Detection and Avoidance Deadlock for Resource Allocation in Heterogeneous Distributed Platforms

Ha Huy Cuong Nguyen¹ and Van Son Le²
¹Department of Information Technology, QuangNam University, Tam Ky City, Viet Nam
²Department of Information Technology, Da Nang University of Education, Da Nang, Viet Nam
¹nguyenhahuycuong@gmail.com, ²levansupham2004@yahoo.com

Abstract— In a heterogeneous distributed platforms large problem is distribute among the processors (such as among the computers of distributed computing system or among the computers of parallel computing system etc.) to make it cost effective and less time consuming instead of a computer system with single processor. There are more general types of resource allocation problems than those we consider here. In this paper, we present an approach for improving detection and avoidance algorithm, to schedule the policies of resource supply for resource allocation on heterogeneous. We propose an algorithm for allocating multiple resources to competing services running in virtual machines platforms.

Index Terms— Cloud Computing, Resource Allocation, Heterogeneous Distributed Platforms, Deadlock Detection and Avoidance Deadlock

I. INTRODUCTION

RECENTLY, there has been a dramatic increase in the popularity of cloud computing systems that rent computing resources on-demand, bill on a pay-as-you-go basis, and multiplex many users on the same physical infrastructure. These cloud computing environments provide an illusion of infinite computing resources to cloud users that they can increase or decrease their resources. In many cases, the need for these resources only exists in a very short period of time.

The increasing use of virtual machine technology in data centers, both leading to and reinforced by recent innovations in the private sector aimed at providing low-maintenance cloud computing services, has driven research into developing algorithms for automatic instance placement and resource allocation on virtualized platforms[1], [2], including our own previous work. Most of this research has assumed a platform consisting of homogeneous nodes connected by a cluster. However, there is a need for algorithms that are applicable to heterogeneous platforms.

Heterogeneity happens when collections of homogeneous resources formerly under different administrative domains are federated and lead to a set of resources that belong to one of several classes. This is the case when federating multiple clusters at one or more geographical locations e.g., grid computing, sky computing.

In this work we propose virtual machine placement and resource allocation deadlock detection algorithms that, unlike previous proposed algorithms, are applicable to virtualized platforms that comprise heterogeneous physical resources. More specifically, our contributions are:

We provide an algorithmic approach to detect deadlock and resource allocation issues in the virtualization platform heterogeneity. This algorithm is in fact more general, even for heterogeneous platforms, and only allowed to allocate minimal resources to meet QoS arbitrary force.

Using this algorithm, we extend previously proposed algorithms to the heterogeneous case.

We evaluate these algorithms via extensive simulation experiments, using statistical distributions of application resource requirements based on a real-world dataset provided by Google. Most resource allocation algorithms rely on estimates regarding the resource needed for virtual machine instances, and do not refer to the issue of detecting and preventing deadlocks. We studied the impact of estimation errors and propose different approaches to mitigate these errors, and identify a strategy that works well empirically.

II. RELATED WORKS

Resource allocation in cloud computing has attracted the attention of the research community in the last few years. Srikantaiah et al. [8] studied the problem of request scheduling for multi-tiered web applications in virtualized heterogeneous systems in order to minimize energy consumption while meeting performance requirements. They proposed a heuristic for a multidimensional bin packing problem as an algorithm for workload consolidation. Garg et al. [10] proposed near optimal scheduling policies that consider a number of energy efficiency factors, which change across different data centers depending on their location, architectural design, and management system. Warneke et al. [11] discussed the challenges and opportunities for efficient data processing in cloud environment and presented a data
processing framework to exploit the dynamic resource provisioning offered by IaaS clouds. Wu et al. [12] propose a resource allocation for SaaS providers who want to minimize infrastructure cost and SLA violations. Addis et al. [13] proposed resource allocation policies for the management of multi-tier virtualized cloud systems with the aim to maximize the profits associated with multiple class SLAs. A heuristic solution based on a local search that also provides availability, guarantees that running applications have developed. Abdelsalem et al. [14] created a mathematical model for power management for a cloud computing environment that primarily serves clients with interactive applications such as web services. The mathematical model computes the optimal number of servers and the frequencies at which they should run. Yazir et al. [15] introduced a new approach for dynamic autonomous resource management in computing clouds. Their approach consists of a distributed architecture of NAs that perform resource configurations using MCDA with the PROMETHEE method. Our previous works mainly dealt with resource allocation, QoS optimization in the cloud computing environment.

There are more general types of resource allocation problems than those we consider here. For instance:

1. We consider the possibility that user might be willing to accept alternative combinations of resources. For example, a user might request elementary capacity CPU, RAM, HDD rather than a specific.

2. We consider the possibility that resources might be shared. In this case, some sharing is typically permitted; for example, two transactions that need only to read an object can be allowed concurrent access to the object.

3. We begin by defining our generalized resource allocation problem, including the deadlock detection problem as an interesting special case. We then give several typical solutions.

III. SYSTEM MODEL RESOURCE ALLOCATION IN HETEROGENEOUS DISTRIBUTED PLATFORMS

Resource allocation in cloud computing has attracted the attention of the research community in the last few years. Cloud computing presents a different resource allocation paradigm than either grids or batch schedulers [2]. In particular, Amazon C2 [10], is equipped to handle may smaller computer resource allocations, rather than a few, large request as is normally the case with grid computing. The introduction of heterogeneity allows clouds to be competitive with traditional distributed computing systems, which often consist of various types of architecture as well.

Like traditional distributed system before we can see a heterogeneous distributed system consists of a set of processes that are connected by a communication network. The communication delay is finite but unpredictable [21], [22].

A). The Application

A heterogeneous distributed program is composed of a set of n asynchronous processes p1, p2,...,pn that communicates by message passing over the communication network. We assume that each process is running on a different processor. The processor does not share a common global memory and communicate solely by passing messages over the communication network. There is no physical global clock in the system to which processes have instantaneous access. The communication medium may deliver messages out of order, messages may be lost garble or duplicated due to timeout and retransmission, processors may fail and communication links may go down. The system can be modeled as a directed graph in which vertices represent the processes and edge represent unidirectional communication channels.

We use the platform graph, for the grid platform. We model a collection of heterogeneous resources and the communication links between them as the nodes and edges of an undirected graph. See an example in Fig. 1 with 8 processors and 11 communication links. Each node is a computing resource (a processor, or a cluster, or node).

![Fig. 1. An example simple platform](image)

A process can be in two states: running or blocked. In the running state (also called active state), a process has all the needed re and is either executing or is ready for execution. In the blocked state, a process is waiting to acquire some resource.

B). The Architecture

The target heterogeneous platform is represented by a directed graph, the platform graph. There are p nodes P1, P2,..., Pn that represent the processors. In the example of figure 1 there at eight processors, hence n = 8.

Each edge represents a physical interconnection. Each edge ej: Pi → Pj is labeled by value cij, which represents the time to transfer a message of unit length between Pi and Pj, in either direction: we assume that the link between Pi and Pj is bidirectional and symmetric. A variant would be to assume two unidirectional links, one in each direction, with possibly different label values. If there is no communication link between Pi and Pj, we let cij = +∞, so that cij < +∞ means that Pi and Pj are neighbors in the communication graph.

C). Wait – For – Graph (WFG)

In distributed systems, the state of the system can be modeled by directed graph, called a wait for graph (WFG) [21] – [25]. In a WFG, nodes are processes and there is a directed edge from node Pi to node Pj if Pi is blocked and is
Waiting for $P_2$ to release some resource. A system is deadlocked if and only if there exists a directed cycle or knot in the WFG.

Let us first of all describe the deadlock condition problem more precisely.

A set $S = \{s_1, s_2, \ldots, s_k\} \subseteq \varepsilon$ of $k > 1$ entities is deadlocked when the following two conditions simultaneously hold:

Each entity $s_i \in S$ is waiting for an event permission that must be generated from another entity in the set;

No entity $s_i \in S$ can generate a permission while it is waiting.

If these two conditions hold, the entities in the set will be waiting forever, regardless of the nature of the permission and of why they are waiting for the "permission"; for example, it could be because $s_i$ needs a resource held by $s_j$ in order to complete its computation.

A useful way to understand the situations in which deadlock may occur is to describe the status of the entities during a computation, with respect to their waiting for some events, by means of a directed graph $W$, called wait-for graph.

Deadlock detection can be represented by a Resource Allocation Graph (RAG), commonly used in operating systems and distributed systems. A RAG is defined as a graph $(V,E)$ where $V$ is a set of nodes and $E$ is a set of ordered pairs or edges $(v_i, v_j)$ such that $v_i, v_j \in V$. $V$ is further divided into a two disjoint subsets: $P = \{p_0, p_1, p_2, \ldots, p_m\}$ where $P$ is a set of processor nodes shown as circles in Figure 1; and $Q = \{q_0, q_1, q_2, \ldots, q_n\}$ where $Q$ is a set of resource nodes shown as boxes in Figure 1. A RAG is a graph bipartite in the $P$ and $Q$ sets. An edge $e_{ij} = (p_i, q_j)$ is a request edge if and only if $p_i \in P$, $q_j \in Q$. The maximum number of edges in a RAG is $m \times n$. A node is a sink when a resource (processor) has only incoming edge(s) from processor(s) (resource(s)). A node is source when a resource (processor) has only outgoing edge(s) to processor(s) (resource(s)). A path is a sequence of edges

$\varepsilon = \{(p_{i_1}, q_{j_1}), (q_{j_1}, p_{i_2}), \ldots, (p_{i_k}, q_{j_k}), (q_{j_k}, p_{i_{k+1}})\}$

where $\varepsilon \in E$. If a path starts from and ends at the same node, then it is a cycle. A cycle does not contain any sink or source nodes.

The focus of this paper is deadlock detection. For our virtual machine resource allocation on heterogeneous distributed platforms deadlock detection implementation, we make three assumptions. First, each resource type has one unit. Thus, a cycle is a sufficient condition for deadlock [3]. Second, satisfies request will be granted immediately, making the overall system expedient [3]. Thus, a processor is blocked only if it cannot obtain the requests at the same time.

All proposed algorithms, including those based on a RAG, have $O(m \times n)$ in the worst case. In this paper, we propose deadlock detection algorithm with $O(\min(m,n))$ based on a new matrix representation. The proposed virtual machine resource allocation on heterogeneous distributed platforms deadlock detection algorithm makes use of ism and can handle multiple requests/grants, making the proposed algorithm faster than the $O(m \times n)$ algorithm [16], [17].

IV. DEADLOCK DETECTION FOR RESOURCE ALLOCATION IN HETEROGENEOUS DISTRIBUTED PLATFORMS

In this section, we will first introduce the matrix representation of a deadlock detection problem. The algorithm is based on this matrix representation. Next, we present some essential features of the proposed algorithm. This algorithm is , and thus can be mapped into a cloud architecture which can handle multiple requests/grants simultaneously and can detect multiple deadlocks in linear time, hence, significantly improving performance.

A). Matrix Representation of a Deadlock Detection Problem

In graph theory, any directed graph can be represented with an adjacency matrix [3]. Thus, we can represent a RAG with an adjacency matrix. However, there are two kinds of edges in a RAG: grant edges, which point from resources to processors, and request edges, which point from processors to resources. To distinguish different edges, we designate elements in the adjacency matrix with three different values as shown in Figure 2. This Figure shows the matrix representation of a given system with processors $p_1, p_2, \ldots, p_m$ and resources $q_1, q_2, \ldots, q_n$. The leftmost column is the processors label column. The top row is the resources label row. If there is a request edge $(p_i, q_j)$ in the RAG, corresponding element in the matrix is $r$. If there is a grant edge $(q_j, p_i)$ in the RAG. The corresponding element in the matrix is $g$. Otherwise, the value of the element is 0.

This variant of the adjacency matrix of a RAG $(V,E)$ can be defined formally as follows:

$M = [m_{ij}]^{m \times n}$. $(1 \leq i \leq m, 1 \leq j \leq n)$, where $m$ is the number of processors and $n$ is the number of resources.

$m_{ij} \in \{r,g,0\}$

$m_{ij} = r, \text{ if } \exists (p_i, q_j) \in E$

$m_{ij} = g, \text{ if } \exists (p_i, q_j) \in E$

$m_{ij} = 0, \text{ if otherwise}$

This matrix provides a template able to represent request and grant combinations. Note that each resource has at most one grant, that is, there is at most one $g$ in a column at any time. However, there is no constraint on the number of requests from each processor.

If there are deadlocks in a system, there must be at least one cycle in its RAG, that is, there must be a sequence of edges, $\varepsilon = \{(p_{i_1}, q_{j_1}), (q_{j_1}, p_{i_2}), \ldots, (p_{i_k}, q_{j_k}), (q_{j_k}, p_{i_{k+1}})\}$, where $\varepsilon \subseteq E$. In the matrix representation, this cycle is mapped into a sequence of matrix elements $\omega = \{m_{i_1,j_1}, m_{i_1,j_2}, \ldots, m_{i_1,j_k}, m_{i_2,j_1}, m_{i_2,j_2}, \ldots, m_{i_2,j_k}, \ldots, m_{k,j_1}, \ldots, m_{k,j_k}\}$ where are requests(r’s) and $m_{i,j}$ are grants (g’s). By this fact, we can detect deadlocks in a system with its adjacency matrix. Next, we will present the new detection algorithm.
B). Deadlock Detection Algorithm

On this basis of the matrix representation, we propose a deadlock detection algorithm. The basic idea in this algorithm is iteratively reducing the matrix by removing those columns or rows corresponding to any of the following cases:

- a row or column of all 0's;
- a source (a row with one or more r's but no g's, or a column with one g and no r's);
- a sink (a row with one or more g's but no r's, or a column with one r's but no g's);

This continues until the matrix cannot be reduced any more. At this time, if the matrix still contains row(s) or column(s) in which there are non-zero elements, then there is at least one deadlock. Otherwise, there is no deadlock. The description of this method is shown in algorithm.

**Algorithm: Deadlock Detection Algorithm**

**Input:** \( P_i \) \( i = 1, \ldots, m \) from IaaS provider \( i \);

**Step 1:** Calculate optimal resource allocation to provide VM. \( x_i^{(CPU)} \), \( x_i^{(RAM)} \) = \( \text{Max}\{U_{IaaS}\} \);

**Step 2:** Computes new resource
- If \( C_j^{CPU} \geq \sum_i x_i^{(CPU)} \), \( C_j^{RAM} \geq \sum_i x_i^{(RAM)} \) then
  - \( r_j^{CPU(n+1)} = \max\{e_i^{CPU(n)} + n(\sum_i x_i^{(CPU)} - C_j^{CPU})\} \);
  - \( r_j^{RAM(n+1)} = \max\{e_i^{RAM(n)} + n(\sum_i x_i^{(RAM)} - C_j^{RAM})\} \);

**Return new resource** \( r_j^{CPU(n+1)} ; r_j^{RAM(n+1)} \)

**Else**

**Step 3:** Initialization
- \( M = \{ m_{ij} \}_{i,j=1}^{m,n} \),

Where \( m_{ij} \in \{r,g,0\}, \ (i = 1, \ldots, m \text{ and } j = 1,\ldots,n) \)
- \( m_{ij} = r \) if \( \exists (p,q) \in E \).
- \( m_{ij} = g \) if \( \exists (p,q) \in E \).
- \( m_{ij} = 0 \), otherwise.

**Step 4:** Remove all sink and sources

**DO**

- Reducible = 0;
  - For each column:
    - if \( (\exists m_{ij} \in \forall k, k \neq i, m_{kj} \in \{m_{ij}, 0\}) \{
      - \( \Lambda_{column} = \Lambda - \{m_{ij} \mid j = 1,2,3,\ldots,m\} \),
      - reducible = 1;
    \}

- For each row:

**UNTIL reducible = 0**;

**Step 5:** Detect Deadlock
- If \( (\Lambda \neq 0) \), then return deadlock exits.
- If \( (\Lambda = 0) \), then return no deadlock exits.

**Output:** new resource \( r_j^{CPU(n+1)} ; r_j^{RAM(n+1)} \)

The following example illustrates how the algorithm works, in each iteration of this algorithm; at least one reduction can be performed if the matrix is reducible. Hence, it takes at most \( \min(m,n) \) iterations to complete the deadlock.

C). Our Algorithm to Avoid Deadlock

In heterogeneous distributed platforms if there are \( n \) processes for each processor and \( m \) resources the algorithm will be as follow:

Step 1: In any of the processes of a processor using one of the resources in a specific time period (i.e., the period-1) then no other requirements for the resources will be allocated for the specified time period.

Step 2: The resource occupied can be used by other processes within a specific time period (i.e., the period-2) only after the release by the process of step 1 by approximately time-1. During the period-2 has no other process will be able to use that resource.

Step 3: In the same way, a process of the processor is determined in step 2 can use an O resources in a specific time period (i.e., the period-3) in the absence the other will be able to use that resource.

Step 4: The resources that were occupied in step 3 can be used by other processes defined in step 1 in a specific time period (i.e., the period-4) only after released by the process of step 3 by approximately-3. During the period-4 has no other process will be able to use that resource.

D). Proof of the Correctness of DDA

**Example 2 State matrix representation**

The system in state shown in Fig. 2 (a) can be representation in the matrix form shown in (b). For the sake of better understanding, we will describe it in the matrix representation, shown in Fig. 3 (c) from now on.
The first case generated the request using a normal distribution of arrival time. This determines the performance of the algorithms to handle the arrival of tasks in an exponentially increasing number of request. There were up 100 request generated and also these requests were assigned in randomly.

The comparative analysis of experimental result can be seen in many times, after task execution, although there were individual time improved DDA algorithm response time was not significantly less than optimal time algorithm, in most cases, improved algorithm is better than the optimal time algorithm, thus validated the correctness and effectiveness.

VI. CONCLUSION

Deadlock is a highly unfavorable situation that can occur in any multiprocessor system. The occurrence of a deadlock can cripple parts of a multiprocessor system. It is necessary to develop control policies that avoid deadlocks by restricting the freedom in resource allocation. In this paper we have presented a fast deadlock detection and deadlock avoidance methodology that is easily applicable to heterogeneous distributed platforms. Once a deadlock is detected, it must be somehow resolved. In a heterogeneous distributed platforms large problem is distribute among the processors (such as among the computers of distributed computing system or among the computers of parallel computing system etc.) to make it cost effective and less time consuming instead of a computer system with single processor. But since the large program is divided into smaller programs chances of occurring deadlock becomes high. So, it is very necessary to apply deadlock avoidance method in multiprocessor system. In this paper we have also provided the algorithm which simple and easy to understand. We can also implement this method everywhere in our computational life as needed.

REFERENCES


