Mike Nelson – SSV article

Hypervisor Host Power Management

These are the days of the "Dynamic" datacenter and a big piece of that has been attributed to virtualization. The hypervisor not only gives us dynamic flexibility with virtual machines, but they also can give us automated control over their use of electrical power within the datacenter.

There are a couple of ways the Hypervisor goes about regulating power usage in their environments. The first is by controlling the hardware components themselves and how they are using and delivering power. The second is monitoring how the environment is using all of its hardware resources and adjusting for capacity and load. Since it can be quite a lengthy topic by itself, I'll only briefly touch on the hardware control functionality and focus mainly on the hypervisor environmental control.

The hardware chipset makers have created settings that allow for what some hardware vendors call "OS Control" or Operating System Control of the power subsystem and all of its options. Pretty much all of the today's modern Operating Systems have functionality built-in to control hardware power settings through this option, including Hypervisors. In the latest CPU's that are on the market, the processor power state, or "P-State" as it is called, is the control switch for how much power is consumed by the onboard processors. In the past versions of these chips, these states where either not directly configurable via the BIOS, or had very limited functionality. The Hypervisor can monitor and moderate the use of power within its environment, and also effectively leverage options such as power saving and redundancy. There may also settings that can be changed that control the load balancing, resiliency, or redundancy of power supplies and power regulator chips on the motherboard. For some more deep technical readings on these, I would suggest IBM's Whitepaper or Intel's posting.

Let's take a look at how some of the more popular Hypervisors provide environmental power control within their host clusters and some of the options available.

VMware vSphere–

VMware allows for the creation of Power Policies, which allow for controlling, monitoring, and adjusting power consumption in the ESXi environment. There are default policies that are included within vCenter, and an option to create a custom policy to your own liking. These policies control the BIOS power options.

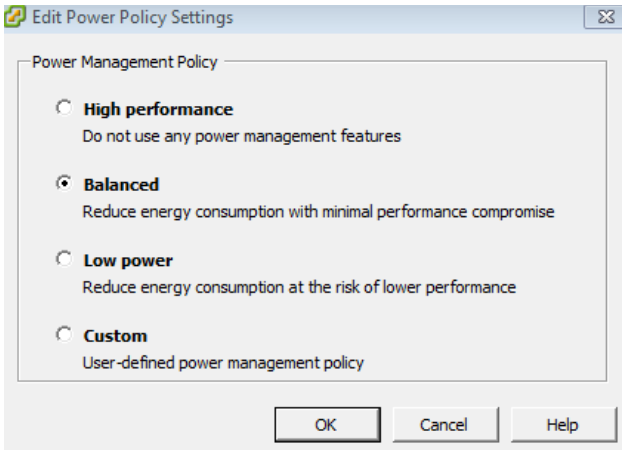

Figure 1 vCenter Power Policy options

In its simplest form, Distributed Power Management, or DPM, is the mechanism that VMware uses to control host resource power usage in its environment. DPM effectively monitors all of the hosts within a Cluster for their resource consumption and their ability to adequately supply necessary resources to their virtual guests. It constantly makes decisions based on this information and determines whether those hosts can be put in a "Stand-By" mode until their resources are needed by the Cluster. Because it can power off hosts, it has to be able to automatically move them to other hosts within the cluster. This is why vSphere Distributed Resource Management (DRS) is a requirement for DPM in vSphere. Basically this means that the machine is powered off, and once the resources are again needed by the Cluster, it sends either a Wake-On-LAN magic packet, uses ILO (HP's "Integrated Lights Out"), or IPMI (Intelligent Platform Management Interface) to initiate a power up of the host.

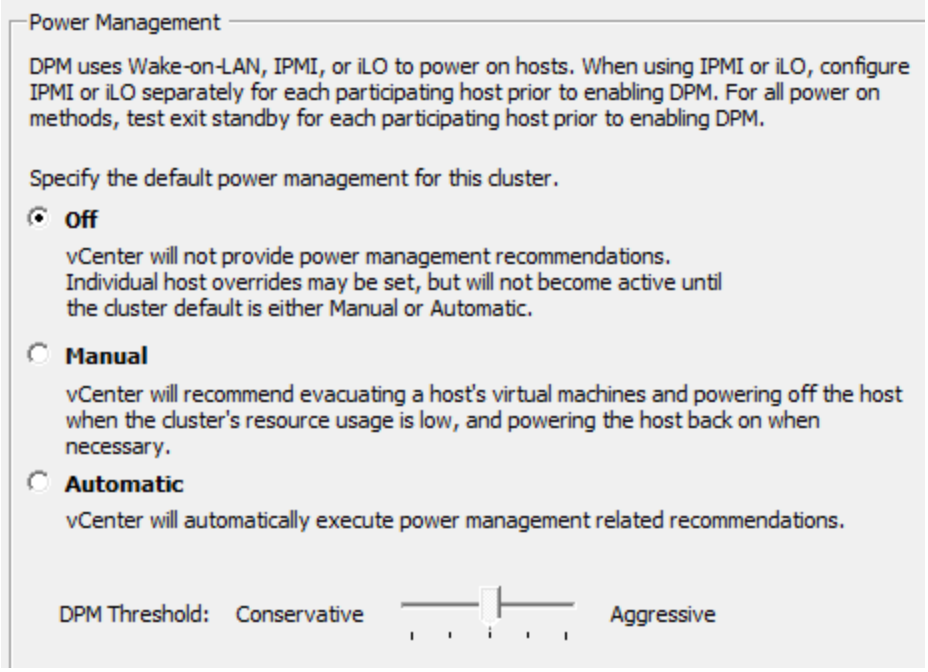

Figure 2 vCenter DPM settings

DPM is disabled by default and should only be used after testing it extensively with your hardware to ensure that it can effectively power hosts on and off. You can reference VMware's well written technical whitepaper on vSphere 5 power Management here.

Microsoft Hyper-V –

The Virtual Machine Manager (VMM), or System Center Virtual Machine Manager (SCVMM) if you run that, is the control point for managing dynamic power optimization in Hyper-V. The Power Optimization option, which requires Dynamic Optimization (which is the same as DRS in vSphere and it used Live Migration to move machines dynamically) be enabled, basically works the same as the vSphere DPM by turning hosts on in the Host Group (i.e. "Cluster") when needed to provide resources and off when they are not. The System Center version of VMM also allows for power control of vSphere and XenServer environments as well as Hyper-V, which is really cool and very convenient for multi-hypervisor datacenters. I highly recommend reading this blog posting by Microsoft MVP Mercelo Sinic on this excellent feature.

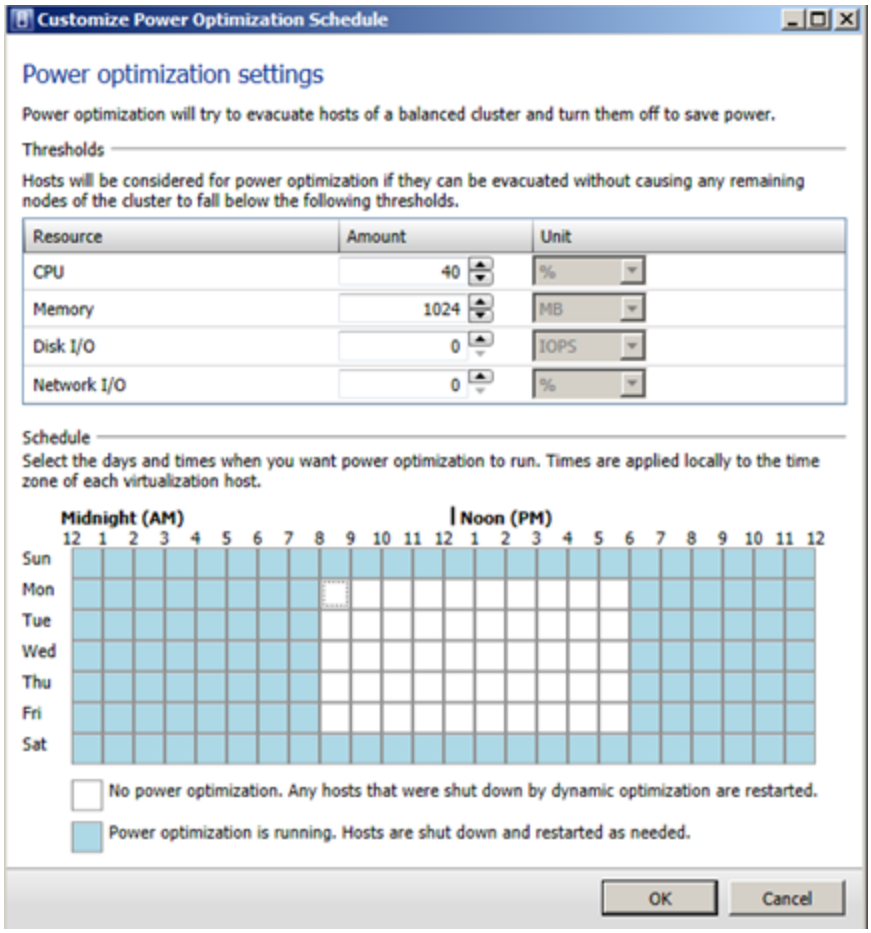

Figure 3 VMM 2012 Power Optimization Settings

Power Optimization can also be scheduled in VMM to only run during non-peak times so as utilize times when low power consumption is anticipated. This is a very good feature addition, one that is only currently available for Hyper-V and vSphere. For more technical information on how to implement these features in Hyper-V, check out this TechNet article.

Citrix XenServer –

Sadly, once again, XenServer trails the other hypervisors in its implementation of intuitive features like power management. The Citrix XenCenter interface provides a basic, and by basic I mean *very* basic, means to configure rudimentary power management within a XenServer Server Pool (again, i.e. "Cluster"). It requires the implementation of a Work Load Balancer server (WLB) in order to provide the dynamic guest movement to other hosts similar to Hyper-V's Live Migration and vSphere's vMotion. It also uses ILO and IPMI to control the hosts power state.

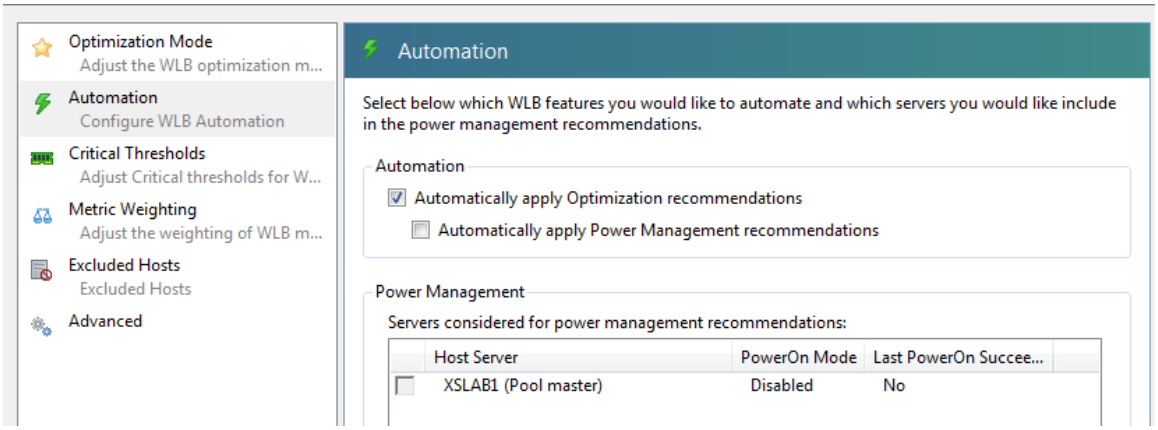

Figure 4Typical XenCenter Power Management Options

The power control company Eaton has created a XenCenter plugin that allows for much greater power control for the datacenter with only XenServer installed, but it is an added cost that is available free in the other hypervisors. If you run XenServer with either vSphere or Hyper-V, I highly recommend you install and use the VMM 2012 to manage all of those hypervisor's, as well as their power management features.

As with all of these feature sets and additional management products that I mentioned for power management, I cannot stress enough that you need to test these mechanisms in your environment before implementing them in production.