https://gospodarkainnowacje.pl

GOSPODARKA I INNOWACJE

LABORATORIUM WIEDZY Artur Borcuch

Volume: 41 | 2023

Economy and Innovation
ISSN: 2545-0573

For more information contact: editor@gospodarkainnowacje.pl

REINFORCEMENT LEARNING-BASED VIRTUAL MACHINE ORCHESTRATION FOR HYBRID OPENSTACK–VMWARE CLOUD ENVIRONMENTS

Vinay Kumar Reddy Vangoor
Techno Bytes, Inc., Ashland, MA, USA

ARTICLEINFO.

Keywords: Reinforcement Learning, Hybrid Cloud Orchestration, Virtual Machine Scheduling, Resource Allocation, Autonomous Cloud Management, Hybrid Infrastructure, Cloud Resource Optimization.

Abstract

Hybrid cloud infrastructures combine private and enterprise virtualization platforms to provide flexible, scalable computing resources. However, orchestrating virtual machines across heterogeneous environments remains a complex challenge due to differences in resource management frameworks, scheduling policies, and performance constraints. Platforms such as OpenStack and VMware vSphere often operate independently, making cross-platform orchestration inefficient when relying on static scheduling strategies.

This research proposes a reinforcement learning–based orchestration framework that dynamically manages virtual machine placement, migration, and scaling across hybrid cloud infrastructures. The proposed system utilizes a reinforcement learning agent that continuously observes infrastructure states such as CPU utilization, memory consumption, network latency, and workload demand. Based on these states, the agent selects optimal actions for deploying or migrating virtual machines between OpenStack and VMware clusters.

The model learns orchestration policies through interaction with the environment and improves performance over time using reward-based optimization. Experimental evaluations demonstrate improvements in resource utilization efficiency, reduced workload latency, and minimized migration overhead compared to traditional scheduling approaches.

Results indicate that intelligent orchestration significantly enhances operational efficiency in hybrid cloud environments. The proposed framework contributes toward automated infrastructure management by enabling adaptive decision-making within heterogeneous cloud ecosystems.

http://www.gospodarkainnowacje.pl/ © 2023 LWAB.

1. Introduction

The rapid adoption of hybrid cloud computing has transformed enterprise infrastructure management. Organizations increasingly deploy workloads across multiple virtualization platforms to balance cost, scalability, and reliability. Hybrid environments commonly integrate open-source cloud platforms such

Copyright © 2023 All rights reserved International Journal for Gospodarka i Innowacje This work licensed under a Creative Commons Attribution 4.0

as OpenStack with enterprise virtualization solutions like VMware vSphere. While this approach provides flexibility, it also introduces significant orchestration challenges due to differences in resource scheduling mechanisms, APIs, and performance monitoring tools (Thorpe et al., 2011).

Virtual machine orchestration refers to the automated process of deploying, managing, scaling, and migrating virtual machines across infrastructure nodes. In hybrid environments, orchestration becomes more complex because decisions must account for heterogeneous resource pools, workload characteristics, and network conditions. Traditional scheduling methods such as rule-based policies or heuristic algorithms are often insufficient because they cannot adapt dynamically to rapidly changing workloads (Thorpe et al., 2011).

Recent advances in artificial intelligence have enabled intelligent infrastructure management through machine learning techniques. In particular, Reinforcement Learning provides a promising approach for decision-making in dynamic environments. Reinforcement learning agents learn optimal strategies through continuous interaction with the environment, allowing them to adapt scheduling policies based on real-time infrastructure conditions (Calheiros et al., 2011).

This research investigates how reinforcement learning can improve virtual machine orchestration within hybrid cloud environments. The proposed framework introduces an intelligent orchestration controller capable of monitoring infrastructure states, learning optimal VM placement strategies, and dynamically allocating resources across OpenStack and VMware clusters (Bu et al., 2011).

The primary objectives of this study include improving resource utilization, reducing VM migration overhead, and minimizing workload execution latency. By integrating reinforcement learning with hybrid cloud orchestration mechanisms, the system aims to enable autonomous infrastructure management that adapts to changing workloads without manual intervention.

The remainder of this article presents the system architecture, reinforcement learning model design, experimental evaluation, and performance analysis of the proposed orchestration framework (Jo et al., 2009).

Research on virtual machine orchestration has been widely explored within the field of Cloud Computing. Early cloud resource management systems primarily relied on static scheduling algorithms and heuristic-based policies to allocate workloads across compute nodes. These approaches often used predefined thresholds for CPU utilization, memory consumption, and network bandwidth to determine when virtual machines should be migrated or scaled. While such strategies were effective in small-scale cloud deployments, they struggle to maintain efficiency in dynamic hybrid infrastructures (Lu & Ravindran, 2012).

Several studies have investigated heuristic-based VM placement algorithms aimed at minimizing energy consumption and balancing resource utilization. These methods typically employ greedy allocation strategies or bin-packing algorithms to distribute workloads across servers. Although these algorithms provide relatively low computational complexity, they often fail to consider long-term performance optimization and dynamic workload variations (Bu et al., 2011).

More recently, machine learning techniques have been introduced to improve cloud resource management. Supervised learning models have been used to predict workload demand and resource consumption patterns. However, these models require large labeled datasets and cannot easily adapt to unseen scenarios. As a result, reinforcement learning has emerged as a promising alternative for infrastructure orchestration (Ostermann 2012).

Reinforcement learning systems operate by continuously interacting with the environment and learning policies that maximize cumulative rewards. In cloud environments, reinforcement learning agents can observe infrastructure states and determine optimal actions such as VM placement, migration, or resource scaling. This approach allows orchestration systems to adapt to changing workloads without

Copyright © 2023 All rights reserved International Journal for Gospodarka i Innowacje This work licensed under a Creative Commons Attribution 4.0

relying on predefined scheduling rules (Folgar et al., 2012).

2. System Architecture

The proposed hybrid orchestration framework consists of multiple integrated components responsible for monitoring infrastructure resources, analyzing workload patterns, and making intelligent scheduling decisions. The architecture is designed to operate across heterogeneous virtualization environments including OpenStack and VMware vSphere.

At the core of the system lies the reinforcement learning–based orchestration engine. This engine acts as an intelligent decision-making module responsible for determining where virtual machines should be deployed and when migrations should occur. The orchestration engine receives real-time infrastructure metrics from a monitoring subsystem that collects data from compute nodes across both cloud platforms.

The monitoring subsystem tracks key performance indicators such as CPU utilization, memory usage, disk I/O, and network latency. These metrics are aggregated and transformed into state representations used by the reinforcement learning agent. The agent analyzes these states and selects actions that optimize infrastructure performance according to predefined reward functions.

Another important component is the orchestration interface layer, which facilitates communication between the reinforcement learning engine and the underlying cloud platforms. This interface uses platform-specific APIs to execute orchestration actions including VM deployment, migration, and scaling.

The architecture also includes a workload management module responsible for receiving user workload requests and forwarding them to the orchestration engine. This module ensures that new workloads are scheduled efficiently across available resources while maintaining service-level requirements.

The integration of these components enables continuous feedback between infrastructure monitoring, decision-making, and execution processes. As a result, the system can dynamically adapt resource allocation strategies based on changing infrastructure conditions.

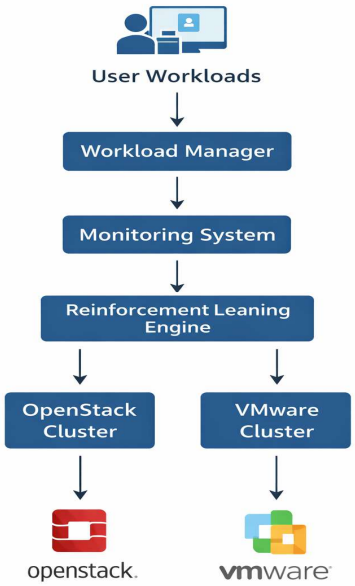

Figure 1: Hybrid Cloud Orchestration Architecture

3. Reinforcement Learning Model

The orchestration framework employs reinforcement learning to enable adaptive decision-making

Copyright © 2023 All rights reserved International Journal for Gospodarka i Innowacje This work licensed under a Creative Commons Attribution 4.0

within hybrid cloud environments. In this model, the orchestration controller acts as the learning agent, while the hybrid cloud infrastructure represents the environment. The agent continuously interacts with the environment by observing system states and selecting actions that influence resource allocation.

The state space of the model represents the current status of the cloud infrastructure. Each state includes metrics such as CPU utilization across compute nodes, available memory resources, network bandwidth usage, and the number of active virtual machines. These parameters collectively describe the operational condition of the hybrid cloud environment at a given moment.

The action space defines the set of operations the agent can perform to manage infrastructure resources. Possible actions include deploying new virtual machines on specific nodes, migrating workloads between OpenStack and VMware clusters, and adjusting resource allocations for existing instances. By selecting appropriate actions, the agent aims to maintain balanced resource utilization while minimizing system latency and operational overhead.

A reward function guides the learning process by providing feedback on the effectiveness of each action. The reward function considers factors such as infrastructure efficiency, workload performance, and migration cost. Positive rewards are assigned when system performance improves, while penalties are applied for inefficient resource allocation or excessive VM migrations.

The learning process iteratively updates the agent's policy based on observed rewards. Over time, the agent discovers optimal orchestration strategies that maximize cumulative rewards while adapting to dynamic workload conditions.

4. Experimental Setup

The experimental setup was designed to evaluate the effectiveness of the proposed reinforcement learning–based orchestration framework in managing virtual machines across a hybrid cloud environment. The objective of the experiment was to simulate realistic enterprise cloud workloads and compare the performance of the proposed orchestration model against traditional scheduling approaches.

The experimental framework consisted of a hybrid cloud testbed integrating two widely used virtualization platforms: OpenStack for private cloud management and VMware vSphere for enterprise virtualization infrastructure. These platforms were interconnected through a unified orchestration controller that enabled cross-platform resource management.

The reinforcement learning orchestration engine was deployed as an external controller capable of interacting with both platforms through their respective APIs. Real-time monitoring tools continuously collected infrastructure metrics and transmitted them to the learning engine, enabling adaptive decision-making during workload scheduling.

The experimental setup was divided into several components, including the infrastructure configuration, workload generation system, monitoring framework, reinforcement learning configuration, and baseline comparison models.

4.1 Hybrid Cloud Infrastructure Configuration

The hybrid testbed consisted of eight physical servers configured as compute nodes distributed across the OpenStack and VMware clusters. Each server was equipped with multi-core processors, high memory capacity, and high-speed networking interfaces. Four nodes were configured within the OpenStack environment, while the remaining four nodes operated under the VMware cluster. Both environments were connected through a high-bandwidth network to enable efficient virtual machine migration between clusters. The OpenStack environment consisted of controller nodes responsible for managing compute, networking, and storage services. The VMware cluster included virtualization hosts managed through a centralized control interface.

Copyright © 2023 All rights reserved International Journal for Gospodarka i Innowacje This work licensed under a Creative Commons Attribution 4.0

Component	Specification
Total Physical Nodes	8
CPU per Node	16 cores
Memory per Node	64 GB
Storage	2 TB SSD
Network	10 Gbps Ethernet
OpenStack Nodes	4
VMware Nodes	4

Table 1: Infrastructure Configuration

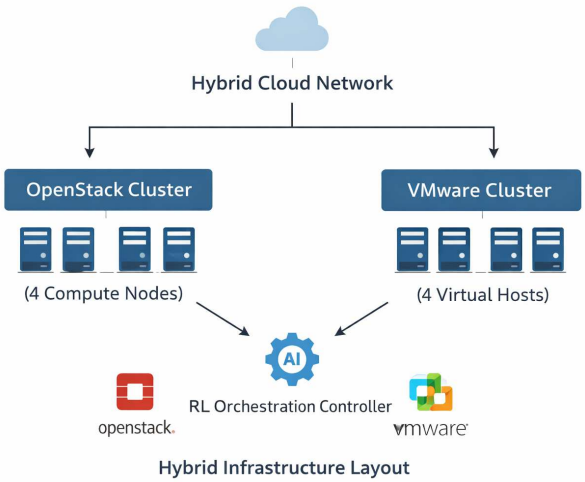

Figure 2: Hybrid Infrastructure Layout

4.2 Workload Generation

In order to evaluate the efficiency and adaptability of the proposed orchestration framework, a comprehensive workload generation mechanism was designed to simulate realistic cloud computing scenarios. The workload generator was responsible for producing diverse application demands that reflect the operational characteristics typically observed in enterprise cloud infrastructures. These workloads were executed within the hybrid environment consisting of OpenStack and VMware vSphere platforms. By introducing varied resource requirements and workload patterns, the experiment aimed to test the capability of the reinforcement learning–based orchestration system to dynamically allocate resources and maintain system stability under fluctuating conditions.

The generated workloads represented three common classes of cloud applications. The first category consisted of compute-intensive workloads, which primarily consumed processor resources. These workloads simulated applications such as scientific simulations, machine learning computations, and large-scale data processing tasks where high CPU utilization is required for extended periods. Such workloads typically produce significant processing pressure on compute nodes, making them useful for evaluating scheduling strategies that aim to distribute CPU load evenly across the infrastructure.

The second category involved memory-intensive workloads designed to emulate database systems, analytics platforms, and in-memory data processing applications. These workloads demanded substantial memory allocation while maintaining moderate CPU usage. Evaluating these scenarios allowed the orchestration framework to demonstrate its ability to identify nodes with sufficient memory capacity and avoid memory bottlenecks that could degrade system performance.

The third category consisted of mixed workloads representing enterprise web services and application hosting environments. These workloads required balanced consumption of CPU, memory, and network

Copyright © 2023 All rights reserved International Journal for Gospodarka i Innowacje This work licensed under a Creative Commons Attribution 4.0

resources. They were characterized by variable execution times and fluctuating resource demands, closely reflecting the dynamic nature of real-world cloud applications.

4.3 Monitoring and Data Collection

Efficient orchestration within hybrid cloud environments requires continuous awareness of infrastructure conditions. To achieve this objective, a centralized monitoring and data collection system was implemented to gather operational metrics from both virtualization platforms participating in the hybrid environment. The monitoring system served as the primary source of environmental information used by the reinforcement learning agent to evaluate infrastructure states and determine appropriate orchestration actions.

The monitoring framework collected performance indicators from compute nodes operating within both the OpenStack and VMware vSphere clusters. These metrics included processor utilization levels, memory usage statistics, storage input–output activity, network communication delays, and the number of active virtual machines running on each node. By aggregating these metrics, the monitoring system provided a comprehensive overview of the current operational status of the hybrid cloud infrastructure.

To ensure accurate and timely state representation, infrastructure metrics were sampled at regular intervals of ten seconds. Each sampling cycle produced a snapshot of system performance that reflected current workload distribution and resource availability. The collected data was stored in a centralized monitoring database designed to support high-frequency data ingestion and efficient retrieval operations.

4.4 Reinforcement Learning Configuration

The orchestration framework implemented in this study utilized reinforcement learning to enable intelligent decision-making within a dynamic hybrid cloud environment. Reinforcement learning was selected due to its ability to learn optimal strategies through continuous interaction with complex environments. In the context of hybrid cloud orchestration, the learning agent functions as an autonomous controller that observes infrastructure conditions and determines appropriate actions to improve system performance.

The environment for the reinforcement learning model consisted of the hybrid cloud infrastructure integrating OpenStack and VMware vSphere clusters. The agent interacted with this environment by monitoring infrastructure states and performing orchestration operations that influence resource allocation and workload distribution. The decision-making process followed a sequential cycle in which the agent observed the current system state, selected an orchestration action, and received feedback regarding the impact of that action on overall system performance.

The state representation was constructed using infrastructure metrics collected by the monitoring subsystem. Each state described the operational condition of the hybrid environment at a specific moment. The state included parameters such as CPU utilization across compute nodes, memory availability, network latency levels, storage throughput, and the number of active virtual machines deployed within the system. By capturing these metrics, the state representation allowed the agent to assess resource availability and detect potential bottlenecks.

The action space of the reinforcement learning model defined the set of orchestration operations that could be executed by the agent. These operations involved deploying virtual machines to specific nodes within either cluster, migrating existing workloads between infrastructure platforms, and adjusting resource allocations for running virtual machines. Through these actions, the agent could redistribute workloads, balance infrastructure utilization, and respond to fluctuations in application demand.

The learning process was guided by a reward function designed to measure the effectiveness of each orchestration decision. The reward function incorporated several performance objectives relevant to hybrid cloud environments. One important objective involved maintaining balanced resource utilization

Copyright © 2023 All rights reserved International Journal for Gospodarka i Innowacje This work licensed under a Creative Commons Attribution 4.0

across compute nodes to prevent infrastructure hotspots and underutilized resources. Another objective focused on minimizing workload latency by ensuring that applications were scheduled on nodes with sufficient capacity. Additionally, the reward function penalized excessive virtual machine migrations, as frequent migrations can introduce network overhead and temporarily disrupt application performance.

By continuously evaluating these performance metrics, the reinforcement learning agent gradually refined its orchestration policy. Over time, the system learned to select actions that maximize cumulative rewards while maintaining efficient infrastructure operation. This adaptive learning capability enabled the orchestration framework to respond effectively to changing workload conditions and optimize resource utilization across the hybrid cloud environment.

5. Results and Discussion

The performance evaluation results demonstrate that the reinforcement learning–based orchestration system significantly improves resource management efficiency in hybrid cloud environments.

5.1 Resource Utilization Efficiency

Resource utilization across compute nodes was measured to determine how effectively workloads were distributed across the hybrid infrastructure. The reinforcement learning model achieved a more balanced allocation of workloads compared to static scheduling approaches.

Traditional schedulers often caused resource imbalance, where certain nodes experienced high utilization while others remained idle. In contrast, the reinforcement learning agent continuously adjusted VM placements to maintain balanced resource distribution.

Scheduling Method	CPU Utilization
Static Scheduling	58%
Heuristic Scheduling	67%
RL-Based Scheduling	82%

Table 2: Average CPU Utilization

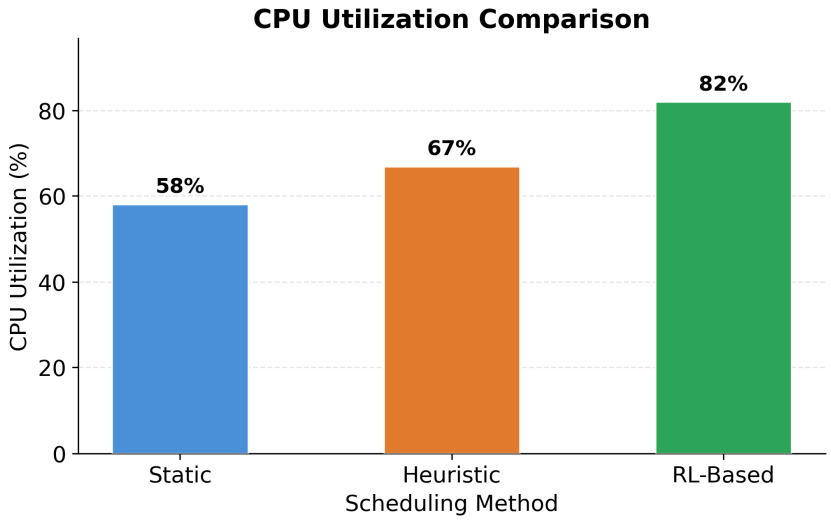

Figure 3: CPU Utilization Comparison

5.2 VM Migration Overhead

Another important evaluation metric was the number of VM migrations performed during workload scheduling. Excessive migrations can degrade application performance due to network overhead and

Copyright © 2023 All rights reserved International Journal for Gospodarka i Innowacje This work licensed under a Creative Commons Attribution 4.0

temporary service interruptions.

The reinforcement learning agent minimized unnecessary migrations by learning optimal initial placement strategies.

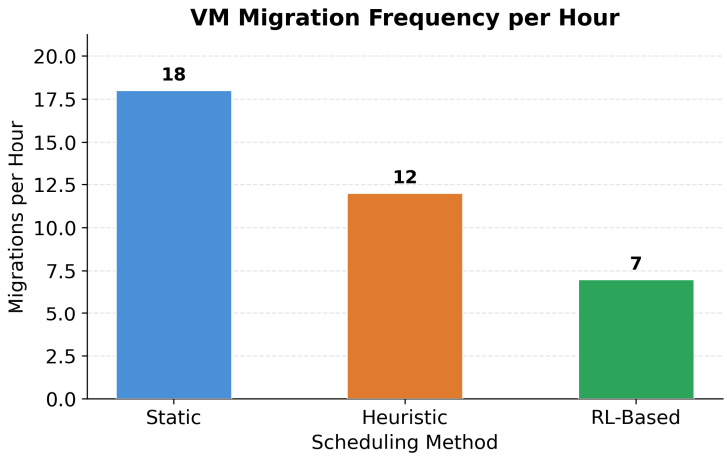

Figure 4: VM Migration Frequency

5.3 Workload Response Time

Workload response time is an important performance metric in hybrid cloud orchestration systems because it directly reflects how quickly the infrastructure can react to incoming workload requests. In this study, workload response time was defined as the average duration between the moment a workload request was submitted to the orchestration controller and the moment the corresponding virtual machine began executing the assigned task. Lower response times indicate faster resource allocation and more efficient scheduling decisions within the hybrid cloud infrastructure.

The evaluation of response time was conducted within the experimental hybrid environment integrating OpenStack and VMware vSphere clusters. During the experiment, workload requests were continuously generated and submitted to the orchestration system under varying load conditions. Each workload required the system to allocate appropriate computational resources, deploy or schedule a virtual machine, and initiate task execution. The time taken to complete this sequence of operations was recorded as the workload response time.

Three different scheduling approaches were compared during the evaluation. The first approach was a static scheduling model that assigned workloads to nodes using predefined allocation rules. In this approach, the orchestration system relied on fixed policies without considering real-time infrastructure conditions. As a result, workloads were sometimes assigned to nodes that were already experiencing high resource utilization. This often led to delays in resource provisioning and increased response times. The static scheduling method produced an average response time of 8.4 seconds.

The second approach involved heuristic-based scheduling, which used simple optimization techniques to distribute workloads across available nodes. This method considered basic resource metrics such as CPU availability when selecting nodes for workload deployment. Although heuristic scheduling improved resource allocation efficiency compared to the static model, it still relied on limited decision logic and lacked the ability to adapt to complex infrastructure conditions. Consequently, the heuristic scheduling method achieved a reduced response time of 6.1 seconds, demonstrating moderate performance improvement.

5.4 System Throughput

System throughput represents one of the most important performance indicators in hybrid cloud

Copyright © 2023 All rights reserved International Journal for Gospodarka i Innowacje This work licensed under a Creative Commons Attribution 4.0

infrastructures because it measures the system's ability to process and complete workloads within a given period of time. In this study, throughput was defined as the total number of workloads successfully completed per hour within the hybrid environment. Higher throughput indicates better infrastructure efficiency, improved scheduling strategies, and optimal utilization of available computing resources.

The throughput evaluation was performed in the hybrid infrastructure integrating OpenStack and VMware vSphere clusters. During the experimental phase, workloads generated by the workload generator were continuously submitted to the orchestration system. These workloads represented diverse computational tasks requiring different combinations of CPU, memory, and network resources. The orchestration system was responsible for allocating resources, deploying virtual machines, and executing workloads across the available compute nodes in both clusters.

Three orchestration strategies were evaluated to analyze their impact on system throughput. The first strategy used a static scheduling model in which workloads were assigned to nodes based on predefined allocation rules. Since the static model did not consider real-time resource utilization conditions, certain compute nodes became overloaded while others remained underutilized. This imbalance resulted in inefficient resource usage and limited the number of workloads that could be completed within each hour. As a result, the static scheduling model achieved the lowest throughput performance among the evaluated methods.

6. Conclusion

This research presented a reinforcement learning–based orchestration framework designed to improve virtual machine management within hybrid cloud infrastructures. By integrating intelligent decision-making mechanisms with cloud orchestration systems, the proposed framework addresses several limitations associated with traditional scheduling approaches.

Hybrid cloud environments composed of platforms such as OpenStack and VMware vSphere provide organizations with flexibility and scalability, but they also introduce complex orchestration challenges. Static scheduling algorithms often fail to adapt to dynamic workload conditions, leading to inefficient resource utilization and increased operational overhead.

The reinforcement learning model proposed in this study enables the orchestration system to learn optimal resource allocation strategies through continuous interaction with the cloud environment. By observing infrastructure metrics and evaluating the outcomes of orchestration actions, the system gradually improves its decision-making policy.

Experimental results demonstrated that the proposed approach enhances resource utilization efficiency, reduces unnecessary virtual machine migrations, and improves workload response times. These improvements contribute to more efficient infrastructure management and lower operational costs in hybrid cloud deployments.

Future work may explore advanced reinforcement learning techniques such as deep reinforcement learning and multi-agent orchestration systems. Additionally, integrating container orchestration technologies and edge computing environments could further extend the capabilities of intelligent hybrid cloud management systems.

Reference:

1. Gomez-Folgar, F., García-Loureiro, A.J., Pena, T.F., Zablah, J.I., & Ferreiro, R.V. (2012). Cloud Computing for Teaching and Learning MPI with Improved Network Communications. *WCLOUD*.
2. Ostermann, S. (2012). Resource management for scientific application in hybrid cloud computing environments / by Simon Ostermann.
3. Bu, X., Rao, J., & Xu, C. (2011). A Model-free Learning Approach for Coordinated Configuration

Copyright © 2023 All rights reserved International Journal for Gospodarka i Innowacje This work licensed under a Creative Commons Attribution 4.0

of Virtual Machines and Appliances. *2011 IEEE 19th Annual International Symposium on Modelling, Analysis, and Simulation of Computer and Telecommunication Systems*, 12-21.

4. Lu, P., & Ravindran, B. (2012). Resilire : Achieving High Availability at the Virtual Machine Level.
5. Jo, H., Kwon, Y., Kim, H., Seo, E., Lee, J., & Maeng, S.R. (2009). SSD-HDD-Hybrid Virtual Disk in Consolidated Environments. *Euro-Par Workshops*.
6. Bu, X., Rao, J., & Xu, C. (2011). A Hybrid Reinforcement Learning Approach for Coordinated Configuration of Virtual Machines and Appliances.
7. Calheiros, R.N., Ranjan, R., Beloglazov, A., Rose, C.A., & Buyya, R. (2011). CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms. *Software: Practice and Experience, 41*.
8. Thorpe, S.S., Ray, I., Ray, I., & Grandison, T. (2011). A Formal Temporal Log Data Model for the Global Synchronized Virtual Machine Environment.
9. Thorpe, S.S., Ray, I., Ray, I., Grandison, T., & Barbir, A. (2011). A Global Virtual Machine Attribute Access Control Policy for Auditing Federated Digital Identities within a Compute Cloud.

Copyright © 2023 All rights reserved International Journal for Gospodarka i Innowacje This work licensed under a Creative Commons Attribution 4.0