Fortuitous Technologies
http://fortuitous.com
512-351-7783

Capacity Planning: A Necessity for Green Computing

Dominique Heger, Ph.D.

1 Introduction

In general terms, green computing can be described as the environmentally responsible usage of computer systems and related resources. Some of the major components of green computing include the implementation of energy-efficient CPU's, server systems, and peripherals such as SAN and NAS subsystems. Further, green computing focuses on reduced resource consumption and the proper disposal of electronic waste (sometimes labeled as E-waste). One of the earliest initiatives toward green computing in the US was a voluntary labeling program known as Energy Star. It was conceived by the Environmental Protection Agency (EPA) in 1992 to promote energy efficiency in hardware components (components of any kinds). Over time, the Energy Star label became a common sight, especially for laptop computers and monitors. While some companies embrace the green computing movement out of some ethical principals, the majority of the organizations will only adopt the concepts of green computing for business reasons such as lowering the IT infrastructure costs. In other words, for these enterprises, the decision will rest with whether the actual investment represents a good return-on-investment (ROI).

Therefore, to build a business case, IT management has to understand the basic concepts of green computing and how the investment can improve their corporate social responsibility and their financial bottom line alike. Based on such a business case, implementing changes that positively impact the environment becomes the right ecological and economical approach. Hence, this report focuses on the basic elements that comprise green computing initiatives for small and large scale organizations, focusing on energy efficiency, reduction in cooling requirements, consolidation, and the importance of performance tuning and capacity planning.

2 Energy Efficient

Independent research studies conducted by IDG and Hewlett-Packard in 2006 revealed that 50 cents is spent on energy related expenses for every dollar of the computer hardware. This value is expected to increase by 54 percent to 71 cents over the next four years (till 2010). While many organizations do not consider the facility costs as part of the ROI and total cost of ownership (TCO) for IT projects, the increasing energy bills may force the issue. Some other study conducted by Hewlett-Packard outlines that simply leaving a computer system up and running consumes electricity and adds to the actual computing costs. Further, the use of screen savers on desktop and laptop systems does not save any energy. The study showed that a typical desktop PC with a flat panel LCD monitor requires about 100 watts (65 watts for the computer and 35 watts for the monitor). If that system is powered on 24x7 for one year, the system will consume 874 kilowatt hours of electricity, enough to release 750 pounds of carbon dioxide into the atmosphere, which is the equivalent of driving 820 miles in an average passenger car. Another study conducted at the Berkley National Laboratory outlined that the electricity demand for IT server systems doubled between 2000 and 2005. The conclusion made was that this is of course not a desirable or sustainable trend. In other words, tremendous savings could be realized by IT data centers if they could deploy more energy-efficient infrastructures and manage the systems in a way that would allow to only run the IT systems necessary to achieve the business goals of the company. In other words, too many companies oversize their IT infrastructure (based on a lack of understanding for performance management and capacity planning), and hence use energy for IT components that do not really contribute to an organizations bottom line. The economics for energy efficiency seems obvious across the board.

3 Reducing Cooling Demands

One artifact of IT technology is that it generates tremendous heat that requires cooling in order to ensure the stability of the equipment. Not only large server systems, but also desktops, laptops, and peripherals (such as IO subsystems, monitors and printers), generate a significant amount of heat that requires additional cooling. The cost of cooling an IT infrastructure are tremendous. Taking into consideration that most organizations are running IT equipment that is not necessary to achieve the actual business goals (oversized IT infrastructures are common across most business areas), significantly reducing the heat load will again result in substantial energy savings.

4 Consolidation

It is a fact that most IT organizations reflect a heterogeneous environment of server, network, and storage equipment that is mostly oversized and hence underutilized, poorly managed, and sometimes over- aged with no retirement planning in place. The buzzword in the industry is consolidation, focusing on reducing the complexity and maintenance cost of the IT infrastructure. It has to be pointed out though that consolidation only makes good environmental sense if the equipment is chosen wisely, is more efficient than the sum of its replaced parts, is operated in a way that optimizes the energy consumption and performance, and is accompanied by a solid retirement plan for the outdated equipment. In other words, a sound design, performance management and capacity planning process has to be in place to realize the potential provided by IT consolidation. Unfortunately, many consolidation projects result in introducing new IT equipment without the departure of the systems that they were intended to replace. Further, the issues of recycling and disposal of the retired equipment is sometimes not addressed at all.

5 Performance Management & Capacity Planning

IT performance management revolves around monitoring and analyzing the workload dependent performance of a system (or a group of systems) and adjusting the logical and physical resources in the application code-path to optimize performance and achieve the throughput and/or response time goals set for the IT solution.

Capacity planning on the other hand is a process that is being used to manage the IT resources. The goal is to ensure that the IT capacity that is available meets the current business requirements. Based on anticipated future growth patterns, capacity planning allows conducting sensitivity studies to determine the future hardware and software needs in a timely and cost effective manner. This is a proactive approach that aims at not oversizing (and hence not over-spending on) hardware and software components. The goal is to have adequate capacity available for the workload at hand, and therefore a rather small headroom (ample capacity) on the systems. Utilizing the available hardware and software components as efficiently and effectively as possible has to be the major goal of any of today's IT departments.

Having said that, the argument made is that green computing and performance management and capacity planning are closely related. In the early days of computer systems, engineers had to be very careful when developing software products. Systems with only a few KB (not MB) of memory and processors with single digit MHz (not GHz) forced the software developers into optimizing the systems very carefully and consciously. Over the

last 20 years, software development became mainstream and the system complexities (such as the OS kernel) became hidden (black box like) to the developers. Most software developers today only have a very limited understanding of the functionality of the kernel and the hardware. Hence, as most IT systems have more capacity than actually needed, a majority of IT departments do not care about actual resource allocation anymore.

Green computing on the other hand again cares very much about actual resource allocation. Via benchmarks, profiling, tracing, and instrumenting application code, the actual footprint of an application can be determined. Based on these workload profiles, application code, OS primitives, as well as hardware components can be tuned to optimize the systems. In other words, if less CPU, memory, network, or IO capacity is required to run the applications, less overall IT equipment is required to achieve the stated business (performance) goals. Therefore, less heat is produced by the IT equipment, which saves energy needed to control the IT environment. This will lower a company's energy bill, streamline the IT equipment, and hence make the systems more manageable. Green computing in conjunction with performance management and capacity planning will become a necessity for IT departments to control cost and stay competitive.

6 Summary

Improving power efficiency is one of the largest challenges faced by companies today, especially in large IT data centers and/or metropolitan areas where power demand is high and the available resources are being pushed to the limit. The issue of power efficiency is not only about escalating utility bills, but also about the question of how to support growth if running out of power capacity becomes more common. It is paramount that organizations today start with a comprehensive assessment and project plan for their IT equipment. Proactive evaluation, analysis, and optimization of the workload demands and IT infrastructures is required to determine the present and future energy consumption needs. The goal is to lower the IT infrastructure costs by lowering power consumption, retiring obsolete equipment, and better utilizing the existing IT infrastructure. Performance and capacity studies aid at achieving the goals, dramatic energy savings becomes a reality, and the ROI will increase significantly.