First USENIX Workshop on Sustainable Information Technology (SustainIT '10)

San Jose, CA
February 22, 2010

WELCOME AND OPENING STATEMENTS

Ethan L. Miller, University of California, Santa Cruz

Summarized by Priya Sehgal <psehgal@cs.sunysb.edu>

Ethan Miller welcomed everyone to the first USENIX workshop on sustainable IT. He began by explaining how sustainability is an important issue, encompassing much more than just power consumption. It is more about electronic waste, conserving natural resources, etc. There were around 45 attendees for this workshop, with one-third from outside the US. Ethan was glad to announce that people from both academia and industry had participated, and he thanked the National Science Foundation (NSF), for the student scholarships, and the industry sponsors, IBM Research and VMware.

INVITED TALK

- ***On the Science of Power Management: Encouraging Sustainability R&D***
  Erez Zadok, Stony Brook University

Summarized by Priya Sehgal (psehgal@cs.sunysb.edu)

Erez Zadok shared the seven key findings from the science of power management (SciPM), a workshop conducted by the NSF in 2009. In this workshop, people from industry, academia, and government identified, prioritized, and recommended promising research directions in the area of power management.

The first finding was to observe systems, i.e., simply measure and analyze what systems are doing, and disseminate the results. This would aid modeling of systems and optimizing them for power and performance. Second, develop useful and clear metrics (e.g., ops/sec, ops/watt). The challenge in metrics is how to account for long-term effects such as e-waste, carbon footprints, longer hardware lifetimes, etc. Third, develop models based on the most significant factors after one observes and develops metrics. Modeling is required at all levels from hardware to software, chip, system, and datacenter.

Optimization at various levels is another important finding. There are many point solutions, but the question is how useful it is to others. Computer systems are generally complex, with multi-dimensions such as reliability, performance, energy, and security. Optimizing for multi-dimensions is challenging, as optimizing one dimension could hurt another and vice versa. For example, rotating the disks faster could improve the performance of I/O but also cost a lot in terms of energy. Thus, there is a need for rigorous analytical techniques such as control theory.

Erez Zadok pointed out that there is very little education on power management within IT. Special graduate and undergraduate courses could help solve this problem, but, for now, it could be integrated with existing system courses. He gave the example of how security education took 15 long years to get into the mainstream of courses and stressed that we should not repeat this mistake in the case of energy literacy.

Cross-disciplinary workshops and scientific community interactions will boost research in the energy and sustainability domains. We need to think beyond just computing and datacenters. There is a lot of development potential for intelligent software and hardware techniques for use in smart buildings, smart power grids, automated transportation and the like.

During the last part of the talk, Erez Zadok discussed some of the research work going on in his lab, the Filesystems and Storage Lab (FSL). Erez and his students have been trying to unravel the intricate interactions among hardware, software, and workloads. The first survey was to address the question of whether compression helps save energy. Some of the results showed an improvement in energy and performance by 10–40%, while in others it was hurt by as much as a factor of 10 to 100. Thus, the impact of compression on both performance and energy is governed by the type of workload and such characteristics as read-write ratios and file or data type. According to other research, Erez and his students found that filesystem performance and energy depend on hardware and software configurations as well as workloads. Varying filesystem configurations could positively impact power/performance from 6–8% up to a 9-fold improvement. The third research project carried out in FSL was to study the mix of NFSv4 clients and servers and their effects on performance. They performed various benchmarks on clients and servers belonging to different platforms such as Linux, FreeBSD, and Solaris and found a performance variation of around 2–3 times.

INVITED TALK

- ***Reduced and Alternative Energy for Cloud and Telephony Applications***
  James Hughes, Huawei Technologies

Summarized by Vasily Tarasov (tarasov@vasily.name)

James Hughes explained that in telcos, most energy per subscriber (57%) is consumed by the base stations, while the datacenters consume only 6%. Although the mobile phone's embodied CO_2 emissions are twice as high as a base station's, a base station's operation CO_2 emissions are 3.5 times higher than a cell phone's. Overall, the ratio of base station to cell phone emissions is approximately 6:5, so optimizations are possible at both sides.

Ethan Miller asked what part of a cell phone consumes the most energy. Somebody from the audience confirmed that nowadays it is the screen.

In the datacenters, 33% of the energy is consumed by the chiller, 30% by the IT equipment, and, surprisingly, 18% by the UPSes. The speaker speculated that it might be possible to eliminate UPSes somehow, but somebody from the audience said that the UPS cleans the power, so if we remove it, we need to build equipment capable of working on dirty power, which might mean higher embodied energy.

Taking into account that cooling consumes a lot of energy, it makes sense to place datacenters in colder areas. In addition, to save on energy transition and distribution (7.2% loss in the US), it is efficient to co-locate datacenters and energy generators. A potential problem in this case is increased network latency. A fibre channel link between San Francisco and Chicago (2,250 miles) has 24ms optical latency. However, the measured latency is 130ms, so there is room for improvement.

In terms of server power consumption distribution, CPU, memory, and PSU consume 30%, 20%, and 18% of energy, respectively. The average server is only 10% utilized, and taking into account that idle energy is very high, it makes sense to consolidate severs. It might require more RAM, though: 25% of RAM power consumption is static; the other 75% is mostly based on access.

Hughes considers Fast Array of Wimpy Nodes (FAWN) as a promising direction, as well as Ceph, a distributed file system. However, the scalability of these technologies needs to be practically proved. Another interesting area is recycling. The average server lives only three years. So, how can we reduce embodied CO_2 emission of a server? And can we reuse waste efficiently?

Erez Zadok asked if one can increase the usage cycle by various software techniques. Hughes answered that software definitely constitutes the largest area for improvement.

INVITED TALK

- ***The Green Cloud: How Cloud Computing Can Reduce Datacenter Power Consumption***
  Anne Holler, VMware

Summarized by Priya Sehgal (psehgal@cs.sunysb.edu)

According to a report submitted by the EPA to Congress in 2007, energy consumption by US datacenters alone amounted to around 1.5% of the total national energy consumption in 2006, and it is expected to double by 2011. Anne Holler's talk envisioned how cloud computing could be exploited to achieve energy conservation in datacenters. Towards the end, though, she presented a few counter-arguments for how cloud computing could even increase datacenter power consumption.

According to NIST's definition of cloud computing, it is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (networks, servers, storage, applications, services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model is composed of five essential characteristics, three service models, and four deployment models. The five characteristics include on-demand self-service, broad network access, resource pooling, rapid elasticity, and measured service. The three service models are Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS), while the four deployment models include private, public, community, and hybrid cloud. Her talk was focused on IaaS, since it is often used in datacenter (DC) virtualization.

Holler presented facts about how VM consolidation saves power. According to some previous reports from the EPA, physical hosts are usually 5–15% utilized but still consume around 60–90% of their peak power. Based on a report from VMware, VM consolidation can achieve energy savings of around 80%. Resource Management (RM) plays a vital role in efficient VM consolidation such that it maintains specified QoS and high overall throughput. Some of the elements of effective DC virtualization RM are dynamic work-conserving allocation instead of static partitioning, rich resource control set (e.g., reservations, shares, limits), RM support across a cluster of nodes with the use of live migration (e.g., VMware VMotion).

The rest of the talk focused on how the cloud model could further increase the power savings in a virtualized DC environment. As the cloud model fosters VM consolidation and provides incentives to reduce operating expenses (OpEx), it looks attractive in terms of power savings. In terms of VM consolidation, cloud provides an aggregation point for workloads that would otherwise be executed on separate DCs. Thus, more pooling of workloads can better utilize multi-core hosts that tend to be more energy-efficient than narrower ones.

In terms of incentives, Holler pointed out that administrators of a private dedicated (non-cloud) DC often are held accountable for high resource availability rather than for power consumption OpEx. So there is a low incentive to reduce power use but significant incentive to over-provision for peak usage. Similarly, users of a private DC are often entitled to waste resources based on their capital expenditure (CapEx) and are not billed for resources their VMs consume. This leads to low incentive to reduce VM resource consumption at steady and peak loads. Cloud providers want low OpEx that can be achieved by high utilization of powered-on computer resources and keeping a few low-cost computers as spares for peak demands. Cloud providers also try to achieve low CapEx through limiting the DC capacity. Thus, cloud providers are more aggressive in reducing OpEx and CapEx than a dedicated DC administrator.

Someone from the audience asked whether Holler was making some assumptions regarding the SLAs in cloud and she answered that in cloud one expects some SLA, depending upon the workload, but they can still take some risks that a dedicated DC would not.

One way cloud providers reduce OpEx is by powering off the hosts when demand is low and waking them up when demand increases (e.g., VMware DRS with DPM enabled), and powering off the less efficient hosts. Another important technique includes using demand prediction and proactive RM: knowing, for example, that there is a steep rise in demand at 8 a.m., while at 10 p.m. it is quiet. Discounting off-peak usage (e.g., running a low-priority MapReduce job at night, when the rates are low) is another viable option. Also, for multi-site providers, moving the workloads to sites providing the cheapest power (VMotion) could help curtail OpEx.

Cloud users also want low OpEx (i.e., a basis to trade CapEx for OpEx). Resource pooling removes CapEx resource entitlement, the measured service model highlights OpEx costs, while rapid elasticity avoids OpEx for peak usage until absolutely necessary. In order to reduce the OpEx, users can characterize computing resources needed for applications and set the QoS SLA appropriately. Users can reduce resource usage of application workloads by right-sizing the VM (e.g., matching the number of virtual CPUs to workload parallelism) and reduce usage when workload is light/idle (e.g., use tick-less OS). Someone from the audience pointed out that tools that could help automatically reduce the OpEx would be very useful, and Holler agreed. There are some challenges in designing these tools—for example, when should the guests go to low power state?

Erez Zadok asked how security in the cloud would impact power consumption. Holler replied that the usefulness of the cloud might decrease in terms of power reduction if we add the dimension of security. She also said that this point could act as a counter-argument for clouds saving energy.

DON'T THROW IT AWAY

Summarized by Vasily Tarasov (tarasov@vasily.name)

- ***Estimating Environmental Costs***
  Kiara Corrigan, Amip Shah, and Chandrakant Patel, Hewlett Packard Laboratories

Kiara Corrigan presented their effort to estimate an IT infrastructure cost in the event of increased electricity prices or additional environmental taxes. Producing a computer involves a long supply chain, and every link in the chain is affected when prices and taxes go up. It is not easy to predict the new price of the final product in such a case, but the authors do so by making use of the Economic Input-Output (EIO) model in conjunction with the Life-Cycle Assessment (LCA) method. The main object in the EIO model is an inter-industry transaction matrix that exhibits the flows of the goods between different industrial sectors.

Using the EIO-LCA model, the authors consider three scenarios: (1) constant electricity prices, no carbon tax; (2) 3% increase in price (industrial rate), $10 carbon tax; (3) 3% increase in price (residential rate), $50 carbon tax.

Estimating the costs of an enterprise IT portfolio—consisting of thousands of laptops, desktops, servers, etc.—their model shows the potential increase in TCO varying from 1% to 12% depending on the scenario.

Someone asked if the model takes disposal costs into account. Corrigan answered that waste is one of the sectors, so the model is capable of accounting for disposal costs. How was the model validated? It was not; at the moment, its purpose is to compare prices between different portfolios but not to project precise costs.

- ***The Green Switch: Designing for Sustainability in Mobile Computing***
  Galit Zadok and Riikka Puustinen

Galit Zadok and Riikka Puustinen presented their methodology for designing energy-efficient mobile devices. Specifically, they concentrated on cell phones during the use phase.

There are 4.6 billion cell phone users in the world at the moment, and the projected number for 2013 is over 6 billion. More smartphones are also appearing: by 2015 all the handsets sold will be smart. An average user replaces a cell phone every 18 months. These parallel trends indicate that the impact of the cell phone industry on the environment is rapidly growing and can go out of control if no appropriate actions are undertaken.

There are certain efforts to reduce the environmental impact of cell phones at the manufacturing and disposal phases. However, less is done at the customer use phase, which, for the iPhone 3GS, for example, accounts for 49% of its greenhouse gas emissions. The Green Switch methodology suggested by the authors allows evaluating at the design stage whether a product or a service fulfills both human and ecological needs. The methodology can be distilled to a checklist that contains human and green appeals. Human appeals evaluate if the product designed is beneficial for the user, convenient to use, has good value, and is socially acceptable. The only green appeal at the moment is the reduction in energy use. As you can see, there are more human-centric appeals than eco-centric ones. This is because sustainable design solutions can have a significant positive impact on the environment only if these solutions are widespread and thus adopted by the mass market.

As an example of the Green Switch methodology, the authors presented the Green Mode concept for mobile phones. The idea is to separate the applications that are frequently used by a specific user from the rarely used ones. When a cell phone runs in the Green Mode, only frequently used services are available. This saves energy by not running background processes of non-used services. When the user wants to have access to all cell phone features, she can switch the phone to the so-called Fat Mode.

Someone asked if the authors have looked at the possibilities of reusing components from old phones in the new phones. The answer was no. The reason why this practice is not widespread nowadays is because manufacturers use proprietary designs and prefer to keep this information secret.

More information about the Green Switch methodology can be found at www.thegreenswitch.org.

- ***Towards Integrated Datacenter Energy Management: An IBM Research Strategic Initiative***
  Jody Glider, IBM Almaden Research Center

  Summarized by Priya Sehgal (psehgal@cs.sunysb.edu)

Jody Glider walked us through some of the initiatives taken by IBM to improve datacenter efficiency. IBM announced Project Big Green in May 2007, which was intended to create a family of energy-saving solutions spanning the entire stack, from hardware to software to facilities based on the large set of monitoring and management tools. In 2009, IBM launched the Data Center Energy Management strategic initiative (DCEM SI), whose goal was to identify and exploit synergies across 15 research laboratories worldwide, a kind of Project Big Green in microcosm.

The mission of DCEM SI is to deliver energy monitoring and management products to the market by working with IBM product and service partners, to develop and showcase these solutions in living datacenters, and to quantify the resulting financial/energy savings. Some of the DCEM SI projects include energy-proportional storage, energy-aware dynamic server consolidation, monitoring and visualization, novel approaches to cooling, and processor microarchitecture/memory design, etc. Out of the many DCEM SI projects, Glider presented the Mobile Measurement Technology project. He also talked about research in storage energy management.

The IBM Mobile Measurement Technology (MMT) project consists of three steps to help save power in a datacenter: (1) capture high resolution temperature, air flow, infrastructure, and layout data; (2) model the datacenter and identify improvement opportunities through optimized algorithms; and (3) manage best practices (realize air transport energy savings, thermodynamic energy savings, etc.) for reduced energy consumption. This project achieved energy savings of around 177kW with the measurements included, and the ROI was around 1–2 months. Since the datacenter can change over time, MMT has evolved from a static to a dynamic solution, which makes use of a base model and deployment of real-time sensors in the datacenter.

Before talking about research in storage energy management, Glider briefly discussed the facts about storage energy consumption. According to a report from StorageIO, storage accounts for up to 37–40% of the energy consumption of all IT components. Energy usage in the storage domain is expected to get worse as storage unit sales outpace server unit sales in the next five years. With the slowing of the per-drive increase in storage capacity and with increasing storage subsystem performance demands, there will be a need for more physical drives. Some of the strategic directions for storage energy efficiency are monitoring, modeling, control, and optimization. Better monitoring will drive control and optimization. One can optimize for better performance vs. more energy savings.

Glider talked briefly about the storage power modeling and estimation work going on in Haifa; people who are interested can contact Kalman Meth (meth@il.ibm.com). In this work Kalman Meth and his team tried to model and estimate the power consumed by a storage subsystem as a function of a particular workload and storage configuration. The total power of the disk is broken down into two main components: fixed power and dynamic power. Fixed power is the power taken up by the rotation of the spindle and the electronics on the disk. Dynamic power is the power required to perform all the I/O, i.e., to perform head seeks and data transfers. The mechanical components draw electricity from aa 12V electrical channel, while electrical components use a 5V electrical channel. Glider showed an interesting graph which showed power consumed by the 5V (electronic) and the 12V, when running various levels of 4K random read workloads on an enterprise disk.

Although the 5V part dominated the overall power consumption and remained constant for all kinds of workloads, the dynamic part was affected by the disk activity and is important in enterprise systems. The interesting thing to note in the case of dynamic power was that initially, as the number of I/O requests increased per second, the dynamic power increased too. But after a certain number of I/O requests per second, dynamic power remained constant. This is because, with a larger number of concurrent I/O requests, the disk controller can effectively reorder the requests to shorten seek times. Power estimation consisted of constructing "Power Tables" formed by a set of pairs denoting seek activity <#seek,power> and data transfer <MBPS, power>. These power tables are later used in linear interpolation to estimate the disk power under a specific workload. This work was validated against the SPC-1 benchmark and it had an estimation error of 2.5%.

Glider presented a table of various energy optimization techniques in the storage subsystem—consolidation, tiering, opportunistic spin down, MAID, adaptive seek speeds, deduplication, etc.—and the potential energy savings achieved from each. Ethan Miller asked whether one could get extra energy savings by combining two or more of these optimization techniques or would it hurt. Glider answered that it depends on the type of the system or workload. For example, a critical airline application would take the least number of such optimizations in order to reduce the response time. On the other hand, an archiving application could make use of one or more such optimizations. Among the combination of techniques, consolidation and tiering can go hand-in-hand, while spin-down can be used where response time is not important.

In terms of tiering, Glider presented various combinations of HDD and SSD and their repercussions on cost, energy, and performance. He briefly explained the "tiering via flash cache" project. The flash cache is a block layer sandwiched between the file system and RAID layers. It caches all disk accesses onto SSD, allowing spin-down of a RAID rank that has been idle for *N* seconds. They obtained good results for a few of the MSR traces; file server trace saved 7% more energy than SAS, while firewall/Web proxy trace consumed 11% less energy than SAS. "Dynamic tiering and consolidation extent migration" (DTAC), another variant of the optimization technique, performs extent-based dynamic placement into tiers of storage after matching the performance requirement with the most appropriate tier. DTAC then consolidates the data and turns off the drives not needed. The results of DTAC looked promising in terms of both performance improvement and energy reduction compared to pure SAS configurations. (Further information on DTAC can obtained from hpucha@us.ibm.com.)

Future work includes pushing storage systems towards more variable energy cost components than fixed ones. Demand response planning is an important direction to focus on. Coordinated energy consumption optimization, i.e., ensuring that optimization in one area does not defeat optimizations elsewhere, and unified modeling and analysis to obtain optimal performance/energy are other interesting research areas that IBM is concentrating on.

Someone from the audience asked if one could replace RAM (volatile memory) with SSD and see its effect on performance and energy. Glider replied that it is not so easy and straightforward to speculate on performance and energy behaviors after replacing RAM with SSD, as the two pieces of hardware have different characteristics and cannot be used interchangeably.

PURE ENERGY

Summarized by Priya Sehgal (psehgal@cs.sunysb.edu)

- ***Power-aware Proactive Storage-tiering Management for High-speed Tiered-storage Systems***
  Kazuhisa Fujimoto, Research Institute of Electrical Communication, Tohoku University; Hirotoshi Akaike, Systems Development Laboratory, Hitachi Ltd.; Naoya Okada, Kenji Miura, and Hiroaki Muraoka, Research Institute of Electrical Communication, Tohoku University

Kazuhisa Fujimoto proposed an energy-efficient and high-speed tiered-storage system that minimizes performance loss, called eHiTs. eHiTs consists of a tiered storage system with high-speed online storage as the first tier and low-powered nearline storage as the second tier. The main idea behind eHiTs is to conserve energy by minimizing online storage capacity and powering off the HDD enclosures of the nearline storage when not needed.

eHiTs was evaluated against an HPC application. The HPC system consists of a supercomputer with an HPC management server. In eHiTs, the online storage tiered with low-powered, high-capacity nearline storage and network-attached storage offers file access to the supercomputer. All the files are always stored on the nearline storage on creation. Jobs executed in the HPC are controlled by the scheduler inside the HPC management server. Each job's script specifies the list of input files and the directory location of the result files. Based on job submission and execution time, eHiTs copies the user volume or the data needed by the job from the nearline to online storage at an appropriate predicted time. After the copy operation, the nearline HDD enclosure is powered off. After the job completes execution, the results are copied to nearline storage. Also, the former volume that was copied to online is copied back and remounted to the user directory. This nearline storage can then be powered off again.

The results obtained from their testbed with 64TB capacity showed that the system was able to conserve up to 16% of energy consumed by an ordinary tiered-storage system with the same capacity.

- ***Towards Energy Proportional Cloud for Data Processing Frameworks***
  Hyeong S. Kim, Dong In Shin, Young Jin Yu, Hyeonsang Eom, and Heon Y. Yeom, Seoul National University

Hyeong Kim investigated the feasibility of using power-save mode (PSM) in cloud computing. He basically tried to answer two questions for data processing frameworks: (1) the feasibility of low-power computers instead of commodity servers, and (2) the practical challenges in enabling PSM. To answer these questions, they evaluated Apache Hadoop on four different classes of machines. They also proposed AnSwer (Augmentation and Substitution), an energy-saving method to reduce energy consumption by augmenting the power-hungry servers with low-powered ones in correct proportion.

The machines evaluated by Kim included two server class machines, called Svr1 and Svr2, and two low-powered machines using Atom CPUs, called Low1 and Low2. He reported the results of running two Apache Hadoop workloads, mainly sort and gridmix, in units of normalized running time and performance/watt. Although Low1 and Low2 exhibited a performance degradation by a factor of 2–3 compared to Svr1 and Svr2, they turned out to be more power-efficient by a factor of 14 to 113. This motivated him to propose AnSwer, which makes use of these low-powered machines.

According to Kim, suspending partial servers leads to inevitable problems such as data loss and performance degradation. Using AnSwer, a rack of low-powered servers can be used for reliability. Whenever a high-powered server is suspended, its replica can be placed at this low-powered remote rack. This ensures reliability at reduced cost. Thus AnSwer uses two techniques, augmentation and substitution. Augmentation reduces the data transfer caused when auxiliary nodes replace existing servers, while substitution reduces the impact on data processing by replacing high-end servers with low-powered ones. Some of the practical challenges which he did not address completely are part of future work: which nodes to choose for partial suspension, and where to migrate lost replicas.

INVITED TALK

- ***Storage Class Memory: A Low-power Storage Opportunity***
  Richard Freitas, IBM Almaden Research Center

Summarized by Vasily Tarasov (tarasov@vasily.name)

Richard Freitas told us about the current status in Storage Class Memory (SCM) technology. SCM is a new class of storage/memory devices that blurs the difference between memory and storage. It is non-volatile, has short access time, is cheap, and does not have moving parts. There are several dozen technologies competing to be the "best" SCM. Flash memory is a widespread example, but Richard believes that in the future the majority of SCM will be represented by the Phase Change Memory—PCM.

Areal density of conventional disk drives keeps growing by approximately 40% a year. However, access time is proportional to the linear density, approximately the square root of areal density. Consequently, disk access time reduces by only 15% a year. Due to physical and cost limitations it is hardly possible that we will have disk drives rotating faster than 15,000 RPM. Consequently, the gap between CPU performance and disk access time continues to widen. Additionally, space and power become larger concerns. To satisfy 2 GIOP/sec one needs 5 million HDDs, which occupy 16,500 sq. ft. and consume 22 megawatts of energy!

SCM is a possible solution for the problems above, and flash memory is the most widespread SCM technology nowadays. It is based on the classical MOS transistor with a redesigned transistor gate. This allows placing or removing charge from the transistor, which corresponds to 0 and 1 values of a bit. Flash provides much lower access times, but has well-known endurance problems. Additionally, the smaller the flash cell, the higher the areal density, which is desirable. But already in certain flash-based products you can almost "name" every electron residing near the gate.

PCM is similar to the technology used nowadays in DVD disks. As it turns out, if you put a piece of DVD-like material between two electrodes, heat it, and then cool it down, depending on how you do it the resistance of the material will be different. Using the difference in resistance one can encode and decode zeros and ones or even multiple values in one PCM cell (which is similar to Flash MLC). The potential problem for PCM is the high current required to heat the cells. Freitas' crystal ball suggests that as early as 2016, the prices on PCM might be comparable to the prices on flash memory. Once SCM memory is available there will be a lot of ways to redesign current computer architecture appropriately.

If it's so fast, should we put it near to CPU, that is, in front of the I/O controller, or should we keep it behind the I/O controller? The presenter thinks both will be appropriate. There might be classes of SCM: very fast and expensive ones that should fit near the DRAM modules, and slower and cheaper SCM that can go behind the I/O controller, next to the disk drives.

IMPROVING THE (DATACENTER) ENVIRONMENT

Summarized by Priya Sehgal (psehgal@cs.sunysb.edu)

- ***Effects of Datacenter Vibration on Compute System Performance***
  Julian Turner, Q Associates

Turner presented an interesting paper in which he tried to determine whether ambient datacenter vibrations affect storage system I/O and performance. This was basically prompted by a YouTube video called the "yell test" performed by Brendan Gregg of Sun Microsystems, in which he demonstrated the adverse impact on I/O when he yelled at a running storage system. Another goal of Turner's presentation was to determine if any performance degradation found could be reduced or eliminated using specially designed anti-vibration rack by Green Platform Corporation (GPC).

In order to evaluate this behavior, Turner ran a number of benchmarks, mainly micro and macro benchmarks from FileBench, on a Sun 7110 array with sixteen 300GB, 10K

RPM SAS disks against two environments. The first environment was a specially constructed sound room with ambient noise less than 40dB and with no source of vibration within two miles. The second environment was a Tier 1 raised floor datacenter. The second environment was characteristic of an enterprise datacenter, with vibrations from compute nodes, A/C equipment, UPS, etc. Also, in the second setup, he ran experiments on two types of racks: a metal CPI rack and a GPC anti-vibration rack (AVR). The Sun Analytics tool was used to capture and report different characteristics of the disks, an important one being disk I/O broken down by latency.

Turner observed that the vibrations did impact random read and writes significantly. Performance numbers on the metal rack in environment 2 were worse than those on AVR and the ideal environment, 1. Performance improvements for random reads ranged from 56% to 246%, while for random writes it ranged from 34% to 88%. Streaming sequential reads and writes experienced a much smaller improvement than its random counterparts. Another important observation made by Turner was about the "latent performance effect," i.e., when a system was moved from, say, a metal rack to an AVR or vice versa the performance numbers remained roughly equivalent to the previous system for some time. Not taking this latency effect into account would lead to significant errors in the benchmarks. In his view, the old state of the system (e.g., cache hits) causes these latent performance effects.

Finally, Turner raised some research questions: (1) how much vibration can be allowed in a datacenter without any performance degradation? (2) how can these vibrations be mitigated? (3) should a datacenter really care about it? Yelling measures 130dB+ and is definitely detrimental, whereas a datacenter operates at 80dB to 95dB, which could still affect performance adversely.

Someone from the audience asked whether they measured the amplitude of the vibrations. Turner replied that they could measure the frequency but not the amplitude.

- ***CFD-Based Operational Thermal Efficiency Improvement of a Production Datacenter***
  Umesh Singh, Amarendra K Singh, Parvez S, and Anand Sivasubramaniam, Tata Consultancy Services Ltd., India

Umesh Singh talked about the application of a computational fluid dynamics model (CFD) in a production datacenter for improving its operational efficiency. The model he described was based on conjugate heat transfer and fluid flow at a datacenter. The authors carried out CFD analysis for a mid-sized (~3000 sq. ft.) datacenter. The model was used for providing design and operational guidelines to improve the energy efficiency of the servers and also to help in ramping up the partially filled datacenter. The guidelines relate to layout and airflow modifications required for an efficient cooling system while avoiding hot-spot formation.

As the first part of the modeling procedure, information regarding the geometry layout of the equipment, relevant thermal loads, and operational conditions was collected. Using the geometrical details, a 3D layout of the datacenter was created. This model was meshed with the hexahedral elements, consisting of hangers, beams, etc. The mesh files were loaded into CFX, software to model air-flow and heat transfer, with necessary boundary conditions incorporated. The computational solution for this problem setup was generated through CFX-Solver and results for different parametric studies were reported.

This model was validated with temperature measurements. Some of the recommendations obtained from this model which helped improving the energy efficiency included optimum placement of tiles in cold aisles and increased supply temperatures from CRAC units. These recommendations resulted in 20% energy savings in a datacenter that was ramping up.

PANEL SESSION

- ***The Present and Future of Sustainability R&D***
  Moderators: Ethan L. Miller, University of California, Santa Cruz; Erez Zadok, Stony Brook University

  Panelists: Kirk Cameron, Virginia Tech; Douglas H. Fisher, National Science Foundation; Dushyanth Narayanan, Microsoft; Amip Shah, HP Labs; Matt E. Tolentino, Intel

  Summarized by Vasily Tarasov (tarasov@vasily.name)

Dushyanth Narayanan (Microsoft) voiced a provocative opinion about the pointlessness of sustainability in IT. If one averages the amount of energy used per person in the world, 40% goes to cars, 30% to jet flights, and only 5% to electronic appliances, which include not only computers, but vacuum cleaners, microwave ovens, etc.! So it is unlikely that there are big sustainability targets in IT. But it makes sense to use IT for optimizing energy use in other areas. Somebody disagreed: optimizations need to be done in every area to make people think "green," and IT has one of the widest proliferations.

Amip Shah from HP Labs pointed out in his speech that sustainability becomes a large-scale problem. Though a lot of studies are concentrated on devices, "the device is not the whole system." If you look at any modern device, there are multiple industrial levels below that allowed creating it, and inter-industry interconnections are very large and complex. Due to these connections, it is possible that improvements in one sector that consumes only 2% of energy (IT) will affect all other industry sectors drastically. Shah concluded that we need to use large system models (some of them already exist, e.g., in economics) to see the whole picture.

Kirk Cameron from Virginia Tech emphasized that different applications exercise different systems differently, e.g., use systems' components unevenly. Consequently, power optimization must be specific to the use scenarios. He showed as an example his tool, MicroMiser, wich allows saving energy on Linux/Windows machines.

Matt Tolentino talked about what Intel does to decrease their footprint on the environment. First, modern CPUs consume 10^-7 less power than decades ago, and they are 10^5 times faster. Intel introduces eco-friendly package materials, reduces the carbon footprint of their factories, cleans wafers using recycled water. Interestingly, Intel is the largest green power consumer. The main question for Intel now is how to bring Intel's green technologies to other industries: how to effectively use it in Smartgrid, build better turbines, etc.

Douglas Fisher from NSF presented a lot of programs and funding opportunities that are for sustainable IT research.

Ethan Miller asked about the possibility of replacing computer components as a measure for increasing IT infrastructure life cycle. In fact, nowadays, people just throw computers away, instead of upgrading them. Amip Shah agreed that this is a big problem.

Somebody asked about solar energy perspectives. Dushan answered that for solar energy to be widespread, a distribution network is needed. Additionally, there is no silver bullet: some devices won't work on solar energy.

Douglas mentioned the use of reduced functionality as a way to save energy. For example, thin clients might be a good way to go.

Tolentino noticed that economic factors are more important than sustainability. Dushyanth answered that revenue optimizations sometimes include environmental factors, especially when the society becomes more environmentally responsible.