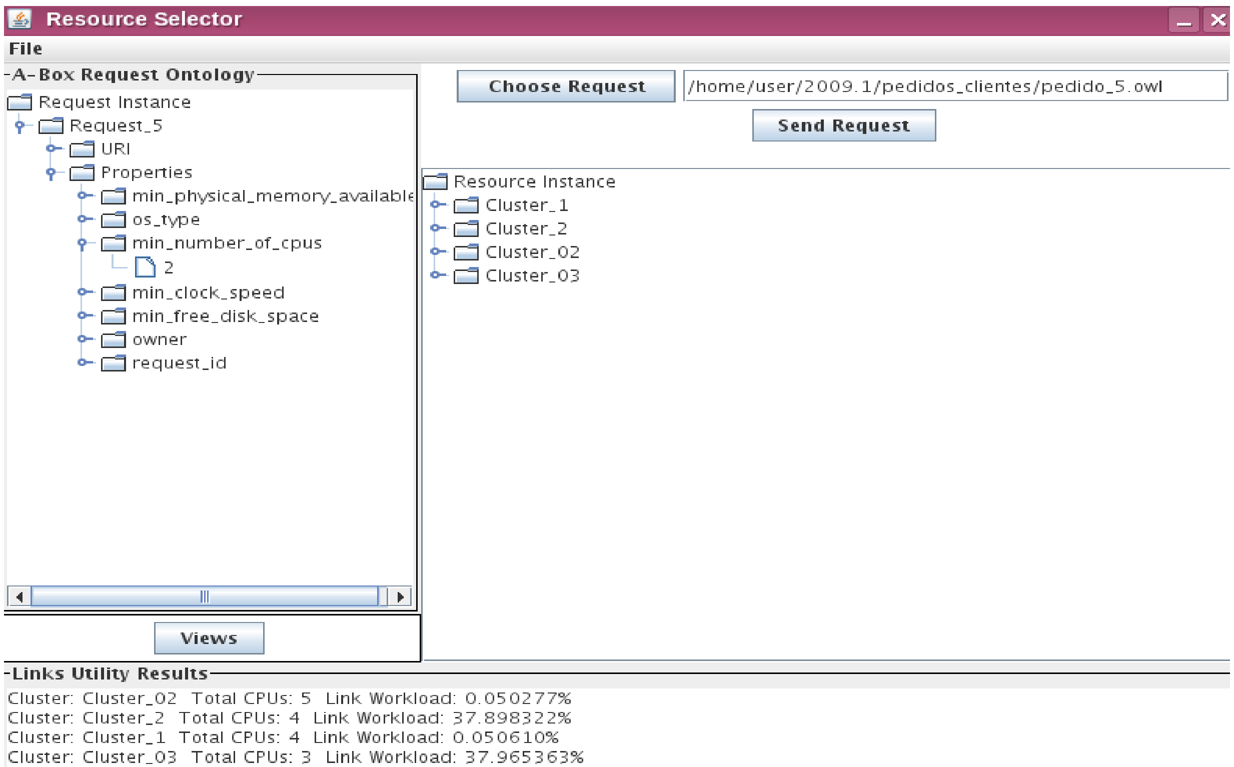


Fig. 6: The first query in the new environment

Experimental Environment and Results

-The third experiment represents a query as it was in the original environment, where *a user requires 6 processors*. Figure 7 shows the result for this query;
- *As expected*, the system replies that *no match was found*. This answer is based upon the *static information* available in the original system.

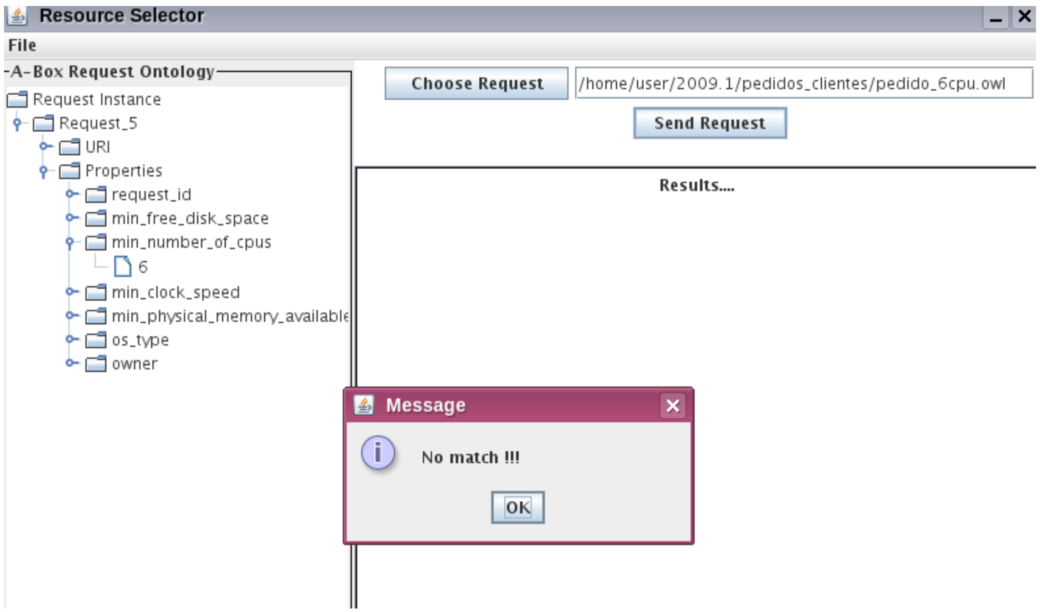


Figure 7: The third experiment without a dynamic information

Experimental Environment and Results

- The next experiment represents the *previous* one considering all *new features* of the new proposal;
- A comprehensive picture of available multi-clusters is presented in Figure 8;
- This figure shows the same third query, however now executed under the new proposal.

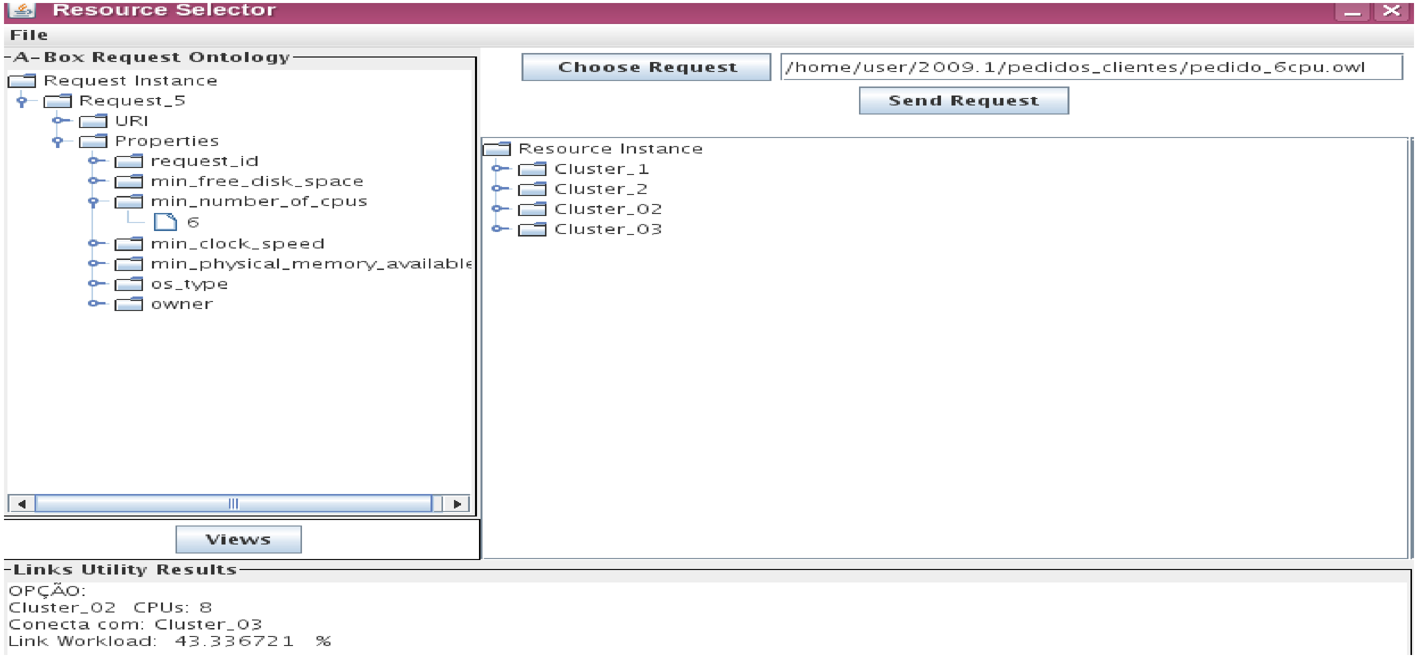


Figure 8: The third experiment under the new environment

Experimental Environment and Results

-The result involves *Cluster_02* and *Cluster_03* which combined reaches the quantity of *8 processors*, sufficient to execute the job;

- It is also shows the *link workload* between these two clusters;

- It is clear from Figure 6 that multi-clusters exist to support the query. The query under the new proposal considers all processors that exist inside all clusters that have one communication link;

- The final result also considers the communication patterns, thus suggesting the job splitting;

-The Result panel shows all clusters that satisfy the query. On the other hand, the *LinksResults* highlights the best calculated solution.

Agenda

- Motivation
- Introduction
- Related Work
- Proposed Approach
- Experimental Environment and Results
- Conclusions and Future Work

UFSC

Western

Conclusion and Future Work

-We have presented *a novel dynamic resource matching approach*, which considers computer resources and communication links characteristics, *to gather nodes from multi-cluster configurations as a grid environment*;

- The idea is to provide *a standard interface* to create a grid environment to execute distributed and parallel applications. The proposal can be understood as an *advance resource* (AR) function to execute application in multi-cluster configurations;

- The contribution adopted the *ontology and fuzzy logic paradigms* for a standard description of resources to *compound a multi-cluster* environment and to monitor resources, and also to dynamically *match end-users*' requests and available computer resources and communication link workloads.

Conclusion and Future Work

- The proposal was tested through the utilization of the *SimGrid simulator*;

- The proposal has reached a successful level, providing interesting multi-cluster configurations options to application programmers;

Conclusion and Future Work

- As a future work we are planning to implement co-reservation and co-allocation algorithms and build an entire *meta-scheduler environment*;

- This meta-scheduler would consider different type of multi-cluster configurations, including Ad Hoc nodes;

- We are also planning a *fault-tolerance module* to provide checkpoint facility to more complex applications;

- Finally, we are also planning a component to automatic execute an application in the *best selected configuration*.

Thank you !

Questions ?

Mario Dantas
mario@inf.ufsc.br