Discovering Resources and Mapping in Large-Scale Distributed Environments

Konstantinos Karaoglanoglou
PhD Student
Department of Informatics
Aristotle University of Thessaloniki
Tuesday, October 13, 2009
Lisbon

Outline

- Introduction
- Depiction of Distributed Environments
- Resource Discovery and Mapping
  - in Dynamic Distributed Environments
  - with a Matchmaking Mechanism
  - in environments prone to Resource Failures
- Future Research
  - towards the Resource Evolution phenomenon
  - regarding Trust Issues

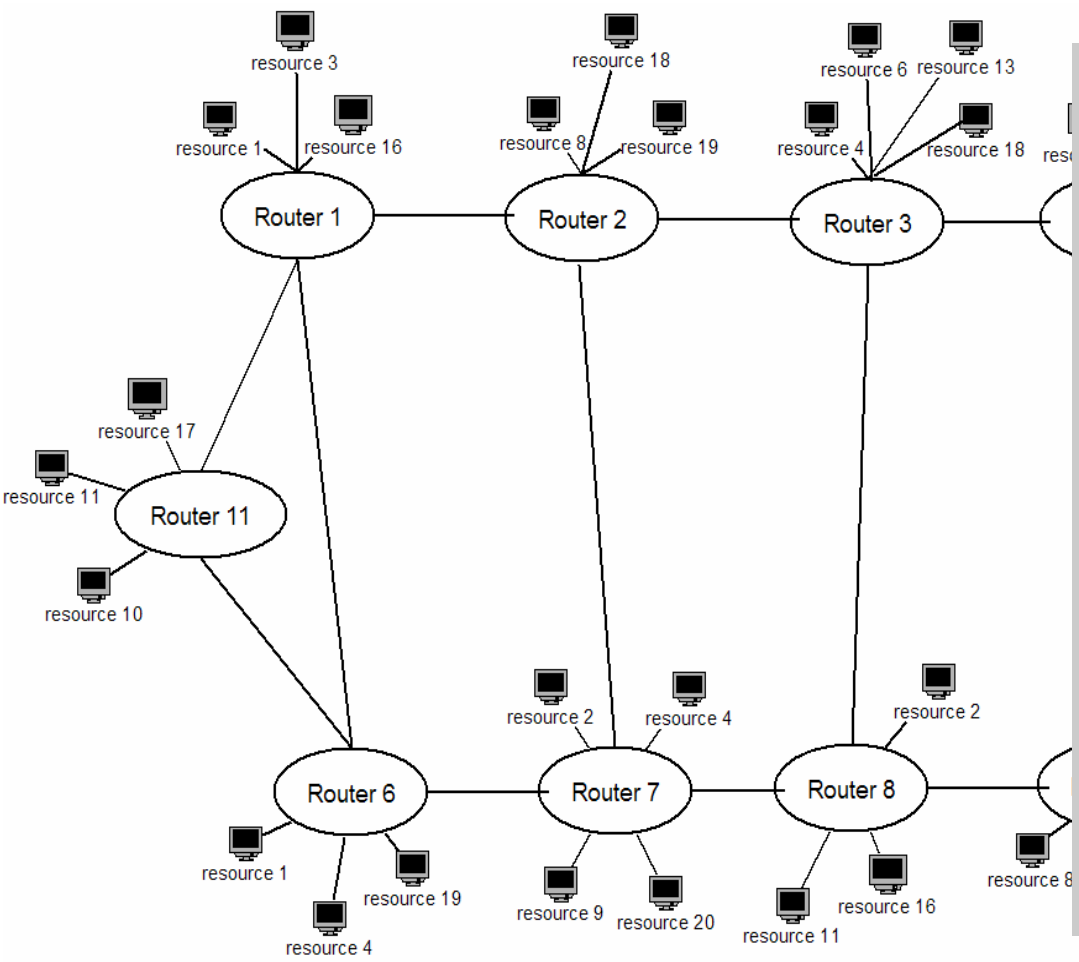
Introduction (1/2)

- The base of distributed systems is the concept of resource-sharing. The types of resources shared in a distributed environment could be desktop systems, clusters, storage-devices, and large data-sets.
- The question is what happens when a remote application or a remote user requests access to a remote resource either to execute a job or to have access in the resource's data?
- A mechanism provided by the distributed infrastructure should be able to discover an appropriate resource for a request.
- Discovering resources in traditional computing environments is relatively easy because the number of shared resources is small and all resources are under central control.

Introduction (2/2)

- In a distributed system there are certain factors that make the resource discovery and mapping problem difficult to solve.
- Some of these factors are: the huge number of resources, the heterogeneity of resources, and the uncertainties met in such systems.
- In our previous efforts, we have dealt with the discovery of resources and mapping, taking into consideration certain particularities of distributed systems, such as:
  - Dynamicity
  - Heterogeneity
  - Resource Failures

Depiction of a Distributed Environment



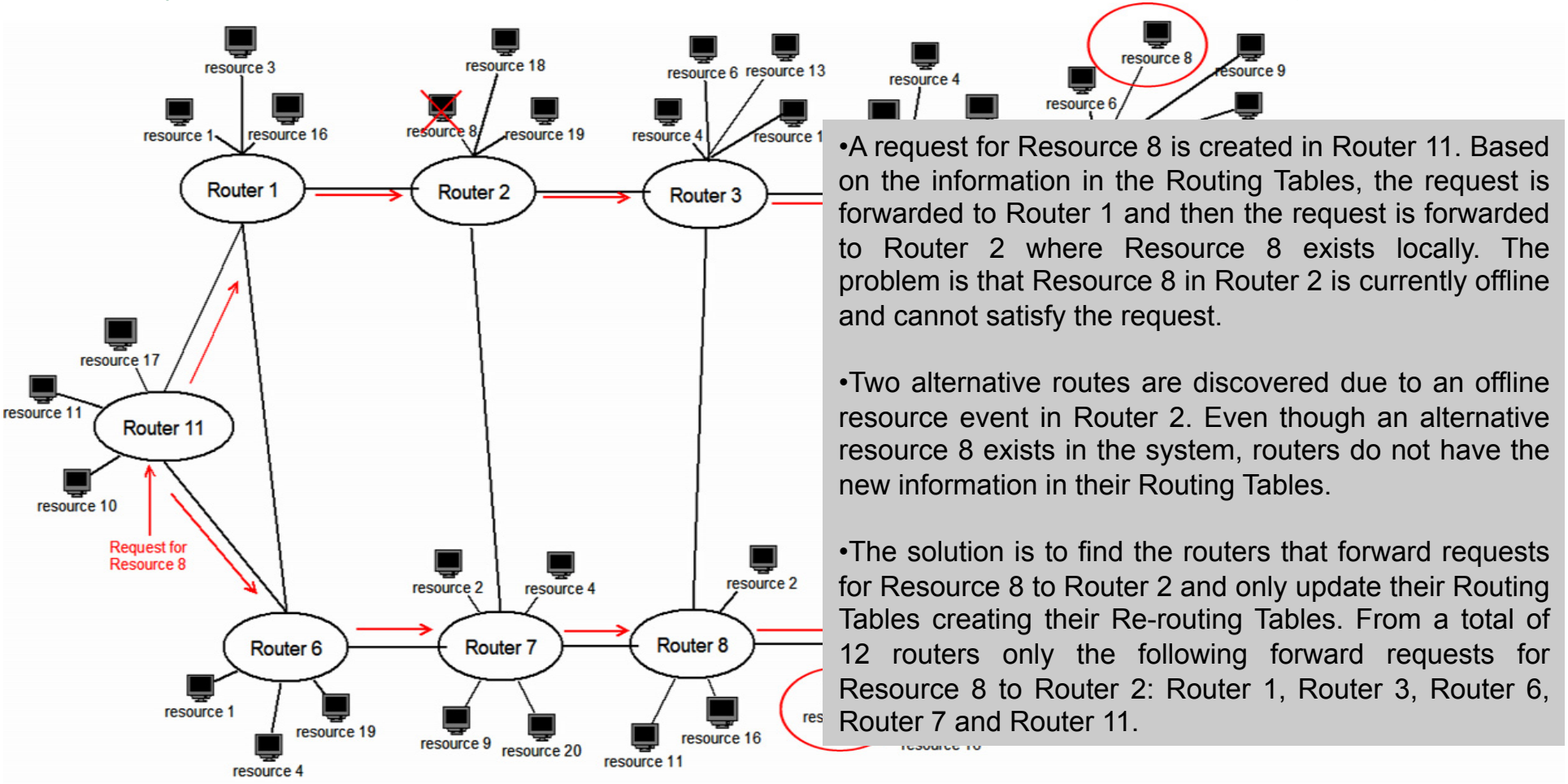
- A distributed system can be seen as an environment comprised by routers and resources.
- Each router is in charge of its local resources and also connects with other routers in the system.
- At some point of time a remote application requests for a resource that can satisfy it.
- The resource discovery mechanism is responsible of discovering the appropriate resource and directing the application until it reaches the appropriate resource.

Konstantinos Karaoglanoglou
<kkaraogl@csd.auth.gr>

Resource Discovery and Mapping in Dynamic Distributed Environments (1/2)

- A distributed system is characterized by resources that can get in an offline state at any time.
- The reason that causes this dynamic behaviour is mainly the distributed ownership. (a) An owner may establish a policy on a workstation that states that a foreign job can be run on a machine only at certain periods of time. (b) An owner may establish a policy on a workstation that states that a foreign job can be run on a machine when the local load is under a certain limit. When the local load reaches that limit, a resource must exit the system and execute local jobs only.
- An effective resource discovery mechanism should be able to overcome the dynamicity of such a system.
- A mechanism, previously proposed, uses Routing Tables in order to efficiently direct the applications to the appropriate resources. The problem is that the Routing Tables mechanism assumes that all resources are permanently online.
- We have enhanced the Routing Tables mechanism with an updating procedure, in order to cover the cases of offline resource events, bound to occur in any distributed environment.

Resource Discovery and Mapping in Dynamic Distributed Environments (2/2)



•A request for Resource 8 is created in Router 11. Based on the information in the Routing Tables, the request is forwarded to Router 1 and then the request is forwarded to Router 2 where Resource 8 exists locally. The problem is that Resource 8 in Router 2 is currently offline and cannot satisfy the request.

•Two alternative routes are discovered due to an offline resource event in Router 2. Even though an alternative resource 8 exists in the system, routers do not have the new information in their Routing Tables.

•The solution is to find the routers that forward requests for Resource 8 to Router 2 and only update their Routing Tables creating their Re-routing Tables. From a total of 12 routers only the following forward requests for Resource 8 to Router 2: Router 1, Router 3, Router 6, Router 7 and Router 11.

Konstantinos Karaoglanoglou
<kkaraogl@csd.auth.gr>

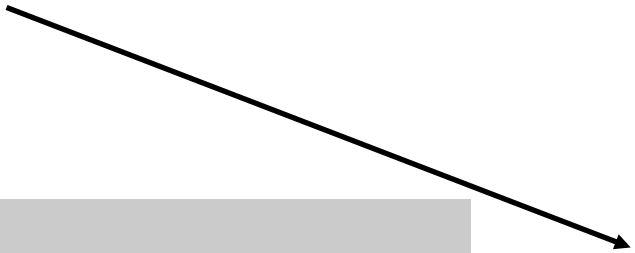
Resource Discovery and Mapping with a Matchmaking Mechanism (1/2)

- A distributed system is composed of heterogeneous networked resources with different technical characteristics.
- When an application requests a resource of specific technical characteristics, an efficient resource discovery mechanism should be able to direct this request to the most suitable request.
- We have enhanced the routers of the distributed system with matchmaking capabilities. Taking into consideration the limited computing capabilities routers have, we kept the matchmaking framework simple.

Resource Discovery and Mapping with a Matchmaking mechanism (2/2)

Architecture: Intel
Operating System: Solaris26
Minimum Disk: 35000
Minimum Memory: 512

Request



Resource Type 1
Architecture: Intel
Operating System: Linux
Available Disk: 20000
Available Memory: 128

Resource Type 2
Architecture: SGI
Operating System: Irix6
Available Disk: 40000
Available Memory: 1024

Resource Type 3
Architecture: Intel
Operating System: Solaris26
Available Disk: 40000
Available Memory: 1024

Resource Type 4
Architecture: Intel
Operating System: Solaris26
Available Disk: 35000
Available Memory: 512

Available Resources

Matchmaking Rules

•The architecture and operating system characteristics of the request must match the architecture and operating system characteristics of the resource.

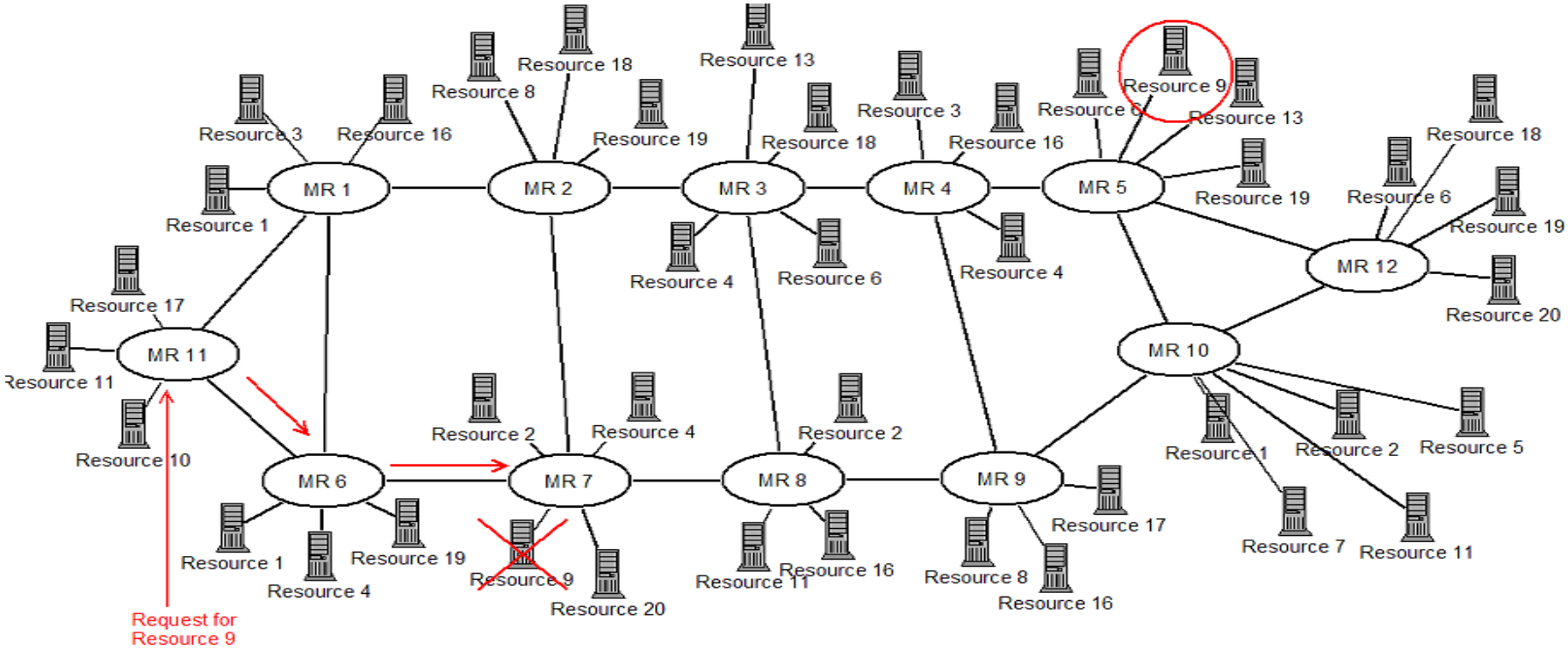
•The minimum disk size required by the request must be smaller or equal to the available disk size of the resource.

•The minimum memory space required by the request must be smaller or equal to the available memory space of the resource.

Resource Discovery and Mapping in Distributed Environments prone to Resource Failures (1/3)

- Due to the huge number of resources in a distributed system, the probability of resource failures is increased.
- Resource failures are unpredictable and common due to hardware faults, software faults or power outages.
- Combining the matchmaking and the re-routing approaches (discussed previously), we have proposed a resource discovery scheme that can guarantee discovering the most suitable resource and then directing an application's request to the appropriate resource in the distributed environment, where resource failures are a common fact.
- Moreover the proposed scheme deals with the phenomenon of consecutive resource failures.

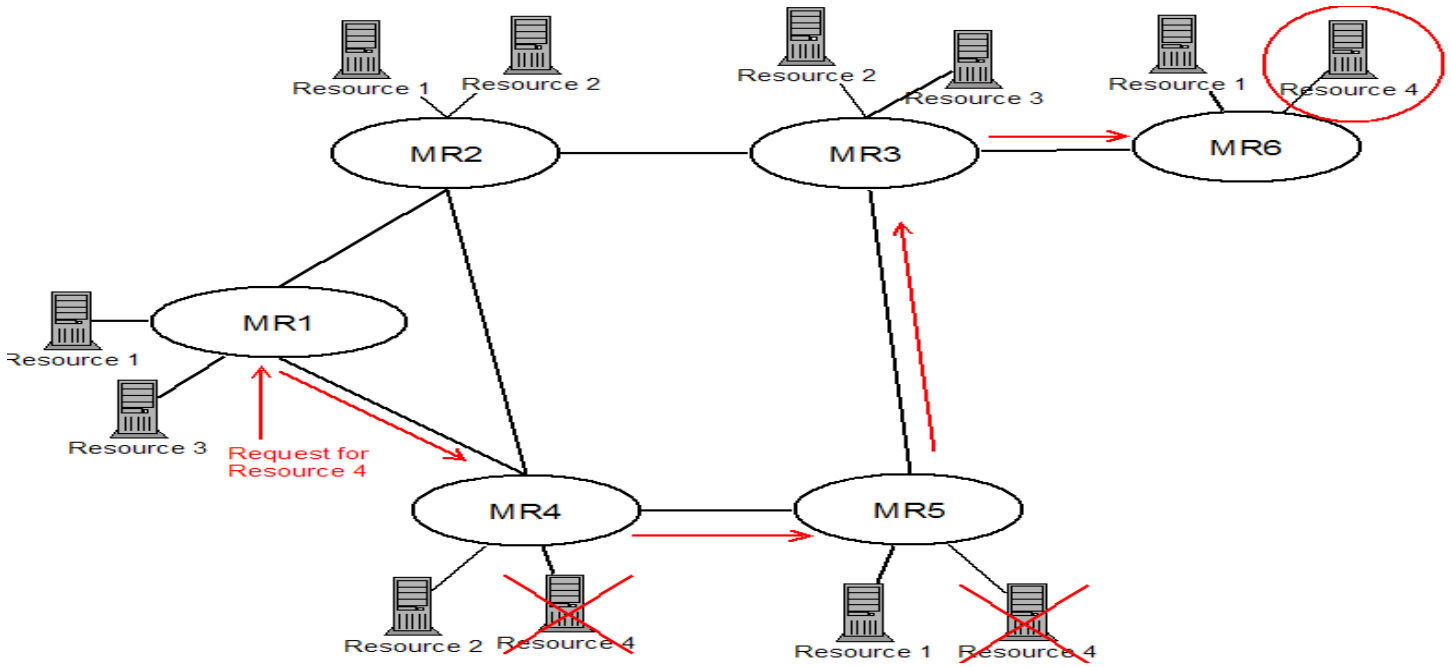
Resource Discovery and Mapping in Distributed Environments prone to Resource Failures (2/3)



An example of a re-routing event due to a resource failure

Konstantinos Karaoglanoglou
<kkaraogl@csd.auth.gr>

Resource Discovery and Mapping in Distributed Environments prone to Resource Failures (3/3)

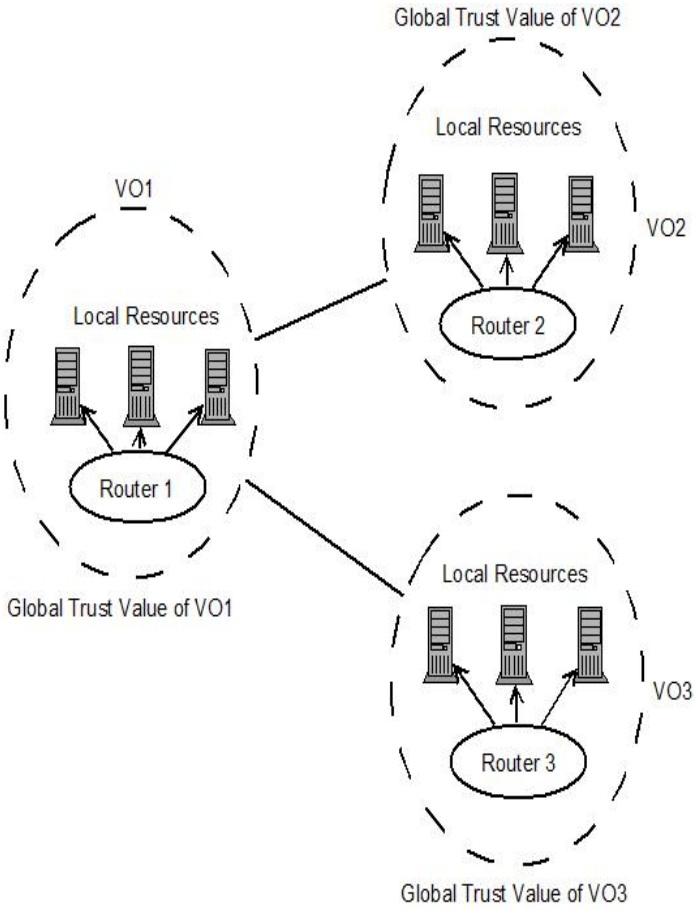


An example of consecutive resource failures

Konstantinos Karaoglanoglou
<kkaraogl@csd.auth.gr>

Work in Progress: Trust Management

- The model suggests that routers and resources comprise virtual organizations (VOs) within a distributed system, where in each VO a router controls locally a number of resources.
- Global trust values are assigned to the system's VOs. These trust values show whether a VO and subsequently its resources are trustworthy or not.
- Furthermore, the under work mechanism also covers the cases of dynamic changes in the trustworthiness of VOs. For instance, VOs that in the past were untrustworthy could now be trustworthy. The proposed mechanism is able of detecting these dynamic changes, so that the directing of requests occurs in an up-to-date way.



Konstantinos Karaoglanoglou
<kkaraogl@csd.auth.gr>

Simulation Tests

- Testing is crucial for evaluating the performance of the implemented resource discovery and mapping mechanisms. All the presented mechanisms have been evaluated via simulation.
- The presented mechanisms have been tested in a number of experiments taken into consideration that a mechanism should perform well in both small and large-scale systems. We use the Grid Graph generator to produce the systems for our simulation needs.
- Grid Graph generator produces the backbone of the systems, meaning the routers. After this we allocate a certain number of resources to each router of the system. Each resource has certain technical characteristics and therefore can satisfy specific requests.
- We use the ability of the Grid Graph generator of producing different topologies for the same size of system.
- During the simulation, requests for random resources are created in random routers. The implemented mechanisms are evaluated by their ability of directing the requests through the system to the appropriate resource that can satisfy them.

Future Research

- We intend to extensively work in the direction of trusted mapping and directing applications to reliable resources in a distributed environment. Apart from the trust management mechanism that is currently under work, we intend to construct a trust metric able to react in the behavioral changes of the entities of a distributed environment.The objective of this research direction is to provide robustness to distributed environments against malicious behaviours.

- Moreover, we intend to deal with efficient mapping mechanisms that take into consideration the “resource evolution” phenomenon, weighing in the technical changes that so commonly occur in the resources of distributed environments.

References

- **Book Chapter**
  - K. Karaoglanoglou, and H. Karatza, "Trust Calculation & Management in P2P and Grid Systems", in the the Handbook of Research on P2P and Grid Systems for Service-Oriented Computing: Models, Methodologies and Applications, Edited by Dr. Nick Antonopoulos, Mr. Georgios Exarchakos, Dr. Maozhen Li and Dr. Antonio Liotta, IGI Global, to appear.

- **Journal Publication**
  - K. Karaoglanoglou and H. Karatza, "Resource Discovery in a Dynamical Grid based on Re-routing Tables", Simulation Modelling Practice and Theory, Elsevier, Volume 16, Issue 6, Pages 704-720, July 2008.

- **Conference Proceedings Publications**
  - K. Karaoglanoglou, and H. Karatza, "Performance Evaluation of a Resource Discovery Scheme in a Grid Environment Prone to Resource Failures", in the 8th International Workshop on Performance Modeling, Evaluation, and Optimization of Ubiquitous Computing and Network Systems (PMEO-UCNS 2009) in conjunction with IEEE International Parallel & Distributed Processing Symposium (IPDPS), Rome, Italy, May 25-29, 2009.

  - K. Karaoglanoglou and H. Karatza, "Resource Discovery in a Grid System based on Matchmaking-Routers", Proceedings of the 12th Panhellenic Conference on Informatics, PCI, IEEE Computer Society, August 28-30, 2008, Samos, Greece, pp. 73-77.

Our Team

- Parallel and Distributed Systems Group
  - Coordinator
    - Prof. Helen Karatza <karatza@csd.auth.gr>
  - Members
    - Stylianos Zikos <szikos@csd.auth.gr>
    - Konstantinos Karaoglanoglou <kkaraogl@csd.auth.gr>
    - Zafeirios Papazachos <zpapazac@csd.auth.gr>
    - Fotis Loukos <fotisl@csd.auth.gr>
    - Georgios Stavrinides <gstavrin@csd.auth.gr>

Thank you!

Konstantinos Karaoglanoglou
<kkaraogl@csd.auth.gr>