A Study of Resource Management for Fault-Tolerant and Energy Efficient Cloud Datacenter

Dong-Ki Kang, Fawaz Alhazemi, Seong-Hwan Kim, and Chan-Hyun Youn(✉)

School of Electrical Engineering, KAIST, Daejeon, Korea
{dkkang, fawaz, s.h_kim, chyoun}@kaist.ac.kr

Abstract. In cloud computing datacenters, the reliability and energy consumption have been studied as main challenges to achieve the reputation of cloud service users and the cost efficiency. To overcome the system fault of the datacenter, VM request load has to be distributed on multiple hosts to minimize the effect to the running cloud applications. Moreover, Dynamic Right Sizing (DRS) which adjusts the number of active hosts and sleep hosts in order to reduce the energy consumption in view of the resource usage cost. To do this, we propose the resource management scheme based on the portfolio diversification which has been studied in economics. The proposed scheme is able to reduce the fault of application significantly by finding the near Pareto optimal solution through Simulated Annealing approach We show the efficiency of our proposed scheme through the simple analytical results.

Keywords: Fault-tolerant · Cloud computing · Energy efficient · Resource management · Load balancing

1 Introduction

Nowadays, cloud computing datacenters are facing the system fault and the power consumption problems. In google datacenter [1], the thousand servers experience the fault at least in the first year per one cluster which consists of 1800 servers in general and the hard disk failure is the main factor of the system fault. Moreover, if the error is occurred in the power distribution unit of the datacenter, then 500 $\sim$ 1000 server machines are disabled during 6 h at least. In general, 50 percents of the cluster are experience overheat.

To solve this problem, many traditional fault-tolerant schemes have employed the data replication mechanism. That is, they replicate the data of the cloud service application to the backup host, and recover the replicated data from the backup host

D.-K. Kang—Please note that the LNICST Editorial assumes that all authors have used the western naming convention, with given names preceding surnames. This determines the structure of the names in the running heads and the author index.

Y. Zhang et al. (Eds.): CloudComp 2015, LNICST 167, pp. 22–29, 2016.
DOI: 10.1007/978-3-319-38904-2_3

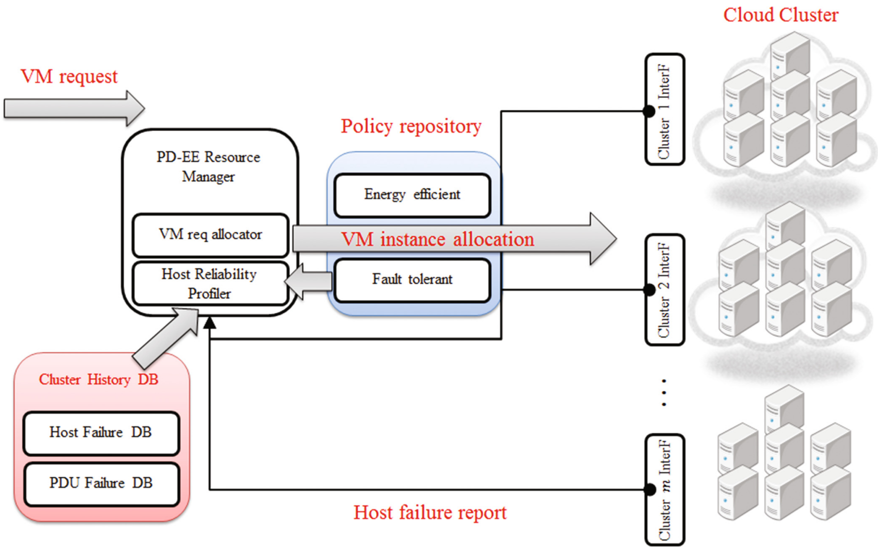


Fig. 1. PD-EE resource manager structure

when the failure is occured in the original host. However, this procedures require the additional host to do replication, this causes the performance overhead by replication process and increases the resource usage cost. In addition, the inconsistency problem may be occurred between the origianl data and replicated data, therefore the synchronization is required consistently. This is inefficient.

In this paper, we propose the dynamic resource balancing algorithm for fault-tolerant resource management without any inefficient replication scheme. Our proposed algorithm is based on the portfolio diversification in economics. In the portfoilo diversification the asset is not invested into the single items but rather is distributed to multiple items, therefore the economic risk is able to be minimized. In this mechanism, the reliablity is increased according the average value and the risk is decreased by reducing the variation value. We find the resource allocation solution with maximized average and minimized variation by searching the near Pareto optimal solution since the allocation solution is kind of the multi objective problems. We apply the Simulated Annealing Procedure as a heuristic algorithm to derive the desirable VM balancing solution. The additional objective for the datacenter resource management is the energy consumption which is a main part of the resource usage cost. The VM instance balancing solution increases the number of active host, the resource usage cost is increased by increasing power consumption. Therefore, it is important to derive the resource management strategy to achieve the reasonable fault-tolerancy with energy efficient resource allocation. To do this, we propose the algorithm satisfying these two objectives. We show that our proposed algorithm outperforms with the random and existing packing schemes through the simulation testbed based on nodeJS servers.

2 System Structure

Figure 1 shows the structure of our proposed Portfolio Diversification based Energy Efficient Resource Manager (PD-EE RM) module. When the VM request is submitted to the interface of the system, PD-EE RM determines the desirable host of the cloud cluster by predefined VM packing scheme and allocate the VM instance through the VM request allocator. In this case, the weight values for the energy efficiency and fault-tolerancy are set in the Policy repository module. In Cluster History DB, the reports of the not-working Power Distribution Unit (PDU) or not-working host in each distributed cluster are written. The reliability of the whole datacenter is derived based on the stored DB in the PE-EE RM module and the parameters for the resource allocation are adjusted according to the derived reliability.

3 Portfolio Diversification Based Resource Management Scheme

In this chapter, we describe the resource management strategy based on the Portfolio diversification. Our proposed scheme is based on following two equations.

$$E[R_x] = \sum\nolimits_{i=1}^{n} w_i E[R_i] = w_1\mu_1 + \cdots + w_n\mu_n \tag{1}$$

$$\begin{aligned} Var(R_x) = E[R_x - E[R_x]]^2 = \sum\nolimits_{i=1}^{n} w_i^2\sigma_i^2 + \sum\nolimits_{i \neq j} w_i w_j Cov\left[R_i, R_j\right] = \\ \sum\nolimits_{i=1}^{n} w_i^2\sigma_i^2 + \sum\nolimits_{i \neq j} w_i w_j \sigma_i \sigma_j \frac{Cov\left[R_i, R_j\right]}{\sigma_i \sigma_j} = \sum\nolimits_{i=1}^{n} w_i^2\sigma_i^2 + \sum\nolimits_{i \neq j} w_i w_j \sigma_i \sigma_j \rho_{ij} \end{aligned} \tag{2}$$

where x is the resource allocation solution, and w_i is the allocated weight for the host i. μ_i represents the average reliability of the host i, and n is the number of physical hosts. $E[R_x]$ is the expected success ratio of the resource allocation solution x. σ_i is the standard deviation of the physical hosts and $Cov\left[R_i, R_j\right]$ is the covariation between the physical host i and the physical host j. ρ_{ij} is the correlation of the physical host i and physical host j, this value is same to $\frac{Cov\left[R_i, R_j\right]}{\sigma_i \sigma_j}$. That is, the average value in Eq. (1) represents the reliability of the cloud resource allocation and as the bigger, the better. The variation value in Eq. (2) represents the risk of cloud resource allocation. For instance, there are two resource allocation solution A and B, and the reliability of A and B is average 80 % in the same. In this case, if the variation of A is 20 and the one of B is 10, then it means that the reliability of B is better than the one of B because the risk of A is bigger than B even the average values are same. When the failure is occurred under A, whether all the VM instances are free from the failure or many VM instances are affected, the moderate effect is not exist. Contrastively, some VM instance are affected when the failure is occured under B, however the effect level is smaller than the case of A. That is, the smaller variation, the better reliability.

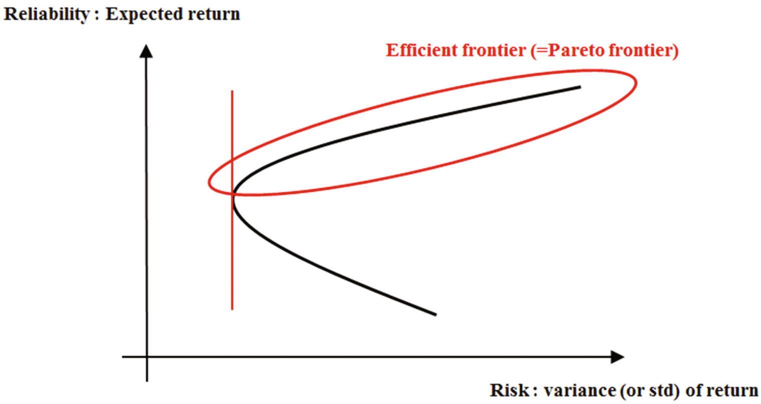


Fig. 2. The curve of average and variation for reliability

Figure 2 shows the performance curve in views of reliability with average and risk as evaluation metrics of the resource allocation solution. The upper part of the curve in graph represents the Pareto frontier of the reliability and the risk. That is, the points within the Pareto frontier represents the optimized resource allocation solutions in view of the fault-tolerant and the points outside the Pareto frontier are not optimized resource allocation solutions.

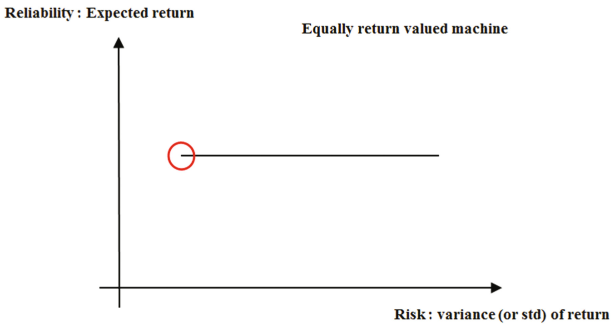


Fig. 3. Reliability and risk curve in the homogeneous host set

Figure 3 shows the performance curve of the homogeneous host set with the same reliability values. In this case, every resource allocation solutions have same fixed avaerage value and only variation values are different. The variation value is minimized when the VM instance workload is distributed perfectly and this solution is the only one optimized solution. Based on Eqs. (1) and (2), we apply the heuristic scheme to find the near optimal resource allocation solution. Our deployed heuristic scheme is based on Simulated Annealing search [3] and this is kind of the local searching techniques. This scheme searches optimal solution by adjusting the searching direction based on the solution searching running time. Our proposed algorithm based on Simulated Annealing is shown in Table 1.

Table 1. Algorithm 1 simulated annealing based VM allocation

Algorithm 1. Simulated Annealing based VM allocation scheme 1. select the random resource allocation solution. 2. Evaluate the reliability and risk by calculating the reliability average and variation of each neighbor solution of the staring solution. 3. Get the weighted sum of reliability and the risk. 4. Get the objective funtion values of each solution and derive the switching probability according to the function values. 5. Select next solution according to the switching probability 6. Update the temperature parameter 7. When the delta E values of all the neighbor slutions are under the predefined threshold value, the searching of the solution is stop and final resource allocation solution is chosen.

Algorithm 1 considers the VM allocation policy with Simulated Annealing scheme. The objective function based on reliability and variation values is defined by using weighted sum approach as follows,

$$f(x) = \alpha \cdot E(R_x) + \beta \cdot Var(R_x) \tag{3}$$

where α is the weight for average value of the solution x and β is the weight for the variation value representing the risk. The switching probability of the neighbor solutions is calculating according to the objective function value as follows,

$$p\left(x_{neigh}\right) = e^{\frac{\Delta E = f\left(x_{neigh}\right) - f\left(x_{origin}\right)}{kT}} \tag{4}$$

$$T_t = z \times T_{t-1}, 0 < z < 1 \tag{5}$$

where k is Bolzman constant. During the solution searching procedure, the temperature value T is decreased and the switching probability to the worse solution is near zero. When the delta E is under the threshold E_{th}, the solution searching is finished and the last chosen solution is the final solution.

4 Analytical Results

In this chapter, we show the efficiency of our proposed algorithm through simple analytical results. Figure 4 shows the the example of the VM instance allocation of the 5 homogeneous host machines with the same success rate and failure rate. The number of allocated VM instances is 5 and the their required flavor types are same. The failure return is represented as the number 0 and the success return is 1, the average reliability of each host is 0.8 and all the values are same. Figure 5 shows the variance and covariance of all the possible host pairs. The variances and the covariances are same since the average reliability of all the homogeneous host is same.

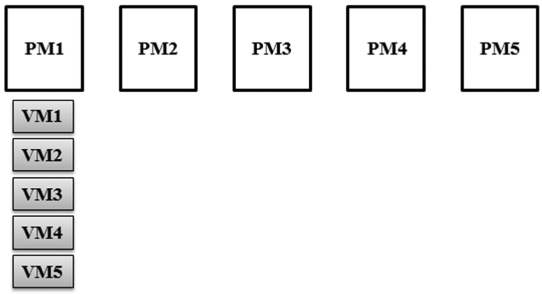
Fig. 4. Five host machines with same success rate

	Variance / covariance				
	PM1	PM2	PM3	PM4	PM5
PM1	$\sigma_1^2 = 0.34$	$\sigma_{12} = 0.16$	$\sigma_{13} = 0.16$	$\sigma_{14} = 0.16$	$\sigma_{15} = 0.16$
PM2	$\sigma_{21} = 0.16$	$\sigma_2^2 = 0.34$	$\sigma_{23} = 0.16$	$\sigma_{24} = 0.16$	$\sigma_{25} = 0.16$
PM3	$\sigma_{31} = 0.16$	$\sigma_{32} = 0.16$	$\sigma_3^2 = 0.34$	$\sigma_{34} = 0.16$	$\sigma_{35} = 0.16$
PM4	$\sigma_{41} = 0.16$	$\sigma_{42} = 0.16$	$\sigma_{43} = 0.16$	$\sigma_4^2 = 0.34$	$\sigma_{45} = 0.16$
PM5	$\sigma_{51} = 0.16$	$\sigma_{52} = 0.16$	$\sigma_{53} = 0.16$	$\sigma_{54} = 0.16$	$\sigma_5^2 = 0.34$

Fig. 5. Variance and covairance with each host pair

Figure 6 shows the case of the all-in VM consolidation. In this case, the weight of host 1 is one and the weights of other hosts are zero since all the VM instances are allocated to the host 1. The average reliability of the allocated solution is 0.8 and the variation is 0.1156. Figure 7 shows the case of equally weighted resource allocation which distributes all the VM instances to the hosts. The number of VM instance is five and the number of hosts is also five, each host has one allocated Vm instance. In this

Case of x_1 : all-in allocation (maximized consolidation)



$$w_i = \frac{|VMs\ on\ PM_i|}{|Total\ VMs|}$$

$$w_1 = \frac{5}{5} = 1 \qquad w_2 = \frac{0}{5} = 0 \qquad w_3 = \frac{0}{5} = 0 \qquad w_4 = \frac{0}{5} = 0 \qquad w_5 = \frac{0}{5} = 0$$

$$E[R_{x_1}] = 1 * (0.8) = 0.8$$

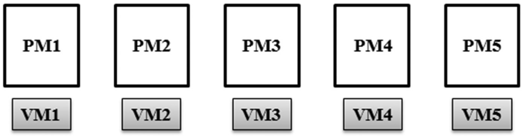
$$Var(R_{x_1}) = (1^2 \times 0.34^2) = 0.1156$$

Fig. 6. Average and variance values of all-in allocation

case, the weight of host is 0.2 and the average relilability is 0.8, these values are same to the case of Fig. 6. However, the variance of the solution is 0.087 and this value is smaller than one in Fig. 6. This means that there is no error in every running VM instances if the failure is occurred from host 2 to host 5 but all the VM instances experience error if the failure is occurred in the host 1 in the case of Fig. 6. In this case, the risk is high. In the case of Fig. 7, there is absolutely the error VM instance if the failure is occurred from host 1 to host 5, but the number of error occurred VM instance is just one. That is, other 4 VM instances do not have any error. In this case, the risk is small.

In conclusion, we derive that the case of Fig. 6 has the better performance than the case of Fig. 7.

Case of x_2 : equally weighted allocation (maximized balancing)



$$w_1 = \frac{1}{5} = 0.2 \quad w_2 = \frac{1}{5} = 0.2 \quad w_3 = \frac{1}{5} = 0.2 \quad w_4 = \frac{1}{5} = 0.2 \quad w_5 = \frac{1}{5} = 0.2$$

$$E[R_{x_2}] = 5 * (0.2 * 0.8) = 0.8$$

$$Var(R_{x_2}) = 5 \times (0.2^2 \times 0.34^2) + 10 \times (0.2 \times 0.2 \times 0.16) = 0.08712$$

Fig. 7. Average and variance values of equally weighted allocation

5 Conclusion

We propose the fault-tolerant VM instance allocation scheme for the cloud computing environment. Our proposed scheme uses the average and variation of the reliability of the allocation solution based on the portfolio diversification. In order to find the near Pareto optimal solution satisfying both of the reasonable average and varaition, the Simulated Annealing heuristic is imployed in the proposed algorithm. In the analytical results, the performance of each solution with simple homogeneous hosts and VM instance submission. In ongoing works, we demonstrate that our proposed algorithm outperforms traditional algorithms through various simulation results.

Acknowledgments. This work was supported by 'Electrically phase-controlled beamforming lighting device based on 2D nano-photonic phased array for lidar' grant from Civil Military Technology Cooperation, Korea.

References

1. Google. www.google.com
2. Belly, Z.Y.: Socially responsible investing and portfolio diversification. J. Financ. Res. **28**(1), 41–57 (2005)

3. Bandyopadhyay, S., Saha, S., Maulik, U., Deb, K.: A simulated annealing-based multiobjective optimization algorithm: AMOSA. IEEE Trans. Evol. Comput. **12**(3), 269–283 (2008)
4. Xiao, Z., Song, W., Chen, Q.: Dynamic resource allocation using virtual machines for cloud computing environment. IEEE Trans. Parallel Distrib. Syst. **24**(6), 1107–1117 (2013)
5. A-Eldin, A., Tordsson, J., Elmroth, E., Kihl, M.: Workload classfication for efficient auto-scaling of cloud resources. Umea University, Sweden (2013)