Saving energy in data centers

with Eco4Cloud and Cisco UCS

How to save energy costs and reduce carbon emissions, complying with worldwide guidelines and certifications, through efficient distribution of workloads in the data center

Raffaele Giordanelli[1], Carlo Mastroianni[1,2], Michela Meo[3], Giuseppe Papuzzo[1] , Andrea Roscetti[4]

(1) Eco4Cloud – www.eco4cloud.com
(2) ICAR-CNR, Institute for High Performance Computing and Networking of the Italian National Research Council, Italy – www.icar.cnr.it
(3) Politecnico di Torino, Italy – www.polito.it
(4) Politecnico di Milano, Italy – www.polimi.it

Saving energy in data centers through workload consolidation

"The physical efficiency is not enough to guarantee the overall efficiency of the data center"

1. Executive summary

In this whitepaper, we first examine two different aspects regarding the energy consumption in data centers: the efficiency of the data center physical infrastructure (Section 2) and the computational efficiency (Section 3), and we highlight the fact that there is room to greatly improve the second aspect. In Section 4, we focus on the consolidation of the workload as an efficient means to increase computational efficiency, since it allows the applications to be clustered in fewer and more efficient servers. In Section 5, we introduce a recent and innovative technique to cope with the consolidation issue. The new technique, based on self-organizing processes and on the distribution of the decisions, has been devised and implemented at ICAR-CNR and at the Politecnico di Torino, and enables several advantages, namely: higher efficiency, energy and cost savings between 30% and 60%, better informed capacity planning, improved quality of service, high scalability. The Eco4Cloud software implements these ideas and techniques. Section 6 illustrates the advantages deriving from the integration of Eco4Cloud with Cisco Unified Computing System (UCS). Finally, Section 7 discusses the urgent need for integrating the consolidation solutions into the software for Data Center Infrastructure Management (DCIM).

All main trends in information technology, namely Cloud Computing, Big Data and Mobile Computing, are based on powerful computing infrastructures. The ever increasing demand for computing resources has led companies and resource providers to build large warehouse-sized data centers, which require a significant amount of power to be operated and hence consume a lot of energy: 120 billion of KWh in the USA alone, as stated by a report published by the USA Environmental Protection Agency [1]. This is enough energy to power more than 11 million average U.S. households. Related carbon emissions are huge as well.

Several concepts and tools have been proposed and implemented at international level to support energy efficiency in data centers including energy efficiency metrics and monitoring, guidelines for best practices, as well as comprehensive management and certification concepts. On the microeconomic side, it is clear that reducing energy consumption and related costs is essential for companies of any size.

2. Physical efficiency of the data center

Only a fraction of the power that enters the data centers is actually used to support the workload, i.e. to execute the applications hosted by servers. A consistent amount of power is used by the facilities for power supply and distribution (for example, UPS and PDU devices), and an even greater amount of power is used to feed the cooling infrastructure.

To assess the efficiency of the physical infrastructure, a widely adopted index is the PUE, Power Usage Effectiveness. If P_{sup}, P_{cool} and P_{comp} are, respectively, the amount of average power (measured in a proper interval of time) used for power supply and distribution, for cooling, and for actual computation (see Figure 1), the PUE is defined as:

$$PUE = \frac{P_{sup} + P_{cool} + P_{comp}}{P_{comp}} \quad (1)$$

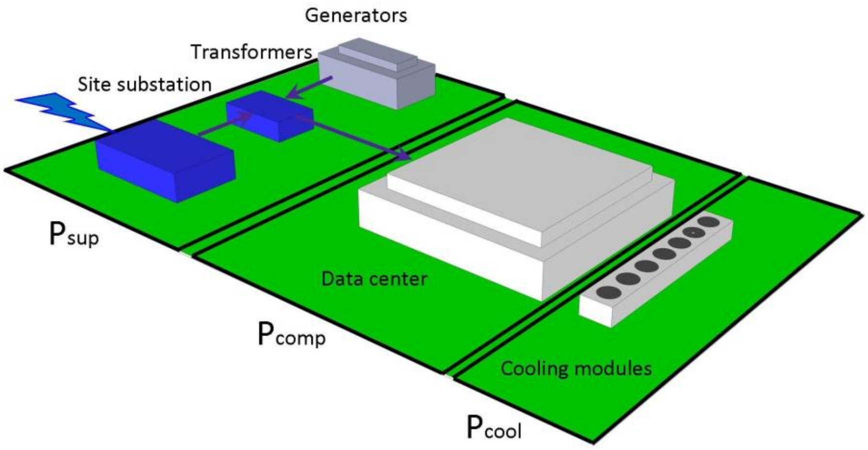


Figure 1. Power usage in a data center.

So, the PUE measures the ratio between the total amount of energy that enters the data center and the energy actually used to operate the IT devices, specifically the servers. In efficient infrastructures, the PUE value should be as low as possible, and a value of 1 would correspond to the maximum theoretical efficiency, as all the energy would be devoted to the IT infrastructure.

A great amount of scientific and industrial efforts have been devoted to improve the physical efficiency, for example by optimizing the cooling system. This allowed to gradually reduce the PUE value: only a few years ago values between 2 and 3 and even higher were common, while today the new data centers built by big companies like Facebook and Google are designed to have PUE values as low as 1.1. Of course the PUE value depends not only on the design accuracy but also on several parameters such as the size of the data centers (bigger data centers have generally lower values of the PUE), the computation workload (the efficiency tends to increase with the workload) and external parameters such as temperature and humidity. Yet, the physical efficiency is not enough to guarantee the overall efficiency of the data center, as discussed in the following section.

3. Computational efficiency of the data center: why PUE is not enough

A good PUE value is not enough to guarantee the global efficiency of the data center, because this metric does not consider the actual utilization of computational resources. In other words, the physical efficiency of the data center should be combined with the efficiency in using its computational resources. Let us consider the following example. In a data center containing 2000 servers the average electrical power entering the system is 1 MW, of which 500 kW are used to operate the servers, 200 kW for the power distribution systems and the UPS devices, and 300 kW for the cooling system. Let us assume that the average utilization of server resources (CPU, RAM etc.) is 30%, a typical value observed in real data centers [2]. This means that the workload carried by the servers is only one third of the overall data center capacity. The PUE can be computed using expression (1) and it is equal to (200+300+500)/500=1000/500=2.

Now let us assume that a better and more efficient strategy for the distribution of the workload to servers is applied, and applications are consolidated on a smaller number of servers. In the new scenario, 1000 servers are assigned more applications and their average utilization increases to 60%. These servers are assigned the entire computational load, while the other 1000 servers are hibernated or switched off. The overall workload executed on the data center is the same as before, equal to 30% of the total data center computing capacity, but the new load distribution leads to considerable energy savings. Indeed, the 1000 active servers consume more energy, because their utilization is higher, but this extra energy consumption is more than compensated by the hibernation of the other 1000 servers, resulting in remarkable overall energy savings. Using the mathematical models devised to relate server utilization to power consumption (specifically, the model published in [3], i.e. the most commonly adopted), it can be estimated that the average power needed to operate the servers is now about 270 kW, instead of 500 kW. In a well-designed and modular data center, the amount of energy needed for power supply and for cooling should be quasi-proportional to the computational load. In a realistic scenario it can be assumed that the average value of P_{sup} can be reduced from 200 kW to 120 kW, while the average value of P_{cool} can be lowered from 300 kW to 180 kW. The measured value of the PUE metric now becomes (120+180+270)/270=2.11.

"The PUE index is useful to measure the efficiency of the physical infrastructure, but does not relate to the computational efficiency"

Using the new strategy, the same workload is executed on the data center, but the data center consumes on average only 570 kW instead of 1MW, with an energy saving equal to 43%, and nearly proportional savings in carbon emissions. But none of these phenomena is captured by measuring the PUE value, which is comparable to the value measured before, and even slightly worse (2.11 vs. 2.0) due to the more than proportional losses that are usually experienced when the total computational load is lower. This example shows that the PUE index is useful to measure the relative efficiency of the physical infrastructure, but does not relate to the total energy consumption or the computational efficiency, strictly related to the workload distribution between servers. Therefore, the PUE cannot be the only parameter for the evaluation of the data center overall efficiency, as experts deemed only a few years ago. Indeed, recent efforts are being devoted to define indices that give the right importance to the computational efficiency and to the green features of data centers. For example, on March 2013 eBay introduced the DSE (Digital Service Efficiency) standard, which measures how many business transactions are completed per kilowatt-hour of energy [4]. The Green Grid initiative has introduced metrics that further elaborate on data center environmental impact, including Water Usage Effectiveness (WUE™), and Carbon Usage Effectiveness (CUE™).

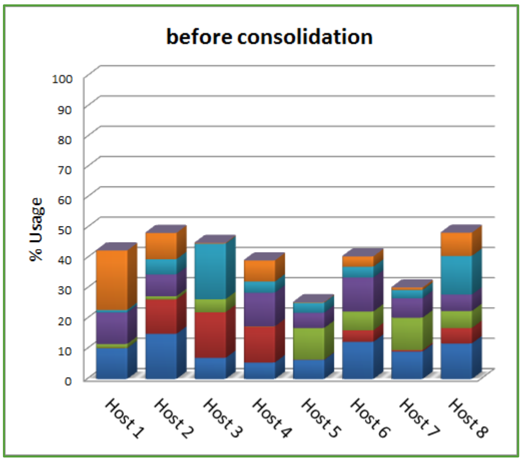
4. Consolidation of the workload

"The objective of consolidation is to cluster the maximum number of VMs onto the minimum number of physical servers"

There are several ways to optimize the use of servers and increase the computational efficiency. For example, if applications are executed on ten year-old servers, the consumed power could be twice the power consumed by newer and more efficient servers. It is also essential to acquire the right hardware for the right type of applications. For example, some applications are CPU-intensive while others are memory-intensive: hardware requirements are clearly different in the two cases. The option of leveraging the proper choice of hardware, however, is only available at the capacity planning stage, when the data center is firstly designed and developed, or when new machines are acquired.

Workload consolidation is a powerful means offered by the virtualization technology to achieve remarkable energy and cost savings at any time during normal data center operation. All the virtualization platforms (for example VMWare, Microsoft Hyper-V, KVM) allow several Virtual Machines (VMs) to be hosted by the same physical server, and provide primitives to move a VM from one server to another in a short time and without service interruption. The objective of consolidation is to cluster the maximum number of VMs onto the minimum number of physical servers, so that the unneeded servers can either be put into a low power state (leading to energy saving and OpEx reduction), or switched-off and devoted to the execution of incremental workload or additional services (leading to CapEx savings, thanks to the reduced need for acquiring additional servers).

Figure 2 shows an example of workload consolidation. Before consolidation (figure on the left), the workload is distributed to 8 servers whose resource utilization is between 20% and 50%. After consolidation (figure on the right), the same VMs are redistributed so that 4 servers take all the load, with their resource usage between 70% and 90%, while the other servers are hibernated in order to save energy.



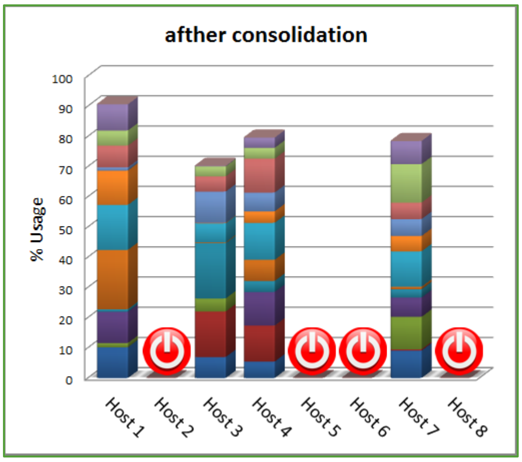


Figure 2. Example of servers usage before consolidation (left) and after consolidation (right). In this simple case consolidation allows half servers to be hibernated.

The reason why workload consolidation has not been widely exploited so far is related to the inherent complexity of the problem. The optimal assignment of VMs to the servers of a data center is analogous to the "Bin Packing Problem", the problem of assigning a given set of items of variable size to the minimum number of bins taken from a given set. Unfortunately it is known that the problem has NP-hard complexity, meaning that it is an "intractable problem" and it would take years or centuries to find the optimal solution even in a data center of relatively small size. To make things even worse, the problem is complicated by two circumstances that need to be considered for the data center case:

- the assignment of VMs should take into account multiple server resources at the same time, for example CPU and RAM; therefore it is formally a "multi-dimensional bin packing problem", much more difficult than the single-dimension problem

"A scalable strategy to effectively approach consolidation is to distribute the intelligence instead of centralizing it, while switching from deterministic algorithms to self-organizing solutions that are able to adapt to the dynamic nature of the problem"

- even when a good assignment has been achieved, the VMs continuously modify their hardware requirements, potentially baffling the previous assignment decisions in a few hours.

5. Innovative solutions for workload consolidation

Owing to technical and administrative reasons, data centers are typically managed with centralized and manual approaches, including the decisions regarding the initial assignment of VMs to servers, and their successive migrations from one server to another. However, such strategies are no longer appropriate today, as the average size of data centers is rapidly increasing and big companies need to manage data centers with thousands or even tens of thousands of servers.

A scalable strategy to effectively approach consolidation in these scenarios is to distribute the intelligence instead of centralizing it in a single point, while switching from deterministic algorithms to self-organizing solutions that are able to adapt to the dynamic nature of the problem. Efficient solutions should adhere to the "emergence" or "swarm intelligence" paradigm, i.e., to the capability of complex systems to let an intelligent behavior emerge from the composition of simple and distributed operations. A classical example is given by ant colonies that are able to build complex and ordered structures due to the simple and local operations of single ants. For their resemblance to the mechanisms of this and other biological systems, solutions based on the emergence behavior are often referred to as "bio-inspired".

The application of these paradigms to data center consolidation consists of achieving an approximate solution (whose efficiency is within a 5% margin with respect to the optimal solution) to the assignment problem due to the combination of simpler decisions taken on each server. For example, each server may decide whether or not to accept a new VM based on the local utilization of resources. Consider the following strategy: (1) servers with low resource utilization ask to migrate local VMs in order to be hibernated and save power; (2) servers with intermediate utilization declare their availability to execute more VMs in order to increase their respective utilization; (3) servers with very high resource utilization ask to migrate/off-load some VMs to other servers in order to avoid possible overload situations that would otherwise deteriorate the Quality of Service perceived by users. When local decisions of single servers about their will to accept or ask for migrations of VMs are collected and proficiently combined at central entities, with the additional support of statistical techniques, it is possible to consolidate the workload efficiently and respond quickly to the changing requirements of VMs.

A bio-inspired strategy of this kind was proposed by a research project carried out by the Institute for High Performance Computing and Networking of the Italian National Research Council (ICAR-CNR) and by the Politecnico di Torino [5]. Research activities lead to an industrial invention filed as an international PCT patent and owned by CNR and University of Calabria, titled "System for Energy Saving in Company Data Centers" [6]. Eco4Cloud is a software tool that implements these ideas and techniques.

Figure 3 shows an example of workload consolidation with Eco4Cloud on a data center with 100 servers. The figure shows the CPU utilization of the servers, which, before applying Eco4Cloud, is between 20% and 40%. At time zero the consolidation algorithm is applied: after VM migrations, some servers increase their respective utilization while others are unloaded and then switched off. At the end of the experiment, 35 servers have taken all the load and the remaining 65 servers have been hibernated, leading to energy savings higher than 50%.

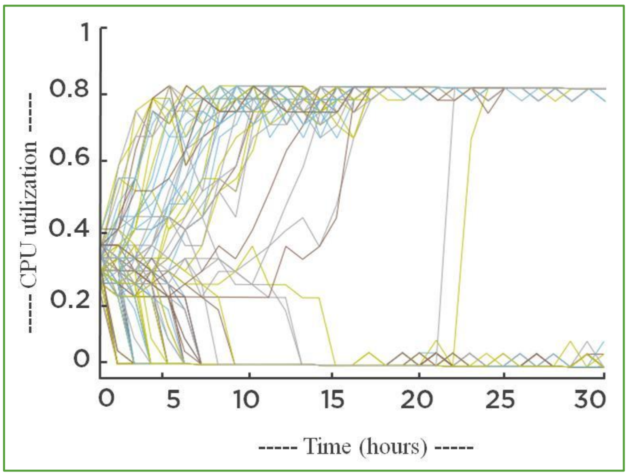


Figure 3. Example of workload consolidation on 100 servers. The workload is consolidated on 35 servers and the other servers are hibernated, leading to great energy savings.

The advantages achievable for a mid/large-scale data center are the following:

1) EFFICIENCY

By using servers at the best of their capacity, efficient consolidation algorithms give the opportunity to execute more tasks (up to 3X) with the same number of physical servers without any impact on the related quality of service.

2) REDUCTION OF DATA CENTER ENERGY BILL

In a virtualized data center, only 20-30% of server capacity is utilized on average. Still, an idle server consumes 65-70% of the power consumed when it is fully utilized. By consolidating the maximum number of VMs on the minimum number of physical servers, it is possible to reduce the overall energy bill up to 60%. Furthermore, by having less energy consumed by physical servers, power cooling gets reduced as well, contributing to indirect additional energy reductions (OpEx).

3) CAPACITY PLANNING

Thanks to the optimal occupancy of physical resources and to the adaptive optimization of inherently variable workloads, CIOs can improve the budget accuracy and capacity planning for their data centers, hence optimizing their CapEx and OpEx.

4) RESPECT Service Level Agreements (reliability, availability, performance)

Adaptive and self-organizing solutions allow to proactively/predictively prevent quality of service deterioration, for example by instructing the move of VMs from servers that are going to be overloaded. VM migrations are gradual, asynchronous and easy to monitor, whereas classical consolidation algorithms may require the concurrent move of hundreds of VMs.

5) SCALABILITY

Adaptive/self-organized algorithms are much more scalable than classical algorithms, thanks to their capacity of taking decisions on the basis of local and readily/easily available information and to the use of statistical data. Therefore, all the aforementioned benefits increase proportionally with the data center size.

6. Integration between Cisco UCS and Eco4Cloud

Cisco Unified Computing System (UCS) is a programmable infrastructure component for the data center. UCS allows the system to be integrated into higher-level, data-center-wide management systems as a single, logical entity. Server, network, and I/O resources are provisioned and configured on demand by applying a service profile to them.

Once Cisco UCS is wired, system resources become part of a flexible pool that can be quickly used for any workload on demand. The system abstracts the configuration and connectivity of server and I/O resources, allowing administration to be automated.

Cisco UCS offers a single unified system, transcending the traditional boundaries of blade chassis and racks. Cisco UCS brings together server, network, and storage access resources to create a physically distributed but centrally managed system. By abstracting the personality, configuration, and connectivity of server and I/O resources, these attributes can be programmed automatically.

The abstraction of resources provided by Cisco UCS is the perfect starting point for the workload consolidation provided by Eco4Cloud. Indeed, UCS creates a homogeneous environment tailored to applications migration and dynamic server hibernation and resume.

Eco4Cloud contributes to boost the power efficiency of Cisco UCS, as explained in the following. The Cisco white paper Power Efficiency Comparison: Cisco UCS 5108 Blade Server Chassis and HP BladeSystem c7000 Enclosure [Unified Computing] [7] compares the performance-to-power ratio of two similar Cisco and HP blade solutions: the results, some of which are reported in Figure 4, show that the Cisco solution consumes less power for all the values of target load. The same study, however, clarifies that power efficiency improves with the target load, and suggests that resource utilization should be in the range from 50 to 70 percent. Specifically, a 70 percent target load is taken as a reference for the comparison. Unfortunately, the resources utilization of data center servers is in general much lower, as the typical utilization range is between 20 and 40 percent. This is due to several reasons: the lack of efficient algorithms for workload distribution in large data centers, the will of data center administrators to over-provision resources, the poor capacity of adapting the applications distribution to the dynamic workload. Eco4Cloud helps to fully exploit the power efficiency of Cisco UCS in that Eco4Cloud consolidates the workload and allows the resources utilization to be increased up to the 70 percent target and even more.

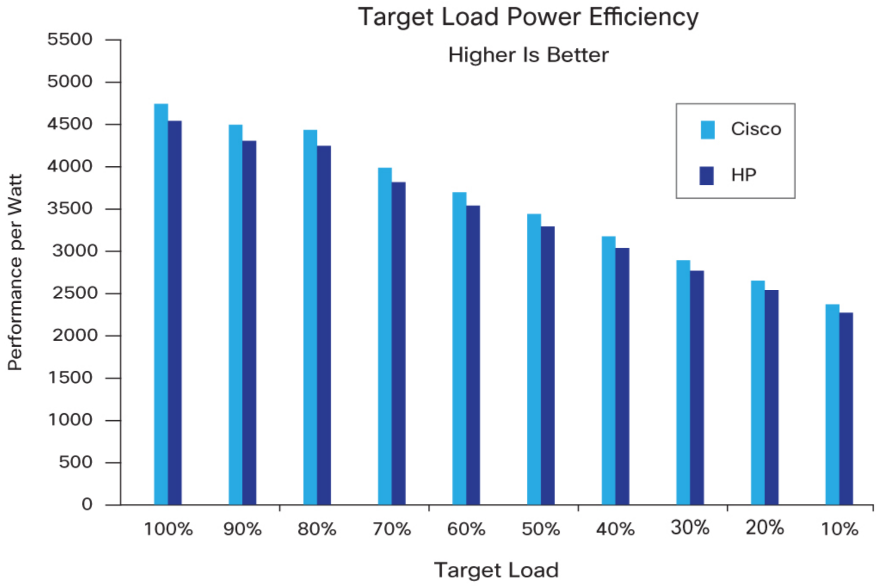


Figure 4. Power efficiency versus the target load for the Cisco and HP blade solutions compared in the Cisco white paper [7].

Eco4Cloud is also useful to boost another technique adopted by Cisco UCS to save energy: "power capping". This technique is described in the Cisco white paper Power Management in the Cisco Unified Computing System: An Integrated Approach [Unified Computing] [8]. Power capping is the capability to limit the power consumption of a system, be it a blade server or a rack server, to some threshold that is less than or equal to the system's maximum rated power. Previously the technique was only implemented in a single server or at most on the servers of the same chassis, but Cisco has extended power capping to "power cap groups", i.e., collections of servers or chassis that share a common power allocation or budget. Any UCS instance can have multiple power cap groups, containing varying numbers of chassis, as shown in Figure 5.

While the technique of power capping can help to optimize the distribution of power and save energy, the difficulty of dynamically adapting the power thresholds to the varying workload increases with size of the data center. It is extremely complex to centrally and deterministically set the power thresholds of all servers and power cap groups when the data center contains hundreds or thousands of servers. In this respect, Eco4Cloud helps to potentiate the benefits of power capping, because it allows the system to adjusting the thresholds only using local information. The threshold values dynamically adapt to the

"DCIM tools should incorporate consolidation strategies/functionalities that allow fewer and more efficient servers to support the workload"

varying workload, and asynchronous VM migrations are performed by Eco4Cloud to respect the constraints and maximize the utilization of servers.

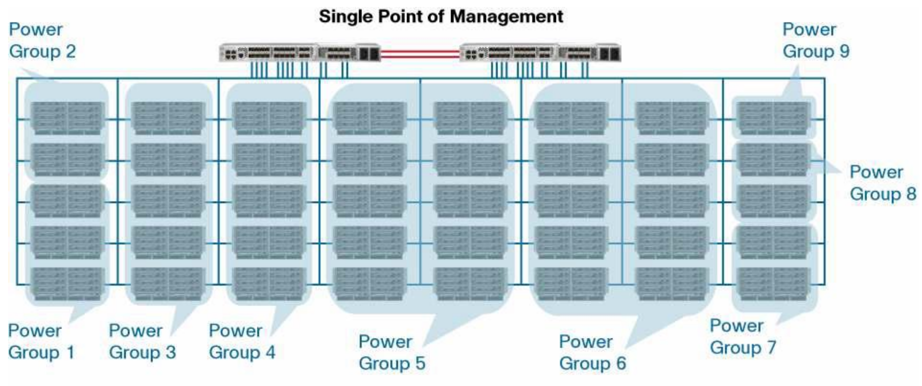


Figure 5. Power cap Groups Across Multiple Chassis.

7. Integrating consolidation solutions into DCIM software

The complexity and strategic importance of data centers has led to the need for more intelligent and automated IT infrastructure management. The tools that enable the data center team to effectively and efficiently operate such complex environments have been grouped into a classification of solutions collectively known as Data Center Infrastructure Management (DCIM).

Due to the increasing impact of energy requirements, next generation DCIM software tools need to monitor, measure, manage and control the energy consumption and the related carbon emissions of all IT equipment and the facility's infrastructure components. While the estimation and control of the power usage effectiveness (PUE) is essential, this is not enough: in a previous section of this document we show that the PUE does not take into account computational efficiency, since this metric does not relate the power consumed by servers to the useful work that they actually execute.

DCIM tools should incorporate consolidation strategies/functionalities that allow fewer and more efficient servers to support the workload, leading to energy savings between 30% and 60% and to comparable reductions of carbon emissions. Consolidation tools should adopt innovative self-organizing and adaptive techniques, as they are proving to perform much better than traditional centralized solutions. It is also essential that the consolidation software smoothly integrates with all the commonly adopted virtualization platforms, for example, VMWare, Microsoft HyperV and KVM, in order to support heterogeneous virtualized data center environments. This can be done by using a uniform and integrated dashboard and exploiting the libraries (APIs) of the underlying virtualization platform(s) to collect information from servers and instruct the assignment and migration of VMs in order to maximize efficiency and minimize energy consumption.

References

[1] Environmental Protection Agency (EPA) Report on Server and Data Center Energy Efficiency, August 2007.

[2] L. A. Barroso and U. Holzle, *"The case for energy-proportional computing"* IEEE Computer, vol. 40, no. 12, pp. 33–37, December 2007.

[3] S. Rivoire, P. Ranganathan, and C. Kozyrakis, "A comparison of high level full-system power models" in Proc. of the 2008 conference on Power aware computing and systems HotPower'08, San Diego, California, December 2008.

[4] eBay white paper, *"Digital Service Efficiency"*, March 2013, available online at http://www.ebaytechblog.com/wp-content/uploads/2013/03/FINAL_DSE-Solution-Paper.pdf

[5] C. Mastroianni, M. Meo, G. Papuzzo, *"Analysis of a Self-Organizing Algorithm for Energy Saving in Data Centers"*. To appear on Proc. of the 9th Workshop on High-Performance, Power-Aware Computing, Boston (MA) USA , May 2013.

[6] International PCT patent: *"System for Energy Saving in Company Data Centers"*, Italian National Research Council (CNR) and University of Calabria. Info and claims on http://patentscope.wipo.int, patent ID: #WO/2013/021407, May 2013.

[7] Cisco white paper Power Efficiency Comparison: Cisco UCS 5108 Blade Server Chassis and HP BladeSystem c7000 Enclosure [Unified Computing]

[8] Cisco white paper Power Management in the Cisco Unified Computing System: An Integrated Approach [Unified Computing]

For more information

- Eco4Cloud: www.eco4cloud.com
- PrimeEnergyIT – Efficient Data Centers: www.efficientdatacenter.org
- Cisco UCS: www.cisco.com/go/ucs