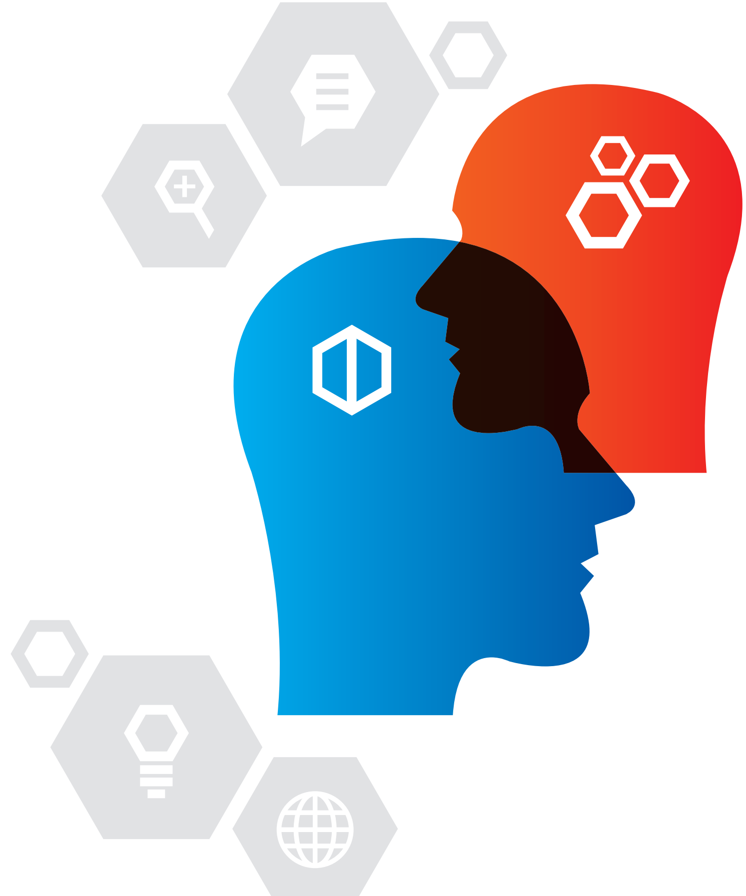
Power Consumption Efficiency



By Eli Karpilovski
Chairman - Cloud Advisory Council
Contributing Editor:
Scott Schultz,
Sr. Strategic Alliance Manager, AMD

Overview

The amount of power used by servers and other Internet infrastructure has become an important issue in recent years. One of the weaknesses of literature on data center power usage has been the lack of credible estimates of the aggregate power used by all servers and associated equipment in U.S. and the world. The estimates on floor area and power densities of data centers are questionable because most companies do not disclose such information and treat it as proprietary. Details of the installed base of servers at various companies are also not disclosed by the companies who track it. As server technologies continue to change rapidly, it is necessary to update power measurements for the affected server models quite frequently.

This white paper describes how to achieve power savings by designing a more efficient data center. It shares the calculations made while designing the servers and infrastructure of a data center. In the past, separate physical networks were required to handle each type of traffic using technologies including LAN, SAN, and interprocess communication (IPC) mechanisms. This article takes into consideration that switches enable I/O consolidation at the rack level, allowing LAN, SAN, and IPC traffic to be carried over the same link between servers and the Access Layer, while using the same driver software, management software, and data center best practices for both LAN and SAN.

Power Consumption Efficiency

A New Sustainability Metric: PCE

The Cloud Advisory Council (CAC) recognizes the importance of establishing metrics for data center su tainability. These metrics and their related processes should help organizations determine if existing data centers can be optimized before considering the building of new ones. For this reason, CAC proposes the use of a new metric, power consumption efficiency (PCE), to address inefficient server utilization associated with data centers. The impact of efficient server usage is emerging as extremely important in the design, location, and operation of current and future data centers. When used in combination with the power usage effectiveness (PUE) metric of Green Grid, data center operators can quickly assess the sustainability of their data centers, compare the results, and determine if energy efficiency and/or sustainability improvements need to be made.

Native I/O Virtualization Solutions

The server virtualization method is a blend of software and hardware-based I/O virtualization. The software-based VMM is still responsible for setting up the I/O resources for the virtual adapters in the VMs and for controlling these resources. This includes all memory mapping and setup operations, and the isolation of I/O resources across the VMs. Once the I/O resources are assigned, the VMs are allowed direct access to the I/O adapter for sending and receiving data, but DMA remapping functions are still controlled by the VMM. In other words, the data path between the VM and the physical I/O adapter bypasses the VMM, especially packet steering functions of virtual switches. However, the control path and memory mapping functions are maintained by the VMM. This requires installing on the VMs drivers that are specific to the physical I/O adapter. The VMM virtual switch may continue to be used for VM-to-VM communication, unless the I/O adapter supports switching between end points that connect to the VMs for the data path.

The native I/O virtualization method has the following advantages and disadvantages:

Advantages:

- I/O data path performance from the VMs is accelerated significantly
- I/O data path latency can be reduced significantly
- CPU utilization is significantly better than soft ware-based I/O virtualization, enabling more VMs per physical server an improving server utilization

Actual CPU Usage (ACU)

ACU can represent various statistics of CPU usage, like the number of VMs per server, IOPS, etc.

In this white paper we will use ACU to indicate the number of VMs per physical server.

VMm = f (C, M, N) where C = CPU, M= Memory, and N=Network
Number of VM(eff) = Min {VM(c)/C * VM(n)/N * VM(m)/M}

Optimal Scenario:
VM(c)/C = VM(n)/N = VM(m)/M
Number of VMopt = Max { VM(c)/C , VM(n)/N , VM(m)/M}

ACU will be limited by the requirements of the desired VMs – for example bandwidth performance per VM.

Theoretical Usage Capabilities per Physical Server (TUCPS)

TUCPS is a constant that is defined by the usability of the server. In this white paper we will define a specific server configuration of CPU, memory and hard disk in order to set the theoretical number of VMs that can be supported per server.

Direct Power Consumption

Data center power consumption can be broken down to information technology (IT) loads (such as servers, disk drives, and network equipment) and infrastructure loads (cooling, fans, pumps, lighting, and uninterruptible power supplies or UPSs). This study focuses on server loads which represent 60-80% of total data center IT loads.

Estimating power use for each server is not an easy task. The power use of electronic equipment varies with hardware configuration, application, and environmental conditions. The power supplies for these servers are normally sized for the maximum loads expected when the server is fully configured. But the actual measured loads observed in typical installations are usually much lower than the rated power of the power supply.

Power Consumption Efficiency

We define the power use per unit for each server based on measured data, on-line server configuration calculators, or estimates from manufacturer specification sheets, and multiplied the maximum measured electricity use, or the maximum rated input power of the power supply, by factors taken from industry experience to estimate typical power use.

Maximum measured electricity use is widely reported by some manufacturers (e.g. IBM) and for others (e.g., HP, Dell) it is possible to calculate it using on-line configuration tools. When we used such tools to estimate maximum measured power, we included the maximum amount of RAM, the largest size and number of hard drives, n+1 redundant power supplies, processor-intensive workloads, and two processors at the fastest clock speeds.

To convert such power estimates for high end servers to typical power use, we multiplied maximum measured power by 66%, which is the rule of thumb accepted by IBM.

Direct power consumption (million kW) is the product of installed base and maximum power use per unit (i.e. with a load factor of 100%).

Power Consumption Efficiency (PCE)

Power Consumption Efficiency (PCE), denotes direct power consumption multiplied by efficiency. Data center efficiency is defined by the following equation:

‘n = VMeff/VMopt

And if we assign X_T to denote the total power used by a physical machine, then the effective power will be:

$X_{eff} = X_T$ * ‘n {0 < ‘n < 1}

Note that the efficiency of a data center will be affected dramatically if ‘n is not close to 1.

Example: The total power for a specific cloud is 1000 [mkW] and the ‘n equal 0.8, the Power Consumption Efficiency is 1200 [w], meaning that the data center power consumption is not efficient as it could potentially be.

PCE Benchmark Tests

In a series of platform capacity tests, we will continually add more instances to the platform while running representative workload samples. The system will be considered to have reached maximum capacity when the cloud administration software no longer allows more instances or when the SYSBENCH indications fall below a minimum performance baseline.

Some cloud coordination tools will attempt to equally distribute the instances across all hypervisors in the cloud. We will disable such tools as the purpose of this test is to examine maximum performance utilization.

1. VM.Capacity.Small

Purpose: This test examines the capabilities of the platform for supporting a large number of small but uniform instances.

Description:

The test operator is to use the cloud platforms' coordination tools to create an increasing number of small instances. The test operator then conducts a set of simultaneous performance benchmarks. The test operator repeats the process until some of the instances can no longer meet the performance threshold.

Test Setup:

The cloud platform must be operational and have maximum resource availability. A single physical hypervisor server should be used. This test assumes that the instances can be selectively scheduled to be located on the same physical hypervisor.

For the purpose of running NETBENCH tests, a Network Test Server should be connected to the same production network as that of the hypervisor under test.

Procedure:

1. Create a TestVM:
   a. x86_64 and install the CentOS 6.1 OS minimum install (TBD – what is minimum install).
   b. Configure a single vNIC network connection such that it can ping the Network Test Server.

Power Consumption Efficiency

2 For each TestVM instance, execute a SYSBENCH test and record the average results across all VMs.

3. Repeat step 2 adding additional VMs to the set of in stances running on the nodes. Continue until the average node performance for SYSBENCH falls below the baseline. Record any starvation or outliers of the SYSBENCH in the additional comments section.

Results:

Observation	Step#	Baseline	Recorded
SYSBENCH 1 Nodes	2	TBD	
SYSBENCH 2 Nodes	2	TBD	
SYSBENCH 4 Nodes	2	TBD	
SYSBENCH 8 Nodes	2	TBD	
SYSBENCH 12Nodes	2	TBD	
SYSBENCH 16 Nodes	2	TBD	
SYSBENCH 20 Nodes	2	TBD	
SYSBENCH 24 Nodes	2	TBD	
SYSBENCH 28 Nodes	2	TBD	
SYSBENCH 32 Nodes	2	TBD	

2. VM.Capacity.Large
   Purpose: This test examines the capabilities of the platform for supporting a large number of large but uniform instances.

Description:

The test operator is to use the cloud platforms' coordination tools to create an increasing number of large instances. The test operator then conducts a set of simultaneous performance benchmarks. The test operator repeats the process until some of the instances can no longer meet the performance threshold.

Test Setup:

The cloud platform must be operational and have maximum resource availability. A single physical hypervisor server should be used. This test assumes that the instances can be selectively scheduled to be located on the same physical hypervisor.

For the purpose of running NETBENCH tests, a Network Test Server should be connected to the same production network as that of the hypervisor under test.

Procedure:

1. Create a TestVM:
   a. x86_64 and install the CentOS 6.1 OS minimum in stall

   b. Configure a single vNIC network connection such that it can ping the Network Test Server.

2. For each TestVM instance, execute a SYSBENCH test and record the average results across all VMs.

3. Repeat step 2 adding additional VMs to the set of instances running on the nodes. Continue until the average node performance for the SYSBENCH falls below the baseline. Record any starvation or outliers of the SYSBENCH in the additional comments section.

Results:

Observation	Step#	Baseline	Recorded
SYSBENCH 1 Nodes	2	TBD	
SYSBENCH 2 Nodes	2	TBD	
SYSBENCH 4 Nodes	2	TBD	
SYSBENCH 8 Nodes	2	TBD	
SYSBENCH 12Nodes	2	TBD	
SYSBENCH 16 Nodes	2	TBD	
SYSBENCH 20 Nodes	2	TBD	
SYSBENCH 24 Nodes	2	TBD	
SYSBENCH 28 Nodes	2	TBD	
SYSBENCH 32 Noes	2	TBD	

Power Consumption Efficiency

Summary

This white paper is a first step for laying out the process for calculating, measuring, and using PCE. The Cloud Advisory Council plans to expand on PCE with further detail. However, in the spirit of industry transparency and urgency, we have released this introductory white paper to encourage an industry discussion dealing with these metrics. Items already under consideration for future white papers include:

1) Detailed process for determining PCE.

2) Measuring PCE for mixed-use buildings.

3) Calculating PCE for other ACUs such as IOPS

and other metrics.

References:

- Estimating Total Power Consumption by Servers in the U.S. and the World
  Jonathan G. Koomey, Ph.D.
  Staff Scientist, Lawrence Berkeley National Laboratory and Consulting Professor, Stanford University

Contact: JGKoomey@stanford.edu
http://www.koomey.com

- Carbon Usage Effectiveness (CUE):
  A Green Grid Data Center Sustainability Metric