

Energy-aware load balancing in distributed data centers

Rodrigo A. C. da Silva¹, Nelson L. S. da Fonseca¹

¹Instituto de Computação – Universidade Estadual de Campinas (UNICAMP)
Campinas - SP - Brazil

rodrigo@lrc.ic.unicamp.br, nelson@ic.unicamp.br

Abstract. *This paper summarizes the dissertation "Energy-aware load balancing in distributed data centers", which proposed two new algorithms for minimizing energy consumption in cloud data centers. Both algorithms consider hierarchical data center network topologies and requests for the allocation of groups of virtual machines (VMs). The Topology-aware Virtual Machine Placement (TAVMP) algorithm deals with the placement of virtual machines in a single data center. It reduces the blocking of requests and yet maintains acceptable levels of energy consumption. The Topology-aware Virtual Machine Selection (TAVMS) algorithm chooses sets of VM groups for migration between different data centers. Its employment leads to relevant overall energy savings.*

Introduction

Data processing and service provisioning have been intensively migrated to the cloud due to the flexibility of obtaining service on demand on a pay per use basis. In infrastructure as a Service (IaaS), users rent virtual machines (VMs) instantiated in data centers, which typically contain thousands of servers. Moreover, cloud providers may own several data centers for the processing of their demand.

A big concern in cloud computing is the energy consumption of data centers. In 2010, for instance, data centers consumed about 1.5% of all the energy consumed in the world. Moreover, energy consumption grows faster than the energy efficiency of servers, storage and network equipment [Koomey 2011]. Therefore, strategies for management of resources need to be developed to decrease the energy consumption. Large energy consumptions affect the environment due to the emissions of pollutant gases and strongly impact the financial costs of service providers. Therefore, decreasing the energy consumption in data centers impacts not only the cloud operation but also Society as a whole.

This paper summarizes the dissertation "Energy-aware load balancing in distributed data centers" [da Silva and da Fonseca 2015b] which proposed a load balancing scheme among the data centers of a provider with the aim of decreasing the energy consumption. Two novel algorithms were proposed. The Topology-aware Virtual Machine Placement (TAVMP) algorithm maps VM groups on a data center in a way to avoid unnecessary activation of servers and network switches. Results show that it is possible to reduce blocking of requests while maintaining acceptable levels of energy consumption. The second algorithm, the Topology-aware Virtual Machine Selection (TAVMS) algorithm chooses sets of groups of virtual machines for potential migration to other data centers when it will decrease the overall energy consumption. Results show that relevant energy savings can be achieved by employing TAVMS. The novelty of these algorithms

relies on the consideration of the data center network topology on resource allocation decisions, the characterization of users' demands as arrivals of groups of virtual machines, and the consideration of realistic inter-VM communications.

The remaining of this paper is organized as follows. Section 2 presents related work. Sections 3 and 4 summarize the first and second algorithms of the dissertation, respectively. Section 5 presents the main results obtained. Section 6 lists the publications of the dissertation. Finally, Section 7 draws conclusions from the dissertation.

Related work

This section briefly describes related work to virtual machine placement and VM migration between data centers. Different approaches have been used to the VM placement problem. The authors of [Beloglazov et al. 2012], for instance, proposed an approach to minimize the energy consumption considering a bin-packing model, using a best fit decreasing algorithm. The solution for the VM placement problem in [Meng et al. 2010] takes into consideration the data center network, proposing a traffic-aware placement algorithm based on the knowledge of the network topology and an estimation of the traffic matrix, matching clusters of virtual machines and servers to minimize VM communications. The TAVMP algorithm differs from proposals in the literature by considering a demand model based on the arrival of groups of virtual machines as well as their pattern of communication. Moreover, the selection of network paths is performed jointly with the decision of mapping VMs onto servers, so that the use of network switches is minimized.

In [Ilyas et al. 2012], the authors proposed an optimization formulation to decide on the migration of a data center workload based on electricity prices in different locations. The algorithm presented in [Mandal et al. 2013] decides to migrate long-lived VMs to sites where renewable energy is available. They show that green energy can replace brown energy with a low increase in the overall consumption. The TAVMS algorithm differs from the literature by the consideration of fine granularity in the decision of workload for migration rather than the whole workload of a data center, allowing fine tuning of energy savings. In addition, it considers the data center network topology in its decisions.

Topology-aware Virtual Machine Placement

The TAVMP algorithm places groups of virtual machines in a data center taking into account the network topology. It aims to place VM groups in small areas of hierarchical data centers so that few network switches are needed to serve the network flows. In order to achieve such goal, the data center is divided in hierarchical areas containing switches and servers. A parameter $j \geq 0$ is used to describe the level of an area in a hierarchical topology. For example, the Fat-tree topology has three levels: racks, PODs and the whole data center, as illustrated in Figure 1.

TAVMP and its notation are presented in Figure 2. The input is the data center network, the group of virtual machines (request) and the highest level of the hierarchy ($j = 2$ in Figure 1). The algorithm divides the topology into small areas, using an auxiliary algorithm to find subgraphs (SUB), Line 1. Recursion is then performed on each area (Line 4), and, when the lowest level is reached, a placement decision is made, specifying which servers will be used to host the virtual machines, as well as the network paths for the inter-VM flows. Another algorithm, named Placement in Current Area (PCA), evaluates

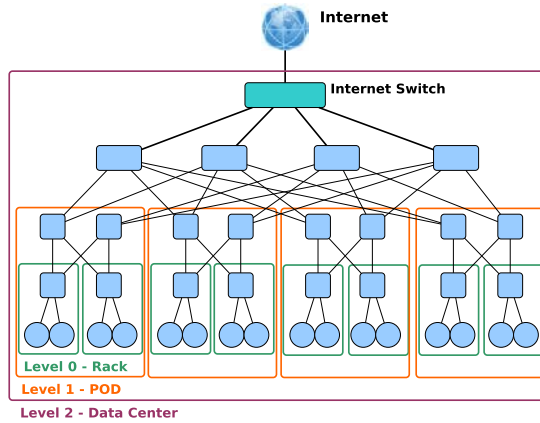


Figure 1. Fat-tree architecture divided into levels. Each circle is a physical server, and each rectangle a switch.

the adequacy of areas for the placement of a group of VMs (Line 7). When the recursion finishes, there can be various choices for the placement of the VMs. The area chosen is the one which involves the minimum energy consumption (Line 5). Furthermore, if none of the areas are suitable for accommodating the group of VMs due to lack of resources, the PCA algorithm is applied to the immediate higher level and, if necessary, to the data center level (Line 7). When TAVMP decides on the final placement of the group of virtual machines, the VMs are instantiated and the status of the network occupancy is updated.

Notation	Description
$\mathcal{D} = (\mathcal{E}, \mathcal{L})$	graph representing a data center area
$\mathcal{E} = \mathcal{H} \cup \mathcal{S}$	\mathcal{H} is the set of servers/hosts and \mathcal{S} is the set of switches
\mathcal{L}	links representing the physical connections
$\mathcal{G} = (\mathcal{V}, \mathcal{F})$	graph representing a VM group
\mathcal{V}	the set of virtual machines
\mathcal{F}	the set of network flows between the VMs

Input: $\mathcal{D} = (\mathcal{E}, \mathcal{L}), \mathcal{G} = (\mathcal{V}, \mathcal{F}), j$

Output: The placement

```

1 subgraphs  $\leftarrow$  SUB( $\mathcal{D}, j$ )
2 if subgraphs  $\neq$   $\emptyset$  then
3    $\forall$  subgraph  $\in$  subgraphs
4     placement  $\leftarrow$ 
       TAVMP(subgraph,  $\mathcal{G}, j - 1$ )
5   Choose placement with smallest power increase
6 if No placement was successful then
7   placement  $\leftarrow$  PCA( $\mathcal{D}, \mathcal{G}$ )
8 return placement

```

Figure 2. The TAVMP algorithm and its notation.

The input of the PCA algorithm is a graph representing an area in the data center topology and information of the demands of the group of VMs. The goal is to choose the servers with resources available to support the VMs computational demands, but minimizing the increase in total energy consumption. Moreover, these servers must be connected by links with sufficient bandwidth to support the communication demands. The PCA algorithm evaluates the placement of one VM at a time, and all the servers are analyzed to find the one on which the energy consumption for that VM workload is minimized. The decision of the paths to accommodate the inter-VM flows is also made in a way that the number of switches visited by a flow is minimized. The PCA algorithm succeeds if all the VMs in the group can be placed on servers, otherwise, the request is rejected.

The input to SUB is the area to be divided and its level, and its output is a set of subgraphs. Figure 1 shows the three levels ($j = 0, 1, 2$) and their subgraphs. Although the SUB algorithm finds subgraphs in the Fat-tree topology, the TAVMP algorithm can accept other topologies as long as SUB is changed to divide the subgraphs according

to the topology. Complete descriptions of the PCA and SUB algorithms can be found in [da Silva and da Fonseca 2015b].

Load balancing in distributed data centers

The TAVMS algorithm performs two tasks: the selection of potential sets of VMs in a data center to be migrated and the negotiation of migration of these workloads to other data centers. The former decision takes into account hierarchical network topologies, while the latter considers the energy effects on the destination data centers (by using the TAVMP algorithm) and the migration downtime. These two tasks are denoted as Selection (SEL) and Negotiation (NEG) algorithms, respectively.

The integration of the algorithms SEL and NEG in TAVMS is explained next. SEL analyzes the PODs in Fat-tree topology discriminating the sets of all virtual machine groups in sequential PODs. This strategy increases the chances of switching off sequential servers and switches, reducing therefore the energy consumption. Its output is used as input to the NEG algorithm, which decides on the migration of the chosen sets by iterating with all the other data centers based on the computation of the energy impact of the migration of each set of VM groups on the overall energy consumption. If the allocation of these VMs will spend less energy in another data center in relation to the source data center, migration is chosen to be performed. The transmission time over the optical infrastructure is calculated, and VMs terminating before this migration time are removed to avoid unnecessary migration. If no migration leads to energy savings, no workload is transferred. The SEL and NEG algorithms are not presented here due to space constraints, their complete description can be found in the dissertation [da Silva and da Fonseca 2015b].

Results

In this section, the main results obtained by the employment of the proposed algorithms are presented. Both algorithms were evaluated by using the Cloudsim Simulator. Energy consumption models were used to account the energy consumption of servers and network switches. The first algorithm was compared to the Power Aware Best Fit Decreasing (PABFD) algorithm from [Beloglazov et al. 2012], while the second was compared to algorithms based on migration of fixed loads, Topology-aware Threshold (TT) and Random Threshold (RT) algorithms, as well as a scenario without migrations (NM).

The parameter K of the Fat-tree topology, which defines the data center size, was varied. Heterogenous servers and VM instances were considered, and the arrival and departure processes of requests were modelled as self-similar series, as suggested in [Han et al. 2013]. Traffic scenarios considered medium (MG) and large (LG) VM groups as well as low (LT), medium (MT) and high (HT) traffic demands. In the second algorithm, four data centers were placed in the National Science Foundation's Network topology. Values for all parameters can be found in the dissertation [da Silva and da Fonseca 2015b].

Energy efficiency and blocking ratio of requests produced by the two algorithms were evaluated. For the first algorithm, energy efficiency was assessed as the total energy spent by a data center divided by the number of accepted VMs. In the second algorithm, it was evaluated as the total energy spent by all data centers. The blocking ratio is the percentage of VM requests which were not placed in relation to the total number of VMs.

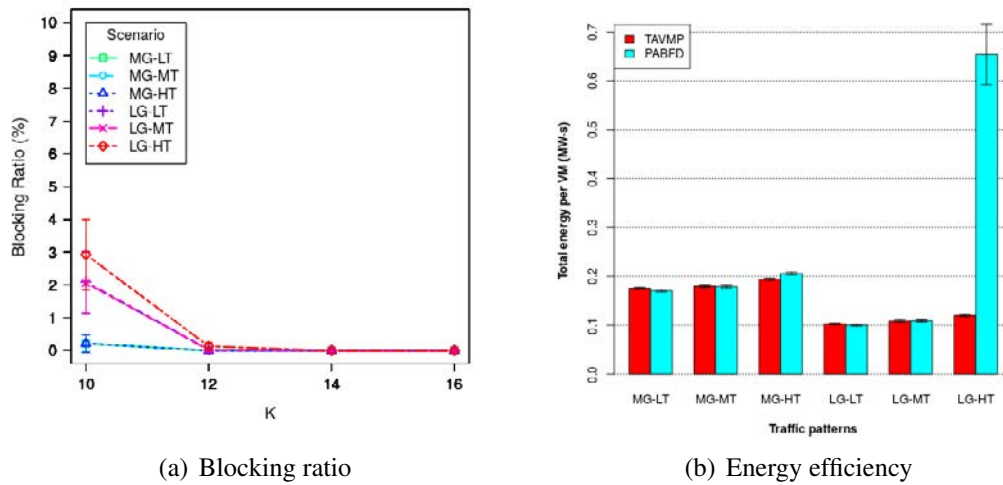


Figure 3. Results for the TAVMP algorithm.

Results for the first algorithm are presented in Figure 3. The blocking ratio of the algorithm is shown in Figure 3(a) for a Hurst parameter value of 0.7. The blocking ratio remains at low levels, even under high traffic scenarios. The blocking ratio produced by the PABFD algorithm for MG-HT remains between 12% and 14%, and for LG-HT it is close to 100%, while the TAVMP algorithm maintains all values under 4%. Figure 3(b) presents the energy efficiency of the TAVMP algorithm for larger data centers ($K = 16$). For low and medium traffic scenarios it remains similar to the PABFD algorithm, while for high traffic demands their VMs use much less energy. This happens due to the VM consolidation in small areas of the topology provided by the TAVMP algorithm, which prevents the formation of network bottlenecks.

Figure 4 presents results for the energy savings obtained by the TAVMS algorithm for data centers with different sizes. TAVMS outperforms the fixed load algorithms in all scenarios. The TT outperforms RT since it takes into account the topology, choosing VMs in close proximity. Moreover, the TAVMS algorithm considers more options of VM groups allocation, considering more precisely the data center network, and consequently saving more energy, ranging from 7.1% to 14% in relation to the NM scenario.

Publications

Results of the dissertation presented in this paper were published in the proceedings of the International Conference on Communications [da Silva and da Fonseca 2015a] (Qualis A2) and in the Journal of Grid Computing [da Silva and da Fonseca 2015c] (impact factor 1.5, Qualis B1). In addition, a third manuscript is under preparation to be submitted to an international journal containing the TAVMS algorithm and associated results.

Conclusion

The dissertation discussed in this paper introduced two algorithms for energy savings in distributed data centers, one for the virtual machine placement and the other to decide on the migration of workload between data centers. The consideration of the network topology and requests as VM groups with traffic demands, which is original consideration

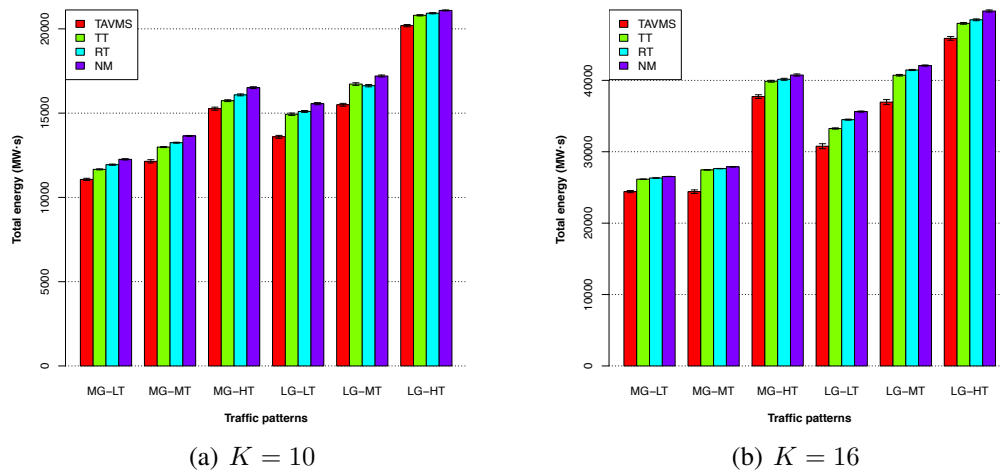


Figure 4. Total energy spent in all data center according to their size.

in this type of algorithms, showed to lead to energy savings while maintaining the quality of the offered services. Future directions of this work include a deeper analysis of traffic patterns of inter-VM communications and also the provisioning of differentiated services.

References

- Beloglazov, A., Abawajy, J., and Buyya, R. (2012). Energy-aware resource allocation heuristics for efficient management of data centers for cloud computing. *Future Generation Computer Systems*, 28(5):755 – 768. Special Section: Energy efficiency in large-scale distributed systems.
- da Silva, R. A. C. and da Fonseca, N. L. S. (2015a). Algorithm for the placement of groups of virtual machines in data centers. In *Communications (ICC), 2015 IEEE International Conference on*.
- da Silva, R. A. C. and da Fonseca, N. L. S. (2015b). Energy-aware load balancing in distributed data centers. Master's thesis, Universidade Estadual de Campinas.
- da Silva, R. A. C. and da Fonseca, N. L. S. (2015c). Topology-aware virtual machine placement in data centers. *Journal of Grid Computing*, pages 1–16.
- Han, Y., Chan, J., and Leckie, C. (2013). Analysing virtual machine usage in cloud computing. In *Services (SERVICES), 2013 IEEE Ninth World Congress on*, pages 370–377.
- Ilyas, M., Raza, S., Chen, C.-C., Uzmi, Z., and Chuah, C.-N. (2012). Red-bl: Energy solution for loading data centers. In *INFOCOM, 2012 Proceedings IEEE*, pages 2866–2870.
- Koomey, J. (2011). Growth in data center electricity use 2005 to 2010. *Oakland, CA: Analytics Press. August*.
- Mandal, U., Habib, M., Zhang, S., Mukherjee, B., and Tornatore, M. (2013). Greening the cloud using renewable-energy-aware service migration. *Network, IEEE*, 27(6):36–43.
- Meng, X., Pappas, V., and Zhang, L. (2010). Improving the scalability of data center networks with traffic-aware virtual machine placement. In *INFOCOM, 2010 Proceedings IEEE*, pages 1–9.