

Virtualisation and dynamic IT loads

JOHANNESBURG – March 15, 2016 – Without question, IT virtualisation – the abstraction of physical network, server, and storage resources – has greatly increased the ability to utilise and scale compute power. Indeed, virtualisation has become the very technology engine behind cloud computing itself. While the benefits of this technology and service delivery model are well known, understood, and increasingly being taken advantage of, their effects on the data centre physical infrastructure (DCPI) are less understood.

This is according to the white paper, *Virtualisation: Optimised Power, Cooling, and Management Maximises Benefits*, by Schneider Electric, the global specialist in energy management and automation. In its research, the company states that virtualised IT loads, particularly in a highly virtualised, cloud data centre, can vary in both time and location. In order to ensure availability in such a system, it's critical that rack-level power and cooling health be considered before changes are made.

The paper demonstrates how the sudden – and increasingly automated – creation and movement of virtual machines require careful management and policies that contemplate physical infrastructure status and capacity down to an individual rack level. Failure to do so could undermine the software fault-tolerance that virtualisation imbues to cloud computing. Fortunately, tools exist today to greatly simplify and assist in doing this.

The research further shows how electrical load on the physical hosts can vary in both time and place as virtual loads are created or moved from one location to another. As the processor computes, changes power state or as hard drives spin up and down, the electrical load on any machine – virtualised or not - will vary. This variation can be amplified when power management policies are implemented, which actively powers machines down and up throughout the day as compute needs change over time. The policy of power capping, however, can reduce this variation.

This is where machines are limited in how much power they can draw before processor speed is automatically reduced. At any rate, since data centre physical infrastructure is most often sized based on a high percentage of the nameplate ratings of the IT gear, this type of variation in power is unlikely to cause capacity issues related to the physical infrastructure particularly when the percentage of

virtualised servers is low.

A highly virtualised environment, such as that characterised by a large cloud-based data centre, however, could as per the white paper study have larger load swings compared to a non-virtualised one. And, unless they are incredibly well-planned and managed, these could be large enough to potentially cause capacity issues or, at least, possibly violate policies related to capacity headroom.

The study also reveals that increasingly, managers are automating the creation and movement of VMs. It is this ability that helps make a virtualised data centre more fault-tolerant. If a software fault occurs within a given VM or a physical host server crashes, other machines can quickly recover the workload with a minimal amount of latency for the user. Automated VM creation and movement is also what enables much of the compute power scalability in cloud computing.

Ironically, however, this rapid and sudden movement of VMs can also expose IT workloads to power and cooling problems that may exist which then put the loads at risk.

Data centre infrastructure management (DCIM) software can monitor and report on the health and capacity status of the power and cooling systems. This software can also be used to keep track of all the various relationships between the IT gear and the physical infrastructure.

“Essential knowledge for good VM management includes knowing, which servers, both physical and virtual, are installed in a given rack, along with understanding each associated power path and cooling system,” says Bruce Grobler, southern Africa’s vice president of the IT business unit at Schneider Electric.

This knowledge is important because without it, it is almost impossible to be sure virtual machines are being created in or moved to a host with adequate and healthy power and cooling resources. The white paper maintains that relying on manual human intervention to digest and act on all the information provided by DCIM software could quickly become an inadequate way to manage capacity, considering the many demands already placed on data centre managers.

“The risk of human error is linked to manual intervention, a main reason for downtime,” adds Grobler.

Human error, as examined in the white paper, is likely to take the form of IT load changes without accounting for the status and availability of power and cooling at a given location. Automating both the monitoring of DCIM information (available rack space, power, and cooling capacity and health) and the implementation of suggested actions greatly reduces the risk.

There is however DCIM software available today that provides real-time, automated management. The two-way communication between the VM manager and DCIM software and the automated action that results from this integration, is what ensures physical servers and storage arrays receive the right power and cooling where and when needed.

A VM is created or moved to a different physical server typically because there are not enough processor, memory, or storage resources available at a given moment and location. But the white paper points out that an effective management system can directly cause VMs to move based also on real time, physical infrastructure capacity and health at the rack level. When DCIM software is integrated with the VM manager, VMs can be safely and automatically moved to areas known to have sufficient power and cooling capacity to handle the additional load.

Conversely, the analysis illustrates how VMs can be moved away from racks that develop power or cooling problems. For example, if there's a power outage at a rack, a cooling fan stops working or there is a sudden loss of power redundancy, the VM manager can be notified of the event and the "at risk" VMs can be moved to a safe and "healthy" rack elsewhere in the data centre. All of this happens automatically in real time without staff intervention.

DCIM software integration with a VM manager is a key capability for ensuring that virtual loads and their physical hosts are protected. In turn, service levels will be more easily maintained and staff will be freed from having to spend as much time physically monitoring the power and cooling infrastructure.

The research demonstrates how this integration becomes even more critical as power and cooling capacities are reduced or rightsized to fit a newly virtualised or consolidated data centre. The less "head room" or excess capacity that exists, the less margin for error there is for placing virtual machines.

Maintaining a highly efficient, leanly provisioned data centre in an environment characterised by frequent and sudden load shifting requires a management system that works automatically in real time with the VM manager.

Also, the white paper highlights that it should not be forgotten that IT policies related to VM management need to be constructed so that power and cooling systems are considered. This must occur in order for the DCIM software integration with the VM manager to work as described above. Policies should set thresholds and limits for what is acceptable for a given application or VM in terms of power and cooling capacity, health, and redundancy.

Virtualising a data centre's IT resources can have certain consequences related to the physical infrastructure, concludes research by Schneider Electric, and if these impacts and consequences are ignored, the broad benefits of virtualisation and cloud computing can be limited or compromised, and in some cases, severely so.