# A Reinforcement Learning Based Management

# Solution for Context-Aware Environments

# Applied on Greeen Datacenters

## Context Data Gathering, Context Representation

## And Fuzzy Logic Policy Negotiation

Daniel Moldovan

June, 22, 2010

Abstract

This report presents a framework for building agent-based autonomic context-aware systems. The solution uses ontology representations for context modeling and policies for rule enforcement. For determining policy violations and quantifying their severity the concept of entropy is defined and used. An extensible data gathering mechanism is provided for retrieving information about context resources and feeding that information in the ontology context representation. An action enforcement mechanism is constructed which uses web service invocation for maximum flexibility. For finding the best sequence of actions which brings the context into an accepted state an action search algorithm based on reinforcement learning is introduced. A method for negotiating between multi-valued conflicting policies is defined and implemented using fuzzy logic. The proposed framework is applied on a simulated smart laboratory in order to assess the overhead and correctness of the reasoning process. The framework is further evaluated in the context of an energy-efficient datacenter management. The datacenter evaluation is performed on both simulated and real-worsld datacenters.

**Table of Contents**

[1 Introduction 1](#_Toc264798540)

[1.1 Motivation 1](#_Toc264798541)

[1.2 Contributions 2](#_Toc264798542)

[1.3 Publications 2](#_Toc264798543)

[1.3. Problem Statement and Goals 3](#_Toc264798547)

[1.3.1. Problem Statement 3](#_Toc264798548)

[1.3.2. Problem Goals 3](#_Toc264798549)

[2.1 Background 6](#_Toc264798550)

[2.1.1 Advanced Object- Oriented Concepts and Principles 6](#_Toc264798551)

[2.1.2 Autonomic Computing Paradigm and Autonomic Systems 7](#_Toc264798552)

[2.1.3 Intelligent Agents and Pervasive Computing 8](#_Toc264798553)

[2.1.4 Artificial Intelligence Learning, Knowledge Representation and Reasoning 9](#_Toc264798554)

[2.1.5 An introduction to Green Computing, Virtualization and Server Consolidation 10](#_Toc264798555)

[3 Related Work 12](#_Toc264798556)

[3.1 An overview of existing negotiation and bargaining solutions 12](#_Toc264798557)

[3.2 An overview of existing self-adapting systems 13](#_Toc264798558)

[4 Analysis and Design 16](#_Toc264798559)

[4.1 The context model 16](#_Toc264798560)

[4.1.1 Context Ontology Representation 16](#_Toc264798561)

[4.1.2 Policies 17](#_Toc264798562)

[4.1.3 Context Entropy 18](#_Toc264798563)

[4.1.4 State reward 18](#_Toc264798564)

[4.1.5 Inter-Independent Resources Group 19](#_Toc264798565)

[4.1.6 RAP instantiation for the self-healing system 21](#_Toc264798566)

[4.1.7 RAP instantiation for the self-adapting datacenter 21](#_Toc264798573)

[4.1.8 Load Distribution Strategy 21](#_Toc264798574)

[4.2 Reinforcement Learning Algorithm 22](#_Toc264798575)

[4.3 QoS-Energy Consumption Negotiation 24](#_Toc264798576)

[4.4 The Conceptual Architecture 26](#_Toc264798577)

[5 Implementation 30](#_Toc264798578)

[5.1 System Components 30](#_Toc264798579)

[5.2 System Input 31](#_Toc264798580)

[5.3 System Output 32](#_Toc264798581)

[5.4 Implementation Technology 33](#_Toc264798582)

[5.5 Ontology Representation Implementation 35](#_Toc264798583)

[5.5.1 Environment Ontology Implementation 35](#_Toc264798584)

[5.5.2 Datacenter Ontology Representation 35](#_Toc264798585)

[5.6 Agents Implementation 37](#_Toc264798586)

[5.6.1 Context Interpreting Agent (CIA) 37](#_Toc264798587)

[5.6.2 Reinforcement Learning Agent (RLA) 38](#_Toc264798588)

[5.6.3 X3D Agent Implementation 39](#_Toc264798589)

[5.6.4 GUI Agent Implementation 41](#_Toc264798590)

[5.7 Utility components 46](#_Toc264798591)

[5.7.1 Fuzzy Logic Negotiator 46](#_Toc264798592)

[5.7.2 Server Management Proxy 47](#_Toc264798593)

[5.7.3 Server Information Gathering Endpoint 48](#_Toc264798594)

[6 Testing and results 50](#_Toc264798595)

[6.1 Self-Healing Environment Case Study 50](#_Toc264798596)

[6.1.1 Algorithm Overhead Evaluation 52](#_Toc264798597)

[6.1.2 Interactive Evaluation 52](#_Toc264798598)

[6.2 Self-Adapting Datacenter Case Study 54](#_Toc264798599)

[6.2.1 Simulated Datacenter Evaluation 55](#_Toc264798600)

[6.2.2 Real Datacenter System Evaluation 58](#_Toc264798601)

[7 Conclusions 66](#_Toc264798608)

[8.1. Future Work 67](#_Toc264798610)

[Bibliography 68](#_Toc264798611)

[Appendix A: List of Acronyms 72](#_Toc264798612)

[Appendix B: Java Source Code 74](#_Toc264798613)

[FuzzyLogicNegotiator.java 74](#_Toc264798614)

[ServerInfoSAXHandler.java 81](#_Toc264798615)

[ResourceMonitorPlotter.java 84](#_Toc264798616)

[ResourceMonitorPieChartPlotter.java 84](#_Toc264798617)

[X3DAgent.java 86](#_Toc264798618)

[DefaultServer.java 97](#_Toc264798619)

[DefaultContextElement.java 111](#_Toc264798620)

[Appendix C: C# Source Code 113](#_Toc264798621)

[ServerMonitor.cs 113](#_Toc264798622)

[ServerInfo.cs 114](#_Toc264798623)

List of Figures

[Figure 1.1: Server Consolidation trough Virtualization 10](file:///C:\Documents%20and%20Settings\Me\Desktop\incercare_intelij\selfoptimizing\Licenta.docx#_Toc264798624)

[Figure 4.1: RAP Ontology Representation 17](#_Toc264798625)

[Figure 4.2: Datacenter Ontology Representation 17](#_Toc264798626)

[Figure 4.3: IIRG Extraction Algorithm Trace 20](#_Toc264798627)

[Figure 4.4: Server Skewing with Tail 22](#_Toc264798628)

[Figure 4.5: Self-Healing system Reinforcement Learning Process 23](#_Toc264798629)

[Figure 4.6: Datacenter Reinforcement Learning Process 24](#_Toc264798630)

[Figure 4.7: Situation which requires negotiation 25](#_Toc264798631)

[Figure 4.8: Desirability-Based Negotiation 26](#_Toc264798632)

[Figure 4.9: System Conceptual Architecture 27](#_Toc264798633)

[Figure 4.10: Agents Interactions 28](#_Toc264798634)

[Figure 5.1: System Components Diagram 30](#_Toc264798635)

[Figure 5.2: System Ontology Representation 31](#_Toc264798636)

[Figure 5.3: Jade Architecture 34](#_Toc264798637)

[Figure 5.4: Jade Asynchronous Message Passing Mechanism 34](#_Toc264798638)

[Figure 5.5: Ontology to Java Code Mapping 35](#_Toc264798639)

[Figure 5.6 : CPU Entity Ontology Representation 36](#_Toc264798640)

[Figure 5.7: DTO Design Pattern SensorAPI Usage 38](#_Toc264798641)

[Figure 5.8: Xj3D Architecture 40](#_Toc264798642)

[Figure 5.9: Environment Monitor GUI 42](#_Toc264798643)

[Figure 5.10: Task Monitor GUI 43](#_Toc264798644)

[Figure 5.11: Server Monitor Window Architecture 44](#_Toc264798645)

[Figure 5.12: Server Monitor Window 44](#_Toc264798646)

[Figure 5.13: Datacenter Management Log 45](#_Toc264798647)

[Figure 5.14: Server-Task Negotiation 47](#_Toc264798648)

[Figure 5.15: Server Management Interface Architecture 48](#_Toc264798649)

[Figure 6.1: 28h Random Values Evaluation 52](#_Toc264798650)

[Figure 6.2: X3D Smart Environment Representation 53](#_Toc264798651)

[Figure 6.3: Face Recognition Policy broken 54](#_Toc264798652)

[Figure 6.4: X3D Datacenter Representation 55](#_Toc264798653)

[Figure 6.5: Deploy Task 1 on Server 1 Action 57](#_Toc264798654)

[Figure 6.6: Wake up Server 3 Action 57](#_Toc264798655)

[Figure 6.7: Deploy Task 2 on Server 3 Action 57](#_Toc264798656)

[Figure 6.8: Deploy Task 3 on Server 3 Action 57](#_Toc264798657)

[Figure 6.9: Case Study Datacenter Structure 58](#_Toc264798658)

[Figure 6.10: Server Technology Mapping 59](#_Toc264798659)

[Figure 6.11: Server Interaction 59](#_Toc264798660)

[Figure 6.12: Initial State 61](#_Toc264798661)

[Figure 6.13: Final State 61](#_Toc264798662)

[Figure 6.14: Running Virtual Machine 62](#_Toc264798663)

[Figure 6.15: Server Monitor Snapshot 63](#_Toc264798664)

List of Tables

[Table 4.1: Resource Dependency Matrix example 19](#_Toc264369918)

[Table 4.2: Self-healing environment RAP mapping 20](#_Toc264369919)

[Table 4.3: Self-Adapting Datacenter RAP Mapping 21](#_Toc264369920)

[Table 6.1: Sensors Values 49](#_Toc264369921)

[Table 6.2: Policies 49](#_Toc264369922)

[Table 6.3: Sensors Associated Actuators 50](#_Toc264369923)

[Table 6.4: Datacenter servers’ characteristics 55](#_Toc264369924)

[Table 6.5: Received tasks SLA requirements 55](#_Toc264369925)

[Table 6.6: Virtual Machines Resource Allocation 59](#_Toc264369926)

List of Listings

[Listing 4.1: IIRG Extraction Algorithm 19](#_Toc264370230)

[Listing 4.2: Reinforcement Learning Algorithm 22](#_Toc264370231)

[Listing 5.1: SWRL QoS Policy Example 32](#_Toc264370232)

[Listing 5.3: Context Management Log Example 33](#_Toc264370233)

[Listing 5.4: CPU Entity Java Representation 36](#_Toc264370234)

[Listing 5.5: SensorAPI functionality 37](#_Toc264370235)

[Listing 5.6: X3D File Shape Entry 41](#_Toc264370236)

[Listing 5.7: FCL specification 46](#_Toc264370237)

[Listing 5.8: Server Management Interface 47](#_Toc264370238)

[Listing 5.9: WIN32 WMI Specific Functions 48](#_Toc264370239)

[Listing 6.1: Alarm Sensor Simulation Web Service 50](#_Toc264370240)

List of Equations

[Equation 4.1: Policy Specification 17](#_Toc264370361)

[Equation 4.2: Entropy Formula 17](#_Toc264370362)

[Equation 4.3: Reward function 18](#_Toc264370363)

[Equation 4.4: Dependency Relation 18](#_Toc264370364)

[Equation 4.5: Center Of Gravity Formula 26](#_Toc264370365)

Chapter 1

# Introduction

This chapter describes the motivation behind the work that is presented in the rest of the report and the contributions this work brings to the selected field.

## Motivation

The ever-growing complexity of IT threatens its growth ]. The need to integrate heterogeneous environments into corporate computing systems and manage them approaches the maximum complexity a human can manage ]. As systems become more interconnected and diverse, system architects are less and less capable of anticipating component interactions, leaving those decisions to be made at runtime.

It is estimated that up to 50% of the total cost of ownership involves recovering or preparing against failures. Also, it has been estimated that IBM adds about 15.000 people per year in order to provide assistance to clients with complex platforms ].

The only option to the complexity problem is autonomic computing. By enhancing systems with autonomic properties the need for human intervention should decrease and the fault recovery time can increase. Also, the need for fault prevention decreases, as fault recovery can take place in real time without any costly human intervention.

Worldwide data centers electricity consumption accounts for almost 2% of the world production and is expected to overcome the 40% of Total Cost of Ownership of worldwide IT by 2012 ]. The US Environmental Protection Agency estimated that in 2006 the servers and datacenters power consumption accounted for 61 billion kWh, about 1.5 % of total U.S electricity consumption and for a cost of $4.5 billion ].

The same organization points out that the average data center is only 30% efficient, with 70% of the electricity lost due to inefficiencies of power and heat dissipation, along with powering cooling equipment ]. The environmental impact of datacenter expansion is of great importance, every server using 7,000 kWh of electricity and indirectly generating four tons of carbon dioxide emissions per year ].

This has lead to the need of finding environmental friendly methods for managing datacenters while maintaining performance. This new research trend has been called Green IT or Green Computing. The philosophy of Green IT is designing and using computer resources in an environmentally friendly way. This work has as aim reducing the overall number of servers used worldwide and further minimizing the number of powered on servers by applying server consolidation.

The main goal of this research is to develop a reinforcement-learning based multi-agent framework for building autonomic systems. This framework should be versatile enough to be used to build any autonomic system, from a smart laboratory to an autonomic datacenter.

## Contributions

This project seeks to provide the necessary tool support for building and managing autonomic context-aware systems. The solution should be applied for building a green datacenter.The necessary support for datacenter virtualization and consolidation should be also provided. Another important contribution made by this work is the development of new negotiation methods between QoS and Power Consumption. There are both theoretical and practical contributions made by this project.

**Theoretical contributions**

* Develop a reinforcement-learning policy-based algorithm for autonomic self-healing environments.
* Adapt the previously mentioned algorithm for data center task management.
* Develop methods for negotiating between QoS requirements and Server Optimum Load Values (implying power consumption).

**Practical contributions**

* + - Create a multi-agent extensible framework for building and managing autonomic context-aware environments.
    - Apply the created framework for building and managing a self-healing room.
    - Use the framework to create an energy-efficient datacenter hosted by a smart self-healing environment.
    - Develop tools for monitoring datacenter resources usage and environment sensors.
    - Develop 3D context representations for better system management.
    - Implement a QoS-Energy negotiation mechanism using Fuzzy Logic.
    - Find and evaluate appropriate technologies for implementing a self-adapting datacenter based on the framework mentioned above.

## Publications

The research conducted for this project has generated the following publications. The report chapters which include material from these publications are listed for each publication. The publications are the result of the combined work of me and Copil Georgiana, author of [ 29 ] and content used from them is also found in [ 29 ].

* **A Reinforcement Learning based Self-healing Algorithm for Managing Context Adaptation** ] presents a reinforcement learning based algorithm use for finding the optimum sequence of actions which enforce some user-specified policies. A proof of concept implementation is also presented, which uses an X3Drepresentation of a smart environment in which a user click on an object to change its state. The algorithm detects if a user-specified policy is broken, searches for the best (having the highest reward) sequence of actions which “repairs” the context and executes them. Material from this paper can be found inChapter 4 and Chapter 6.
* **An Autonomic Algorithm for Energy Efficiency in Service Centers** ] is an adaptation of the algorithm presented in the previous paper to handle datacenter task management. The old algorithm is still used for monitoring the datacenter environment. Both algorithms are used in a proof of concept implementation which uses a simulated datacenter in which tasks are dynamically added and removed, triggering consolidation actions. Material from this paper can be found inChapter 4 and Chapter 6.



## Problem Statement and Goals

### Problem Statement

The purpose of this report is to present a framework for building autonomic systems based on intelligent agents which uses ontologies for context representation and policies for goal description. This solution must provide a general extensible platform for building context-aware self-adapting systems applicable to any field. The framework should provide a uniform mechanism of data gathering, context representation and policy enforcement. The policy enforcement engine should detect when the context is in a state which violates at least one policy. It should use a reinforcement learning algorithm to search for the best course of action which brings the context back in an accepted state. The expected reward of being in the resulting state must be the highest between reachable acceptable states. The reward mechanism must be flexible and should capture accurately the differences in severity between different unacceptable states based not only on the number of broken policies, but also on the distance to the closest acceptable state and the number of policy sub rules which are violated. For completing the environment management, the proposed framework must provide a flexible and extensible mechanism for actuator management. The framework should be used o build a real-life green context-aware datacenter.

### Problem Goals

1. Create a suitable context representation
2. Build a reinforcement learning based context management solution
3. Apply the context management solution to a self-healing environment
4. Extends and adapt the context management solution for building a green datacenter.

In order to achieve this goals there are some important sub-goals to be met.

1. **Create a suitable context representation**
   1. Study and use the RAP ]context model
   2. Discover the context object types and their relationships
   3. Build a context ontology representation using Protégé ].
2. **Build a context management suite**
   1. Create an information gathering mechanism:

* Create an information gathering module for the System Under Test(SUT) endpoint
* Create a context information consumption mechanism for querying the SUT information gathering modules.
  1. Create an extensible actuator management mechanism:
* Insert the actuators in the context ontology representation
* Design a self-adapting algorithm which can handle in real time actuator list modifications.
  1. Build a logging mechanism for better context monitoring and debugging
* Create a PDF log writer and a log display window
  1. Build a distributed context management framework based on mobile agents
* Study intelligent mobile agents
* Define agents and their behaviors
* Choose an inter-agent communication mechanism
  1. Build a 3D context representation:
* Study the available 3D technologies
* Find a synchronization method between the 3D representation and the real context
* Make the representation interactive: interactions with the 3D context must generate actuator actions.
  1. Provide additional visual system information:
* Provide a real-time plot representing the algorithm running time
* Provide a visual representation of the context policies, sensors and their values

1. **Apply the context management solution to a smart environment**
   1. Create web services for the smart environment sensors
   2. Define policies, actuators and associated actions
   3. Evaluate the context management solution using broken context patterns
   4. Extensively evaluate the system using the interactive 3D representation
2. **Extends and adapt the context management solution for server cluster management**
   1. Create web services for interacting with the server operating system
   2. Adapt the context representation

* Split the policies into two categories : Energy policies and QoS policies
* Augment the context representation with information particular to datacenters
  1. Find techniques for improving the energy efficiency in datacenters
* Find a mechanism for dynamic task migration
* Find a mechanism for improving server consolidation
  1. Study datacenter management technologies
  2. Provide system information visually:
* Adapt the existing 3D representation
* Implement server resource monitors
* Create a visual representation of the tasks(both deployed and undeployed) for better task tracking
  1. Apply the system on a real world server cluster
  2. Develop a workload generating tool for extensive system evaluation.

Chapter 2

## Background

This chapter provides an overview of the background theory in the area of Autonomic Systems, Artificial Intelligence, Self-Adapting Systems and Green Computing specific to the requirements of this project.

The following topics are discussed:

* Advanced Object-Oriented Concepts and Principles
* An introduction to the Autonomic Computing Paradigm and Autonomic Systems
* An introduction to Intelligent Agents and Pervasive Computing
* An overview of Artificial Intelligence Learning, Knowledge Representation and Reasoning
* An introduction to Green Computing, Virtualization and Server Consolidation

### Advanced Object- Oriented Concepts and Principles

There are high level concepts and principles used in this report that need to be clarified in order to understand how and why they were applied.

**General Responsibility Assignment Software Principles (GRAPS)** are well defined guidelines for assigning responsibility to classes and objects [ 33 ]. There are five main principles which must be enforced in any system which wants to benefit from a good design:

* **Information Expert** principle states that any responsibility should be assigned to the entity which has the necessary information to realize that responsibility. This principle helps to distribute responsibilities in a logical manner and makes further system modifications easier because each responsibility is realized by the entity which has all the information, not some other unrelated entity.
* The **Creator** principle is of great importance because the task of object creation is present in absolutely every object-oriented system. This principle states that an entity should be responsible of creating instances of another entity if it contains, uses closely or has the necessary information to create the instance.
* **Low Coupling** is a responsibility assignment principle needed by any system. By respecting this principle the impact of system change is greatly reduced, thus decreasing the cost of performing any change.

**Dependency Inversion Principle** is a principle which states that components should not depend on low-level components, but on abstractions [ 35 ] . In a naive object oriented approach, the dependency would be on the classes located at the bottom of the class hierarchy tree (leaf subclasses). This type of dependency reduces the generality of the system which uses it and increases the coupling between classes. By changing the dependency from the low-level classes to the abstract super-classes or interfaces the dependency relation is *inversed*, the dependency now being to the top of the class hierarchy tree. This inversion allows for a greater amount of flexibility, each abstract dependency being able to use any concrete class that subclasses the abstract entity used in the relation.

**Separation of Concerns** is the process of separating a program in distinct features which overlap as little as possible [ 34 ].

**Open/Closed Principle** states that a system should be open for change but closed to modification. A more concrete definition is that the system should allow change without needing source code modifications to achieve that change [ 36 ]. In order to achieve this feature, other principles like *dependency of inversion*, *low coupling* and *high cohesion* must be respected.

### Autonomic Computing Paradigm and Autonomic Systems

**Autonomic Computing Paradigm** is a computing paradigm in which systems can manage themselves. This paradigm has been inspired by the human nervous system and has as goal creating applications that can manage themselves according to high-level goals given by humans ].

**Autonomic Systems** are self-governing autonomous systems. Such a system, given some goals to reach or policies to enforce will manage itself and take the appropriate actions to enforce the policies or reach the specified goals without any human intervention.

IBM frequently cites four properties of autonomic systems: Self-**C**onfiguration, Self-**O**ptimization, Self-**H**ealing and Self-**P**rotection ]. Such systems are referred to as Self-Star systems or **CHOP** systems from the four autonomic properties. IBM envisions that these properties will at first be treated separately but in time will merge into one general Self-Management concept.

**Self-Configuration** is desired as a property of every autonomic system because installing, configuring and integrating different vendor components into today’s systems is time consuming and error prone. Following high-level policies the system should adjust and configure automatically and transparently from the user’s point of view.

**Self-Optimization** need arises from the increasing complexity of today’s systems, involving hundreds of parameters which need to be manually tuned. A self-optimizing system will continuously seek methods of improving its performance and efficiency.

**Self-Healing** is very important because the error detection process performed by humans is inefficient and takes a long time. The ability of a system to detect errors and recover from them automatically would greatly improve the fault recovery time.

**Self-Protection** enables a system to detect attacks and recover from them. The automatic attack detection gives a faster response time, thus minimizing the damage done by the attack and the recovery time.

### Intelligent Agents and Pervasive Computing

Pervasive computing, also named as **everywhere computing** or **ubiquitous computing** T.Cioara, I.Anghel, I.Salomie, C.Georgiana and **M.Daniel:** **An Autonomic Algorithm for Energy Efficiency in Service Centers**, 2010 IEEE International Conference on Computer Communication and Processing (ICCP10), submitted

[ 11] is a computing paradigm in which information processing has been integrated in everyday life by means of small networked processing devices. These devices link communications and computing infrastructure to everyday life settings and commonplace tasks.

An agent are defined as an entity which perceives its environment through sensors and performs actions on that environment through actuators ].

#### Agent Environment

By their nature, the environments are roughly classified by ] in:

**Fully observable and partially observable environments:** As the name suggests, in fully observable environments the agent has all the needed information, while in partially observable not everything about the context is known.

**Deterministic and Stochastic environments**: In a deterministic environment the next state is determined entirely based on the current state and the action executed by the agent.

**Episodic and Sequential environments**: in an episodic environment there are independent episodes, while in sequential ones the next world state is always influenced by the previous one.

**Static and Dynamic environments**: In a static environment the world does not change while the agent searches for what to do.

**Discrete and Continuous environments**: Discrete environments have a finite number of states.

**Single-Agent and Multiple-Agent environments**: In multiple-agent environments agents compete for a set of resources.

#### Agent Types

Based on the complexity of the agent’s reasoning process the agents can be classified in 5 categories ]:

**Simple reflex agents:** Act only based on the current input data, ignoring the context history. The agent’s actions are determined by simple rules. Such an agent can succeed only in fully observable environments.

**Model-based agents:** This type of agents maintains an internal model of the world in order to be able to reason over partially observable environments. The model holds previously observed information that now is not observable anymore and reacts just as the simple reflex agent to input data.

**Goal-based agents:** Because in complex situations just reacting will not lead to the desired outcome this type of agent has appeared. These agents know some **goal states** and will try to reach those states.

**Utility–based agents:** Knowing about just the goal does not provide a measure of degree of how expensive is a set of actions the concept of utility is introduced. A **utility function** is used to map each an environment state to a desirability value, thus the agent being capable of comparing two courses of action and choose the one with the highest utility.

**Learning agents:** Are the most complex and can operate in unknown environments. These agents learn about the surrounding world as they go and continuously adapt themselves.

### Artificial Intelligence Learning, Knowledge Representation and Reasoning

#### Knowledge Representation

Any intelligent agent, even reflex agents use knowledge about the environment in their reasoning process. The process of knowledge engineering is a daunting one due to the complexity of the surrounding world.

**Ontologies** are the state of the art mechanism for knowledge representation. Ontology can be defined as a set of concepts from a particular domain together with the relationships between those concepts. These concepts are grouped in categories by the use of inheritance, thus providing a flexible and extensible means of representing the surrounding world.

#### Learning

The main idea behind learning is to use the gathered world data for improving the decision process, allowing an intelligent agent to improve its behavior trough study of its own experience. The goal of machine learning is for the intelligent system to be able to recognize complex patterns and make intelligent decisions based on them. For achieving this goal, machine learning has to borrow methods and concepts from other fields such as statistics, probability theory, data mining, pattern recognition and others.

Based on the type of feedback available, machine learning is classified in three categories: supervised, unsupervised and reinforcement learning.

**Supervised learning** is based on having the correct output for each input state. Mostly used in completely observable environments, in this type of learning the agent knows all possible outcomes from all possible actions and such it always knows what the best actions are.

**Unsupervised learning** is the opposite of supervised learning as it involves no knowledge of the system’s output for a given input. An unsupervised agent cannot learn what to do, because it has no information about what is desirable and what is not.

**Reinforcement learning** is a combination of the previous presented methods. The agent is not fully supervised nor is it left clueless. Instead, it receives a valuation function called reinforcement which indicates if a behavior is acceptable or not. The function value is the reinforcement, a negative value representing a penalty and a positive value an encouragement. Most important in this type of learning is the representation of the learned data and the accuracy with which the reinforcement function classifies the situation the agent encounters.

#### Reasoning

Reasoning in computer science can be regarded as the process of finding in the input data support for some concept. The most developed research area in this direction is in automated theorem proving algorithms, which are used in finding support for theorems based on the input data. Other research area as reasoning under uncertainty is of major importance in designing intelligent agent systems because the real world is not discrete and conclusions need to be drawn without knowing every outcome of every action.

### An introduction to Green Computing, Virtualization and Server Consolidation

Green Computing or Green IT is a computing paradigm that refers to environmental sustainable IT. All hardware and software components must use as little energy as possible and must have as small as possible impact on the environment. Another reason why this trend is important is because implementing its concepts reduces the total cost of ownership by reducing the energy cost.

#### Virtualization and Server Consolidation

A major problem in today’s datacenters is under usage of resources ] ]. Due to the lack of dynamic datacenter management based on the current or incoming load each datacenter must have all servers online in the event that the traffic increases. The need to be prepared for the largest traffic possible leads to the need of having a large number of servers. These two facts combined generate a problem called **server sprawl**, a situation in which multiple, under-utilized servers take up more space and consume more resources than can be justified by their workload.

**Server Consolidation** () has as goal solving this problem by finding means of reducing the number of idle servers. There are several methods for server consolidation, some based on more powerful servers, some on virtualization.

**Virtualization** is a technique for running several operating systems on a single system. By employing virtualization many servers can be transferred to dedicated virtual machines and run on a more powerful system (Fig 1), thus increasing server consolidation. In virtualized environments, the virtual machines run on a “virtual machine OS (operating system)” called a **hypervisor** which manages the resource allocation and distribution among the virtual machines. A major advantage of this approach is that a virtual machine’s resources can be modified depending on the load. Also, with modern hypervisors, virtual machines can be migrated from one server to another without any visible downtime to increase server utilization or in the event of a hardware failure.

Figure 1.1: Server Consolidation trough Virtualization

#### Datacenter Load Distribution

**Load skewing** is a datacenter load distribution technique which involves several tasks. The entire datacenter load is fitted onto as few servers as possible. One or more servers are kept as **tails** in order to accommodate future workload. The rest of the servers are sent to hibernation.

**Throttling** is the mechanism of reducing the performance of underused hardware components and thus reducing the energy consumption. This mechanism is very useful in datacenters where the workload is evenly distributed among servers.

Chapter 3

# Related Work

The related work this paper is based on can be classified into two categories:

* Existing negotiation and bargaining solutions
* Existing self-adapting systems

## An overview of existing negotiation and bargaining solutions

An improvement to server consolidation is the use of negotiation techniques to find a tradeoff between power consumption and QoS. Although QoS requirements are of most importance, in some case a tradeoff is needed due to hardware failure or even unjustified power consumption. For example if a decrease in 5% of CPU requirements for the entire datacenter would allow, after virtual machine rearrangement, for one server to be turned off or send to low power state, then a negotiation technique can be used to find the best value to decrease the CPU for each virtual machine.

The research area of distributed negotiation or multi-agent negotiation has a lot of work associated to it. From this work is clear that negotiation is an important issue in almost any distributed or intelligent system, being present from web services negotiation to grid resource allocation and with the help of this paper, in datacenter consolidation efforts.

Web service discovery and invocation benefit from negotiation as described in ], where a tradeoff between QoS and Cost of Service is achieved trough exchange of less desirable tokens with more desirable ones between parties. As presented in ], very important in web service discovery is the process of finding the most appropriate Web Service providers for a specific Web Service requestor. For being able to conduct real world business by automated web service composition, a negotiation mechanism must exist and it must ensure an optimal “deal” for both sides. This approach introduces *logrolling*, a negotiation technique used mostly in politics, where one person can trade his vote in the exchange of a vote received for his law proposition. The negotiation framework is based on tokens, split in two categories: Quality of Service (QoS) tokens and Cost of Service (CoS) tokens. Logrolling is combined with a token-based negotiation framework augmented with token weights representing each token’s importance. Using this solution, the provider and supplier exchange tokens until both sides reach a consensus, a state in which the utility functions of both parties are the same.

Another approach to multiple-party negotiation is presented in ] under the form of a logic programming framework. This approach, instead of negotiating the existing proposal focuses on creating a counter proposal, ignoring the process of deciding on the utility of a proposal. The system is centered on a knowledge-base which contains all the necessary information about creating a counter-proposal. The proposal is generated using Abductive Logic Programming ], capable of dealing with the unknown goals of the other negotiating party. In case the counter-proposal is rejected, another proposal is generated from the previous one by relaxing the knowledge-base search criteria. This automated negotiation approach solution can be built on top of existing answer set solvers, thus providing a clear path to concrete results.

The research work in the field of multi-issue negotiation can be split into three categories after their understanding of the best negotiation result: minimum loss, maximum gain, and the more general utility function maximization. An approach that fits in the first category is ] which describes negotiation as searching for envy-free states in multi-agent environments. The negotiation process is viewed as a resource allocation one for better understanding of the presented concepts. Given that each agent involved in the resource allocation process has a valuation function to indicate its resource preference, this work focuses on proving that in an envy-free state the resources are efficiently distributed. It proves that envy-free states can be reached if they exist and that resource allocation efficiency and envy-freenes are compatible.

The work presented in ] fits in the second category, searching for joint gains as negotiation result in multiple independent issue negotiation. This approach focuses on “creating value” instead of minimizing loss. For coping with the tendency of agents to hide their intentions in order to obtain as much profit as possible from a trade, an impartial mediator is introduced. Agents disclose secret information to the mediator, which based on that information, tries to achieve a fair deal. Another key element of this solution is that it computes a Pareto-optimal set of outcomes (outcome in which no improvement can be made for a party without worsening of the outcome for the other party). The Pareto-optimal set is further inspected for finding fair situations in which the inter-agent resource distribution is appropriate. Other than providing a generic multi-issue multi-agent negotiation framework, ] also presents utility maximization methods and compares them with respect to their manipulation susceptibility for use in environments where agents tend to misrepresent their utility.

An improvement over existing negotiation techniques is brought by ] which describes a involving multiple interdependent issues, as encountered in many real-world scenarios. A Distributed Mediator Protocol is defined for finding Pareto-optimal agreement points. In order to have an efficient solution, a genetic algorithm is used to find multiple Pareto-optimal agreement points in the nonlinear negotiation space. The genetic algorithm is compared with two other search methods: simulated annealing and hill climbing. Another algorithm called Direct Search is defined which maximizes Nash products ] without finding Pareto-optimal agreements. The Direct Search algorithm is compared against the Distributed Mediator Protocol and an approximated fairness concept is introduced for maximizing Nash products.

## An overview of existing self-adapting systems

The work in the area of self-adapting systems is usually centered on the use of ontologies for representing context information and policies for representing goals. This leads to the need of having a reliable mechanism for gathering data from the surrounding environment and representing it in a manner that supports reasoning. Such a mechanism is presented in ] under the form of an event-driven publish/subscribe architecture for data gathering, processing and event creation. Events are created based on the input data and are used by subscribers to monitor different areas or activities within the surrounding context. The events are fed to the event processing system through a series of event streams. Adding semantical information to an event, each event is assign to a context. This approach also defines a mechanism for fusing similar events from the same context and combining the information given by the last event with the old information held in the knowledge base. Aside from event processing, an alert mechanism is presented, used to inform about specific situations detected based on generated events. The mechanism is applied to real world monitoring situations, like video monitoring of a room. The major contribution brought by this work is the event contextualization mechanism used to map events to contexts, thus providing a semantic event space which can be further extended and applied to various domains.

Another approach to context information gathering is presented in ]. This work focuses on the self-configuring aspect of context-aware systems in the area of pervasive computing T.Cioara, I.Anghel, I.Salomie, C.Georgiana and **M.Daniel:** **An Autonomic Algorithm for Energy Efficiency in Service Centers**, 2010 IEEE International Conference on Computer Communication and Processing (ICCP10), submitted

[ 11 ]. It provides a dynamically reconfigurable fault tolerant context management system. The system is based on a context model containing both information about the required context information and context information metadata. The use of sensor description standards is advocated, as they support opportunistic discovery and integration of sensors and thus adding flexibility to the overall system. By relying on the IEEE 1451 smart sensor interface and SensorML ] sensor description framework , this approach allows smart sensors to advertise and describe themselves to higher level management systems. The proposed approach is tested for a rescue crew scenario having a central management unit and several sensors mounted on each person. Testing results demonstrated the adaptability of the system from using sensor description standards.

One important research direction in context-aware system is in Multi-Agent Systems. Such a system is presented in ] for management of power and performance in datacenters. The purpose of this paper is to demonstrate practically that agents can be used for implementing a coherent automated datacenter management system. The system is centered on three agents: Performance Agent, Power Agent and Coordination Agent. The Performance Agent is responsible for load distribution among servers. The Power Agent is responsible for power consumption monitoring and setting power caps. Also, this agent uses a drastic power consumption reduction mechanism by turning off and on servers depending on the datacenter workload. Finally, the Coordination Agent handles the communication between the other two agents and uses predefined utility functions and policies in sending control signals to the system. The presented system was tested on a datacenter hosting IBM blade servers running Linux. Although the experimental results are promising, the technology is not there to help in implementing such a radical power management. A server running Linux takes 5-10 minutes to start after being turned off, an amount of time which makes it difficult to implement such an approach in real-world datacenters.

Under the pressure to make computing eco friendly an important research branch in self-adapting systems is creating energy-aware systems that are capable of reducing power consumption and maximize resource usage. One such approach is ] , which presents a dynamic load management system for virtualized datacenters. Employing virtualization, this solution allows a single server to be shared by multiple services, thus improving server consolidation and resource utilization. Also, virtualization allows for on demand task resource allocation based on the datacenter workload. For maximizing the energy consumption, this solution, as the one presented before exploits the idea of turning the servers on and off depending on the datacenter load. Although turning off servers has the visible advantage of lower power consumption, the process of waking up a server implies a large amount of time, as discovered and described in ]. In order to minimize the impact of the server wake-up time, a limited look ahead control mechanism is introduced. Also, the server switching costs are considered. From the experimental results, using server virtualization together with a lookahead needed for anticipating datacenter state results in an average of 26% power reduction while maintaining QoS requirements.

Very important in self-adapting systems is the system’s ability to learn the proper actions that solve some problematic situation. One school of thought is focused on applying reinforcement learning algorithms to smart environments. Reinforcement Learning advocates view this type of learning as a more “human” way of searching for solutions. Also, it seems to be more natural to think of goals as states with high rewards. Such an approach is presented in ] which demonstrates that reinforcement learning can be used successfully in autonomic systems. The reward is computed based on the multi-attribute system state and represents how desirable is for the system to be in that particular state. The selected learning algorithm is State-Action-Reward-State-Action (SARSA) and it is compared with a Goal-Action Attribute Model technique ] and. The solution is tested using a series of simulated models and the result is that reinforcement learning is feasible for the management of autonomic systems, but future research is needed because this type of approach does not provide a high performance solution

Chapter 4

# Analysis and Design

The system’s core algorithm and associated concepts where developed by me and Copil Georgiana, author of ], for the **CONSENS** research project

[ 32 ] . The results of the research in the context of this project are described in ].Due to these facts, the **reinforcement based self-healing algorithm** and associated concepts descriptions will appear both in this report and in ]. The entire andsubchapters are joint development effort, elements contained in it being also described in ].

In this chapter both the core components of the autonomic system framework and the design decisions needed for applying the framework to build a self-healing laboratory and a self-adapting datacenter are presented.

## The context model

The context model used by the self-adapting solution presented in this report is based on the RAP (Resources, Actors, Policies) model presented in ]. This model was chosen because is general and powerful enough to be used in any self-adapting system. Any context can be represented as a collection of Resources, Actors that interact with those resources and a set of Policies which control the state of the resources and actors and of their interaction.

### Context Ontology Representation

The core ontology representation only extends the Resources element of the RAP model with Sensor and Actuator concepts. represents the core ontology structure. The distinction between resources is needed as the system needs to know what element returns context information and what element can influence the context. This is important because the system iterates trough actuators attached to a sensor which does not respect some policy and tries the available actions in order to see if some improvement over the current situation can be made.



Figure 4.1: RAP Ontology Representation

Because the datacenter context needs more specific information, another ontology representation is defined based on the RAP model and used together with the previous basic representation. The datacenter ontology representation is illustrated in . The previous representation is used for environmental representation and the next ontology is used to represent the datacenter physical and software components. Also, due to the fixed number of actions possible in a datacenter, the generality of the representation is reduced by creating a concrete class for each possible action. The reduction in generality is motivated by an increase in efficiency by eliminating the search for all possible actions at each step in the action selection algorithm.



Figure 4.2: Datacenter Ontology Representation

### Policies

The policies enforce some high-level conditions and consist of a set of acceptable sensors values. In order to have a fine-grained view over any context situation and be able to differentiate between the severity of two different context situations, weights are used. Each policy has a weight associated to it which represents the importance of the policy in the overall context. Furthermore, each resource associated to a policy has a weight representing the importance of the sensor in that particular policy. The formal policy definition is defined in .

Equation 4.1: Policy Specification

|  |
| --- |
| [, |
| , where:  - : range of acceptable values for Resource i |

### Context Entropy

For understanding the concept of context state entropy the state must be defined. A context state is a snapshot of the execution environment taken at a specific moment in time.

For evaluating the degree of context state deviation from the acceptable value the concept of context state entropy is used ]. The context state entropy is used to represent the desirability of the system to be in a particular state. Based on the definition of the policy, the entropy is defined as the product between the policy weight and the sum of products between the associated resources weight and their distance from the acceptable value. The Entropy formula is defined in .

Equation 4.2: Entropy Formula

|  |
| --- |
|  |
| , where:  - : entropy  - **:** the importance of i-th policy  - **:** the importance of resource j in policy i  - **:** is a value marking the degree to which resource j respects policy I (deviation). |

### State reward

For the self-healing environment management algorithm the reward is represented by the inverse of the entropy. The states with the highest reward are the states with the smallest entropy, as resulting from .

Equation 4.3: Reward function

|  |
| --- |
|  |
| , where:  - R**:** the state reward  - **:** the state entropy |

For the self-adapting datacenter management algorithm the reward is computed using a more fine-grained formula described in ].

### Inter-Independent Resources Group

The time required to find the best sequence of actions that brings the system in an accepted state depends on the number of policies and on the number of resources attached to a policy. For improving the performance of the system the concept of Inter-Independent Resources Group (IIRG) is introduced. In order to understand the IIRG concept the dependency relation must be defined first. A resource is dependent on another resource when a change in the value of one of the resources triggers a change in the value of the other resource. Furthermore, the dependency relation is of two types: direct dependency and indirect dependency. Direct dependency between two resources and occurs when a change in the value of triggers a change in ’s value. Indirect dependency between two resources and occurs when a change in ’s value triggers a chain of changes in the values of several resources () which eventually affects . Both dependency types are formalized in .

Equation 4.4: Dependency Relation

|  |
| --- |
|  |
| , where:  - : Resource X  - : implies, triggers  - **:** dependency between Resource X and Resource Y |

The dependency relation is represented under the form of a dependency matrix with has as entries the length of the dependency between the resources (1 for direct dependency and path length +1 for indirect dependency). A “0” in the dependency matrix means no dependency. The dependency matrix representation was chosen because it allows for fast direct dependency verification and also it provides support for indirect dependency checking by allowing the use of graph cycle detection algorithms for dependency chain detection. A dependency matrix example is shown in .

Table 4.1: Resource Dependency Matrix example

|  |  |  |  |
| --- | --- | --- | --- |
|  | R1 | R2 | R3 |
| R1 | 0 | 0 | 0 |
| R2 | X | 0 | 0 |
| R3 | Y | 0 | 0 |

From the dependency relation defined above it results that two resources are inter-independent if there is no direct dependency between them or if the length of the indirect dependency chain is over a certain limit. The certain limit statement is included because the influence of the resources can drop as the dependency chain increases in length, reaching a state where the dependency is negligible and thus can be ignored. By separating the resources in IIRGs solutions for context repair can be searched in parallel over the IIRGs, thus reducing the system’s search space and improving the running time. The pseudo code of the IIRG extraction algorithm is presented in .

|  |
| --- |
| Listing 4.1: IIRG Extraction Algorithm |
|  |

For a better understanding of the IIRG extraction process, Figure 4.3 presents a trace on an example where it is assumed that the resources weights follow this rule : weight(R2) >weight (R3) >weight (R1).



Figure 4.3: IIRG Extraction Algorithm Trace

### RAP instantiation for the self-healing system



For using the RAP model in implementing a self-healing environment, the domain model elements must be mapped onto real world entities. The mapping is represented in .

Each sensor can have associated one or more actuators which have an influence on it. Each actuator has at least one action defined, each action having its effect represented by a **+ value (increment action), - value (decrement action)** or **value (set value action).**

Table 4.2: Self-healing environment RAP mapping

|  |  |  |
| --- | --- | --- |
| <R,A,P> entity | Self-Adapting context entity | Details |
| Resources | **Sensors** | * Smart room sensors |
| **Actuators** | * Components capable of influencing the environment |
| Actors | **Actions** | * Each actuator has associated actions |
| Policies | **User specified policies** | * Acceptable sensor values combinations for the environment |

### RAP instantiation for the self-adapting datacenter

While the previous RAP instantiation is still useful in the self-adapting datacenter for environmental control, a new mapping is needed for datacenter workload management. Due to the reduced number of possible datacenter actions (excepting environmental control) the new mapping is more concrete and includes five concrete action instances: Deploy task, Move task, Send server to hibernate, Wake up server. The reduction in representation generality is compensated by the increase in representation efficiency, reducing the action search to a set of five predefined instances. The RAP mapping for the datacenter scenario is presented in .

Table 4.3: Self-Adapting Datacenter RAP Mapping

|  |  |  |
| --- | --- | --- |
| <R,A,P> entity | Self-Adapting context entity | Details |
| Resources | **Servers** | * All datacenter servers |
| Actors | **Actions** | * Deploy task * Move task * Send server to hibernate * Wake up server |
| Policies | **Energy policies** | * Servers green performance indicators ranges |
| **QoS policies** | * Task requested resources ranges |

### Load Distribution Strategy

For optimizing the energy consumption for a datacenter is not sufficient to have an efficient workload distribution strategy. The main reduction in datacenter energy consumption is obtained by sending to a low power state (hibernation or sleep mode) as many servers as possible. This leads to the problem of finding a mechanism for reducing the system response delay introduced by the server wake-up time. In a real-world datacenter scenario, when a task is required to be run the task can’t wait for a server to power up. In order to solve this problem server skewing with tails is employed. This approach keeps one or more servers are kept online for accommodating future loads based on workload predictions. Also based on this prediction a **load threshold** is computed for each tail and when that threshold is reached a new tail server is brought online. The idea behind the threshold is to be computed dynamically based on workload prediction. For maximum efficiency, the time it takes for a server to wake up should be the time it takes for the tail server to become fully used after the load threshold is reached. An example of a situation in which the load threshold is reached, thus triggering a wake up of another server as tail is presented in .



Figure 4.4: Server Skewing with Tail

## Reinforcement Learning Algorithm

For finding the best sequence of actions to bring the system in an acceptable state an algorithm based on reinforcement learning was developed. The advantage of basing a reinforcement learning based approach is the generality provided by this type of learning. Any real world situation can be represented in terms of <action, reward> pairs. Also, by manipulating the expected reward the mobile agent that uses this learning type can be guided to take any desired course of action without actually specifying the actions he needs to take. The adaptability of such an intelligent agent is a major advantage in building a generic platform for building autonomic systems.

The desired context states are described using policies. Based on these policies the algorithm computes the entropy value for the current context state using the formula defined in 4.1.4 State reward. If the entropy value is above 0 (or above a user-defined threshold) the search for the repairing sequence of actions is triggered. For each broken policy, the algorithm takes each broken resource and tries to change its value in an acceptable one by executing actions on associated actuators. Each associated action executed is simulated, the resulting state’s entropy is computed and the state is placed in the states priority queue, sorted by entropy. The next state to be expanded is the top of the queue, which holds the best state found so far. The pseudo code of the generic reinforcement learning algorithm is presented in .

|  |
| --- |
| Listing 4.2: Reinforcement Learning Algorithm |
|  |

Due to the generality of the reinforcement learning approach, the previously described algorithm can be applied easily both to a self-healing smart environment and to a self-adapting datacenter. Also, the algorithm expands the search space in a breadth-first based manner, but always expanding the node which has the highest reward associated. This search mechanism allows for finding the best solution in a less amount of time. In the reinforcement learning process is represented.



Figure 4.5: Self-Healing system Reinforcement Learning Process

Due to the real-time performance required from real world datacenters, for applying the algorithm to a datacenter, the algorithm action search must be stopped at any given moment in time and the best action sequence found so far to be used. This constraint comes from the real-time performance demands of a datacenter. In the case that a very complex scenario with an increased number of servers and a large number of tasks that need to be run takes place, the algorithm might take too long to find the best solution. In this case the best solution found so far needs to be used in order to respect the QoS requested from the datacenter. The new flow of the algorithm is presented in .



Figure 4.6: Datacenter Reinforcement Learning Process

## QoS-Energy Consumption Negotiation

In the real world sometimes situations in which some policies conflict can appear. For example one policy should request a value or range of values for some resource and another policy could request another range of values for the same resource at the same time. In this case the conflict needs to be solved somehow. A good conflict solving mechanism is negotiation, in which each party involved in the dispute gives up something until everyone is satisfied and a consensus is reached.

The negotiation problem was studied intense for the self-adapting datacenter implementation case because here the conflicting policies problem is the most pertinent. As described in , each server has associated green performance indicators. These indicators specify the optimum server load for maximizing the performance-per-watt ratio. The performance indicators are specified as ranges. For example, a server could have an optimum CPU load of 60%-80%. The task resource requirements are also specified in ranges, for example for CPU 500 MHz-800 MHz. A situation in which the task requires more from the server that he is willing to give can appear easily. In such a situation choosing to enforce one policy or another based on some policy importance is not the best choice because an energy efficient datacenter which respects QoS requirements is needed and this implies that both energy and QoS policies need to be enforced.

In a situation where negotiation is required is presented. In such a situation the maximum resources required by a task can’t be allocated, but there is room to fit the task in the required resources range. A basic approach would be to allocate the minimum resources requested by the task. But this might lead to situations in which the maximum optimum server load is not reached, but some task performs in a degraded mode with the minimum resources needed to run. For avoiding this kind of situation, a tradeoff mechanism is needed. The tradeoff solution presented here follows the reinforcement approach taken to system management in this report.



Figure 4.7: Situation which requires negotiation

**Desirability** is defined here as the satisfaction of the system to be in a particular state. This concept can be applied both to server optimum load indicators and to task requested resources range. Desirability information can be added in the context by using a bidimensional function, one dimension for desirability and the other for the actual value. Such a function is presented in . The desirability is represented in the Y axis, while the corresponding values are on the X axis.



Figure 4.8: Desirability-Based Negotiation

The negotiation process is based on Fuzzy Logic. The negotiation is done by taking the requested task range and the available server range and checking if the two ranges overlap. If so, the overlapping range is extracted. A desirability function is assigned to the overlapping range based on the QoS and Energy Consumption importance. The result is another bidimensional function. In the desirability of the task is assigned to the overlapping range in order to achieve a greater QoS. Finally, the negotiation result is the **Center Of Gravity** of the area given by the previously generated bidimensional function using .

Equation 4.5: Center Of Gravity Formula

|  |
| --- |
|  |

## The Conceptual Architecture

The system’s architecture is a multi-agent one. The reason for choosing such architecture is its mobility, extensibility and separation of concerns it provides. These reasons make multi-agent systems very good candidates for implemented distributed systems. After establishing the agent interaction mechanism, each agent can be developed separately and also it can reside anywhere in the distributed system. Also, any agent can change location inside the distributed system without affecting the system’s functionality.

Having an agent-based architecture, the system consists of six agents: Context Model Administering Agent (CMAA), Context Interpreting Agent (CIA), GUI Agent, X3D Agent, Task Management Agent (TMA) and Reinforcement Learning Agent (RLA), which is the most important agent and implements the core functionality needed for this autonomic system building framework. The CMAA reads an xml context description file and loads the information into the shared ontology. The context description file contains xml definitions of ontology instances. CIA periodically queries the context sensor data and refreshes the ontology context information. The GUI Agent is responsible for the graphical representation of the system status: system log, sensor values, etc. RLA is the most important agent. It detects when the context is broken and searches for actions to bring it back in an acceptable state using a reinforcement learning based algorithm applied on the common context ontology representation. The X3D Agent is responsible for the 3D interactive context representation. The 3D representation is sued both for environment simulation and for environment control, allowing the user to interact with the virtual environment and evaluate the autonomic mechanism which manages it. The TMA is used only in the self-adapting datacenter extended version of the proposed framework and is responsible for dynamic task creation. This agent is very important as it allows the user to create a task at runtime and such it provides a good testing mechanism and also brings more dynamicity to the system. The agent-based conceptual architecture is illustrated in Figure 4.9.



Figure 4.9: System Conceptual Architecture

One important issue in any multi-agent systems is responsibility assignment. The responsibility assignment process was guided by the General Responsibility Assignment Software Principles (GRASP) [ 33 ]. Conform to the *Creator* principle, the CMA Agent creates all the other agents because it holds the necessary information needed to instantiate the agents. The CMA agent loads the ontology description, creates ontology representation models and during the agent instantiation process it sends those models to the agents. The GUI Agent respects the *High Cohesion* principle as it is concerned only with displaying information received from the other agents. If the *Information Expert* pattern would have been used here it would have violated the *High Cohesion* principle, because each agent would have had also the responsibility of displaying the context information which it contains. The same principle, *High Cohesion*, was applied when assigning the responsibilities to the other agents. The other GRASP principle left unmentioned, *Low Coupling,* is respected implicitly by using a multi-agent architecture. Because the agents communicate only trough a well defined mechanism over the network, without direct method invocation, any agent can be replaced without affecting the other agents as long as it returns the expected data.

Another important issue in multi-agent systems is the agent interaction mechanism. After assigning responsibilities between agents it must be decided which agent communicates with which agent and what type of information to the agents exchange. In Figure 4.10 agents interactions are represented. The Context Interpreting Agent captures data from the environment sensors, populates the context ontology representation with that data and informs the X3D Agent about the modification and sends the new data. This notification is needed as the X3D agent does not use the ontology representation. Instead, it receives pairs of <sensor name, value> from the CIA and RLA and uses that information to refresh the 3D representation. This mechanism allows the X3D Agent to be used independently on the underlying ontology representation, bringing a needed plus in system flexibility. By adhering to the *Low Coupling* and *High Cohesion* principles, the X3D Agent minimizes the effect a change in the underlying ontology has on the system.



Figure 4.10: Agents Interactions

The Reinforcement Learning Agent (RLA) in turn sends notifications to the GUI Agent and X3D Agent in order to refresh the displayed data for both agents. The GUI Agent needs messages sent to it because it handles the logging of the RLA actions and it cannot have that information retrieved by any other means. The X3D Agent receives information from the RLA which is needed for synchronization of the 3D representation with the underlying ontology used by the RLA.

In turn, the X3D Agent notifies the RLA when some interaction with the 3D objects takes place. Because the X3D Agent does not interact with the ontology used by the RLA, it cannot change the ontology values directly so it must inform the RLA of what interaction has place. This is conforming to the *Information Expert GRASP* principle because the entity which has the information needed to execute an action executes the action. In this case this expert entity is the RLA and after receiving the notification from the X3D agent it enforces the action on the environment actuators

The Task Management Agent (TMA) interacts with the RLA because it does not have the possibility of creating a physical datacenter task. So it sends a create task notification together with the additional information needed about the new task to the RLA. The RLA creates an ontology instance for that task and executes a create task action on the datacenter.

Chapter 5

# Implementation

The current chapter focuses on the development and implementation of the proposed agent- based context management framework. It gives a closer look at the technologies used to implement the context management components presented in the previous chapter, and then it explains how these technologies were used to obtain the proposed goals.

## System Components



Figure 5.1: System Components Diagram

Only the components developed by me (highlighted in with green) will be presented in detail. The other components are detailed in ] by their creator and will be mentioned and properly referenced when needed in order to maintain a certain mental flow of the system description.

The “Environment Self-Healing Module” (highlighted half with blue and half with green) was developed by me and the Copil Georgiana, the author of ] for the CONSENS research project ] and as such it is a joint development effort, also described in ].

## System Input

For any knowledge-based system to be able to reason about the surrounding world it is necessary to have a mechanism of describing that world. The selected mechanism is **Ontology Representation**.

The **context** is described using the two previous ontology representations ( and ). Both the datacenter and environment ontology representations are built using Protégé ] and populated with instances. The instances represent real world physical elements. A closer look of the ontology representations used by this architecture is presented in Figure 5.2.



Figure 5.2: System Ontology Representation

The **system acceptable states** are defined by policies specified under the form of **SWRL (Semantic Web Rule Language)**  ] rules. By specifying the policies in SWRL the system uses **Pellet** ] OWL2 Reasoner for policy evaluation. The SWRL rules evaluation triggers at each change in the ontology variables, providing a reliable mechanism for context evaluation. An example of a SWRL QoS Policy definition is presented in Listing 5.1. The SWRL rule is simplified in the listing mentioned previously in order not to overload the representation. In the example for each property a value is specified, when in the real SWRL representation there is a value range. An example of a SWRL rule definition is illustrated in Listing 5.1.

|  |
| --- |
| Listing 5.1: SWRL QoS Policy Example |
| J:\cercetare\paper suceava\deliverables\source images\swrl_QoS_policy.png |

The SWRL rule example from Listing 5.1 enforces several conditions for *Task* with *name* “Task\_1”:

* Task HAS-PROPERTY *associatedInfo*
* Property *associatedInfo* has property *Received*
* Property *Received* has properties *cores, cpu, memory and storage*
* Property *cores*  has value 1
* Property *cpu*  has value 3000
* Property *memory*  has value 2048
* Property *storage*  has value 200

If the conditions presented above are met, the *QosPolicy\_1* instance has its *respected* property set to true. Otherwise the *respected* property is set to *null*.

The **world information** is accessed using **ASP.NET Web Services** ]which provide the interface between world datacenter sensors and resources. This approach was chosen because Java ] does not provide such a tight coupling with the underlying operating system. C# **Error! Reference source not found.** ] provides access to the underlying operating system’s functions and thus allows access and real time monitoring of system’s hardware.

## System Output

The most important output provided by the system is the **context repairing sequence of actions**. Having two system management components, one for environment self-healing and one for datacenter self-adapting, there are separate outputs for each component. The repairing sequence of actions is as the name suggests, a sequence of actions which brings the system in an acceptable state from a state in which some policies where broken. The found actions are executed on the corresponding targets.

As secondary output the system stores the <context state, action sequence> pair for future use and creates a Pdf file containing the context monitoring and repair log, for future analysis. An example of a context management log is presented in Listing 5.2. In the presented example three policies where broken (the yellow entries in the log). The resulting repairing action resulted is represented with blue. The initial state in which some policies where broken is represented with red and contains all the environment sensors values, while the states in which no policy is broken are rendered with green.

|  |
| --- |
| Listing 5.2: Context Management Log Example |
| J:\cercetare\licenta\pdf_log.png |

## Implementation Technology

For implementing the agent-based architecture the Java Agent Development Framework (JADE) ] was used. JADE is a software framework fully implemented in Java [ 52 ] which simplifies the implementation of multi-agent systems through a middle-ware that complies with the FIPA (Foundation for Intelligent Physical Agents) specifications ]. Another description of Jade would be : JADE is an enabling technology, a middleware for the development and run-time execution of peer-to-peer applications which are based on the agents paradigm and which can seamless work and interoperate both in wired and wireless environment ].

There are several advantages for using JADE. JADE is based on a peer-to-peer architecture so each peer is able to initiate a communication or be subject to a request, providing a fully distributed system. Also, it implements the agent paradigm, applying concepts from artificial intelligence to distributed systems. Agents are active entities, that can refuse a request and which are loosely coupled, thus providing a flexible infrastructure.

JADE supports agent migration. Each instance of the JADE run-time is called a *container.* The set of all containers is called a *platform* and the platform provides a layer that hides the complexity of the underlying hardware. This enables the agents to migrate from one container to the other in real time.



Figure 5.3: Jade Architecture

Figure 5.3 presents the basic Jade architecture. Each Jade instance is called a *container*, since it contains agents [ 47 ]. The set of all Jade containers is called a platform. As represented in the architecture, Jade can run on almost any computing device, from a J2ME [ 53 ] compatible mobile phone to pocket pc’s and desktops. The ability to host a Jade container in any environment brings a great degree of application mobility, every agent being able to migrate from one container to another seemingly. For example, one Jade agent could migrate from a desktop to a user’s mobile phone when the user leaves the room and return to the desktop when the user returns.

Jade provides an asynchronous inter-agent communication mechanism applied in Figure 5.4 for two agents, Ai and A2. The message passing process consists of three steps. In the first step, the source agent wraps the data it wants to send in the proper message format. The second step is the actual sending of the message. The message is placed in the message queue of the destination agent. The third step takes place at a later time, when the destination agent decides to inspect its message queue, extracts the first message contained in it and processes it.

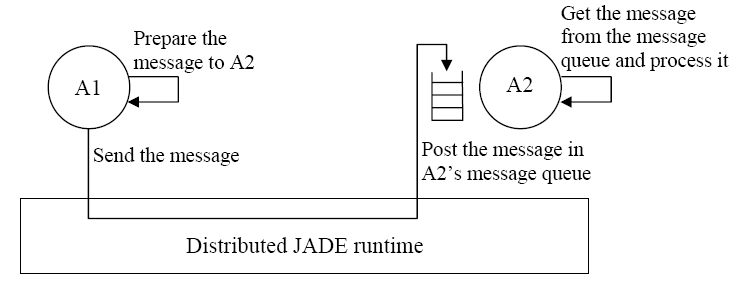


Figure 5.4: Jade Asynchronous Message Passing Mechanism

Jade agents can communicate using various protocols, but for this implementation Java Serialization was used. Another important communication protocol supported by Jade is FIPA SL language [ 50 ] . By using FIPA SL languages two agents communicate using a common ontology as common knowledge. This would have coupled all the agents to the context ontology representation and thus would have reduces the generality of the approach and would have increase the impact a change in the ontology would have had on the system. This is the reason behind using Java Serialization in inter-agent communication.

## Ontology Representation Implementation

### Environment Ontology Implementation

Based on the .owl environment ontology description, a *JenaOWLModel* [ 44 ] is generated and accessed with the specific Jena API. This is further detailed in ] by its author.

### Datacenter Ontology Representation

From the datacenter .owl ontology file description an *OWLModel* [ 40 ] is created using Protégé OWL Loader ]. Using the Protégé ] Ontology Editor tool for generating Protégé-OWL Java code a class structure respecting the ontology hierarchy and containing all the ontology properties is generated. This hierarchy of generated classes has as core the *OWLModel* created from the ontology file. An example of an ontology mapped to java class hierarchy is illustrated in Figure 5.5.



Figure 5.5: Ontology to Java Code Mapping

This is a more useful ontology representation because having Java classes to work with instead of just an underlying ontology model, the ontology concepts can be enhanced with Java specific properties such as *Serializable* or *Cloneable* , providing a better integration of the ontology representation with the rest of the regular Java code. Another major advantage of using Java classes instead of directly accessing the *OWLModel* or *OntModel* [ 45 ] is that all the properties values can be set directly in the ontology entity slot without using the higher *OntModel* functionality. This proved to be an advantage in the experimental results phase when by using a home-brewed policy evaluation mechanism a major improvement in performance was obtained. This improvement was obtained if the properties values were not set also on the *OntModel* and such the Pellet SWRL evaluation mechanism did not trigger at every property change event.

In Figure 5.6 a *CPU* class ontology representation is shown. The *CPU* entity is a subclass of *Component* and has two properties: *associatedCore* and *numberOfCores.* There are cardinality restrictions on the properties: there must be at least one *associatedCore* value associated to each instance of *CPU* and must be exactly one *numberOfCores* value.



Figure 5.6 : CPU Entity Ontology Representation

|  |
| --- |
| Listing 5.3: CPU Entity Java Representation |
|  |

The java code mapping of the *CPU* entity is presented in Listing 5.3. It can be remarked that for each of the properties defined for the *CPU* entity setter and getter methods where generated. Also it can be observer that for the *numberOfCores* property, due to the “exactly one” cardinality restriction, the *getNumberOfCores()* method return *int*, while for the *associatedCore* property which can have more values a *Collection* is returned by the corresponding getter method.

## Agents Implementation

### Context Interpreting Agent (CIA)

The Context Interpreting Agent is responsible for synchronizing the ontology representations with the real context. The agent is implemented by the *CIAgent* class found in contextawaremodel.agents package.

Upon creation, the agent receives from the Context Management Agent described in ] the ontology models for the environment and datacenter. These models are used by the single behavior attached to the agent: *ReceiveMessagesCIABehaviour*. The *ReceiveMessagesCIABehaviour* extends *CyclicBehaviour* and handles the messages received by *CIA*, informing it of newly created instances. For each new instance of *Sensor* and *Server* the behavior adds the instance to the *SensorAPI*.

The **SensorAPI** is based on the API described in ].It implements a pooling data gathering mechanism. At a certain time interval the web service associated to the world element (*Sensor* or *Server)* is accessed and the new data is used to update the ontology representations. This API provides two static methods used for registering a *Server* or a *Sensor* to the pooling mechanism.

|  |
| --- |
| Listing 5.4: SensorAPI functionality |
|  |

The *SensorAPI* functionality shown in Listing 5.4 contains both the mechanism needed for monitoring sensors with discrete values and servers which have a lot of information associated to them. For providing a flexible and low-coupled server information gathering mechanism the Data Transfer Object design pattern [ 37 ] was used. By abiding to this pattern, a *ServerDto* class has been defined with its only purpose to store server information.

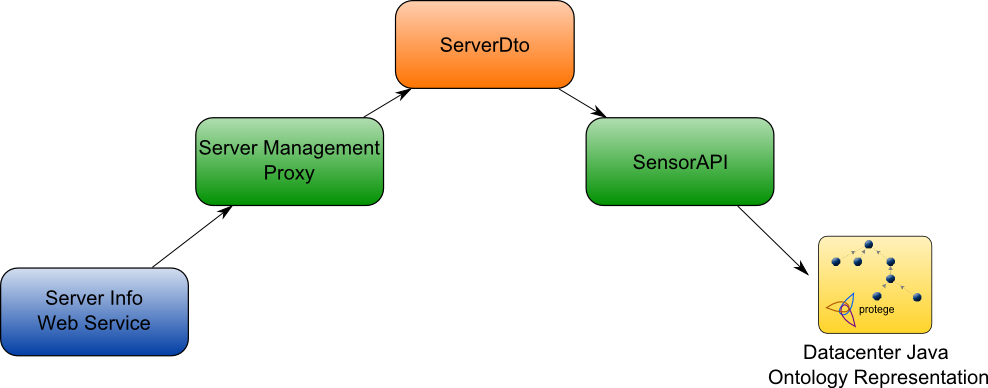


Figure 5.7: DTO Design Pattern SensorAPI Usage

As illustrated in Figure 5.7, the server information is gathered from a web service by a specific *ServerManagementProxy* instance. Depending on the web service type and implicitly on the format of the web service invocation response, the proxy processes the web service response and based on its information creates a *ServerDto* instance. From here on, no matter the source of the server information, only the DTO is transmitted, thus having a uniform mechanism for accessing server information. One the *ServerDto* instance is created, the *SensorAPI* will use it and refresh the ontology data. By employing a object just for data transfer, the particularities of the server information gathering endpoint are hidden from the rest o the application. These particularities are known only by the concrete server management proxy. A change in the server information gathering endpoint would only require an implementation of a different proxy for server management. The rest of the system would be unaffected. This leads to a low coupling between the real world information structure and the internal information representation, providing flexibility to the environment information gathering mechanism.

### Reinforcement Learning Agent (RLA)

This agent represents the core of the infrastructure. It contains the ontology reasoning and solution search processes. The implementation of the agent is available in the *ReinforcementLearningAgent* class located in the *contextawaremodel.agents* package.

Upon initialization, the agent receives from the CMA an *OntModel* together with its underlying Protégé *OWLModel* for both environment and datacenter ontology representations*.* The *OntModel* is used by Pellet ] for evaluating SWRL ] rules, while the *OwlModel* is used in conjunction with a Protégé ] generated *Ontology Factory* for ontology management.

The two main components of this agent, the **Environment Self-Healing Module** and the **Datacenter Self-Adapting Module** are implemented as behaviors extending *TickerBehaviour* . The environment management module is implemented in *contextawaremodel.agents.behaviours* in *ReinforcementLearningEnvironmentManagementBehaviour* and the datacenter management module in *ReinforcementLearningDataCenterManagementBehavior*.

#### ReinforcementLearningEnvironmentManagementBehaviour

This behavior implements the algorithm presented in ] in the context of maintaining the state of the environment within acceptable parameters. The implementation is a joint development effort done in the context of the CONSENS research project ] and is also described in ].

Extending *TickerBehaviour* , the ReinforcementLearningEnvironmentManagementBehaviour implements the *onTick()* method which is called at a specific time interval, interval specified when creating a new instance of the behavior. The *onTick()* method contains the business logic needed for the self-healing system. First the context entropy is computed using . If the computed entropy is different than 0 then the reinforcement action search begins.

For each broken policy, for each broken resource from the policy and for each actuator associated to the policy all the possible actions are simulated and the corresponding states’ rewards are computed. Each action simulation has as effect the creation of a new state which is placed in a list of states, sorted after reward. If no resulting state had the entropy 0, the search continues with the next best state (the state with the highest reward from the states list) and repeats the actions described above. If a sequence of actions which can repair the context exists it will be found and enforced with the help of the *Action Enforcement* *Unit*.

#### ReinforcementLearningDataCenterManagementBehavior

This behavior extends the algorithm presented in ] with enhanced reward computation and elements specific to datacenter management. This behavior is presented in detail in ] by its developer.

#### Logger Unit

This component has the role of logging the context state, the broken policies and the actions taken to repair the context. The log is saved under the form of a PDF file using IText ]. IText is a library used to generate PDF files. The advantage over other libraries is that it supports several mechanisms for PDF creation: java Graphics can be used to “draw” the PDF or specific library classes can be used to create a document structure under the form of paragraphs. An example of a log generated by this component is illustrated in Listing 5.2/

#### Memory Store Unit

This component stores <context state, actions> pairs for future use. When the context is evaluated, if some policy is broken it is checked if the current state hasn’t been encountered before and if it has the action sequence associated to it is executed. Described in detail in ] by its author.

### X3D Agent Implementation

The implementation is available in the X3DAgent class located in the *contextawaremodel.agents* package.

For providing a realistic context representation X3D ] was chosen. X3D is a scalable and open software standard for defining and communicating real-time, interactive 3D content for visual effects and behavioral modeling. X3D provides both the XML-encoding and the Scene Authoring Interface (SAI) to enable both web and non-web applications to incorporate real-time 3D data, presentations and controls into non-3D content. X3D is the successor to the Virtual Reality Modeling Language (VRML). It improves upon VRML with new features, advanced APIs, additional data encoding formats, stricter conformance, and a componentized architecture. The advantage of using X3D is that it provides a simple and easy to use mechanism for defining 3D scene navigation and interaction.

Xj3D ] is a toolkit for VRML97 and X3D content written completely in Java. Although at first Xj3D was based entirely on Java-3D, it has moved from that state and now it provides a faster and less CPU intensive rendering engine. The architecture of the Xj3D is presented in Figure 5.8.



Figure 5.8: Xj3D Architecture

The context representation is described in an .x3d XML file which is loaded by XJ3D and rendered. In the X3D description file every object is wrapped into at least one *Transform* node which specifies the object scale, rotation and translation. Inside the *Transform* node a *Shape* node represents the 3D shape. The *Shape* contains the object’s *Material* and definition. The *TouchSensor* node is used for capturing user input. And finally, everything is wrapped in a *Scene* node.

An example of an X3D node description is presented in Listing 5.5. In the example a *box* is defined and a material called *boxColor*having RGB values 0.5, 0.5, 0.5 (1 representing 255) attached to it. The box is translated on the Y axis by 13 pixels. Also a *TouchSensor* is attached to the box. By clicking on the box a touch event will be generated. The touch event can me caught and processed using an *X3DEventListener* attached to the *BoxTouchSensor*.

|  |
| --- |
| Listing 5.5: X3D File Shape Entry |
|  |

The role of the X3D Agent is to provide a meaningful context representation which can be used for simulating various scenarios or for real-time monitoring of the surrounding environment. More details about the simulated contexts are presented in the Experimental Results chapter.

### GUI Agent Implementation

The implementation is available in the GUIAgent class in the *contextawaremodel.agents* package. The GUI Agent is responsible of the creation and management of the user interface elements used for information output. There are four main user interface components handled by this agent: Environment Monitor, Tasks Monitor, Server Monitors and Logging Window.

#### Environment Monitor

This component is responsible for displaying the environment sensors values, the broken environment policies and the actions taken by the self-healing algorithm. The environment monitor window, represented in , has three main areas: current state area, broken policies area and context repair area.

In the **current state** area the context sensors and their values are displayed. The **broken policies** area displays the name of each broken policy, while the **context repair** area displays the context repair actions. The context repair actions information is displayed in three columns. The first column containing the name of the context repair action. The second column contains the name of the sensor which will be affected by this action, more specific the name of the sensors which has associated the actuator on which this action will be executed. The third and last column contains the value with which the sensor will be influenced. In the effect is the *Decrement* of the value for the *TemperatureSensorI* sensor by 5 degrees.



Figure 5.9: Environment Monitor GUI

#### Tasks Monitor

This component is responsible for displaying the pending tasks queue together with information about the requested resources and is implemented in the *TasksQueueMonitor* class located in the *contextawaremodel.gui.resourceMonitor.taskMonitor* package.

In Figure 5.10 the Task Monitor window is represented together wish several tasks. The monitor window displays the tasks information in two ways under the form of a tab panel. On the first tab just a list with all the tasks monitored by this monitor is displayed (Figure 5.10 left). The second tab displays the entire information available for each task. More exactly, for each task the requested resources ranges and received resources values are displayed. For better visibility, each task receives a randomly assigned color (lighter than the black used for the task names).



Figure 5.10: Task Monitor GUI

#### Server Monitors

This component is the most complex context monitoring user interface module used in this application because it displays information about the running tasks, the total resource usage and the resource usage per task. The component implementation can be found in the *contextawaremodel.gui.resourceMonitor.serverMonitorPlotter.impl* package, the component representing a composition of different classes.

For managing the complexity of the Server Monitor components layout, the components are organized in a tree-like hierarchy, with smaller components being included in larger ones and so on until the entire monitor is created. The architecture of the Server Monitor is presented in .

The Server Monitor has three main components: a Tasks Monitor, a Total Resource Usage Monitor and a Resource Usage Per Task Monitor. The Tasks Monitor consists of the component described at 5.6.4.2 Tasks Monitor. The Total Resources Usage Monitor is a XY Chart plotter which creates a plot of the total resource usage history. The Resource Usage Per Task Monitor creates a Pie chart with a slice for each task running on the monitored server. Each slice represents the amount of resources used by that task. Also, the free resources and resources used by the server’s operating system are shown.

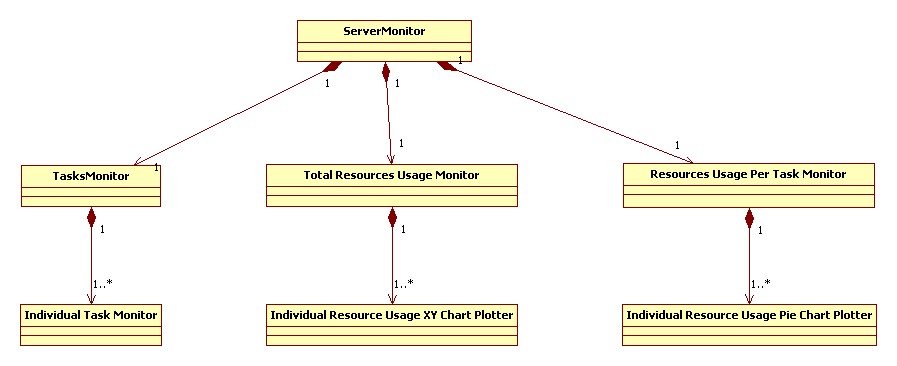


Figure 5.11: Server Monitor Window Architecture

In Figure 5.12 a server monitor snapshot is illustrated. In the snapshot the monitor displays information about a server having two cores and which does not run any virtual machines. In the left of the image is the running tasks display window, in the center is the total resources usage window and in the right hand side is the resource monitor who shows individual resources usage, in this case just the OS and Free.



Figure 5.12: Server Monitor Window

#### Logging Window

The logging window is responsible for displaying periodically the context state, the broken policies and the actions taken for bringing the context in an acceptable state. The component implementation can be found in the *LoggerGUI* class located in the *logger* package.

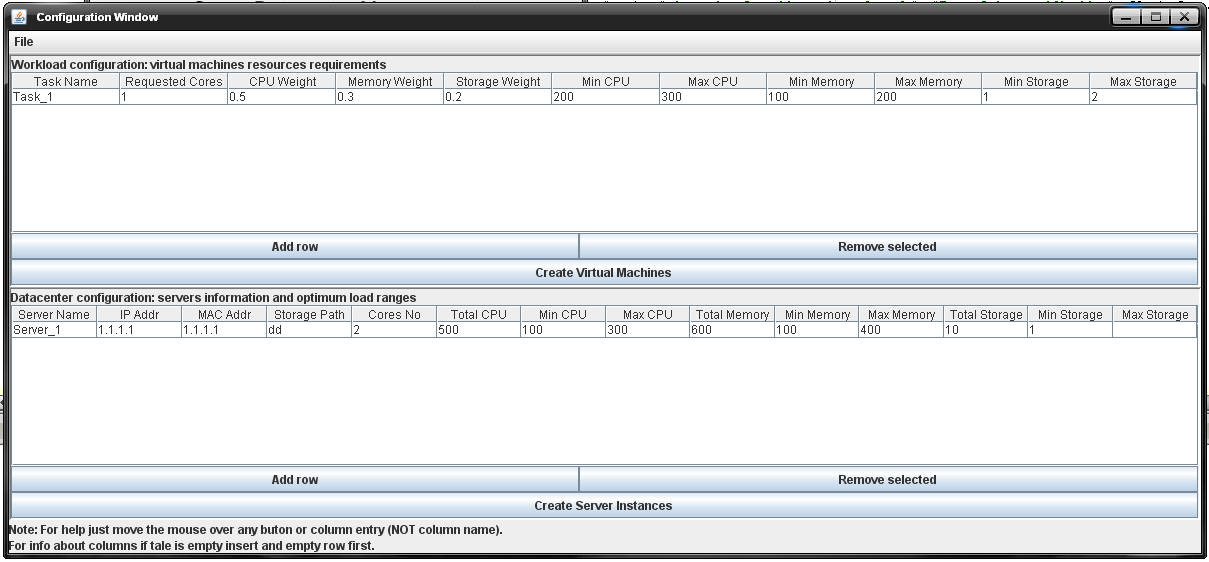
This component is important because it provides the context management history and makes it easier to trace the system. Two log windows are provided, one for the environment self-healing and one for the datacenter self-adapting systems. An example of a log is shown in Figure 5.13.



Figure 5.13: Datacenter Management Log

#### Datacenter Servers and Workload Definition Window

This window is used for simulating different sysyem evalaution scenarios. The servers in the datacenter are defined in terms of network location, their resources and optimum load ranges. The workload is defined as a series of virtual machines. Each virtual machine is defined in terms of importance for each resource and the acceptable resources usage ranges. The window also provides the support for saving and loading separately workload and datacenter servers configurations.



## Utility components

### Fuzzy Logic Negotiator

The component is implemented in the *FuzzyLogicNegotiator* class in the *negotiator.impl* package. The Fuzzy Logic Negotiator is the implementation of the negotiation mechanism described in . The implementation of the *FuzzyLogicNegotiator* relies on JFuzyLogic ] , which implements the Fuzzy Control Language(FCL) specification ]. Theis negotiator is used in the self-adapting datacenter scenario for QoS-Energy Consumption tradeoff.

JFuzzyLogic is a package written in Java which provides a mechanism for loading and dynamic management of FCL files. Using JFuzzyLogic the user can define any membership function, modify values or ranges and evaluate the fuzzy rules described in the input file. For deffuzzification JFuzzyLogic implements entirely the FCL specification, supporting Center of Gravity, Centre of Area, Left Most Maximum, Right Most Maximum and Centre of Gravity for Singletons. In Fuzzy Logic, each value has associated a membership value. The membership represents the degree to which the value belongs to the specified range. JFuzzyLogic was chosen for implementing the negotiation mechanism because it makes it easy to define bidimensional functions where one dimension represents the membership of the other dimension to a certain fuzzy value. Also, the provided deffuzzification mechanism implements the computation of the center of gravity in an efficient manner.

In this negotiation mechanism the two negotiating parties are a Server and a Task. The server wants to keep its load inside some optimal parameters, while the task wants more resources than the server is prepared to offer. In this case, the server desirability is decreasing as the load he would need to take increases. The same is for the task. In in the left half the two negotiation parties offers are represented: the server’s with pink and the task’s offer with gray. In this case the desirability function for the server is maximal at 1.500, meaning that the server would be most happy if that would be the result of the negotiation. But the minimum amount requested by the task is 1.650 and it would be happiest to receive 1700. So, the area where the two offers overlap is extracted: 1650 – 1700. Because the task satisfaction is the main priority, the desirability of the task is used for building the negotiation area function, as shown at the right half in . After the new fuzzy function is defined based on the overlapping ranges of both offers and the desirability function of the task, the center of gravity of the resulting area is computed by JFuzzyLogic. The center of gravity represents the result of the negotiation.

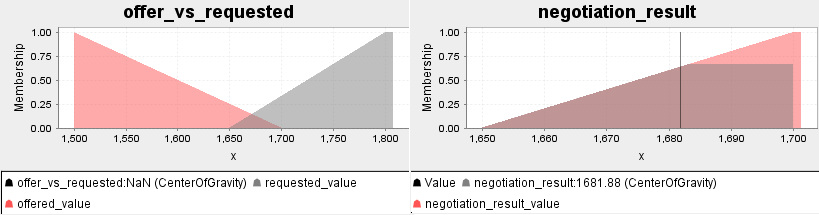


Figure 5.14: Server-Task Negotiation

JFuzzyLogic uses a .fcl file in which the user specifies the variables and their fuzzification and deffuzzification mechanisms. Also, each file has at least one rule block inside which if-then-else conditions are used to assign values to the output variables depending on the input variables. For implementing the negotiation mechanism described above, a (value, desirability) pair is defined for each inflexion point of each offer as shown in . In this example a trapezoidal desirability function is defined, with low and high values being less desirable than values from 1500 to 1700.

|  |
| --- |
| Listing 5.6: FCL specification |
|  |

### Server Management Proxy

For connecting the system described in this report to a real-world datacenter an interface between this system and the datacenter is needed. Also, the interface should have a high degree of flexibility.

In order to support different *ServerManagementProxy* subclasses, needed when porting the system from one datacenter to another, the Factory Method Design Pattern is used to provide a flexible and transparent mechanism for proxy instantiation. The pattern is realized using the *ProxyFactory* class located in the *contextawaremodel.worldInterface.datacenterInterface.proxies.impl* package and is illustrated in .



Figure 5.15: Server Management Interface Architecture

The *ServerManagementProxyInterface* interface specifies the functionality needed for managing a datacenter and represents the contract which every concrete server management proxy must implement, as presented in .

|  |
| --- |
| Listing 5.7: Server Management Interface |
|  |

### Server Information Gathering Endpoint

This component was created to provide real-time server resources usage data. The implementation was developed for servers running Windows Server 2088 R2, using operating system specific functions.

For gathering server data the WIN32 Classes ] provided trough the Windows Management Instrumentation(WMI) ] where used. WMI offers the means to access system information using Win32 classes from any .NET supported language. C# was chosen as the host language due to its support for web services and tight coupling with the underlying operating system. The reason Java wasn’t used for information gathering is that the platform independency of the language makes it hard to query hardware specific information. C# does not have this drawback, being able to access both hardware and software information about the host system. Another advantage of using C# is the easiness with which functionality can be exposed under the form of web services.

The web service created for exposing the functionality of the WMI server information gathering module are hosted on Microsoft IIS Application Server ] and accessed through HTTP ].

|  |
| --- |
| Listing 5.8: WIN32 WMI Specific Functions |
|  |

Chapter 6

# Testing and results

The concepts and algorithm described in Related Work where applied on a simulated smart room, on a simulated datacenter and on real servers.

## Self-Healing Environment Case Study

This case study involves applying the self-healing reinforcement learning algorithm described in to a smart environment. It has as goal to demonstrate the versatility and correctness of the autonomic framework described in this report by using it to manage a smart laboratory.

For this study the autonomic framework is described in this report is used to build a management system for a smart environment having a computer, a camera with face recognition, a light source and an alarm (Figure 2 and Figure 3). Each environment component has attached a sensor for monitoring its state. Also humidity, temperature and room person count sensors have been added to increase the environment’s complexity. Each sensor has a set of possible values encoded using integers for internal representation as described in . There are three policies which the management system has to enforce: *light policy,* *face* *recognition policy* and *temperature and humidity* policy (

Table 6.2). For interacting with the environment, actuators have been associated to the sensors they have an effect on as shown in .

Table 6.1: Sensors Values

|  |  |  |
| --- | --- | --- |
| Sensor | Possible values | Value encoding |
| Temperature | Z | Any integer |
| Humidity | [0-100]% | [0-100] |
| Light | {ON, OFF} | {1,0} |
| Face Recognition | {Professor, Student, Unknown} | {0,1,2} |
| Computer State | {ON, OFF} | {1,0} |
| Alarm State | {ON, OFF} | {1,0} |
| Room State | {Empty, Not Empty} | {0,1} |

Table 6.2: Policies

|  |  |
| --- | --- |
| Policy | Accepted values combination for policy |
| Temperature And Humidity |  |
| Face Recognition |  |
| Light |  |

Table 6.3: Sensors Associated Actuators

|  |  |  |
| --- | --- | --- |
| Sensor | Actuator | Available actions |
| Temperature | Air Conditioning Unit | {Decrease by 5 , Decrease by 2 } |
| Heater | { Increase by 5 } |
| Humidity | Humidity Controller | {Increase by 3 %,Decrease by 3 %} |
| Light | Light Controller | {Turn ON, Turn OFF} |
| Computer State | Computer Controller | {Turn ON, Turn OFF} |
| Alarm State | Alarm Controller | {Turn ON, Turn OFF} |
| Face Recognition | - | - |
| Room State | - | - |

Two tests have been performed on this scenario. The first test is meant to show the overhead of the learning mechanism used in terms of running time. The second test is an interactive system evaluation in which all the possible combinations of broken policies are tested by a human participant using a 3D interactive environment representation.

In order to have a scenario as realistic as possible, C# ASP.NET Web Services where used to simulate the environment sensors. One web service was built for each environment sensor. This approach allows the implementation to be easily switched to use real sensors accessible by web services and thus providing the possibility of aplying the system on real environments. Listing 6.1 presents an example of a simulated sensor: the Alarm sensor. The functionality exposed by the web service for a sensor is always get and set sensor value, allowing the web services to be called uniformly.

|  |
| --- |
| Listing 6.1: Alarm Sensor Simulation Web Service |
|  |

When an action is performed on one of the represented actuators, the actuator influences the sensor values by changing the target web service value according to the specified effect.

### Algorithm Overhead Evaluation

For the first system evaluation, the context was continuously and randomly broken for 28 hours. During this time all sensors received random values as following: for Temperature from 15 to 25, for Humidity from 15 to 35, 0 or 1(“OFF”, “ON”) for Light, Room State, Computer State and Alarm State sensors and 0, 1, 2 (“Professor”, “Student”, “Unknown”) for the Face Recognition Sensor. After the algorithm found a solution the random values assignment was made. The running time of the algorithm is illustrated below in .



Figure 6.1: 28h Random Values Evaluation

Each spike in the plot represents a situation in which an unknown context state has been encountered and a search for the best sequence of actions was performed. In the first 10000 seconds, almost all the running times of the action selection algorithm are larger than 50 seconds. As the knowledge base is being populated with learned knowledge the action search running time decreases, reaching only three running times greater than 50 seconds in the time interval [50000, 70000]. Also an overall reduction in the number and height of the peaks is visible because at each step the algorithm checks if it doesn’t already know the best sequence of actions for the context that it arrived in. If this is not the case, results will be stored in association with the current context. Considering that the number of possible sensor combinations for the chosen ranges is 22.481.940, the self-healing mechanism behaves quite well in rapidly finding and taking the needed actions for fixing the broken context.

### Interactive Evaluation

For the second test an interactive X3D context representation was built using the X3DAgent. The 3D context representation is illustrated in . The representation allows the user to interact with each of the environment actuators by mouse click events. By allowing the user to perform actions on the environment actuators and not on the sensors themselves, the representation can be used to apply the self-healing solution on real-world systems.

In this representation there is one actuator associated to each sensor, even for *FaceRecognition* and *RoomState* sensors because in order to fully evaluate the system their values should be changeable to. For altering the environment, the user has to click on the corresponding actuator and the actuator action will be performed. The actuators are represented as follows:

* The *RoomState* sensor has as actuator the green floor. By clicking it the state of the sensor changes to “Room empty: TRUE “or “Room empty: FALSE “. This representation for the actuator was chosen because it can be associated to a pressure sensor.
* The *AirConditioningUnit* is represented in the upper left part of the context screenshot from Figure 6 having the shape of a regular air conditioning unit.
* The *Heater* is represented in the right of the image, under the Alarm.
* The *AlarmController* is represented as a red ball in the right hanbd side of the image.
* For controlling the *ComputerState* the suer must click on the green box representing the computer cenbtral unit, located at the right of the computer screen visible in the left of the image.
* The *LightController* is represented under the form of a white box representing a spotlight in the upper part of the image.
* Finally, the *FaceRecognitionSensor*’s value can be changed by performing a click event on the surveilance camera representation.

For the *AirConditioningUnit* actuator only the *Descrease* by 5 degrees action is available during simulation. For a sensor which has only discrete values, an action performed on its actuator representation will change its state to the next state in the set of possible states associated to it.



Figure 6.2: X3D Smart Environment Representation

Using this simulated environment the output of the system is checked by a human for every possible input scenario. The user modifies the environment such that some policies are broken and then monitors the system’s output sand checks if the suggested sequence of actions is the best for bringing the environment in an acceptable state.

For better understanding, a trace of the environment repair process is presented on the following case: a professor enters in the room, the Computer is “OFF” and the Alarm is “ON”, breaking the Face Recognition Policy. For each resource which doesn’t have an acceptable value for the current policy the learning algorithm simulates executing each action attached to the actuators of the current resource and generates a new context. The first resource which is broken is the Computer, so the algorithm explores every action which brings it to a state different from the current state. Here only one action is possible, so the Computer will be “Set” to “ON” (setting it to “OFF” would not change its state so this action is not considered). For each resulting state we continue with the next sensor that breaks the policy, the Alarm sensor. Similar to the previous sensor, the Alarm can only be set to “OFF”. By performing these actions the Face Recognition Policy is fixed and the context is brought to an acceptable state. This situation is illustrated in the left part of the figure representing the broken context situation while the right hand side of the image contains the monitor described in 5.6.4.1 Environment Monitor, which displays the set of actions which were found to repair the context.

**

Figure 6.3: Face Recognition Policy broken

## Self-Adapting Datacenter Case Study

For this case study the framework described in this report was used to build an energy-efficient datacenter. This is meant to further prove the versatility of the solution proposed in this report and to show that self-adaptability and self-healing can be considered isomorphic. The isomorphism between the two self-\* properties can be deduced from the fact that they can be both realized with the same domain model and the same property enforcement protocols.

A second environment ontology representation shown in which contains datacenter specific information has been added and used in the datacenter self-adapting process. The original ontology representation was used for environment self-healing.

The self-adapting extension was tested first on a simulated datacenter and after the solution proved correct on a small home-brewed datacenter containing two servers and a shared storage.

### Simulated Datacenter Evaluation

An X3D representation of a datacenter was built, containing 5 servers and a central control unit arranged in a spoke and arrow layout (). The datacenter has one Temperature and one Humidity sensor. The datacenter cooling system is represented as a giant fan in the middle of the datacenter. The servers are represented as wireframe towers with a solid base. The base color represents the state of the server: gray means *offline* and blue means *online*. The tasks are represented by purple boxes. Each server has a power meter attached which shows the power consumption. For this test the power consumption for a server was considered to be 10\* number of running tasks + 10 Watts.



Figure 6.4: X3D Datacenter Representation

For this test all the tasks are considered having the same priority. Also, the learning algorithm is run for the first time so no action memory exists. For the real case, if the current context has a sequence of actions associated to it in the memory, that sequence of actions is executed or added to the existent path of actions reached so far. For each task, a policy is generated, containing the specified values for the QoS attributes.

The negotiation part described in was not yet plugged in for this test, thus the system was tested on a workload and hardware infrastructure for which at least one deployment strategy which fully respects both QoS and Energy policies exists. Otherwise the algorithm would return the best path it has found and at the next context evaluation time it will search again for ways of repairing the context.

#### Simulated Datacenter Hardware infrastructure

Concerning the hardware infrastructure, the simulated datacenter contains 5 servers described in the table below (). For each server a policy and associated SWRL rule containing key performance indicators are generated and used in the optimization process.

Table 6.4: Datacenter servers’ characteristics

|  |  |  |  |
| --- | --- | --- | --- |
| Nr | CPU | Memory | Storage |
| 1 | 1 x 3000 MHz | 2048 DDR2-1066 | 250 GB @ 7200 rpm |
| 2 | 2 x 3000 MHz | 2048 DDR2-800 | 250 GB @ 7200 rpm |
| 3 | 4 x 2000 MHz | 4096 DDR3-800 | 146 GB @ 10000 rpm |
| 4 | 4 x 2260 MHz | 6144 DDR3-1333 | 146 GB @ 15000 rpm |
| 5 | 8 x 2000 MHz | 8192 DDR3-1600 | 300 GB @ 15000 rpm |
|  |  |  |  |

#### Simulated Datacenter Workload

For simulating a real world scenario we have chosen 5 tasks which are received by the datacenter and are considered to have an infinite life span. The tasks are received in the order they are in the table below ().

Table 6.5: Received tasks SLA requirements

|  |  |  |  |
| --- | --- | --- | --- |
| Task Nr | Task Requirements | | |
| CPU | Memory | Storage |
| 1 | 3000 MHz | 2048 MB | 200 MB |
| 2 | 2 x 1500 MHz | 512 MB | 400 MB |
| 3 | 2 x 2000 MHz | 1024 MB | 256 MB |
| 4 | 8 x 512 Mhz | 240 MB | 128 MB |
| 5 | 3 x 2000 MHz | 4096 | 300 GB |

As initial context the Server 1 is active and the other servers are in a low power state. The basic flow of this example scenario is the following: (i) task 1 is received and is deployed on server 1 which reaches its load threshold and a wake up search which concludes by activating server 2; (ii) task 2 is received and deployed on server 2; (ii) task 3 is received; there is no room to deploy so server 3 is activated; all the task rearrangement and deployment options are analyzed and the solution is to move task 2 on server 3, deploy task 3 on the same server and send server 2 to low power state.

#### System Response Trace

At context evaluation time, the self-adapting algorithm notices that the entropy is larger than zero therefore one or more broken policies exist. The reinforcement learning algorithm will evaluate all possible paths and simulate taking the one with the largest reward. It will reach to a result composed of four actions: deploy task 1 on server 1(), wake up server 3(), deploy task 2 on server 3() and deploy task 3 on server 3(). After this sequence of actions is returned by the reinforcement learning algorithm, the actions are taken in this exact order, bringing the system to an accepted state.



Figure 6.5: Deploy Task 1 on Server 1 Action



Figure 6.6: Wake up Server 3 Action



Figure 6.7: Deploy Task 2 on Server 3 Action



Figure 6.8: Deploy Task 3 on Server 3 Action

After finding a new action plan for a context, the pair (context, action plan) is stored for further reference. The reinforcement learning algorithm will always check first if this situation hasn’t been encountered before and if so, it will take the learned actions.

#### Algorithm Overhead

After running the algorithm a few times it became clear that the overhead generated by the **Pellet** **OWL 2 Reasoner for Java** ] when evaluating SWRL rules is unacceptable in a real-time application, the rules being evaluated at each change in any ontology value. So SWRL rules where disabled and internal reasoning was implemented using Java. This lead to a decrease in running time of up to 4 times, depending on the type of processor the global loop runs on.

### Real Datacenter System Evaluation

The self-adapting extension was also tested on a home-made datacenter running virtual machines.



#### Hardware Infrastructure

The datacenter architecture, as presented in consists of a server cluster (2 servers in this test), a shared storage used by the cluster, environment sensors and a computer running the autonomic system controller built using the framework presented in this report. The shared storage is used to host the virtual machines in order to be accessible from any server in the task migration process. For this there are only two environment sensors, *Temperature* and *Humidity* sensors, which are simulated as described in Self-Healing Environment Case Study.



Figure 6.9: Case Study Datacenter Structure

#### Software Architecture

The software architecture of the evaluation datacenter’s servers is presented in . The servers in the datacenter are running Microsoft Windows Server 2008 R2 ] as operating system. The hypervisor installed on the servers is Hyper-V Server R2 ]. For this evalauton 2 server where used, each server having a 2 GHz Dual Core processor, 2 GB of RAM and a 150 GB hard drive.



Figure 6.10: Server Technology Mapping

For interacting with the datacenter C# WMI classes where used both for gathering server resources usage data and to access the hyper-v server running on the servers in order to move, import and destroy virtual machines dynamically. The C# classes functionality was exposed trough ASP.NET web services, published on Microsoft IIS Application Server [ 68 ]. The web services are called from within the *Hyper-VManagementProxy* class located in the *contextawaremodel.worldInterface.datacenterInterface.proxies.impl* package and which implements the *ServerManagementProxyInterface*. The Hyper-V management proxy and the WMI Hyper-V interaction mechanism are described in detail in ] by its author. In Figure 6.11 the interaction between the autonomic controller and the server endpoints is shown. A concrete implementation of the *ServerManagementProxyInterface* is used to call the web services publiched on the IIS Server.The web services in turn call the specific functionality on the two C# endpoints, Server Monitor and Hyper-V Manager.



Figure 6.11: Server Interaction

#### Case Study Workload

For the ase study, the datacenter workload consists of three virtual machines running Windows XP ] where created using the Hyper-V Management GUI. Due to the lack of time, no dynamic virtual machine resource allocation mechanism was built. Given this fact, the virtual machines physically receive the minimum resources amount specified in when deployed. But inside the control logic, specified ranges are used and each task deployed without negotiation is marked to use the maximum amount of resources required.

For this test the negotiation mechanism presented in 5.7.1 Fuzzy Logic Negotiator was plugged in the controller. The negotiation mechanism was used to negotiate the allocation of resources between tasks when no other solution was possible, after every task reallocation strategy was tested. Such a situation appears when every server in the datacenter is running and can’t accommodate any new tasks while maintaining its optimum load, but could accommodate more tasks if its load would exceed the optimum limits. In this case the negotiation mechanism is used to find a tradeoff between the optimum load and the resources required by the task. After negotiation, the server receives a new optimum load value.

The task requirements were chosen based on the datacenter server’s configuration to use less RAM in order to be accommodated on the servers. Windows Server 2008 R2 together with Hyper-V Server, IIS Server and other programs use around 1GB of RAM of the 2 available on the test servers, so the amount of RAM the virtual machine use was a limitation.

Table 6.6: Virtual Machines Resource Allocation

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Task Nr | Task Requirements | | | | | | |
| CPU (Mhz) | | Memory (MB) | | Storage (GB) | | |
| Min | Max | Min | Max | Min | max | |
| 1 | 100 | 500 | 100 | 300 | 1 | | 2 |
| 2 | 2 x 300 | 2 x 500 | 100 | 300 | 1 | | 2 |
| 3 | 2 x 100 | 2 x 200 | 50 | 400 | 1 | | 2 |

The virtual machines described by Table 6.6 where created using the graphical user interface of the Hyper-V Manager. Each virtual machine has a lightweight Windows XP version installed as operating system. For each virtual machine a java is set to run when the operating system loads, program which performs some computations in an infinite loop, using the virtual machine CPU at 100%.

#### Evaluated Cases

Using these virtual machines two evaluations were performed. In the first, by cloning the virtual machines, workload test sequences were generated randomly. A random number of test sequences of random length scheduled to execute (to be received by the system to be deployed) at random intervals of time where used to test the datacenter management system under real world loads.

The second evalaution involved testing if the system allocates the resources accordingly and if the virtual machines deployed by the application are usable. This test only involved the three virtual machines described in Table 6.6. In this case all the three machines can be fitted to only one server. The system chose to wake up Server\_2. Initially there are no virtual machines managed by the Hyper-V Server running on Server\_2, as Figure 6.12 illustrates.

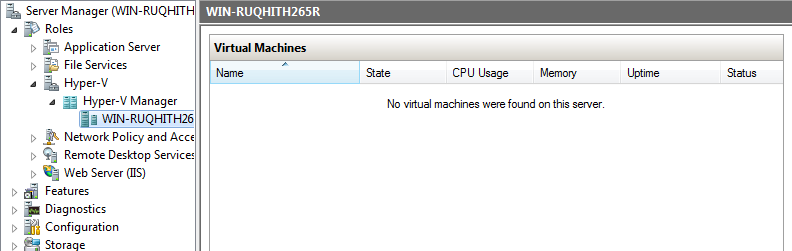


Figure 6.12: Initial State

The system deploys all three virtual machines on Server\_2 and the three virtual machines can be seen running on the Hyper-V Server and using resources from Server\_2. Figure 6.13 shows the state of the Hyper-V Server.

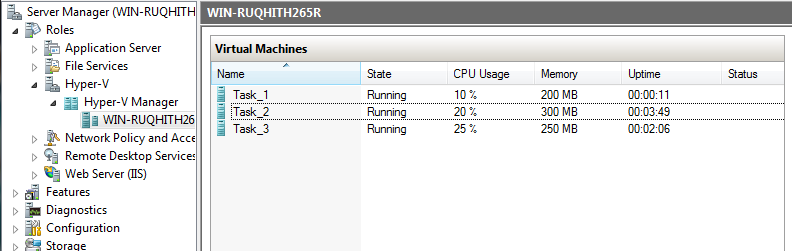


Figure 6.13: Final State

In order to test if the virtual machines that are running haven’t been corrupted by the deployment mechanism each machine was inspected. For each machine it was proven that it indeed run a Windows XP operating system and had a Java program using the virtual machine CPU. In Figure 6.14 a connection to the virtual machine named Task\_2 is made in order to see what it is indeed running in that virtual machine. From the image it can be observed that the virtual machine runs Windows and that a java program was started from the command line and is still running. The fact that the program is still running is obvious from the command line, which is hanging, waiting for the program to complete. When the task completes a C:\ entry will appear in the command line.

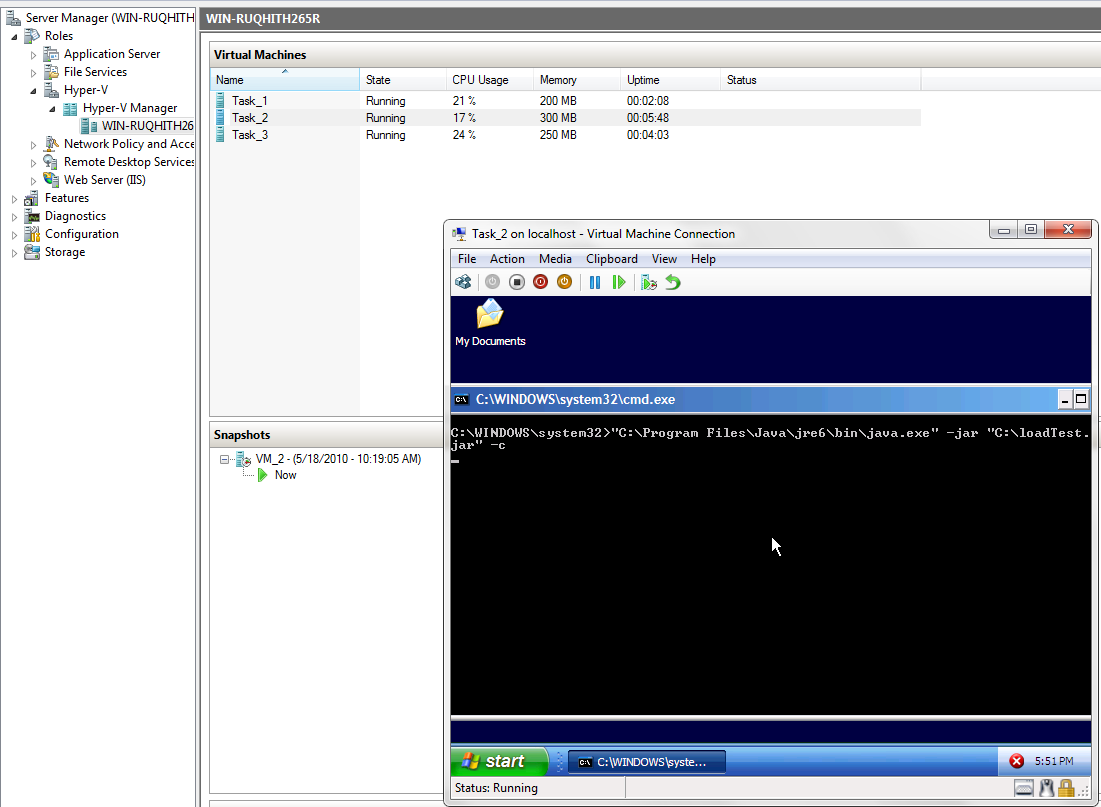


Figure 6.14: Running Virtual Machine

On the next page Figure 6.15 illustrates a snapshot of the server monitor after the three tasks where deployed. It can be seen in the total resources usage the intense usage of the CPU. Also, the Memory monitor captured the increase in memory usage after deploying the last task onto the server. It can be noticed that the Hyper-V Server did not allocated all the memory it was requested. This is standard Hyper-V behavior. The requested resources are reserved, but the virtual machine decides if it uses them or not.

In the lower part of the monitor window for each resource it is represented the amount each tasks uses. It can be noticed that Core\_22 hosts all three tasks, while Core\_21 runs only two tasks. This is because Task-1 is a single – core virtual machine. The usage can be also seen for the server memory and storage.

In the left of the monitor image the running task monitor is displaying the name of the running virtual machines.

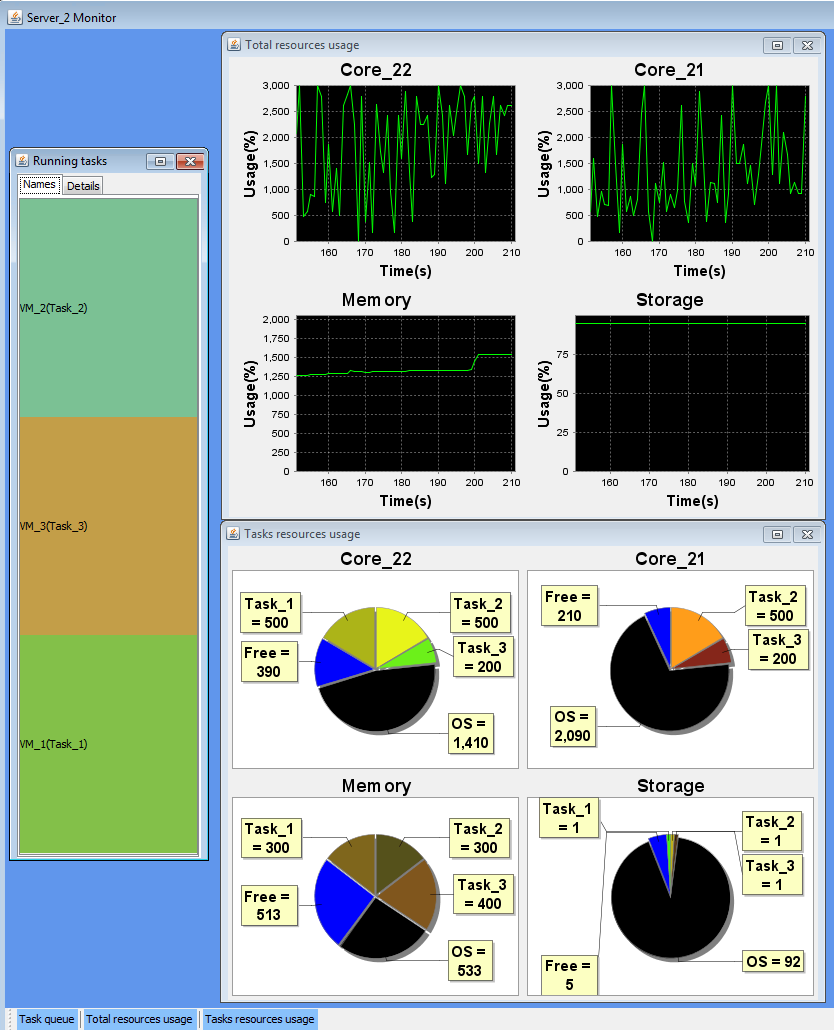


Figure 6.15: Server Monitor Snapshot

#### Encountered Issues

There were several problems encountered when applying the autonomic system’s management solution to a self-adapting datacenter due to the drawbacks of existing technologies.

##### Server wake up time problem

The server wake up time turned out to be a problem. As the authors of [ 23 ] also encountered for Linux servers, a server running Windows Server 2008 takes between 5 to 10 minutes to start. This means that is not feasible to turn off unused servers because the amount of power they use when starting and the time it takes to start such a server can greatly reduce the response time of the datacenter. Also, there is no currently available mechanism for turning on a Windows Storage Server that was turned off.

The solution to the problem was to send the server in “Sleep” mode. A server in sleep still consumes energy, but only a small amount, but it recovers from sleep in less than 1 minute.

Following the “send to sleep” solution, the server wake up mechanism consists of sending a magic packet containing 16 times the server’s MAC address the datacenter router which forwards it to port 9 of the server. By enabling Wake On LAN from the server network card, the event of receiving the magic packet will wake up the server from the sleep state.

##### Send server to sleep problem

Hyper-V Server disables the Sleep and Hibernate options of the Windows Server before running. Also, the Sleep and Hibernate options can’t be activated once turned off by Hyper-V without a system restart.

For solving this problem, a registry modification was made in order to ban Hyper-V Server from running automatically when the system started. Further, the send *server to low power state* action was implemented in two steps. In the first step the server is restarted and a small program is set to run on startup. When the server restarts, the program set to run on startup sends the server to sleep. This solution is affected by first problem described in 7.2.2.1.1 Server wake up time problem, the time needed for a Windows Server 2008 system to boot. But this would not represent a problem if a load prediction mechanism would be employed that could estimate that the server sent to sleep would not require to be brought online in the next 10 minutes.

##### Dynamic task migration problem

Hyper-V supports real-time task migration with minimum downtime only under special conditions. More exactly, the shared storage used by the server cluster must support Internet Small Computer Interface (ISCSI) [ 72 ] with Persistent Reservation. This reservation is implemented in ISCI-3 standard and allows reserving persistently a storage area for a particular server. Unfortunately, the only Software Storage Servers which implement this functionality and can be included in a Windows Server 208 R2 cluster are servers running Windows Storage Server 2008. Windows Storage server 2008 is available only through Original Equipment Manufacturer licensing, preinstalled with built Storage Servers.

To solve this problem, the task migration was, just as “send server to sleep operation”, performed in two steps: in the first step a snapshot of the virtual machine is exported to the shared storage and in the second step that snapshot is imported onto the desired server by the Hyper-V Server running on it. The drawback of this implementation is that exporting and importing a virtual machine with a hard drive usage of 500 MB takes up to 4 minutes, delay time observed in the real datacenter case study.

##### Virtual Machines Resource Allocation Issue

Even if for a virtual machine it was set that is must receive 50% of one CPU core and the task manager of the windows installed inside the virtual machine showed that the virtual machine CPU is loaded at 100%, Hyper-V reports that the virtual machine uses 2-3% of the total CPU time. This is due to some optimizations done internally by Hyper-V Server and for the time being this issue was left unsolved.

##### Maximum Virtual Machines Number

The Windows Server 2008 R2 version used for the servers in the datacenter was the Enterprise edition, which allows only 4 virtual machines to be run on a server. This means that when our application tried to deploy and start a 5’Th virtual machine the result was an error coming from the operating system.

Chapter 7

# Conclusions



The solution to building autonomic systems proposed in this report has been built together with Copil Georgiana ]. The solution is an extensible agent-based framework that can be used to create and manage autonomic context-aware systems. The framework uses a reinforcement learning based learning mechanism in which the system decides what course of action to take based on the expected reward. The framework also contains the necessary support for context data gathering, data reasoning and action enforcement.

For the reinforcement learning mechanism a means of computing the expected reward based on the context entropy was presented. Also, a context evaluation mechanism based on based policies and weights is presented.

The presented framework was first used to build an autonomic self-healing laboratory. The resulting system was extensively evaluated both in terms of overhead and correctness. The framework was extended with some domain-specific information in order o be applied on a self-adapting datacenter, where it controlled both the smart environment which housed the datacenter and the datacenter workload distribution. By using this approach to enforce energy policies the power consumption of a datacenter is reduced.

A negotiation mechanism to be used in computing QoS-Energy tradeoffs was presented. The mechanism uses bidimensional functions as negotiation offers, one dimension representing the desirability of the negotiating party to offer the value on the other dimension. The negotiation mechanism was implemented efficiently using a fuzzy logic library and used for the energy-aware datacenter

The approach to datacenter workload representation was the use of virtual machines, which provide dynamic migration from one server to another and give an independent environment for the client to use. Another reason behind the virtual machines approach is that a client can connect remotely to his virtual machine and tailor it after its specific needs, installing or uninstalling software. This gives the client the ability to personalize and tailor the virtual machine while keeping his data isolated from the other clients.

3D context representation where built both for the smart room and the energy efficient datacenter. The smart room 3D representation allowed user interaction and could have been used to monitor and manage a real environment.

The datacenter management system was applied on a homemade datacenter containing two virtual machines servers and one shared storage server. For interacting with the hypervisor the hypervisor API exposed trough WMI was in turn exposed trough ASP.NET web services invoked from the java application endpoint. The servers’ resource utilization was recorded using the WIN32 API exposed trough WMI, in turn as with hyper-v exposed trough ASP.NET web services.

By mapping the same domain model and the same solutions search algorithm on both a self-healing and a self-adapting environment, the two autonomic systems properties can be considered isomorphic and solvable by the same processes.

The presented solution is a general framework which can be used to build any autonomic system just by modifying the domain model ontology representation and adding domain specific policies. This demonstrates that a reinforcement-learning approach is very versatile because any real world system can be mapped to an environment described by policies and reward functions.

## Future Work

Regarding the energy-aware datacenter, the current implementation lacks the support to create an install an operating system in a virtual machine dynamically by the user. Also, the live migration mechanism should be used for task migration. For this a suitable storage server needs to be found and the Hyper-V functionality needed for real-time task migration discovered and used. Another future aspect is to install the Datacenter edition of Windows Server 2008 R2, which allows an unlimited number of virtual machines per server. Also regarding live datacenter case study, power meters should be used to measure the power consumption of the datacenter with and without the self-adapting controller.

Another improvement over the existing platform is to integrate Open Sim [ **74** ] as means of context representation and interaction. Open Sim simulator gives the possibility to define detailed interactive 3D environments which the user can navigate using a virtual character. Such a representation would provide a good mechanism for monitoring and managing autonomic environments.

# Bibliography

[ 1 ] D.A.Menasce, J.O.Kephart: **Autonomic Computing**

[ 2 ] D.M.Chess, J.O.Kephart: **The Vision of Autonomic Computing**

[ 3 ] R.Sterrit, M.Hinchey: **Why Computer-Based Systems Should be Authomatic**

[ 4 ] M.Parashar, S.Hariri: **Autonomic Computing: An Overview**

[ 5 ] Green **Active Management of Energy in IT Service centres:** Performance Computing Center Stuttgart (HLRS), <http://www.hlrs.de/organization/av/isis/research/distributed-systems/games/>

[ 6 ] **VMware Whitepaper**: How VMware Virtualization Right-sizes IT Infrastructure to Reduce Power Consumption, <http://www.vmware.com/files/pdf/WhitePaper_ReducePowerConsumption.pdf>, pages 4

[ 7 ] **EPA** [Data Center**Report to Congress -FINAL 7-25-07**](http://www.federalnewsradio.com/pdfs/EPADatacenterReporttoCongress-August2007.pdf)**,** <http://www.federalnewsradio.com/pdfs/EPADatacenterReporttoCongress-August2007.pdf>, Pages 6

[ 8 ] M. Poniatowski: **Foundations of Green IT: Consolidation, Virtualization, Efficiency, and ROI in the Data Center,** Prentice Hall, 2009, ISBN-10: 0137043759, ISBN-13: 9780137043750, 978-0137043750, pages xiii

[ 9 ] T.Cioara, I.Anghel, I.Salomie, M.Dinsoreanu, C.Georgiana and **M.Daniel**: **A Reinforcement Learning based Self-healing Algorithm for Managing Context Adaptation,** First International Workshop on Communication, Collaboration and Social Networking in Pervasive Computing Environments (PerCol 2010)

[ 10 ] T.Cioara, I.Anghel, I.Salomie, C.Georgiana and **M.Daniel:** **An Autonomic Algorithm for Energy Efficiency in Service Centers**, 2010 IEEE International Conference on Computer Communication and Processing (ICCP10), submitted

[ 11 ] Ubiquitous Computing, <http://en.wikipedia.org/wiki/Ubiquitous_computing>

[ 12 ] S.Russell, P.Norvig : **Artificial Intelligence A Modern Approach**, Second Edition, Prentice Hall,ISBN-10: 0137903952, ISBN-13: 978-0137903955

[ 13 ] P.C.K. Hung, H. Li: **Web Services Discovery based on the Trade-off between Quality and Cost of Service: A Token-based Approach**

[ 14 ] C. Sakama, K. Inoue: **Negotiation by Abduction and Relaxation**

[ 15 ] Abductive Logic Programming, <http://en.wikipedia.org/wiki/Abductive_logic_programming>

[ 16 ] Y.Chevaleyre, U.Endriss, S.Estivie, N.Maudet : **Reaching Envy-free States in Distributed Negotiation Settings**

[ 17 ] Q.B.Vo, L.Padgham: **Searching for Joint Gains in Automated Negotiations Based on Multi-criteria Decision Making Theory**

[ 18 ] K.Fujita, T.Ito, .Klein: **Finding Nash Bargaining Solutions for Multi-issue Negotiations: A Preliminary Result**

[ 19 ]D. Baker, D. Georgakopoulos, M. Nodine and A. Cichocki: **From Events to Awareness**

[ 20 ] P.Hu, J.Indulska, R,Robinson: **An Autonomic Context Management System for Pervasive Computing**

[ 21 ] IEEE 1451, <http://ieee1451.nist.gov/intro.htm>

[ 22 ] SensorML, <http://vast.nsstc.uah.edu/SensorML/>

[ 23 ] R.Das, J.o.Kephart, C.Lefurgy, G.Tesauro, D.W.Levine, H.Chan **: Autonomic Multi-Agent Management of Power and Performance in Data Centers**

[ 24 ] G.Chen, W.He, J.Liu,S.Nath,L.Rigas,L.Xiao,F.Zhao: **Energy-Aware Server Provisioning and Load Dispatching for Connection-Intensive Internet Services**

[ 25 ] D.Kusic,J.Kephart, J.E.Hanson, N.Kandasamy and G.Jiang:**Power and Performance Management of Virtualized Computing Environments Via Lookahead Control**

[ 26 ] M.Amoui, M.Salehie, S.Mirarab, LTahvildari : **Adaptive Action Selection in Autonomic Software Using Reinforcement Learning**

[ 27 ] M.Salehieand,L.Tahvildari: **A weighted voting mechanism for action selection problem in self-adaptive software.**

[ 28 ] I.Salomie, T.Cioara, I.Anghel, M. Dinsoreanu : **RAP – A Basic Context Awareness Model**

[ 29 ] C. Georgiana: **A Multi-Agent Reinforcement Learning Based Autonomic**

**Context Management Solution**:Reasoning, Nash Equilibrium Policy Negotiation and Action Enforcement, diploma thesis, Technical University of Cluj-Napoca, 2010

[ 30 ] S.C.Patrasciuc: **An Agent-Based Context Management Solution (1)**, diploma thesis, Technical University of Cluj-Napoca, 2010

[ 31 ] Nash Bargaining Game, <http://en.wikipedia.org/wiki/Nash_bargaining_game>

[ 32 ] Distributed Systems Research Laboratory, <http://dsrl.coned.utcluj.ro/>

[ 33 ] General Responsibility Assignment Software Principles, <http://en.wikipedia.org/wiki/GRASP_(object-oriented_design)>

[ 34 ]Separation of Concerns, <http://en.wikipedia.org/wiki/Separation_of_concerns>

[ 35 ] Dependency Inversion Principle, <http://en.wikipedia.org/wiki/Dependency_inversion_principle>

[ 36 ] Open/Closed principle, <http://en.wikipedia.org/wiki/Open/closed_principle>

[ 37 ] Data Transfer Object (DTO) design pattern, <http://en.wikipedia.org/wiki/Data_transfer_object>

[ 38 ] Protégé, <http://protege.stanford.edu/>

[ 39 ] Protégé Owl Loader, <http://protege.stanford.edu/protege/3.4/docs/api/owl/edu/stanford/smi/protegex/owl/ProtegeOWL.html>

[ 40 ] OWLModel interface, <http://protege-owl.sourceforge.net/javadoc/edu/stanford/smi/protegex/owl/model/OWLModel.html>

[ 41 ] SWRL, <http://www.w3.org/Submission/SWRL/>

[ 42 ] Pellet, <http://clarkparsia.com/pellet/>

[ 43 ] Jena, <http://jena.sourceforge.net/>

[ 44 ] JenaOWLModel, <http://protege.stanford.edu/protege/3.4/docs/api/owl/edu/stanford/smi/protegex/owl/jena/JenaOWLModel.html>

[ 45 ] OntModel, <http://jena.sourceforge.net/javadoc/com/hp/hpl/jena/ontology/OntModel.html>

[ 46 ] JADE**,** <http://jade.tilab.com/>

[ 47 ] JADE A White Paper**,** <http://jade.tilab.com/papers/2003/WhitePaperJADEEXP.pdf>

[ 48 ] Jade Tutorial: Jade programming for beginners, <http://jade.tilab.com/doc/tutorials/JADEProgramming-Tutorial-for-beginners.pdf>

[ 49 ] FIPA, <http://www.fipa.org/>

[ 50 ] FIPA SL Languages, <http://www.fipa.org/specs/fipa00008/SC00008I.html>

[ 51 ] ASP.NET Web Services

[ 52 ] Java, <http://en.wikipedia.org/wiki/Java_(programming_language)>

[ 53 ] Java Mobile Edition, <http://java.sun.com/javame/index.jsp>

[ 54 ] Java Serialization, <http://en.wikipedia.org/wiki/Serialization>

[ 55 ] C#, <http://en.wikipedia.org/wiki/C_Sharp_(programming_language)>

[ 56 ] IText , <http://itextpdf.com/>

[ 57 ] X3D, <http://www.web3d.org/x3d/>

[ 58 ] XJ3D, <http://www.xj3d.org/>

[ 59] HTTP, <http://en.wikipedia.org/wiki/Hypertext_Transfer_Protocol>

[ 60] Fuzzy Logic, <http://en.wikipedia.org/wiki/Fuzzy_logic>

[ 61 ] JFuzzyLogic, <http://jfuzzylogic.sourceforge.net/html/index.html>

[ 62 ] FCL, <http://www.fuzzytech.com/binaries/ieccd1.pdf>

[ 63 ] Factory Method Design Pattern, <http://en.wikipedia.org/wiki/Factory_method_pattern>

[ 64 ] Windows Server 2008 R2, <http://www.microsoft.com/windowsserver2008/en/us/default.aspx>

[ 65 ] Windows Storage Server 2008 , <http://www.microsoft.com/windowsserver2008/en/us/wss08.aspx>

[ 66 ] Microsoft Hyper-V Server, <http://www.microsoft.com/hyper-v-server/en/us/default.aspx>

[ 67 ] Windows Management Instrumentation, <http://msdn.microsoft.com/en-us/library/aa394582(v=VS.85).aspx>

[ 68 ] Microsoft IIS, <http://www.iis.net/>

[ 69 ] WIN32 WMI Classes, <http://msdn.microsoft.com/en-us/library/aa394084(VS.85).aspx>

[ 70 ] Hyper-V WMI Provider, <http://msdn.microsoft.com/en-us/library/cc136992(VS.85).aspx>

[ 71 ] Microsoft Windows XP, <http://www.microsoft.com/windows/windows-xp/default.aspx>

[ 72 ] Small Computer Interface, <http://en.wikipedia.org/wiki/SCSI>

[ 73 ] Persistent reservation, <http://en.wikipedia.org/wiki/SCSI>

[ 74 ]Open Simulator, <http://opensimulator.org/wiki/Main_Page>

# Appendix A: List of Acronyms

|  |  |
| --- | --- |
| **3D** | Three-Dimensional |
| **API** | Application Programmer Interface |
| **CHOP** | Self-(Configuring, Healing, Optimizing, Protecting) |
| **CIA** | Context Interpreting Agent |
| **CMAA** | Context Model Administering Agent |
| **CONSENS** | Context Aware Autonomic Systems |
| **CPU** | Central Processing Unit |
| **DTO** | Data Transfer Object |
| **FCL** | Fuzzy Control Language |
| **GRASP** | General Responsibility Assignment Principles |
| **GUI** | Graphical User Interface |
| **IEEE** | Institute of Electrical and Electronics Engineer |
| **IIRG** | Inter-Independent Resources Group |
| **IIS** | Internet Information Services |
| **ISCSI** | Internet Small Computer Interface |
| **IT** | Information Technology |
| **J2ME** | Java Mobile Edition |
| **JADE** | Java Agent DEvelopment Framework |
| **kWh** | Kilowatts per hour |
| **LAN** | Local Area Network |
| **MAC** | Media Access Control |
| **MB** | Mega Bytes |
| **MHz** | Mega Hertz |
| **OS** | Operating System |
| **OWL** | Ontology Web Language |
| **PDF** | Portable Document Format |
| **QoS** | Quality Of Service |
| **RAP** | Resources, Actors, Policies |
| **RGB** | Red, Green, Blue |
| **RLA** | Reinforcement Learning Agent |
| **SUT** | System Under Test |
| **SWRL** | Semantic Web Rule Language |
| **TMA** | Task Management Agent |
| **VRML** | Virtual Reality Modeling Language |
| **WMI** | Windows Management Instrumentation |

# Appendix B: Java Source Code

|  |
| --- |
| FuzzyLogicNegotiator.java |

1 **package** negotiator.impl;

2

3 **import** greenContextOntology.\*;

4 **import** negotiator.Negotiator;

5 **import** net.sourceforge.jFuzzyLogic.FIS;

6 **import** net.sourceforge.jFuzzyLogic.FunctionBlock;

7 **import** net.sourceforge.jFuzzyLogic.membership.MembershipFunction;

8 **import** net.sourceforge.jFuzzyLogic.membership.MembershipFunctionPieceWiseLinear;

9 **import** net.sourceforge.jFuzzyLogic.membership.Value;

10 **import** net.sourceforge.jFuzzyLogic.rule.LinguisticTerm;

11 **import** net.sourceforge.jFuzzyLogic.rule.Variable;

12 **import** org.antlr.runtime.RecognitionException;

13

14 **import** java.util.Collection;

15 **import** java.util.HashMap;

16 **import** java.util.Map;

17

18 */\*\**

19  *\** ***@author*** *Moldovan Daniel*

20  *\*/*

21 **public class** FuzzyLogicNegotiator **implements** Negotiator {

22

23 **private** FIS fis;

24 **private** FunctionBlock functionBlock;

25

26 **private** MembershipFunction requestedRangeMembershipFunction;

27

28 **private** Variable serverRange;

29 **private** Variable requestedRange;

30 **private** Variable negotiatedRange;

31 **private** Variable serverAndRequested;

32

33 **private** LinguisticTerm serverRangeValue;

34 **private** LinguisticTerm requestedRangeValue;

35 **private** LinguisticTerm negotiatedRangeValue;

36 **private** LinguisticTerm serverValue;

37 **private** LinguisticTerm requestedValue;

38

39 **protected** FuzzyLogicNegotiator(String fuzzyControlLanguageFile) {

40

41 **try** {

42 fis = FIS.load(fuzzyControlLanguageFile, **true**);

43 } **catch** (RuntimeException e) {

44

45 System.err.println(**"FuzzyLogicNegotiator creation failed : Can't load file: '"**

46 + fuzzyControlLanguageFile + **"'."**);

47 e.printStackTrace();

48 **return**;

49 }

50 *// Error while loading?*

51 **if** (fis == **null**) {

52 System.err.println(**"FuzzyLogicNegotiator creation failed : Can't load file: '"**

53 + fuzzyControlLanguageFile + **"'."**);

54 **return**;

55 }

56

57 functionBlock = fis.getFunctionBlock(**"negotiator"**);

58

59

60 requestedRangeMembershipFunction =

61 functionBlock.getVariable(**"requested\_range"**)

62 .getLinguisticTerm(**"requested\_range\_value"**).getMembershipFunction();

63

64 serverRange = functionBlock.getVariable(**"server\_range"**);

65 requestedRange = functionBlock.getVariable(**"requested\_range"**);

66 negotiatedRange = functionBlock.getVariable(**"negotiated\_range"**);

67 serverAndRequested = functionBlock.getVariable(**"server\_and\_requested"**);

68

69 serverRangeValue = serverRange.getLinguisticTerm(**"server\_range\_value"**);

70 requestedRangeValue = requestedRange.getLinguisticTerm(**"requested\_range\_value"**);

71 negotiatedRangeValue = negotiatedRange.getLinguisticTerm(**"negotiated\_range\_value"**);

72 serverValue = serverAndRequested.getLinguisticTerm(**"server\_value"**);

73 requestedValue = serverAndRequested.getLinguisticTerm(**"requested\_value"**);

74

75 requestedRangeMembershipFunction = requestedRangeValue.getMembershipFunction();

76

77 }

78

79 */\*\**

80  *\* Uses JFuzzyLogic to compute the tradeoff between what a server is*

81  *\* willing to offer and what a task requests in terms of resurces usage*

82  *\**

83  *\** ***@param*** *server - the server which can accomodate the task but is not willing to*

84  *\** ***@param*** *task - the task to negotiate it's acomodation on the server*

85  *\** ***@return*** *[negotiated CPU, negotiated Memory, negotiated Storage]*

86  *\* map containing the negotiated values for the CPU, memory and storage*

87  *\*/*

88

89 **public** Map<String, Double> negotiate(Server server, Task task) {

90

91 Map<String, Double> negotiatedValues = **new** HashMap<String, Double>();

92 *//TODO: eventually to reintroduce variables for each of the 3 elements in the file*

93 Collection<Core> cores = server.getAssociatedCPU().getAssociatedCore();

94 Memory memory = server.getAssociatedMemory();

95 Storage storage = server.getAssociatedStorage();

96 RequestedTaskInfo requestedTaskInfo = task.getRequestedInfo();

97

98 **int** requestedCores = requestedTaskInfo.getCores();

99

100 **for** (Core core : cores) {

101

102 **int** maxCPU = core.getMaxAcceptableValue();

103 **int** usedCPU = core.getUsed();

104 **int** totalCPU = core.getTotal();

105 **int** minRequestedCPU = requestedTaskInfo.getCpuMinAcceptableValue();

106 **int** maxRequestedCPU = requestedTaskInfo.getCpuMaxAcceptableValue();

107

108 *//TODO: also negotiate in order to give more resources than needed to the task*

109 *//negotiate CPU max optimum value*

110 **if** ((maxCPU - usedCPU < maxRequestedCPU && minRequestedCPU <= totalCPU - usedCPU)

111 && (maxCPU < totalCPU)) {

112

113 **int**[] values;

114 **int**[] membership = **new int**[]{1, 0};

115 **int** count = 0;

116 Value[] serverValues;

117 Value[] membershipValues;

118

119 *//convert available CPU ranges in Value type*

120 values = **new int**[]{maxCPU, totalCPU};

121 count = values.length;

122 serverValues = **new** Value[count];

123 membershipValues = **new** Value[count];

124 **for** (**int** i = 0; i < count; i++) {

125 serverValues[i] = **new** Value(values[i]);

126 membershipValues[i] = **new** Value(membership[i]);

127 }

128

129 *//create a new membership function for the CPU*

130 MembershipFunction serverCpuMembershipFunction =

131 **new** MembershipFunctionPieceWiseLinear(serverValues, membershipValues);

132

133 serverRangeValue.setMembershipFunction(serverCpuMembershipFunction);

134

135

136 **if** (totalCPU > usedCPU + maxRequestedCPU) {

137 negotiatedRangeValue

138 .getMembershipFunction().setParameter(2, usedCPU + maxRequestedCPU);

139 } **else** {

140 negotiatedRangeValue

141 .getMembershipFunction().setParameter(2, totalCPU);

142 }

143

144 negotiatedRangeValue

145 .getMembershipFunction().setParameter(0, usedCPU + minRequestedCPU);

146 negotiatedRangeValue.getMembershipFunction().setParameter(1, 0);

147

148 negotiatedRangeValue.getMembershipFunction().setParameter(3, 1);

149

150 requestedRangeMembershipFunction.setParameter(0, usedCPU + minRequestedCPU);

151 requestedRangeMembershipFunction.setParameter(2, usedCPU + maxRequestedCPU);

152

153 serverValue.setMembershipFunction(serverCpuMembershipFunction);

154 requestedValue.setMembershipFunction(requestedRangeMembershipFunction);

155

156 *//serverAndRequested.setValue();*

157 serverAndRequested.setUniverseMin(maxCPU);

158 serverAndRequested.setUniverseMax(totalCPU);

159

160 negotiatedRange.setUniverseMin(maxCPU);

161 negotiatedRange.setUniverseMax(totalCPU);

162

163 *//After modifying the output variable membership function a new FIS*

164 *// has to be created in order to evaluate the rules properly*

165 *//Otherwise the value from the FIS creation time is used for*

166 *//deffuzification of the output variable*

167 FIS finalFuzzyInferenceSystem = **null**;

168 **try** {

169 finalFuzzyInferenceSystem = FIS.createFromString(fis.toString(), **true**);

170 } **catch** (RecognitionException e) {

171 System.err.println(e.getMessage());

172 e.printStackTrace();

173 **return null**;

174 }

175

176

177 finalFuzzyInferenceSystem.setVariable(**"server\_range"**, usedCPU + minRequestedCPU + 1);

178 finalFuzzyInferenceSystem.setVariable(**"requested\_range"**, usedCPU + minRequestedCPU + 1);

179

180 finalFuzzyInferenceSystem.evaluate();

181

182 finalFuzzyInferenceSystem

183 .setVariable(**"server\_and\_requested"**, usedCPU + maxRequestedCPU - 1);

184 finalFuzzyInferenceSystem.evaluate();

185

186 **double** negotiatedCPU =

187 finalFuzzyInferenceSystem

188 .getFunctionBlock(**"negotiator"**)

189 .getVariable(**"negotiated\_range"**).getValue();

190 negotiatedValues.put(NEGOTIATED\_CPU, negotiatedCPU - core.getUsed());

191

192 core.setMaxAcceptableValue(

193 (**int**) finalFuzzyInferenceSystem.getFunctionBlock(**"negotiator"**)

194 .getVariable(**"negotiated\_range"**).getValue());

195

196 *//only negotiate for requested number of cores*

197 requestedCores--;

198 **if** (requestedCores == 0) {

199 **break**;

200 }

201 }

202 }

203

204

205 **int** maxMemory = memory.getMaxAcceptableValue();

206 **int** usedMemory = memory.getUsed();

207 **int** totalMemory = memory.getTotal();

208 **int** maxRequestedMemory = requestedTaskInfo.getMemoryMaxAcceptableValue();

209 **int** minRequestedMemory = requestedTaskInfo.getMemoryMinAcceptableValue();

210

211 *//negotiate Memory max optimum value*

212 **if** ((maxMemory - usedMemory < maxRequestedMemory)

213 && (minRequestedMemory <= totalMemory - usedMemory)

214 && (maxMemory < totalMemory)) {

215

216 **int**[] values;

217 **int**[] membership = **new int**[]{1, 0};

218 **int** count = 0;

219 Value[] serverValues;

220 Value[] membershipValues;

221

222 **int** negotiatedMemoryMaxRange =

223 (totalMemory > usedMemory + maxRequestedMemory) ?

224 usedMemory + maxRequestedMemory : totalMemory;

225

226 *//convert available memory ranges in Value type*

227 values = **new int**[]{maxMemory, negotiatedMemoryMaxRange};

228 count = values.length;

229 serverValues = **new** Value[count];

230 membershipValues = **new** Value[count];

231 **for** (**int** i = 0; i < count; i++) {

232 serverValues[i] = **new** Value(values[i]);

233 membershipValues[i] = **new** Value(membership[i]);

234 }

235

236 *//create a new membership function*

237 MembershipFunction serverMemoryMembershipFunction =

238 **new** MembershipFunctionPieceWiseLinear(serverValues, membershipValues);

239

240 *//set the new membership function for the server\_memory variable*

241 serverRangeValue.setMembershipFunction(serverMemoryMembershipFunction);

242 negotiatedRangeValue.setMembershipFunction(serverMemoryMembershipFunction);

243

244 serverRange.setValue(maxMemory + 1);

245 requestedRange.setValue(usedMemory + minRequestedMemory + 1);

246

247 requestedRangeMembershipFunction.setParameter(0, usedMemory + minRequestedMemory);

248 requestedRangeMembershipFunction.setParameter(2, usedMemory + maxRequestedMemory);

249

250 negotiatedRange.setUniverseMin(maxMemory);

251 negotiatedRange.setUniverseMax(totalMemory);

252 serverAndRequested.setUniverseMin(maxMemory);

253 serverAndRequested.setUniverseMax(totalMemory);

254

255 *//After modifying the output variable membership function a*

256 *//new FIS has to be created in order to evaluate the rules properly*

257 *//Otherwise the value from the FIS creation time is used for deffuzification*

258 *//of the output variable*

259 FIS finalFuzzyInferenceSystem = **null**;

260 **try** {

261 finalFuzzyInferenceSystem = FIS.createFromString(fis.toString(), **true**);

262 } **catch** (RecognitionException e) {

263 **for** (**int** i = 0; i < count; i++) {

264 System.out.println(**"Server "** + serverValues[i]

265 + **", Membership "** + membershipValues[i]);

266 }

267 System.err.println(e.getMessage());

268 e.printStackTrace();

269 **return null**;

270 }

271

272

273 finalFuzzyInferenceSystem.setVariable(**"server\_range"**, usedMemory + 1);

274 finalFuzzyInferenceSystem

275 .setVariable(**"requested\_range"**, usedMemory + minRequestedMemory + 1);

276

277

278 finalFuzzyInferenceSystem.evaluate();

279

280 **double** negotiatedMemory =

281 finalFuzzyInferenceSystem

282 .getFunctionBlock(**"negotiator"**).getVariable(**"negotiated\_range"**).getValue();

283 negotiatedValues.put(NEGOTIATED\_MEMORY, negotiatedMemory - memory.getUsed());

284

285 memory.setMaxAcceptableValue(

286 (**int**) finalFuzzyInferenceSystem

287 .getFunctionBlock(**"negotiator"**).getVariable(**"negotiated\_range"**).getValue());

288

289 }

290

291 **int** maxStorage = storage.getMaxAcceptableValue();

292 **int** usedStorage = storage.getUsed();

293 **int** totalStorage = storage.getTotal();

294 **int** minRequestedStorage = requestedTaskInfo.getStorageMinAcceptableValue();

295 **int** maxRequestedStorage = requestedTaskInfo.getStorageMaxAcceptableValue();

296

297 *//negotiate Storage max optimum value*

298

299 **if** ((maxStorage - usedStorage < maxRequestedStorage)

300 && (minRequestedStorage <= totalStorage - usedStorage)

301 && (maxStorage < totalStorage)) {

302

303 **int**[] values;

304 **int**[] membership = **new int**[]{1, 0};

305 **int** count = 0;

306 Value[] serverValues;

307 Value[] membershipValues;

308

309 **int** negotiatedStorageMaxRange =

310 (totalStorage > usedStorage + maxRequestedStorage) ?

311 usedStorage + maxRequestedStorage : totalStorage;

312

313 *//convert available Storage ranges in Value type*

314 values = **new int**[]{maxStorage, negotiatedStorageMaxRange};

315 count = values.length;

316 serverValues = **new** Value[count];

317 membershipValues = **new** Value[count];

318 **for** (**int** i = 0; i < count; i++) {

319 serverValues[i] = **new** Value(values[i]);

320 membershipValues[i] = **new** Value(membership[i]);

321 }

322

323 *//create a new membership function*

324

325 MembershipFunction serverStorageMembershipFunction =

326 **new** MembershipFunctionPieceWiseLinear(serverValues, membershipValues);

327

328 *//set the new membership function for the server\_storage variable*

329 serverRangeValue.setMembershipFunction(serverStorageMembershipFunction);

330 negotiatedRangeValue.setMembershipFunction(serverStorageMembershipFunction);

331

332 requestedRangeMembershipFunction.setParameter(0, usedStorage + minRequestedStorage);

333 requestedRangeMembershipFunction.setParameter(2, usedStorage + maxRequestedStorage);

334

335 negotiatedRange.setUniverseMin(maxStorage);

336 negotiatedRange.setUniverseMax(totalStorage);

337

338 serverAndRequested.setUniverseMin(maxStorage);

339 serverAndRequested.setUniverseMax(totalStorage);

340

341 *//After modifying the output variable's membership function*

342 *//a new FIS has to be created in order to evaluate the rules properly*

343 *//Otherwise the value from the FIS creation time is used for*

344 *// deffuzification of the output variable*

345 FIS finalFuzzyInferenceSystem = **null**;

346 **try** {

347 finalFuzzyInferenceSystem = FIS.createFromString(fis.toString(), **true**);

348 } **catch** (RecognitionException e) {

349 System.err.println(e.getMessage());

350 e.printStackTrace();

351 **return null**;

352 }

353

354

355 finalFuzzyInferenceSystem.setVariable(**"server\_range"**, usedStorage + 1);

356 finalFuzzyInferenceSystem

357 .setVariable(**"requested\_range"**, usedStorage + minRequestedStorage + 1);

358

359 finalFuzzyInferenceSystem.evaluate();

360

361 **double** negotiatedStorage =

362 finalFuzzyInferenceSystem

363 .getFunctionBlock(**"negotiator"**)

364 .getVariable(**"negotiated\_range"**).getValue();

365 negotiatedValues.put(NEGOTIATED\_STORAGE, negotiatedStorage - storage.getUsed());

366 System.out.println(**"Negotiated for "** + storage.getLocalName()

367 + **" from "** + storage.getMaxAcceptableValue() +

368 **" to "** + negotiatedStorage);

369 storage.setMaxAcceptableValue(

370 (**int**) finalFuzzyInferenceSystem

371 .getFunctionBlock(**"negotiator"**)

372 .getVariable(**"negotiated\_range"**).getValue());

373 }

374

375 **return** negotiatedValues;

376 }

377 }

|  |
| --- |
| ServerInfoSAXHandler.java |

1 **package** contextawaremodel.worldInterface.datacenterInterface.xmlParsers;

2

3 **import** contextawaremodel.worldInterface.dtos.ServerDto;

4 **import** contextawaremodel.worldInterface.dtos.StorageDto;

5 **import** org.xml.sax.Attributes;

6 **import** org.xml.sax.SAXException;

7 **import** org.xml.sax.helpers.DefaultHandler;

8

9 **import** java.text.NumberFormat;

10 **import** java.text.ParseException;

11 **import** java.util.ArrayList;

12 **import** java.util.List;

13

14 */\*\**

15  *\* Class used to parse the XML returned by the Server Monitor C# endpoint*

16  *\* and construct a ServerDto object from that information*

17  *\*/*

18 **public class** ServerInfoSAXHandler **extends** DefaultHandler {

19 */\**

20  *XML Example*

21  *<ServerInfo xmlns="http://www.SelfOptimizingDatacenter.edu/">*

22  *<TotalCPU>int</TotalCPU>*

23  *<CoreCount>int</CoreCount>*

24  *<FreeCPU>*

25  *<int>int</int>*

26  *<int>int</int>*

27  *</FreeCPU>*

28  *<Storage>*

29  *<Storage>*

30  *<Name>string</Name>*

31  *<Size>int</Size>*

32  *<FreeSpace>int</FreeSpace>*

33  *</Storage>*

34  *<Storage>*

35  *<Name>string</Name>*

36  *<Size>int</Size>*

37  *<FreeSpace>int</FreeSpace>*

38  *</Storage>*

39  *</Storage>*

40  *<TotalMemory>int</TotalMemory>*

41  *<FreeMemory>int</FreeMemory>*

42  *</ServerInfo>*

43  *\*/*

44 **private static final** String TOTAL\_CPU = **"TotalCPU"**;

45 **private static final** String CORE\_COUNT = **"CoreCount"**;

46

47 **private static final** String FREE\_CPU = **"FreeCPU"**;

48 **private static final** String FREE\_CPU\_VAL = **"int"**;

49

50 **private static final** String STORAGE = **"Storage"**;

51 **private static final** String STORAGE\_NAME = **"Name"**;

52 **private static final** String STORAGE\_SIZE = **"Size"**;

53 **private static final** String STORAGE\_FREE\_SPACE = **"FreeSpace"**;

54

55 **private static final** String TOTAL\_MEMORY = **"TotalMemory"**;

56 **private static final** String FREE\_MEMORY = **"FreeMemory"**;

57

58 **private boolean** inTotalCPU = **false**;

59 **private boolean** inCoreCount = **false**;

60 **private boolean** inFreeCPUVal = **false**;

61

62 **private boolean** inStorageName = **false**;

63 **private boolean** inStorageSize = **false**;

64 **private boolean** inStorageFreeSpace = **false**;

65

66 **private boolean** inTotalMemory = **false**;

67 **private boolean** inFreeMemory = **false**;

68

69 **private** String text;

70

71 **private final** NumberFormat numberFormat = NumberFormat.getIntegerInstance();

72

73 ServerDto serverDto;

74 List<Integer> freeCPUValues;

75 List<StorageDto> storageInfo;

76 StorageDto currentStorage;

77

78 **public** ServerInfoSAXHandler() {

79 serverDto = **new** ServerDto();

80 freeCPUValues = **new** ArrayList<Integer>();

81 storageInfo = **new** ArrayList<StorageDto>();

82 text = **""**;

83 }

84

85 **public** ServerDto getServerDto() {

86 serverDto.setStorage(storageInfo);

87 **return** serverDto;

88 }

89

90 @Override

91 **public void** startElement(String uri, String localName, String qName, Attributes attributes)

92 **throws** SAXException {

93

94 **if** (localName.equals(TOTAL\_CPU)) {

95 inTotalCPU = **true**;

96 } **else if** (localName.equals(CORE\_COUNT)) {

97 inCoreCount = **true**;

98 } **else if** (localName.equals(FREE\_CPU)) {

99 } **else if** (localName.equals(STORAGE)) {

100 currentStorage = **new** StorageDto();

101 } **else if** (localName.equals(TOTAL\_MEMORY)) {

102 inTotalMemory = **true**;

103 } **else if** (localName.equals(FREE\_MEMORY)) {

104 inFreeMemory = **true**;

105 } **else if** (localName.equals(FREE\_CPU\_VAL)) {

106 inFreeCPUVal = **true**;

107 } **else if** (localName.equals(STORAGE\_NAME)) {

108 inStorageName = **true**;

109 } **else if** (localName.equals(STORAGE\_SIZE)) {

110 inStorageSize = **true**;

111 } **else if** (localName.equals(STORAGE\_FREE\_SPACE)) {

112 inStorageFreeSpace = **true**;

113 }

114 }

115

116 **public void** endElement(String namespaceURI, String localName, String qualifiedName)

117 **throws** SAXException {

118

119 **if** (localName.equals(TOTAL\_CPU)) {

120 **try** {

121 serverDto.setTotalCPU(numberFormat.parse(text).intValue());

122 } **catch** (ParseException e) {

123 e.printStackTrace();

124 }

125 inTotalCPU = **false**;

126 } **else if** (localName.equals(CORE\_COUNT)) {

127 **try** {

128 serverDto.setCoreCount(numberFormat.parse(text).intValue());

129 } **catch** (ParseException e) {

130 e.printStackTrace();

131 }

132 inCoreCount = **false**;

133 } **else if** (localName.equals(FREE\_CPU)) {

134 serverDto.setFreeCPU(freeCPUValues);

135 } **else if** (localName.equals(FREE\_CPU\_VAL)) {

136 **try** {

137 freeCPUValues.add(numberFormat.parse(text).intValue());

138 } **catch** (ParseException e) {

139 e.printStackTrace();

140 }

141 inFreeCPUVal = **false**;

142 } **else if** (localName.equals(STORAGE)) {

143 storageInfo.add(currentStorage);

144 } **else if** (localName.equals(TOTAL\_MEMORY)) {

145 **try** {

146 serverDto.setTotalMemory(numberFormat.parse(text).intValue());

147 } **catch** (ParseException e) {

148 e.printStackTrace();

149 }

150 inTotalMemory = **false**;

151 } **else if** (localName.equals(FREE\_MEMORY)) {

152 **try** {

153 serverDto.setFreeMemory(numberFormat.parse(text).intValue());

154 } **catch** (ParseException e) {

155 e.printStackTrace();

156 }

157 inFreeMemory = **false**;

158 } **else if** (localName.equals(STORAGE\_NAME)) {

159 inStorageName = **false**;

160 currentStorage.setName(text);

161 } **else if** (localName.equals(STORAGE\_SIZE)) {

162

163 **try** {

164 currentStorage.setSize(numberFormat.parse(text).intValue());

165 } **catch** (ParseException e) {

166 e.printStackTrace();

167 }

168 inStorageSize = **false**;

169 } **else if** (localName.equals(STORAGE\_FREE\_SPACE)) {

170 **try** {

171 currentStorage.setFreeSpace(numberFormat.parse(text).intValue());

172 } **catch** (ParseException e) {

173 e.printStackTrace();

174 }

175 inStorageFreeSpace = **false**;

176 }

177 }

178

179 **public void** characters(**char**[] ch, **int** start, **int** length) **throws** SAXException {

180

181 **if** (inTotalCPU || inCoreCount

182 || inFreeCPUVal || inTotalMemory || inFreeMemory

183 || inStorageName || inStorageFreeSpace || inStorageSize) {

184 text = **new** String(ch, start, length);

185 }

186 }

187 }

188

|  |
| --- |
| ResourceMonitorPlotter.java |

1 **package** contextawaremodel.gui.resourceMonitor.resourceMonitorPlotter;

2

3 **import** javax.swing.\*;

4 **import** java.awt.\*;

5

6 */\*\**

7  *\* Abstract class used as base class for any graphical resource monitor*

8  *\*/*

9 **public abstract class** ResourceMonitorPlotter {

10

11 **protected** String resourceName;

12

13 **protected static final** Font LABEL\_FONT = **new** Font(**"sherif"**, Font.BOLD, 15);

14

15 **protected** JPanel graphPanel;

16

17 **protected** ResourceMonitorPlotter(String resourceName) {

18 **this**.resourceName = resourceName;

19 graphPanel = **new** JPanel();

20 }

21

22 **public abstract void** setCurrentValue(Object currentValue);

23

24 **protected abstract void** setup(**int** minimumValue, **int** maximumValue);

25

26 **public** JPanel getGraphPanel() {

27 **return** graphPanel;

28 }

29 }

30

|  |
| --- |
| ResourceMonitorPieChartPlotter.java |

1 **package** contextawaremodel.gui.resourceMonitor.resourceMonitorPlotter.impl;

2

3 **import** contextawaremodel.gui.resourceMonitor.resourceMonitorPlotter.ResourceMonitorPlotter;

4 **import** org.jfree.chart.ChartFactory;

5 **import** org.jfree.chart.ChartPanel;

6 **import** org.jfree.chart.JFreeChart;

7 **import** org.jfree.chart.plot.PiePlot;

8 **import** org.jfree.data.general.DefaultPieDataset;

9

10 **import** java.awt.\*;

11 **import** java.util.Map;

12 **import** java.util.Random;

13

14 */\*\**

15  *\* Pie chart used for displaying the resources used by each task*

16  *\* on a server together with the OS used and Free amount under the form of pie slices*

17  *\*/*

18 **public class** ResourceMonitorPieChartPlotter **extends** ResourceMonitorPlotter {

19

20 **private** JFreeChart chart;

21 **private** Random random;

22 **private int** oldDatasetSize = 0;

23

24 **public** ResourceMonitorPieChartPlotter(String resourceName, **int** minimumValue, **int** maximumValue) {

25 **super**(resourceName);

26 random = **new** Random();

27 setup(minimumValue, maximumValue);

28

29 }

30

31 */\*\**

32  *\** ***@param*** *currentValue Map<Integer,String> representing <value,label> pairs*

33  *\*/*

34 @Override

35 **public void** setCurrentValue(Object currentValue) {

36 DefaultPieDataset dataset = **new** DefaultPieDataset();

37

38 Map<String, Integer> values = (Map<String, Integer>) currentValue;

39 **for** (Map.Entry<String, Integer> entry : values.entrySet()) {

40 dataset.setValue(entry.getKey(), entry.getValue());

41 }

42

43 PiePlot plot = (PiePlot) chart.getPlot();

44 plot.setDataset(dataset);

45

46 **if** (values.size() == oldDatasetSize) {

47 **return**;

48 }

49

50 oldDatasetSize = values.size();

51

52 plot.setSectionPaint(dataset.getIndex(**"Free"**), Color.BLUE);

53 plot.setSectionPaint(dataset.getIndex(**"OS"**), Color.BLACK);

54

55 **for** (**int** i = 0; i < values.size(); i++) {

56 plot.setExplodePercent(i, 0.025);

57 }

58

59 */\*\**

60  *\* Find the representation pie slices different than OS and Free*

61  *\* and assign random colours*

62  *\*/*

63 **for** (Object key : dataset.getKeys()) {

64 **if** (!key.equals(**"Free"**) && !key.equals(**"OS"**)) {

65 plot.setSectionPaint(dataset.getIndex((String) key),

66 **new** Color(25 + random.nextInt(230), 25 + random.nextInt(230), 25));

67 }

68 }

69 }

70

71 @Override

72 **protected void** setup(**int** minimumValue, **int** maximumValue) {

73 DefaultPieDataset dataset = **new** DefaultPieDataset();

74 chart = ChartFactory.createPieChart(

75 resourceName, *// chart title*

76 dataset, *// data*

77 **false**, *// include legend*

78 **true**,

79 **false**

80 );

81

82 PiePlot plot = (PiePlot) chart.getPlot();

83 plot.setLabelFont(LABEL\_FONT);

84 plot.setNoDataMessage(**"No data available"**);

85 graphPanel = **new** ChartPanel(chart);

86 }

87 }

88

|  |
| --- |
| X3DAgent.java |

1 */\**

2  *\* Agent used for creating 3D context representations*

3  *\*/*

4 **package** contextawaremodel.agents;

5

6 **import** contextawaremodel.GlobalVars;

7 **import** contextawaremodel.agents.behaviours.BasicX3DBehaviour;

8 **import** jade.core.Agent;

9 **import** org.web3d.x3d.sai.\*;

10

11 **import** javax.swing.\*;

12 **import** java.awt.\*;

13 **import** java.awt.event.ActionEvent;

14 **import** java.awt.event.ActionListener;

15 **import** java.util.ArrayList;

16 **import** java.util.HashMap;

17 **import** java.util.Map;

18

19 **public class** X3DAgent **extends** Agent {

20

21 **public static final** String ADD\_TASK\_COMMAND = **"addTask"**;

22 **public static final** String REMOVE\_TASK\_COMMAND = **"removeTask"**;

23 **public static final** String MOVE\_TASK\_COMMAND = **"moveTask"**;

24 **public static final** String SEND\_SERVER\_TO\_LOW\_POWER\_COMMAND = **"sendToLowPower"**;

25 **public static final** String WAKE\_UP\_SERVER\_COMMAND = **"wakeUpServer"**;

26 **public static final** String SET\_TEMPERATURE\_COMMAND = **"setTemperature"**;

27 **public static final** String SET\_HUMIDITY\_COMMAND = **"setHumidity"**;

28

29 **public static final float** POWER\_METER\_LABEL\_TRANSLATION = 0.7f;

30 **public static final float** SENSOR\_LABEL\_TRANSLATION = 0.4f;

31

32 **public static final float**[] SENSOR\_LABEL\_COLOR = **new float**[]{0.5f, 0, 0};

33 **public static final float**[] TASK\_LABEL\_COLOR = **new float**[]{0, 1, 0};

34 **public static final float**[] POWER\_LABEL\_COLOR = **new float**[]{0, 0, 0.6f};

35 **public static final float**[] ACTIVE\_WIRE\_COLOR = **new float**[]{0, 0.6f, 0};

36 **public static final float**[] ACTIVE\_SERVER\_COLOR = **new float**[]{0.3451f, 0.7804f, 0.8824f};

37 **public static final float**[] INACTIVE\_SERVER\_COLOR = **new float**[]{0.5f, 0.5f, 0.5f};

38

39

40 **private** JFrame frame;

41

42 **private** X3DScene mainScene;

43 **private** ArrayList<X3DNode> taskLabels;

44 **private** Map<String, X3DNode> tasks;

45 **private** Map<String, MFString> objectLabels;

46 **private** ArrayList<String> wiresIndexes;

47

48 **private** Timer changeWiresBackTimer;

49 **private** Timer fanAnimationTimer;

50 **private** Timer sensorAnimationTimer;

51 **private float** fanIncrement = 0.1f;

52 **private** ExternalBrowser x3dBrowser;

53 **private int** activeServersNo = 6;

54

55 **private** X3DNode attentionArow;

56 **private** X3DNode inverseAttentionArow;

57

58 **private** ActionListener sensorsAnimationListener = **new** ActionListener() {

59

60 **public void** actionPerformed(ActionEvent e) {

61 X3DNode sensor\_1 = mainScene.getNamedNode(**"SensorTube\_01\_XFORM"**);

62 X3DNode sensor\_2 = mainScene.getNamedNode(**"SensorTube\_02\_XFORM"**);

63 SFRotation rotation\_1 = (SFRotation) sensor\_1.getField(**"rotation"**);

64 SFRotation rotation\_2 = (SFRotation) sensor\_2.getField(**"rotation"**);

65 **float**[] values = **new float**[4];

66 rotation\_1.getValue(values);

67 values[3] += 0.1f;

68 rotation\_1.setValue(values);

69 **float** temp = values[3] \* -1;

70 rotation\_2.getValue(values);

71 values[3] = temp;

72 rotation\_2.setValue(values);

73 }

74 };

75

76

77 **private** ActionListener fanAnimationListener = **new** ActionListener() {

78

79 **public void** actionPerformed(ActionEvent e) {

80 X3DNode fan = mainScene.getNamedNode(**"Fan\_XFORM"**);

81 SFRotation rotation = (SFRotation) fan.getField(**"rotation"**);

82 **float**[] values = **new float**[4];

83 rotation.getValue(values);

84 values[3] += fanIncrement;

85 rotation.setValue(values);

86 }

87 };

88

89 **public void** setFanSpeed(**float** speed) {

90 fanIncrement = speed;

91 }

92

93

94 @Override

95 **protected void** takeDown() {

96 **super**.takeDown();

97 frame.setVisible(**false**);

98 frame.dispose();

99 }

100

101 */\*\**

102  *\* Method called when the agent is created.*

103  *\*/*

104 @Override

105 **protected void** setup() {

106

107 System.out.println(**"[X3DAgent] : Hellooo ! "**);

108 frame = **new** JFrame(**"X3D vizualization"**);

109 frame.setDefaultCloseOperation(JFrame.DISPOSE\_ON\_CLOSE);

110 Container contentPane = frame.getContentPane();

111

112 *// Setup browser parameters*

113 HashMap requestedParameters = **new** HashMap();

114 requestedParameters.put(**"Xj3D\_ShowConsole"**, Boolean.FALSE);

115 requestedParameters.put(**"Xj3D\_FPSShown"**, Boolean.TRUE);

116 requestedParameters.put(**"Xj3D\_NavbarShown"**, Boolean.TRUE);

117

118 *// Create an SAI component*

119 X3DComponent x3dComp = BrowserFactory.createX3DComponent(requestedParameters);

120

121 taskLabels = **new** ArrayList<X3DNode>();

122 objectLabels = **new** HashMap<String, MFString>();

123 tasks = **new** HashMap<String, X3DNode>();

124 wiresIndexes = **new** ArrayList<String>();

125

126 frame.setSize(500, 500);

127 frame.setVisible(**true**);

128

129 *// Add the component to the UI*

130 JComponent x3dPanel = (JComponent) x3dComp.getImplementation();

131 contentPane.add(x3dPanel, BorderLayout.CENTER);

132

133 *// Get an external browser*

134 x3dBrowser = x3dComp.getBrowser();

135

136 **boolean** useXj3D = **true**;

137

138 **if** (x3dBrowser.getName().indexOf(**"Xj3D"**) < 0) {

139 System.out.println(**"Not running on Xj3D, extended functions disabled"**);

140 useXj3D = **false**;

141 }

142

143

144 *// Create an X3D scene by loading a file*

145 mainScene = x3dBrowser.createX3DFromURL(**new** String[]{GlobalVars.X3D\_SCENE\_FILE});

146

147 **if** (mainScene == **null**) {

148 System.err.println(**"X3D file "** + GlobalVars.X3D\_SCENE\_FILE + **" not found "**);

149 **return**;

150 }

151

152 *// Replace the current world with the new one*

153 x3dBrowser.replaceWorld(mainScene);

154

155 **if** (!useXj3D) {

156 **return**;

157 }

158

159 attentionArow = mainScene.getNamedNode(**"AttentionArrow\_XFORM"**);

160 inverseAttentionArow = mainScene.getNamedNode(**"InverseAttentionArrow\_XFORM"**);

161

162 *//spin the UTCN crest*

163 ActionListener actionListener = **new** ActionListener() {

164

165 **public void** actionPerformed(ActionEvent e) {

166 X3DNode sigla = mainScene.getNamedNode(**"Sigla\_UTCN\_XFORM"**);

167 X3DNode tube = mainScene.getNamedNode(**"Server\_0\_tube\_XFORM"**);

168

169 SFRotation rotation = (SFRotation) sigla.getField(**"rotation"**);

170 SFRotation rotationTube = (SFRotation) tube.getField(**"rotation"**);

171

172 **float**[] values = **new float**[4];

173 rotation.getValue(values);

174 values[3] += 0.1;

175

176 rotation.setValue(values);

177 rotationTube.setValue(values);

178 }

179 };

180

181

182 Timer timer = **new** Timer(50, actionListener);

183 timer.start();

184

185 **for** (**int** i = 1; i <= 5; i++) {

186 addObjectLabel(**"Watts: "** + 1 \* 10, **"PowerMeterGroup\_0"** + i,

187 POWER\_LABEL\_COLOR, POWER\_METER\_LABEL\_TRANSLATION);

188 }

189

190

191 fanAnimationTimer = **new** Timer(100, fanAnimationListener);

192

193 sensorAnimationTimer = **new** Timer(50, sensorsAnimationListener);

194 sensorAnimationTimer.start();

195

196 addObjectLabel(**"Temperature : 32"**, **"SensorSphere\_01"**,

197 SENSOR\_LABEL\_COLOR, SENSOR\_LABEL\_TRANSLATION);

198 addObjectLabel(**"Humidity : 21"**, **"SensorSphere\_02"**,

199 SENSOR\_LABEL\_COLOR, SENSOR\_LABEL\_TRANSLATION);

200

201 **final float**[] initialWireColor = **new float**[]{0.8784f, 0.5608f, 0.3412f};

202

203 ActionListener setWiresColorBack = **new** ActionListener() {

204

205 **public void** actionPerformed(ActionEvent e) {

206 **for** (String index : wiresIndexes) {

207 setWireColor(index, initialWireColor);

208 }

209 }

210 };

211

212 changeWiresBackTimer = **new** Timer(400, setWiresColorBack);

213 changeWiresBackTimer.start();

214

215 addBehaviour(**new** BasicX3DBehaviour(**this**));

216

217 }

218

219 */\*\**

220  *\* Sets the color of the wire associated to server serverName to*

221  *\* the specified color*

222  *\** ***@param*** *serverName the server the wire is connected to*

223  *\** ***@param*** *color the new color of the wire*

224  *\*/*

225 **private void** setWireColor(String serverName, **float**[] color) {

226 String[] elements = serverName.split(**"\_"**);

227 X3DNode material =

228 mainScene.getNamedNode(**"Wire\_0"** + Integer.parseInt(elements[1]) + **"\_MAT"**);

229 SFColor diffuseColor = (SFColor) material.getField(**"diffuseColor"**);

230 diffuseColor.setValue(color);

231 }

232

233 */\*\**

234  *\* Adds a text label to the other side of the object. The side initially not visible*

235  *\* when the scene is loaded.*

236  *\** ***@param*** *textLabel the label of the object*

237  *\** ***@param*** *objectName the name of the X3D node to which the label is applied*

238  *\** ***@param*** *color the color of the label*

239  *\** ***@param*** *translation the translation of the text relative to the object*

240  *\*/*

241 **private void** addInverseObjectLabel(String textLabel, String objectName,

242 **float**[] color, **float** translation) {

243

244 **if** (!objectLabels.containsKey(objectName + **"\_Inverse"**)) {

245

246 X3DNode transform = mainScene.createNode(**"Transform"**);

247

248 X3DNode shape = mainScene.createNode(**"Shape"**);

249 X3DNode label = mainScene.createNode(**"Text"**);

250 X3DNode appearance = mainScene.createNode(**"Appearance"**);

251 X3DNode material = mainScene.createNode(**"Material"**);

252 X3DNode fontStyle = mainScene.createNode(**"FontStyle"**);

253

254 MFString justify = (MFString) fontStyle.getField(**"justify"**);

255 SFFloat size = (SFFloat) fontStyle.getField(**"size"**);

256

257 justify.set1Value(0, **"MIDDLE"**);

258 justify.set1Value(1, **"MIDDLE"**);

259 size.setValue(0.15f);

260 SFNode fontStyleAttribute = (SFNode) label.getField(**"fontStyle"**);

261 fontStyleAttribute.setValue(fontStyle);

262

263 SFColor diffuseColor = (SFColor) material.getField(**"diffuseColor"**);

264 SFFloat ambientIntensity = (SFFloat) material.getField(**"ambientIntensity"**);

265

266 diffuseColor.setValue(color);

267 ambientIntensity.setValue(1);

268

269 SFNode appearanceMaterial = (SFNode) appearance.getField(**"material"**);

270 appearanceMaterial.setValue(material);

271

272 SFNode shapeAppearance = (SFNode) shape.getField(**"appearance"**);

273 shapeAppearance.setValue(appearance);

274

275 MFString string = (MFString) label.getField(**"string"**);

276 string.clear();

277 string.insertValue(0, textLabel);

278

279 SFNode shapeGeometry = (SFNode) shape.getField(**"geometry"**);

280 shapeGeometry.setValue(label);

281

282 MFNode newTransformChildren = (MFNode) transform.getField(**"children"**);

283 newTransformChildren.append(shape);

284

285 SFVec3f serverTranslation = (SFVec3f) transform.getField(**"translation"**);

286 **float**[] translationValues = **new float**[3];

287

288 X3DNode powerMeter = mainScene.getNamedNode(objectName + **"\_XFORM"**);

289 SFVec3f powerMeterTranslation = (SFVec3f) powerMeter.getField(**"translation"**);

290 powerMeterTranslation.getValue(translationValues);

291 translationValues[1] += 0.2;

292 translationValues[0] -= translation;

293

294 serverTranslation.setValue(translationValues);

295

296 SFRotation serverRotation = (SFRotation) transform.getField(**"rotation"**);

297 **float**[] rotationValues = **new float**[]{0, 1, 0, 0f};

298 serverRotation.setValue(rotationValues);

299 mainScene.addRootNode(transform);

300 objectLabels.put(objectName + **"\_Inverse"**, string);

301 } **else** {

302 MFString string = objectLabels.get(objectName + **"\_Inverse"**);

303 *//string.clear();*

304 **if** (string.size() == 0) {

305 string.insertValue(0, textLabel);

306 } **else** {

307 string.set1Value(0, textLabel);

308 }

309 }

310 }

311

312 */\*\**

313  *\* Adds a text label to the visible side of the object. The side initially visible*

314  *\* when the scene is loaded.*

315  *\** ***@param*** *textLabel the label of the object*

316  *\** ***@param*** *objectName the name of the X3D node to which the label is applied*

317  *\** ***@param*** *color the color of the label*

318  *\** ***@param*** *translation the translation of the text relative to the object*

319  *\*/*

320 **public void** addObjectLabel(String textLabel, String objectName,

321 **float**[] color, **float** translation) {

322

323 **if** (!objectLabels.containsKey(objectName)) {

324 X3DNode transform = mainScene.createNode(**"Transform"**);

325 X3DNode shape = mainScene.createNode(**"Shape"**);

326 X3DNode label = mainScene.createNode(**"Text"**);

327 X3DNode appearance = mainScene.createNode(**"Appearance"**);

328 X3DNode material = mainScene.createNode(**"Material"**);

329 X3DNode fontStyle = mainScene.createNode(**"FontStyle"**);

330

331 MFString justify = (MFString) fontStyle.getField(**"justify"**);

332 SFFloat size = (SFFloat) fontStyle.getField(**"size"**);

333

334 justify.set1Value(0, **"MIDDLE"**);

335 justify.set1Value(1, **"MIDDLE"**);

336 size.setValue(0.15f);

337 SFNode fontStyleAttribute = (SFNode) label.getField(**"fontStyle"**);

338 fontStyleAttribute.setValue(fontStyle);

339

340 SFColor diffuseColor = (SFColor) material.getField(**"diffuseColor"**);

341 SFFloat ambientIntensity = (SFFloat) material.getField(**"ambientIntensity"**);

342

343 diffuseColor.setValue(color);

344 ambientIntensity.setValue(1);

345

346 SFNode appearanceMaterial = (SFNode) appearance.getField(**"material"**);

347 appearanceMaterial.setValue(material);

348

349 SFNode shapeAppearance = (SFNode) shape.getField(**"appearance"**);

350 shapeAppearance.setValue(appearance);

351

352 MFString string = (MFString) label.getField(**"string"**);

353 string.clear();

354 string.insertValue(0, textLabel);

355

356 SFNode shapeGeometry = (SFNode) shape.getField(**"geometry"**);

357 shapeGeometry.setValue(label);

358

359 MFNode newTransformChildren = (MFNode) transform.getField(**"children"**);

360 newTransformChildren.append(shape);

361

362 SFVec3f serverTranslation = (SFVec3f) transform.getField(**"translation"**);

363 **float**[] translationValues = **new float**[3];

364

365 X3DNode powerMeter = mainScene.getNamedNode(objectName + **"\_XFORM"**);

366 SFVec3f powerMeterTranslation = (SFVec3f) powerMeter.getField(**"translation"**);

367 powerMeterTranslation.getValue(translationValues);

368 translationValues[1] += 0.2;

369 translationValues[0] += translation;

370

371 serverTranslation.setValue(translationValues);

372

373 SFRotation serverRotation = (SFRotation) transform.getField(**"rotation"**);

374 **float**[] rotationValues = **new float**[]{0, 1, 0, 3f};

375 serverRotation.setValue(rotationValues);

376

377 mainScene.addRootNode(transform);

378 objectLabels.put(objectName, string);

379 addInverseObjectLabel(textLabel, objectName, color, translation);

380 } **else** {

381 MFString string = objectLabels.get(objectName);

382 **if** (string.size() == 0) {

383 string.insertValue(0, textLabel);

384 } **else** {

385 string.set1Value(0, textLabel);

386 }

387

388 addInverseObjectLabel(textLabel, objectName, color, translation);

389 }

390 }

391

392 */\*\**

393  *\**

394  *\** ***@param*** *taskName the name of the task to be used as label*

395  *\** ***@param*** *taskTransform the transform of the task. Each task is placed above the*

396  *\* previous task so it is necesary to know the transform off the target task*

397  *\** ***@param*** *color the color of the label*

398  *\*/*

399 **private void** addLabelToTask(String taskName, X3DNode taskTransform, **float**[] color) {

400

401 X3DNode transform = mainScene.createNode(**"Transform"**);

402 X3DNode shape = mainScene.createNode(**"Shape"**);

403 X3DNode label = mainScene.createNode(**"Text"**);

404 X3DNode appearance = mainScene.createNode(**"Appearance"**);

405 X3DNode material = mainScene.createNode(**"Material"**);

406 X3DNode fontStyle = mainScene.createNode(**"FontStyle"**);

407

408 MFString justify = (MFString) fontStyle.getField(**"justify"**);

409 SFFloat size = (SFFloat) fontStyle.getField(**"size"**);

410

411 justify.set1Value(0, **"MIDDLE"**);

412 justify.set1Value(1, **"MIDDLE"**);

413 size.setValue(0.15f);

414 SFNode fontStyleAttribute = (SFNode) label.getField(**"fontStyle"**);

415 fontStyleAttribute.setValue(fontStyle);

416

417 SFColor diffuseColor = (SFColor) material.getField(**"diffuseColor"**);

418 SFFloat ambientIntensity = (SFFloat) material.getField(**"ambientIntensity"**);

419

420 diffuseColor.setValue(color);

421 ambientIntensity.setValue(1);

422

423 SFNode appearanceMaterial = (SFNode) appearance.getField(**"material"**);

424 appearanceMaterial.setValue(material);

425

426 SFNode shapeAppearance = (SFNode) shape.getField(**"appearance"**);

427 shapeAppearance.setValue(appearance);

428

429 MFString string = (MFString) label.getField(**"string"**);

430 string.clear();

431 string.insertValue(0, taskName);

432

433 SFNode shapeGeometry = (SFNode) shape.getField(**"geometry"**);

434 shapeGeometry.setValue(label);

435

436 MFNode newTransformChildren = (MFNode) transform.getField(**"children"**);

437 newTransformChildren.append(shape);

438

439

440 SFRotation serverRotation = (SFRotation) transform.getField(**"rotation"**);

441 **float**[] rotationValues = **new float**[]{0, 1, 0, 3};

442 serverRotation.setValue(rotationValues);

443

444 SFVec3f serverTranslation = (SFVec3f) transform.getField(**"translation"**);

445 **float**[] translationValues = **new float**[3];

446

447 translationValues[1] += 0.1;

448 translationValues[0] += 0.15;

449 translationValues[2] -= 0.15;

450 serverTranslation.setValue(translationValues);

451

452 MFNode taskTransformChildren = (MFNode) taskTransform.getField(**"children"**);

453 taskTransformChildren.append(transform);

454 mainScene.addRootNode(taskTransform);

455 taskLabels.add(taskTransform);

456 }

457

458 */\*\**

459  *\* Creates a new task representation and adds it to the 3D scene*

460  *\** ***@param*** *taskName the name of the new task to be used as label*

461  *\** ***@param*** *serverName*

462  *\** ***@param*** *taskNumber*

463  *\*/*

464 **public void** addTask(String taskName, String serverName, **int** taskNumber) {

465

466 wiresIndexes.add(serverName);

467 setWireColor(serverName, ACTIVE\_WIRE\_COLOR);

468 changeWiresBackTimer.restart();

469

470 X3DNode newShape = mainScene.createNode(**"Shape"**);

471 X3DNode newBox = mainScene.createNode(**"Box"**);

472 X3DNode newTransform = mainScene.createNode(**"Transform"**);

473 X3DNode appearance = mainScene.createNode(**"Appearance"**);

474 X3DNode material = mainScene.createNode(**"Material"**);

475

476 SFNode appearanceField = (SFNode) newShape.getField(**"appearance"**);

477 SFVec3f boxSize = (SFVec3f) newBox.getField(**"size"**);

478 boxSize.setValue(**new float**[]{0.2699f, 0.08436f, 0.2801f});

479

480 SFColor diffuseColor = (SFColor) material.getField(**"diffuseColor"**);

481 SFFloat ambientIntensity = (SFFloat) material.getField(**"ambientIntensity"**);

482

483 diffuseColor.setValue(**new float**[]{0.5529f, 0.02745f, 0.2275f});

484 ambientIntensity.setValue(1);

485

486 SFNode appearanceMaterial = (SFNode) appearance.getField(**"material"**);

487 appearanceMaterial.setValue(material);

488 appearanceField.setValue(appearance);

489

490 SFNode shape\_geometry = (SFNode) (newShape.getField(**"geometry"**));

491 shape\_geometry.setValue(newBox);

492 MFNode newTransformChildren = (MFNode) newTransform.getField(**"children"**);

493 newTransformChildren.append(newShape);

494

495 SFVec3f translation = (SFVec3f) newTransform.getField(**"translation"**);

496

497 X3DNode serverTransform = mainScene.getNamedNode(serverName + **"\_XFORM"**);

498 SFVec3f serverTranslation = (SFVec3f) serverTransform.getField(**"translation"**);

499 **float**[] translationValues = **new float**[3];

500

501 serverTranslation.getValue(translationValues);

502 **float** heightIndex = -0.2f;

503 heightIndex += 0.1f \* (taskNumber - 1);

504

505 translationValues[1] += heightIndex;

506 translation.setValue(translationValues);

507

508 SFRotation newTransformRotation = (SFRotation) newTransform.getField(**"rotation"**);

509 SFRotation serverRotation = (SFRotation) serverTransform.getField(**"rotation"**);

510 **float**[] rotationValues = **new float**[]{0, 1, 0, 3};

511 serverRotation.getValue(rotationValues);

512 newTransformRotation.getValue(rotationValues);

513

514 addLabelToTask(taskName, newTransform, TASK\_LABEL\_COLOR);

515 tasks.put(taskName, newTransform);

516

517 *//place attention arrow*

518 String[] elements = serverName.split(**"\_"**);

519 X3DNode planeTransform = mainScene.getNamedNode(**"Server\_"** + elements[1] + **"\_XFORM"**);

520 addAttentionArrow(planeTransform);

521 }

522

523 */\*\**

524  *\* Removes the representation for task taskName from the scene*

525  *\** ***@param*** *taskName the name of the task to be removed from the scene*

526  *\*/*

527 **public void** removeTask(String taskName) {

528 X3DNode task = tasks.remove(taskName);

529 addInverseAttentionArrow(task);

530 mainScene.removeRootNode(task);

531 }

532

533 */\*\**

534  *\* Places an attention arow above the server and changes the color of the server*

535  *\* platform to gray*

536  *\** ***@param*** *serverName the name of the target server*

537  *\*/*

538 **public void** sendServerToLowPower(String serverName) {

539 String[] elements = serverName.split(**"\_"**);

540 X3DNode material = mainScene.getNamedNode(**"ServerPlane\_0"** + elements[1] + **"\_MAT"**);

541 SFColor diffuseColor = (SFColor) material.getField(**"diffuseColor"**);

542 diffuseColor.setValue(INACTIVE\_SERVER\_COLOR);

543 activeServersNo--;

544 fanAnimationTimer.setDelay(100 / activeServersNo);

545 fanAnimationTimer.restart();

546

547 *//place attention arrow*

548 X3DNode planeTransform = mainScene.getNamedNode(**"Server\_"** + elements[1] + **"\_XFORM"**);

549 addAttentionArrow(planeTransform);

550

551 }

552

553 */\*\**

554  *\* Places an attention arow above the server and changes the color of the server*

555  *\* platform to blue*

556  *\** ***@param*** *serverName the name of the target server*

557  *\*/*

558 **public void** wakeUpServer(String serverName) {

559 String[] elements = serverName.split(**"\_"**);

560 X3DNode material = mainScene.getNamedNode(**"ServerPlane\_0"** + elements[1] + **"\_MAT"**);

561 SFColor diffuseColor = (SFColor) material.getField(**"diffuseColor"**);

562 diffuseColor.setValue(ACTIVE\_SERVER\_COLOR);

563 activeServersNo++;

564 fanAnimationTimer.setDelay(100 / activeServersNo);

565 fanAnimationTimer.restart();

566

567 *//place attention arrow*

568 X3DNode planeTransform = mainScene.getNamedNode(**"Server\_"** + elements[1] + **"\_XFORM"**);

569 addAttentionArrow(planeTransform);

570 }

571

572 */\*\**

573  *\* Adds an attention arow pointing down for a certain time above the specified node*

574  *\** ***@param*** *targetNode the node above which an attention arrow will appear*

575  *\*/*

576 **private void** addAttentionArrow(X3DNode targetNode) {

577 *//place attention arrow*

578 SFRotation planeRotation = (SFRotation) targetNode.getField(**"rotation"**);

579 SFRotation arrowRotation = (SFRotation) attentionArow.getField(**"rotation"**);

580 SFVec3f planeTranslation = (SFVec3f) targetNode.getField(**"translation"**);

581 SFVec3f arrowTranslation = (SFVec3f) attentionArow.getField(**"translation"**);

582 SFVec3f arrowScale = (SFVec3f) attentionArow.getField(**"scale"**);

583 arrowScale.setValue(**new float**[]{1, 1, 1});

584 **float**[] arrowTranslationValues = **new float**[4];

585 **float**[] arrowRotationValues = **new float**[4];

586 planeTranslation.getValue(arrowTranslationValues);

587 planeRotation.getValue(arrowRotationValues);

588 arrowTranslationValues[1] += 0.5;

589 arrowRotation.setValue(arrowRotationValues);

590 arrowTranslation.setValue(arrowTranslationValues);

591

592 Timer timer = **new** Timer(2000, **new** ActionListener() {

593

594 **public void** actionPerformed(ActionEvent e) {

595 SFVec3f arrowScale = (SFVec3f) attentionArow.getField(**"scale"**);

596 arrowScale.setValue(**new float**[]{0, 0, 0});

597 }

598 });

599

600 timer.setRepeats(**false**);

601 timer.start();

602 }

603

604 */\*\**

605  *\* Adds an attention arow pointing up for a certain time above the specified node*

606  *\** ***@param*** *targetNode the node above which an attention arrow will appear*

607  *\*/*

608 **private void** addInverseAttentionArrow(X3DNode targetNode) {

609 *//place attention arrow*

610 SFRotation planeRotation = (SFRotation) targetNode.getField(**"rotation"**);

611 SFRotation arrowRotation = (SFRotation) inverseAttentionArow.getField(**"rotation"**);

612 SFVec3f planeTranslation = (SFVec3f) targetNode.getField(**"translation"**);

613 SFVec3f arrowTranslation = (SFVec3f) inverseAttentionArow.getField(**"translation"**);

614 SFVec3f arrowScale = (SFVec3f) inverseAttentionArow.getField(**"scale"**);

615 arrowScale.setValue(**new float**[]{1, 1, 1});

616 **float**[] arrowTranslationValues = **new float**[4];

617 **float**[] arrowRotationValues = **new float**[4];

618 planeTranslation.getValue(arrowTranslationValues);

619 planeRotation.getValue(arrowRotationValues);

620 arrowTranslationValues[1] += 1;

621 arrowRotation.setValue(arrowRotationValues);

622 arrowTranslation.setValue(arrowTranslationValues);

623

624 Timer timer = **new** Timer(2000, **new** ActionListener() {

625

626 **public void** actionPerformed(ActionEvent e) {

627 SFVec3f arrowScale = (SFVec3f) inverseAttentionArow.getField(**"scale"**);

628 arrowScale.setValue(**new float**[]{0, 0, 0});

629 }

630 });

631

632 timer.setRepeats(**false**);

633 timer.start();

634 }

635

636 */\*\**

637  *\* Rotates the circle located around of the sensor sphere representation*

638  *\** ***@param*** *sensorTubeName name of the sensor to be animated*

639  *\*/*

640 **public void** animateSensor(String sensorTubeName) {

641 X3DNode node = mainScene.getNamedNode(sensorTubeName);

642 **final** SFRotation arrowRotation = (SFRotation) node.getField(**"rotation"**);

643 **final float**[] values = **new float**[4];

644 arrowRotation.getValue(values);

645 values[0] = 1;

646 values[1] = 0;

647 arrowRotation.setValue(values);

648

649 Timer timer = **new** Timer(2000, **new** ActionListener() {

650

651 **public void** actionPerformed(ActionEvent e) {

652 values[0] = 0;

653 values[1] = 1;

654 arrowRotation.setValue(values);

655 }

656 });

657

658 timer.setRepeats(**false**);

659 timer.start();

660

661 }

662

663 }

|  |
| --- |
| DefaultServer.java |

1 **package** greenContextOntology.impl;

2

3 **import** com.hp.hpl.jena.ontology.OntModel;

4 **import** contextawaremodel.worldInterface.datacenterInterface.proxies.ServerManagementProxyInterface;

5 **import** edu.stanford.smi.protege.model.FrameID;

6 **import** edu.stanford.smi.protegex.owl.model.OWLModel;

7 **import** edu.stanford.smi.protegex.owl.model.RDFProperty;

8 **import** greenContextOntology.\*;

9

10 **import** java.util.Collection;

11 **import** java.util.Iterator;

12

13 */\*\**

14  *\* Generated by Protege-OWL (http://protege.stanford.edu/plugins/owl).*

15  *\* Source OWL Class: http://www.owl-ontologies.com/Datacenter.owl#Server*

16  *\**

17  *\** ***@version*** *generated on Sun Mar 07 13:11:11 EET 2010*

18  *\*/*

19 **public class** DefaultServer **extends** DefaultResource

20 **implements** Server {

21

22 **private** ServerManagementProxyInterface proxy;

23

24 **public** DefaultServer(OWLModel owlModel, FrameID id) {

25 **super**(owlModel, id);

26 }

27

28 **public** DefaultServer() {

29 }

30

31 *// Property http://www.owl-ontologies.com/Datacenter.owl#serverMacAddress*

32 **public** String getServerMacAddress() {

33 **return** (String) getPropertyValue(getServerMacAddressProperty());

34 }

35

36

37 **public** RDFProperty getServerMacAddressProperty() {

38 **final** String uri = **"http://www.owl-ontologies.com/Datacenter.owl#serverMacAddress"**;

39 **final** String name = getOWLModel().getResourceNameForURI(uri);

40 **return** getOWLModel().getRDFProperty(name);

41 }

42

43

44 **public boolean** hasServerMacAddress() {

45 **return** getPropertyValueCount(getServerMacAddressProperty()) > 0;

46 }

47

48

49 **public void** setServerMacAddress(String newServerMacAddress) {

50 setPropertyValue(getServerMacAddressProperty(), newServerMacAddress);

51 }

52

53 *// Property http://www.owl-ontologies.com/Datacenter.owl#serverWakeUpPort*

54 **public int** getServerWakeUpPort() {

55 **return** getPropertyValueLiteral(getServerWakeUpPortProperty()).getInt();

56 }

57

58

59 **public** RDFProperty getServerWakeUpPortProperty() {

60 **final** String uri = **"http://www.owl-ontologies.com/Datacenter.owl#serverWakeUpPort"**;

61 **final** String name = getOWLModel().getResourceNameForURI(uri);

62 **return** getOWLModel().getRDFProperty(name);

63 }

64

65

66 **public boolean** hasServerWakeUpPort() {

67 **return** getPropertyValueCount(getServerWakeUpPortProperty()) > 0;

68 }

69

70

71 **public void** setServerWakeUpPort(**int** newServerWakeUpPort) {

72 setPropertyValue(getServerWakeUpPortProperty(), **new** java.lang.Integer(newServerWakeUpPort));

73 }

74

75 *// Property http://www.owl-ontologies.com/Datacenter.owl#associatedCPU*

76 **public** CPU getAssociatedCPU() {

77 **return** (CPU) getPropertyValueAs(getAssociatedCPUProperty(), CPU.**class**);

78 }

79

80 **public** RDFProperty getAssociatedCPUProperty() {

81 **final** String uri = **"http://www.owl-ontologies.com/Datacenter.owl#associatedCPU"**;

82 **final** String name = getOWLModel().getResourceNameForURI(uri);

83 **return** getOWLModel().getRDFProperty(name);

84 }

85

86 **public boolean** hasAssociatedCPU() {

87 **return** getPropertyValueCount(getAssociatedCPUProperty()) > 0;

88 }

89

90 **public void** setAssociatedCPU(CPU newAssociatedCPU) {

91 setPropertyValue(getAssociatedCPUProperty(), newAssociatedCPU);

92 }

93

94 *// Property http://www.owl-ontologies.com/Datacenter.owl#associatedMemory*

95 **public** Memory getAssociatedMemory() {

96 **return** (Memory) getPropertyValueAs(getAssociatedMemoryProperty(), Memory.**class**);

97 }

98

99 **public** RDFProperty getAssociatedMemoryProperty() {

100 **final** String uri = **"http://www.owl-ontologies.com/Datacenter.owl#associatedMemory"**;

101 **final** String name = getOWLModel().getResourceNameForURI(uri);

102 **return** getOWLModel().getRDFProperty(name);

103 }

104

105 **public boolean** hasAssociatedMemory() {

106 **return** getPropertyValueCount(getAssociatedMemoryProperty()) > 0;

107 }

108

109 **public void** setAssociatedMemory(Memory newAssociatedMemory) {

110 setPropertyValue(getAssociatedMemoryProperty(), newAssociatedMemory);

111 }

112

113 *// Property http://www.owl-ontologies.com/Datacenter.owl#associatedStorage*

114 **public** Storage getAssociatedStorage() {

115 **return** (Storage) getPropertyValueAs(getAssociatedStorageProperty(), Storage.**class**);

116 }

117

118 **public** RDFProperty getAssociatedStorageProperty() {

119 **final** String uri = **"http://www.owl-ontologies.com/Datacenter.owl#associatedStorage"**;

120 **final** String name = getOWLModel().getResourceNameForURI(uri);

121 **return** getOWLModel().getRDFProperty(name);

122 }

123

124 **public boolean** hasAssociatedStorage() {

125 **return** getPropertyValueCount(getAssociatedStorageProperty()) > 0;

126 }

127

128 **public void** setAssociatedStorage(Storage newAssociatedStorage) {

129 setPropertyValue(getAssociatedStorageProperty(), newAssociatedStorage);

130 }

131

132 *// Property http://www.owl-ontologies.com/Datacenter.owl#isInLowPowerState*

133 **public boolean** getIsInLowPowerState() {

134 **return** getPropertyValueLiteral(getIsInLowPowerStateProperty()).getBoolean();

135 }

136

137 **public** RDFProperty getIsInLowPowerStateProperty() {

138 **final** String uri = **"http://www.owl-ontologies.com/Datacenter.owl#isInLowPowerState"**;

139 **final** String name = getOWLModel().getResourceNameForURI(uri);

140 **return** getOWLModel().getRDFProperty(name);

141 }

142

143 **public boolean** hasIsInLowPowerState() {

144 **return** getPropertyValueCount(getIsInLowPowerStateProperty()) > 0;

145 }

146

147 **public void** setIsInLowPowerState(**boolean** newIsInLowPowerState, OntModel ontModel) {

148 setPropertyValue(getIsInLowPowerStateProperty(), **new** java.lang.Boolean(newIsInLowPowerState), ontModel);

149 }

150

151 *// Property http://www.owl-ontologies.com/Datacenter.owl#runningTasks*

152 **public** Collection getRunningTasks() {

153 **return** getPropertyValuesAs(getRunningTasksProperty(), Task.**class**);

154 }

155

156 **public** RDFProperty getRunningTasksProperty() {

157 **final** String uri = **"http://www.owl-ontologies.com/Datacenter.owl#runningTasks"**;

158 **final** String name = getOWLModel().getResourceNameForURI(uri);

159 **return** getOWLModel().getRDFProperty(name);

160 }

161

162 **public boolean** hasRunningTasks() {

163 **return** getPropertyValueCount(getRunningTasksProperty()) > 0;

164 }

165

166 **public** Iterator listRunningTasks() {

167 **return** listPropertyValuesAs(getRunningTasksProperty(), Task.**class**);

168 }

169

170 */\*\**

171  *\* It gives to the new task the maximum between the requested and available*

172  *\* resources and updates the server's <code>used</code> resources*

173  *\**

174  *\** ***@param*** *newRunningTasks the task to be deployed on the server*

175  *\*/*

176 **public void** addRunningTask(Task newRunningTasks, OntModel model) {

177

178 RequestedTaskInfo requestedSLA = newRunningTasks.getRequestedInfo();

179 ReceivedTaskInfo receivedSLA = newRunningTasks.getReceivedInfo();

180

181 CPU cpu = **this**.getAssociatedCPU();

182 Collection<Core> cores = cpu.getAssociatedCore();

183 Iterator<Core> coresIterator = cores.iterator();

184 **int** coreCount = requestedSLA.getCores();

185 **int** availableCores = cores.size();

186

187 coreCount = (coreCount > availableCores) ? availableCores : coreCount;

188 **int** index = 0;

189 **while** (coresIterator.hasNext() && coreCount > 0) {

190 Core core = coresIterator.next();

191 *//TODO : to be modified for more core flexibility*

192 **int** availableCore = core.getTotal() - core.getUsed();

193 **int** requestedCPU = requestedSLA.getCpuMaxAcceptableValue();

194 **if** (requestedCPU > availableCore) {

195 index++;

196 **continue**;

197 }

198 coreCount--;

199 receivedSLA.setCpuReceived(requestedCPU, model);

200 core.setUsed(core.getUsed() + requestedCPU);

201 receivedSLA.addReceivedCoreIndex(index);

202 index++;

203 }

204

205 receivedSLA.setCores(receivedSLA.getReceivedCoreIndex().size(), model);

206

207 Memory memory = **this**.getAssociatedMemory();

208 **int** availableMemory = memory.getTotal() - memory.getUsed();

209 **int** requestedMemory = requestedSLA.getMemoryMaxAcceptableValue();

210 **int** receivedMemory = (requestedMemory < availableMemory) ? requestedMemory : availableMemory;

211 receivedSLA.setMemoryReceived(receivedMemory, model);

212 memory.setUsed(memory.getUsed() + receivedMemory);

213

214 Storage storage = **this**.getAssociatedStorage();

215 **int** availableStorage = storage.getTotal() - storage.getUsed();

216 **int** requestedStorage = requestedSLA.getStorageMaxAcceptableValue();

217 **int** receivedStorage = (requestedStorage < availableStorage) ? requestedStorage : availableStorage;

218 receivedSLA.setStorageReceived(receivedStorage, model);

219 storage.setUsed(storage.getUsed() + receivedStorage);

220

221 *//add task to ontology*

222 Collection tasks = getRunningTasks();

223 removePropertyValue(getRunningTasksProperty(), tasks);

224 tasks.add(newRunningTasks);

225

226 *//setPropertyValue(getRunningTasksProperty(), newRunningTasks);*

227 addPropertyValue(getRunningTasksProperty(), newRunningTasks);

228

229 }

230

231 */\*\**

232  *\* Adds a task to be run on this server for which the resources had been*

233  *\* negotiated because the server must exceed its optimum load paremeters in order*

234  *\* to run this task*

235  *\** ***@param*** *newRunningTasks the task to be run on the server*

236  *\** ***@param*** *model the ontModel of the underlying ontology*

237  *\** ***@param*** *negotiatedCPU*

238  *\** ***@param*** *negotiatedMemory*

239  *\** ***@param*** *negotiatedStorage*

240  *\*/*

241 **public void** addNegotiatedTasks(Task newRunningTasks, OntModel model,

242 **int** negotiatedCPU, **int** negotiatedMemory, **int** negotiatedStorage) {

243 RequestedTaskInfo requestedSLA = newRunningTasks.getRequestedInfo();

244 ReceivedTaskInfo receivedSLA = newRunningTasks.getReceivedInfo();

245

246 **if** (negotiatedCPU == 0) {

247 negotiatedCPU = requestedSLA.getCpuMaxAcceptableValue();

248 }

249 **if** (negotiatedMemory == 0) {

250 negotiatedMemory = requestedSLA.getMemoryMaxAcceptableValue();

251 }

252 **if** (negotiatedStorage == 0) {

253 negotiatedStorage = requestedSLA.getStorageMaxAcceptableValue();

254 }

255

256 CPU cpu = **this**.getAssociatedCPU();

257 Collection<Core> cores = cpu.getAssociatedCore();

258 Iterator<Core> coresIterator = cores.iterator();

259 **int** coreCount = requestedSLA.getCores();

260 **int** availableCores = cores.size();

261

262 coreCount = (coreCount > availableCores) ? availableCores : coreCount;

263 **int** index = 0;

264 **while** (coresIterator.hasNext() && coreCount > 0) {

265 Core core = coresIterator.next();

266 *//TODO : to be modified for more core flexibility*

267 **int** availableCore = core.getTotal() - core.getUsed();

268 **if** (negotiatedCPU > availableCore) {

269 index++;

270 **continue**;

271 }

272 coreCount--;

273 receivedSLA.setCpuReceived(negotiatedCPU, model);

274 core.setUsed(core.getUsed() + negotiatedCPU);

275 receivedSLA.addReceivedCoreIndex(index);

276 index++;

277 }

278

279 receivedSLA.setCores(receivedSLA.getReceivedCoreIndex().size(), model);

280

281 Memory memory = **this**.getAssociatedMemory();

282 **int** availableMemory = memory.getTotal() - memory.getUsed();

283 **int** receivedMemory = (negotiatedMemory < availableMemory) ? negotiatedMemory : availableMemory;

284 receivedSLA.setMemoryReceived(receivedMemory, model);

285 memory.setUsed(memory.getUsed() + receivedMemory);

286

287 Storage storage = **this**.getAssociatedStorage();

288 **int** availableStorage = storage.getTotal() - storage.getUsed();

289 **int** receivedStorage = (negotiatedStorage < availableStorage) ? negotiatedStorage : availableStorage;

290 receivedSLA.setStorageReceived(receivedStorage, model);

291 storage.setUsed(storage.getUsed() + receivedStorage);

292

293 *//add task to ontology*

294 Collection tasks = getRunningTasks();

295 removePropertyValue(getRunningTasksProperty(), tasks);

296 tasks.add(newRunningTasks);

297

298 addPropertyValue(getRunningTasksProperty(), newRunningTasks);

299 }

300

301 **public void** addRunningTask(Task newRunningTasks) {

302 addPropertyValue(getRunningTasksProperty(), newRunningTasks);

303 }

304

305 */\*\**

306  *\* Removes the specified task from the server*

307  *\** ***@param*** *oldRunningTasks the task to be removed from the server*

308  *\** ***@param*** *model*

309  *\*/*

310 **public void** removeRunningTask(Task oldRunningTasks, OntModel model) {

311

312

313 ReceivedTaskInfo receivedSLA = oldRunningTasks.getReceivedInfo();

314

315 CPU cpu = **this**.getAssociatedCPU();

316 Collection<Core> cores = cpu.getAssociatedCore();

317 Iterator<Core> coresIterator = cores.iterator();

318

319 Collection receivedCoresIndexes = receivedSLA.getReceivedCoreIndex();

320 receivedSLA.setCores(0, model);

321 **int** coreCount = cores.size();

322 **for** (**int** i = 0; i < coreCount; i++) {

323 Core core = coresIterator.next();

324 **if** (receivedCoresIndexes.contains(i)) {

325 **int** receivedCPU = receivedSLA.getCpuReceived();

326 core.setUsed(core.getUsed() - receivedCPU);

327 receivedSLA.removeReceivedCoreIndex(i);

328 }

329

330 }

331

332 receivedSLA.setCpuReceived(0, model);

333

334 Memory memory = **this**.getAssociatedMemory();

335 memory.setUsed(memory.getUsed() - receivedSLA.getMemoryReceived());

336 receivedSLA.setMemoryReceived(0, model);

337

338 Storage storage = **this**.getAssociatedStorage();

339

340 storage.setUsed(storage.getUsed() - receivedSLA.getStorageReceived());

341 receivedSLA.setStorageReceived(0, model);

342

343 *//remove task from ontology*

344 *//getRunningTasks().remove(oldRunningTasks);*

345 removePropertyValue(getRunningTasksProperty(), oldRunningTasks);

346

347 }

348

349 **public void** removeRunningTasks(Task oldRunningTasks) {

350

351 *//remove task from ontology*

352 removePropertyValue(getRunningTasksProperty(), oldRunningTasks);

353 }

354

355 *// Property http://www.owl-ontologies.com/Datacenter.owl#serverIPAddress*

356 **public** String getServerIPAddress() {

357 **return** (String) getPropertyValue(getServerIPAddressProperty());

358 }

359

360

361 **public** RDFProperty getServerIPAddressProperty() {

362 **final** String uri = **"http://www.owl-ontologies.com/Datacenter.owl#serverIPAddress"**;

363 **final** String name = getOWLModel().getResourceNameForURI(uri);

364 **return** getOWLModel().getRDFProperty(name);

365 }

366

367

368 **public boolean** hasServerIPAddress() {

369 **return** getPropertyValueCount(getServerIPAddressProperty()) > 0;

370 }

371

372

373 **public void** setServerIPAddress(String newServerIPAddress) {

374 setPropertyValue(getServerIPAddressProperty(), newServerIPAddress);

375 }

376

377 *// Property http://www.owl-ontologies.com/Datacenter.owl#virtualMachinesPath*

378

379 **public** Collection getVirtualMachinesPath() {

380 **return** getPropertyValues(getVirtualMachinesPathProperty());

381 }

382

383

384 **public** RDFProperty getVirtualMachinesPathProperty() {

385 **final** String uri = **"http://www.owl-ontologies.com/Datacenter.owl#virtualMachinesPath"**;

386 **final** String name = getOWLModel().getResourceNameForURI(uri);

387 **return** getOWLModel().getRDFProperty(name);

388 }

389

390

391 **public boolean** hasVirtualMachinesPath() {

392 **return** getPropertyValueCount(getVirtualMachinesPathProperty()) > 0;

393 }

394

395

396 **public** Iterator listVirtualMachinesPath() {

397 **return** listPropertyValues(getVirtualMachinesPathProperty());

398 }

399

400 */\*\**

401  *\**

402  *\** ***@param*** *newVirtualMachinesPath the physycal path of the virtual machine files*

403  *\*/*

404 **public void** addVirtualMachinesPath(String newVirtualMachinesPath) {

405 addPropertyValue(getVirtualMachinesPathProperty(), newVirtualMachinesPath);

406 }

407

408 */\*\**

409  *\**

410  *\** ***@param*** *oldVirtualMachinesPath the physycal path of the virtual machine files*

411  *\*/*

412 **public void** removeVirtualMachinesPath(String oldVirtualMachinesPath) {

413 removePropertyValue(getVirtualMachinesPathProperty(), oldVirtualMachinesPath);

414 }

415

416 */\*\**

417  *\** ***@param*** *newVirtualMachinesPath the physycal path of the virtual machines files*

418  *\*/*

419 **public void** setVirtualMachinesPath(Collection newVirtualMachinesPath) {

420 setPropertyValues(getVirtualMachinesPathProperty(), newVirtualMachinesPath);

421 }

422

423 */\*\**

424  *\* Checks if a server has resources to run atask while respecting its optimum*

425  *\* load parameters*

426  *\** ***@param*** *task task to be accomodated on the server*

427  *\** ***@return*** *true if there are enough resources to satify the task's*

428  *\* requirements and false otherwise*

429  *\*/*

430 **public boolean** hasResourcesFor(Task task) {

431

432 RequestedTaskInfo requestedSLA = task.getRequestedInfo();

433

434

435 CPU cpu = **this**.getAssociatedCPU();

436 Collection cores = cpu.getAssociatedCore();

437 **int** requestedCores = requestedSLA.getCores();

438 **if** (cores.size() < requestedCores) {

439 **return false**;

440 }

441 **for** (Object coreInst : cores) {

442

443 Core core = (Core) coreInst;

444 **if** (core.getUsed() + requestedSLA.getCpuMaxAcceptableValue() > core.getMaxAcceptableValue()) {

445 **continue**;

446 } **else** {

447 requestedCores--;

448 }

449 }

450

451 **if** (requestedCores > 0) {

452 **return false**;

453 }

454

455 Memory memory = **this**.getAssociatedMemory();

456 **if** (memory.getUsed() + requestedSLA.getMemoryMaxAcceptableValue() > memory.getMaxAcceptableValue()) {

457 **return false**;

458 }

459

460 Storage storage = **this**.getAssociatedStorage();

461

462 **if** (storage.getUsed() + requestedSLA.getStorageMaxAcceptableValue() > storage.getMaxAcceptableValue()) {

463 **return false**;

464 }

465

466 **return true**;

467 }

468

469 */\*\**

470  *\* Check if server can run the specified task if it exeeds the optimum load values*

471  *\** ***@param*** *task the task to be run on the server*

472  *\** ***@return*** *true if the server has enough resources in the case it exceeds its*

473  *\* optimal load parameters and false otherwise*

474  *\*/*

475 **public boolean** hasResourcesToBeNegotiatedFor(Task task) {

476

477 RequestedTaskInfo requestedSLA = task.getRequestedInfo();

478

479

480 CPU cpu = **this**.getAssociatedCPU();

481 Collection cores = cpu.getAssociatedCore();

482 **int** requestedCores = requestedSLA.getCores();

483 **if** (cores.size() < requestedCores) {

484 **return false**;

485 }

486 **for** (Object coreInst : cores) {

487

488 Core core = (Core) coreInst;

489 **if** (core.getUsed() + requestedSLA.getCpuMinAcceptableValue() > core.getTotal()) {

490 **continue**;

491 } **else** {

492 requestedCores--;

493 }

494 }

495

496 **if** (requestedCores > 0) {

497 **return false**;

498 }

499

500 Memory memory = **this**.getAssociatedMemory();

501 **if** (memory.getUsed() + requestedSLA.getMemoryMinAcceptableValue() > memory.getTotal()) {

502 **return false**;

503 }

504

505 Storage storage = **this**.getAssociatedStorage();

506

507 **if** (storage.getUsed() + requestedSLA.getStorageMinAcceptableValue() > storage.getTotal()) {

508 **return false**;

509 }

510

511 **return true**;

512 }

513

514 **public void** setRunningTasks(Collection newRunningTasks, OntModel model) {

515 setPropertyValues(getRunningTasksProperty(), newRunningTasks);

516 **for** (Object task : newRunningTasks) {

517 addRunningTask((greenContextOntology.Task) task, model);

518 }

519 }

520

521 **public void** setRunningTasks(Collection newRunningTasks) {

522 setPropertyValues(getRunningTasksProperty(), newRunningTasks);

523 **for** (Object task : newRunningTasks) {

524 addRunningTask((greenContextOntology.Task) task);

525 }

526 }

527

528 *// Property http://www.owl-ontologies.com/Datacenter.owl#serverName*

529 **public** String getServerName() {

530 **return** (String) getPropertyValue(getServerNameProperty());

531 }

532

533 **public** RDFProperty getServerNameProperty() {

534 **final** String uri = **"http://www.owl-ontologies.com/Datacenter.owl#serverName"**;

535 **final** String name = getOWLModel().getResourceNameForURI(uri);

536 **return** getOWLModel().getRDFProperty(name);

537 }

538

539 **public boolean** hasServerName() {

540 **return** getPropertyValueCount(getServerNameProperty()) > 0;

541 }

542

543 **public void** setServerName(String newServerName) {

544 setPropertyValue(getServerNameProperty(), newServerName);

545 }

546

547 *// Property http://www.owl-ontologies.com/Datacenter.owl#webService*

548 **public** String getWebService() {

549 **return** (String) getPropertyValue(getWebServiceProperty());

550 }

551

552 **public** RDFProperty getWebServiceProperty() {

553 **final** String uri = **"http://www.owl-ontologies.com/Datacenter.owl#webService"**;

554 **final** String name = getOWLModel().getResourceNameForURI(uri);

555 **return** getOWLModel().getRDFProperty(name);

556 }

557

558 **public boolean** hasWebService() {

559 **return** getPropertyValueCount(getWebServiceProperty()) > 0;

560 }

561

562 **public void** setWebService(String newWebService) {

563 setPropertyValue(getWebServiceProperty(), newWebService);

564 }

565

566 @Override

567 **public** String toString() {

568

569 String description;

570 description = **"Server "** + **this**.getName().split(**"#"**)[1] + **"\n"**;

571 description += **"Inactive = "** + **this**.getIsInLowPowerState() + **"\n"**;

572 Collection cores = **this**.getAssociatedCPU().getAssociatedCore();

573 Iterator iterator = cores.iterator();

574

575 description += **"Cores "** + cores.size() + **"\n"**;

576

577 **while** (iterator.hasNext()) {

578 Core core = (Core) iterator.next();

579 description += **"CPU used "** + core.getUsed() + **" total "** + core.getTotal()

580 + **" range [ "** + core.getMinAcceptableValue()

581 + **".."** + core.getMaxAcceptableValue() + **" ]\n"**;

582 }

583

584 description += **"Memory used "** + **this**.getAssociatedMemory().getUsed()

585 + **" total "** + **this**.getAssociatedMemory().getTotal()

586 + **" range [ "** + **this**.getAssociatedMemory().getMinAcceptableValue()

587 + **".."** + **this**.getAssociatedMemory().getMaxAcceptableValue()

588 + **" ]\n"**;

589

590 description += **"Storage used "** + **this**.getAssociatedStorage().getUsed()

591 + **" total "** + **this**.getAssociatedStorage().getTotal()

592

593 + **" range [ "** + **this**.getAssociatedStorage().getMinAcceptableValue()

594 + **".."** + **this**.getAssociatedStorage().getMaxAcceptableValue() + **" ]\n"**;

595

596 **return** description;

597 }

598

599 */\*\**

600  *\* To check if a task is already present and such avoid adding it twice*

601  *\** ***@param*** *task*

602  *\** ***@return*** *true if the task runs on this server and false otherwise*

603  *\*/*

604 **public boolean** containsTask(Task task) {

605 Collection tasks = **this**.getRunningTasks();

606 **for** (Object o : tasks) {

607 Task t = (Task) o;

608 **if** (t.getName().equals(task.getName())) {

609 **return true**;

610 }

611 }

612 **return false**;

613 }

614

615 */\*\**

616  *\* Checks if all the optimal resources load parameters are respected*

617  *\** ***@return*** *true if the optimum load is respected and false otherwise*

618  *\*/*

619 **public boolean** areOptimumValuesRespected() {

620 CPU cpu = **this**.getAssociatedCPU();

621 **for** (Object o : cpu.getAssociatedCore()) {

622 Core core = (Core) o;

623 **int** usedCore = core.getUsed();

624 **if** (usedCore < core.getMinAcceptableValue() || usedCore > core.getMaxAcceptableValue()) {

625 **return false**;

626 }

627 }

628 Memory memory = **this**.getAssociatedMemory();

629 **int** usedMemory = memory.getUsed();

630 **if** (usedMemory < memory.getMinAcceptableValue() || usedMemory > memory.getMaxAcceptableValue()) {

631 **return false**;

632 }

633

634 Storage storage = **this**.getAssociatedStorage();

635 **int** usedStorage = storage.getUsed();

636 **if** (usedStorage < storage.getMinAcceptableValue() || usedStorage > storage.getMaxAcceptableValue()) {

637 **return false**;

638 }

639

640 **return true**;

641 }

642

643 */\*\**

644  *\* Reclaim previousely distributed resources*

645  *\** ***@param*** *model*

646  *\*/*

647 **public void** collectPreviouslyDistributedResources(OntModel model) {

648 Collection runningTasks = getRunningTasks();

649 **for** (Object t : runningTasks) {

650 Task task = (Task) t;

651 removeRunningTask(task, model);

652 }

653

654 **for** (Object t : runningTasks) {

655 Task task = (Task) t;

656 addRunningTask(task, model);

657 }

658

659 }

660

661 */\*\**

662  *\* Distribute unused resources among running tasks*

663  *\** ***@param*** *model*

664  *\*/*

665 **public void** distributeRemainingResources(OntModel model) {

666 Collection runningTasks = getRunningTasks();

667 **if** (runningTasks.size() == 0) {

668 **return**;

669 }

670 CPU cpu = getAssociatedCPU();

671

672 **for** (Object task : runningTasks) {

673 **for** (Object object : cpu.getAssociatedCore()) {

674 Core core = (Core) object;

675

676 *//TODO: Modify. For now the empty cpu-s are assigned in order to the running tasks*

677 **if** (core.getUsed() == 0) {

678 *//if core used in optimum parameters continue*

679 **continue**;

680 } **else** {

681 ReceivedTaskInfo received = ((Task) task).getReceivedInfo();

682 received.setCores(received.getCores() + 1, model);

683 core.setUsed(core.getMinAcceptableValue());

684 *//in order to continue and assign the next empty core to the next task*

685 **break**;

686 }

687 }

688 }

689

690 Memory memory = getAssociatedMemory();

691 **int** usedMemory = memory.getUsed();

692 **int** memoryMinAcceptableValue = memory.getMinAcceptableValue();

693 **if** (usedMemory < memoryMinAcceptableValue) {

694 **int** remaining = memoryMinAcceptableValue - usedMemory;

695 **int** partition = remaining / runningTasks.size();

696 **if** (partition == 0) {

697 partition = 1;

698 }

699 **for** (Object task : runningTasks) {

700 ReceivedTaskInfo received = ((Task) task).getReceivedInfo();

701 received.setMemoryReceived(received.getMemoryReceived() + partition, model);

702 }

703 memory.setUsed(remaining);

704 }

705

706 Storage storage = getAssociatedStorage();

707 **int** usedStorage = storage.getUsed();

708 **int** storageMinAcceptableValue = storage.getMinAcceptableValue();

709 **if** (usedStorage < storageMinAcceptableValue) {

710 **int** remaining = storageMinAcceptableValue - usedStorage;

711 **int** partition = remaining / runningTasks.size();

712 **if** (partition == 0) {

713 partition = 1;

714 }

715 **for** (Object task : runningTasks) {

716 ReceivedTaskInfo received = ((Task) task).getReceivedInfo();

717 received.setStorageReceived(received.getStorageReceived() + partition, model);

718 }

719 storage.setUsed(remaining);

720 }

721 }

722

723 */\*\**

724  *\**

725  *\** ***@param*** *resources amount of extra resources to be given*

726  *\** ***@param*** *task the task which receives the extra resources*

727  *\** ***@param*** *model ontModel of the underlying ontology*

728  *\*/*

729 **public void** giveMoreResourcesToTask(**double**[] resources, Task task, OntModel model) {

730

731 ReceivedTaskInfo receivedTaskInfo = task.getReceivedInfo();

732

733 Collection<Integer> receivedCores = receivedTaskInfo.getReceivedCoreIndex();

734 Object[] cores = **this**.getAssociatedCPU().getAssociatedCore().toArray();

735 **for** (Integer index : receivedCores) {

736 Core core = (Core) cores[index];

737 core.setUsed(core.getUsed() + (**int**) resources[0]);

738 }

739 receivedTaskInfo.setCpuReceived(((Core) cores[0]).getUsed() + (**int**) resources[0], model);

740

741 Memory memory = **this**.getAssociatedMemory();

742 **int** newMemoryValue = (memory.getUsed() + (**int**) resources[1]);

743 memory.setUsed(newMemoryValue);

744 receivedTaskInfo.setMemoryReceived(newMemoryValue);

745

746 Storage storage = **this**.getAssociatedStorage();

747 **int** newStorageValue = (storage.getUsed() + (**int**) resources[2]);

748 storage.setUsed(newStorageValue);

749 receivedTaskInfo.setStorageReceived(newStorageValue);

750

751 }

752

753 **public void** removeExtraResourcesGivenToTask(Task task, OntModel model) {

754 removeRunningTask(task, model);

755 addRunningTask(task, model);

756 }

757

758 **public void** changeOptimumCPURange(**int** max) {

759 *//TODO: optimum ranges for a specific core can be negotiated so implement method that does that*

760 CPU cpu = getAssociatedCPU();

761 Collection<Core> cores = cpu.getAssociatedCore();

762 **for** (Core core : cores) {

763

764 core.setMaxAcceptableValue(max);

765 }

766

767 }

768

769 **public void** changeOptimumMemoryRange(**int** max) {

770 Memory memory = getAssociatedMemory();

771 memory.setMaxAcceptableValue(max);

772

773 }

774

775 **public void** changeOptimumStorageRange(**int** max) {

776 Storage storage = getAssociatedStorage();

777 storage.setMaxAcceptableValue(max);

778 }

779

780 */\*\**

781  *\* Resets the optimum server load values to their initial values*

782  *\*/*

783 **public void** resetOptimumValues() {

784 CPU cpu = getAssociatedCPU();

785 Collection<Core> cores = cpu.getAssociatedCore();

786 **for** (Core core : cores) {

787 core.restoreDefaultOptimumValues();

788 }

789 getAssociatedMemory().restoreDefaultOptimumValues();

790 getAssociatedStorage().restoreDefaultOptimumValues();

791 }

792

793 **public** ServerManagementProxyInterface getProxy() {

794 **return** proxy;

795 }

796

797 **public void** setProxy(ServerManagementProxyInterface proxy) {

798 **this**.proxy = proxy;

799 }

800 }

801

|  |
| --- |
| DefaultContextElement.java |

1 **package** greenContextOntology.impl;

2

3 **import** com.hp.hpl.jena.ontology.Individual;

4 **import** com.hp.hpl.jena.ontology.OntModel;

5 **import** com.hp.hpl.jena.rdf.model.Property;

6 **import** com.hp.hpl.jena.rdf.model.RDFNode;

7 **import** edu.stanford.smi.protege.model.FrameID;

8 **import** edu.stanford.smi.protegex.owl.model.OWLModel;

9 **import** edu.stanford.smi.protegex.owl.model.RDFProperty;

10 **import** edu.stanford.smi.protegex.owl.model.impl.DefaultOWLIndividual;

11 **import** edu.stanford.smi.protegex.owl.swrl.exceptions.SWRLFactoryException;

12 **import** edu.stanford.smi.protegex.owl.swrl.model.SWRLFactory;

13 **import** edu.stanford.smi.protegex.owl.swrl.model.SWRLImp;

14 **import** greenContextOntology.ContextElement;

15

16 */\*\**

17  *\* Base class for every class which represents a datacenter ontology entity*

18  *\*/*

19 **public class** DefaultContextElement **extends** DefaultOWLIndividual

20 **implements** ContextElement {

21

22 **public** DefaultContextElement(OWLModel owlModel, FrameID id) {

23 **super**(owlModel, id);

24 }

25

26

27 **public** DefaultContextElement() {

28 }

29

30 */\*\**

31  *\* Sets the property value both on the OWL model and on the ONT model*

32  *\* to trigger SWRL rule evaluation*

33  *\** ***@param*** *rdfProperty the slot of the entity*

34  *\** ***@param*** *o the value of the property to be inserted in slot rdfProperty*

35  *\** ***@param*** *ontModel the ont model on which this property will also be set*

36  *\*/*

37 **public void** setPropertyValue(RDFProperty rdfProperty, Object o, OntModel ontModel) {

38 **super**.setPropertyValue(rdfProperty, o);

39

40 *//to be commented to avoid SWRL rule evaluation*

41 Individual targetIndividual = ontModel.getIndividual(**this**.getName());

42

43 Property targetProperty = ontModel.getProperty(rdfProperty.getName());

44 **if** (targetIndividual.getPropertyValue(targetProperty) != **null**) {

45 targetIndividual.removeAll(targetProperty);

46 }

47 targetIndividual.setPropertyValue(targetProperty, ontModel.createLiteralStatement(

48 targetIndividual, targetProperty, o).getLiteral().as(RDFNode.**class**));

49 }

50

51 **public final void** deleteInstance(OntModel ontModel, SWRLFactory swrlFactory)

52 **throws** SWRLFactoryException {

53 **super**.delete();

54

55 *//to be commented to avoid SWRL rule evaluation*

56 *//remove instance from underlying Ont model*

57 Individual i = ontModel.getIndividual(getName());

58 i.remove();

59

60 *//remove swrl rule associated to task*

61 SWRLImp rule = swrlFactory

62 .getImp(**"http://www.owl-ontologies.com/Datacenter.owl#QoS\_Policy\_"**

63 + **this**.getName().split(**"\_"**)[1] + **"\_swrl\_rule"**);

64 rule.delete();

65 }

66

67 }

68

69

# Appendix C: C# Source Code

|  |
| --- |
| ServerMonitor.cs |

using System;

using System.Collections.Generic;

using System.Linq;

using System.Text;

using System.Xml;

using System.Diagnostics;

using System.Management;

using System.Threading;

using ServerMonitorUtils;

namespace ServerManagement

{

public class ServerMonitor

{

public static ServerInfo CollectServerInfo()

{

ServerInfo serverInfo = new ServerInfo();

List<Storage> storageInfo = new List<Storage>();

List<int> freeCPU = new List<int>();

////get host computer core count using WMI

ManagementObjectSearcher mgmtObjects = new

ManagementObjectSearcher("Select \* from Win32\_ComputerSystem");

foreach (var item in mgmtObjects.Get())

{

serverInfo.CoreCount =

Convert.ToInt32(item["NumberOfLogicalProcessors"]);

break;

}

//get free and total clock speed for each core

mgmtObjects = new ManagementObjectSearcher("Select \* from

Win32\_Processor");

foreach (var item in mgmtObjects.Get())

{

serverInfo.TotalCPU = Convert.ToInt32(item["MaxClockSpeed"]);

break;

}

mgmtObjects = new ManagementObjectSearcher("Select \* from

Win32\_PerfFormattedData\_PerfOS\_Processor where NOT Name = '\_Total' ");

foreach (ManagementObject item in mgmtObjects.Get())

{

freeCPU.Add((int)((100 –

Convert.ToDouble(item["PercentProcessorTime"])) / 100 \*

serverInfo.TotalCPU ));

}

serverInfo.FreeCPU = freeCPU;

//get host computer RAM info

mgmtObjects = new ManagementObjectSearcher("Select \* from

Win32\_OperatingSystem");

foreach (ManagementObject item in mgmtObjects.Get())

{

serverInfo.TotalMemory =

Convert.ToInt32(item["TotalVisibleMemorySize"]) / 1024; // to store in MB

serverInfo.FreeMemory = Convert.ToInt32(item["FreePhysicalMemory"]) / 1024;

break;

}

//find host computer logical drives size and free space in KB

mgmtObjects = new ManagementObjectSearcher("Select \* from

Win32\_LogicalDisk ");

foreach (ManagementObject item in mgmtObjects.Get())

{

int size = (int)(Convert.ToInt64(item["Size"]) / (1024 \* 1024 \* 1024));

if (size == 0)

{

continue;

}

Storage storage =

new Storage(item["Name"].ToString(),

size,

(int)(Convert.ToInt64(item["FreeSpace"]) / (1024 \* 1024 \* 1024)));

storageInfo.Add(storage);

}

serverInfo.Storage = storageInfo;

return serverInfo;

}

}

}

|  |
| --- |
| ServerInfo.cs |

using System;

using System.Collections.Generic;

namespace ServerManagement{

public class ServerInfo

{

public int CoreCount{get;set;}

public int TotalCPU;

public List<int> FreeCPU { get; set; }

public List<Storage> Storage { get; set; }

public int TotalMemory { get; set; }

public int FreeMemory { get; set; }

}

public class Storage

{

public String Name { get; set; }

public int Size{get;set;}

public int FreeSpace { get; set; }

public Storage(String name, int size, int freeSpace){

Name = name;

Size = size;

FreeSpace = freeSpace;

}

public Storage()

{

}

}

}