Sustainable Computing REGISTER

Monthly newsletter of the IEEE Computer Society Special Technical Community on Sustainable Computing

Providing quick access to timely information on sustainable computing.

http://www.linkedin.com/groups/IEEE-STC-on-Sustainable-Computing-4092681
http://www.facebook.com/groups/STC.Sustainable.Computing/
http://stc-sustainable-computing.ieee.net/members

Volume 1 Issue 5 May 2012

Contents

The IEEE STC on Sustainable Computing

From the Chair's Desk:

How will we deal with the waste generated by retired computing devices?

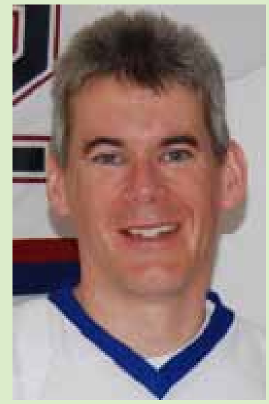
Energy is one of the most valuable and scarce resources available to humanity. Some utilization of massive energy consumption is an escalating threat to the environment as well as the economy. The explosive growth in computing is leading to rapidly increasing consumption of precious natural energy resources such as oil and coal, increasing the looming danger of an energy shortage, as well as rare minerals and metals used in computer construction causing pollution and driving up costs. In addition to energy and thermal problems, computers are built with rare materials: metals, rare earths, and plastics that are seldom recycled, pollute when used in fabrication, and increasing fill landfills. Displays, monitors, disks, and other peripherals are often as bad as or worse than processors and memory boards.

Some estimates show that there will be more than a billion computers (in some form or the other) around the world in a few years. If the life expectancy of a computer is 3 years, then how and where are these computers or their parts going to be dumped after they are no longer usable, and more importantly, how is the pipeline of this waste going to be handled, given its environmental hazardous impacts? These are going to be some very important questions that computer scientists and engineers will soon have to ask themselves.

Sustainable computing encompasses considerably more than energy-aware computing. It also means computing for sustainability that includes using computers to facilitate better sustainable environments. It also includes enhancing the generation, distribution, harnessing and utilization of natural energy resources such as electricity, water, wind, etc., and safe guarding the planet from some of the haz-

Co-Chair: Ishfaq Ahmad
U. of Texas at Arlington

Co-Chair: Martin Arlitt
HP Labs/U. of Calgary



Officers

Secretary/Treasurer: Giuliano Casale
Imperial College

Conferences Chair: Diwakar Krishnamurthy
University of Calgary

Academic Chair: Niklas Carlsson
Linköping University

Membership Chair: Anirban Mahanti
NICTA

Communications Chair: Abishek Chandra
University of Minnesota

Policies and Procedures: Stephen Dawson
SAP

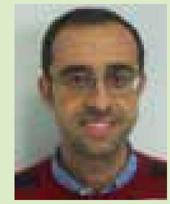
Industry Chair: Canturk Isci
IBM

Newsletter Editor: Christopher Stewart
The Ohio State University

Webmaster: Raoufehsadat Hashemian
University of Calgary

Resource and Energy Management

By Danilo Ardagna, Politecnico di Milano



European Research on Green ICT

A lot of effort has been spent in Europe to develop GreenICT solutions. During the seventh framework programme (the current funding programme created by the European Union in order to support and encourage research in Europe since 1984) up to now 10 projects have been funded among STReP(Specific Targeted Research Projects), IP(Integrated Project), PPP (Public-Private Partnership), and NoE (Network of Excellence) with a total cost of 124.8€ MLN and a total European Union contribution of 84.4€ MLN. I report below the main goals of the previous and running funded projects:

- Geysers (Generalised architEcture for dYnamic infraStructure sERvices): The objective of the project is to allow Optical network infrastructure providers and IT providers to compose logical infrastructures and rent them out to virtual operators. This is achieved with virtualization techniques through mediation layer to control the infrastructures and to optimize the energy consumption.
- Econet (low Energy COnsumption NETworks): The project deals with green networking for wireline including access, devices and metro segment. The techniques adopted are related to power scaling and idle mode of systems, boards and devices. Path engineering based on low consumption is also applied. The technology proposed is based on a green adaptation layer as general control layer for energy aware services.
- Fit4Green (Federated IT for a sustainable environmental impact): FIT4Green is a STREP project in the ICT for Energy unit, funded within the FP7 program, challenge 6, objective 6.3. FIT4Green's goal is to develop energy aware optimization policies for service centers which will reduce the energy consumption of their ICT infrastructure, without compromising compliance with SLA and QoS metrics. The FIT4Green approach will be potentially applicable to any type of service center with any automation framework.
- All4Green (Active collaboration in data centre ecosystem to reduce energy consumption and GHG emissions): All4Green is a STREP in the ICT for Energy unit, funded within the FP7 program, challenge 6, objective 6.3. All4Green broadens the scope of energy savings to the full ecosystem in which service centers operate. The novel approach proposed by All4Green specifically fosters collaboration between all entities in this ecosystem with the common goal of saving energy and emissions: ICT users deploying services in the service center, electrical power providers, and service centers cooperating in a federated way.
- OPTIMIS (Optimized Infrastructure Services): OPTIMIS aims at creating a dependable ecosystem of Cloud providers and consumers that altogether become the foundation for an efficacious operation of services and infrastructures. Trust, risk, eco-efficiency and cost are the main factors considered in the project for optimizing deployment and execution of applications in the Cloud.
- FI-WARE (Future Internet Core Platform): FI-WARE is an FP7 Future Internet PPP project aiming at developing an innovative infrastructure for cost-effective creation and delivery of services. The underlying infrastructure is Cloud based and the project provides directives for the implementation of efficient and interoperable systems at the full Cloud stack.
- GAMES (Green Active Management of Energy in IT Service centres): GAMES is an FP7 project aiming at developing a set of methodologies, software tools and services, and innovative metrics for the energy efficiency design and management of next generation service centers operating in Cloud computing environments.

Continued on page 5

Sustainable Computing Register

PICK OF THE MONTH
MAY 2012

<u>Power Budgeting for Virtualized Data Centers</u> by Harold Lim, Aman Kansal, and Jie Liu published in the Proceedings of the Usenix Annual Technical Conference.

From large Internet companies to medium enterprises, virtualization technology is now a key component in almost all data centers. However, as we know, power concerns are equally pervasive. The pick of month provides an accounting framework to determine the power needs of each virtual machine within a datacenter based on 1) the workload it supports and 2) the hardware it uses. Then, by adjusting the hardware online, the pick of the month achieves its title, power budgeting. The economic value of a power budget can hardly be overstated, as we will discuss in the interview. This month, Aman Kansal from Microsoft Research will share his reflections on this paper.

Christopher Stewart: Congratulations. Your paper was the first pick of the month selected by popular vote of the STC membership. In chatting with voters, I think people were wooed by the strong practical and technical merit of paper. That said, let's begin this interview with a slightly silly, off-the-wall question (*smile*). Is there an analogy between "Power Budgeting for Virtualized Data Centers" and power budgeting for virtualized people? I multi-task between well encapsulated (virtualized) tasks a lot. Is there a life lesson in this paper?

Aman Kansal: As far as raw Joules are concerned, for people the typical problem is being able to burn more energy rather than conserve it (*smile*). Probably the lesson for life is that when stretched for resources, be ready to throttle down lower priority tasks selectively without compromising on what really matters.

Christopher Stewart: Well said. Your paper is about power efficiency, a first cousin to energy efficiency. Looking at the trends in power production, data center growth, and systems design, which metric do you think is more important over the next 5 years? 10 years?

Aman Kansal: Currently, power efficiency is more important because when consuming energy at large scale, it's the capacity cost that matters. The actual consumption is paid for at deeply discounted prices. In the coming 5 or so years, this will not change since it costs the utilities a lot more to build new capacity and to fire up "peaker" power plants than to produce extra energy from existing capacity. In the next ten years or beyond, if the penetration of renewable energy resources increases, energy will be essentially free and only power efficiency will matter. At a solar power plant, it does not cost much extra to produce more energy once a solar panel is installed, but new panels do cost more and then it is only the power capacity that matters. Energy efficiency will still remain relevant as it indirectly helps improve power efficiency as well. For instance, better energy management can help reduce the size of battery storage required.

Christopher Stewart: Following up on that question, there are two approaches to power efficiency: minimize power while keeping throughput the same or maximizing throughput while keeping power the same. Do you see more practical value in one approach versus the other?

Aman Kansal: If a data center pays a penalty for every MW of capacity exceeded but the increased workload [generates profit]..., then the first approach of minimizing power at desired throughout would be better.... For current data centers, the power capacity is a hard limit.... Hence the latter approach is more relevant.

Christopher Stewart: Let's dive into the technical details. If I am correct, your work depends on power metering at the VM level. However, some datacenter managers simply don't trust today's power meters. Or more precisely, their customers don't trust them. Is there an incremental way to apply your approach?
Aman Kansal: Very much so. All customers that do not trust the power metering and capping mechanisms can simply be treated as high priority applications that will not be capped. The data center can internally measure their power and cap other [applications] that are willing to reduce power in exchange for discounts....

Cloud design also facilities a way to build customer trust. In cloud platforms, a key design principle is to design applications using scale-out techniques as opposed to scale-up. What this means is that each customer has multiple parallel instances of their VMs for each stage or tier of their [application]. Customers who wish to test the waters before diving in, can conduct A/B tests by opting in for power budgeting for only a small fraction of their VM instances. As and when the customers see a concrete cost advantage without undesirable performance consequences, they may increase their opt-in rate.

Christopher Stewart: Your tests were conducted on an 11-node rack. As you note in Section 2, nonlinear effects could arise when many racks are connected to a large uninterruptible power supply (UPS). For example, UPS efficiency degrades nonlinearly as power needs decrease. Would extending your power budgeting approach at this scale represent a research contribution, to you, or is it just an engineering effort?
Aman Kansal: There is definitely an engineering effort involved in setting up a much larger test bed. Our techniques are designed for a large scale system, and that is one reason why a hierarchy of controllers was used. It is difficult to claim whether new research problems will or will not be faced, if testing at scale, without actually doing it. The non-linearity of the UPS should be handled by the feedback mechanisms to a large extent.

We did some follow on work in investigating the network and power actuator latencies that might affect the scalability of our solution and largely speaking those are not hard challenges.

Christopher Stewart: Aman, thank you for taking time to discuss your paper with me. Congratulations on your research success and best wishes on your future work.

Disclaimer: Comments in this article reflect the personal views of the interviewed author only. These views may not reflect the views of other authors, affiliated institutions, or the publishing organization.

Nominations are Open!

Each issue of the Sustainable Computing Register features a Pick of the Month, a research publication or industry project that has significantly advanced the field of sustainable computing. The goal is to increase awareness within our community about high-impact, transformative research. Members can submit worthy papers and industry projects by emailing me. Submissions endorsed by 2 STC-SC officers will advance to public vote on Facebook. *By visiting our Facebook page, all members can vote for their favorite paper.* At the end of each month, the paper with the most votes will become a Pick of the Month (provided the authors agree to be interviewed).

Requirements for nominees:

- The paper must have been published in a peer-reviewed, research forum.
- The paper must be related to sustainable computing, e.g., energy efficiency, renewable-powered computing, smart grid, life cycle of ICT, smart buildings, etc.
- The paper must have been published in the last 2 years.
- Industry projects must have shown significant practical impact or intellectual contribution.

Energy-Efficient Hardware

By Fan Dongrui

Godson-D: Data Processing Unit for High-Volume Throughput Computing

Handheld terminals, like smart phones and tablets, are fashionable. Users worldwide enjoy the new, fancy applications supported on such terminals, especially applications integrated with backend cloud servers, e.g., Siri and micro blog. Compared with traditional scientific computing, these applications have totally different requirements for servers. For starters, these applications need servers that provide timely responses to a high-volume of independent user requests. More fundamentally, these applications do not execute many floating point instructions. Despite these differences, the processing units used in scientific computing centers are very similar to those used in data centers. The result is low energy efficiency and poor utilization of resources.

I recently had the opportunity to analyze some high-throughput applications running in cloud data centers. I would like to share a few of interesting observations from my study. First, L2 and L3 miss rate cache can be surprisingly high because of massive amount of independent requests conflicting with each other for the limited on-chip memory. Data access patterns were often irregular, making it hard to find temporal or spatial locality. Even though on-chip memory was poorly utilized, the on-chip memory still occupied die area and consumed a lot of energy. Another interesting observation was that, because the processor cores always wait for data from off-chip memory, it incurred low utilization of aggressive out-of-order components of processing cores, such as the reorder buffer and load store queue. Once again, these lightly used components stilled used a lot of space and power. Finally, it is tempting to think that memory bandwidth would bottleneck such applications. But my analysis found that the provided bandwidth is always more than enough for both on and off chip structures. This surprising result arose because there was only limited data sharing among processing cores and the complex relationship of data structures in each independent request. So more processing cores and hardware threads instead of large shared memory and aggressive scheduling logic, should be integrated into the chip to efficiently utilize the on-chip and off-chip bandwidth.

State Key Laboratory of Computer Architecture, Institute of Computing Technology, Chinese Academy of Sciences, is conducting research on designing high-volume throughput computing systems. The goal is to design a high-volume throughput processor -- Godson-D (D means DPU, that is, High-Volume Data Processing Unit). Through analyzing the special features of cloud services for high-volume throughput requests, and detecting the performance bottlenecks and the architectural limitations of commercial processors, Godson-D will exploit an energy-efficient and scalable micro-architecture for high-volume throughput computing.

European Research on Green ICT (continued from page 3)

- Reservoir (Resources and services virtualisation without barriers): Reservoir was the first FP7-ICT funded collaborative project dealing with Horizontal Federation for distributed cloud service centers and energy efficiency. Reservoir established a reference Cloud architecture pointing towards the role of the third party, the service providers, between the end-users and the infrastructure providers.
- Manthycore (Provide infrastructure resources and IP networks as a service): The Mantychore FP7 project has evolved from previous research projects MANTICORE and MANTICORE II. The main objective is to create a pilot study for best practices and guidelines to follow when building low carbon networks.
- TREND (Towards Real Energy-efficient Network Design): TREND integrates the activities of major European players in networking to quantitatively assess the energy demand of current and future telecom infrastructures, and to design energy-efficient, scalable and sustainable future networks.

STC Updates

By Giuliano Casale, Imperial College

Membership: 120

Report from Secretary/Treasurer (Giuliano Casale):

- Collected officers' activity reports and prepared monthly STC report.

Report from Conferences Chair (Diwakar Krishnamurthy):

- Continued to update list of venues based on officer inputs.
- Solicited collaboration from the general chairs of IEEE NOCS 2012, CCGRID 2012, and IEEE ISSST 2012.

Report from Academic Chair (Niklas Carlsson):

- Working with the industry chair on a potential short-feature that may help showcase academic and industry groups/people.
- Next month we will also work on a student award.

Report from Membership Chair (Anirban Mahanti):

- Our membership has doubled since March. We currently have 139 members.
- The officers have improved the membership joining/tracking process. A membership drive to further increase uptake is currently being evolved.

Report from Communications Chair (Abhishek Chandra):

- Continued to identify conferences, workshops and journals relevant to sustainable computing.
- Prepared a spreadsheet with information about upcoming call for papers and call for participation, for inclusion in the monthly newsletter and website.

Report from Policies and Procedures Chair (Stephen Dawson):

- Documenting currently implemented STC-SC processes and gathering requirements for future processes.

Report from Industry Chair (Canturk Isci):

- Finalized the drafts of four industry features. Coordinating with the project leads for potential revisions. Worked with Niklas on industry/academy group picks format and candidates.

Report from Information Officers (Danilo Ardagna, Guillaume Jourjon):

- Contributed material for newsletter and blogs.

Report from the Newsletter Editor (Christopher Stewart):

- Published the April newsletter
- Looking for new contributors to the newsletter

Upcoming Events

By Abhishek Chandra, University of Minnesota

The following venues are all requesting submissions on subtopics related to sustainable computing or IT for sustainability.

Conference, Workshop & Symposium Call For Papers

Short Name	Main Topic	Location	Dates	Abstracts Due	Papers Due	Notification
SustKDD	*Data Mining for Sustainability*	*Beijing, China*	*Aug. 12, 2012*		*May 23, 2012*	*Jun. 4, 2012*
Middleware 2012	Middleware Systems	Montreal, Canada	Dec. 3-7, 2012	May 18, 2012	May 25, 2012	Aug. 10, 2012
SOCC'12	Cloud Computing	San Jose, CA	Oct. 14-17, 2012	Jun. 8, 2012	Jun. 15, 2012	Sep. 3, 2012
HotPower'12	Power-Aware Computing and Systems	Hollywood, CA	Oct. 7, 2012		July 13, 2012	Aug. 24, 2012

Journal and Special Issue Call For Papers

Journal	Papers Due	Notification
Sustainable Computing	*(Open)*	

Conference, Workshop & Symposium Call for Participation

Short Name	Main Topic	Location	Dates
NOCS'12	*Networks-on-Chip*	*Lyngby, Denmark*	*May 9-11, 2012*
CCGrid 2012	*Cluster, Grid and Cloud Computing*	*Ottawa, Canada*	*May 13-16, 2012*
ISSST'12	*Sustainable Systems and Technology*	*Boston, MA*	*May 16-18, 2012*
IGCC'12	*Green Computing*	*San Jose, CA*	*Jun. 5-8, 2012*
Greenmetrics'12	*Sustainable computing*	*London, UK*	*Jun. 11, 2012*
IPDPS'12	Parallel and Distributed Systems	Shanghai, China	May 21-25, 2012
ICDCS'12	Distributed Computing Systems	Macau, China	Jun. 18-21, 2012
Cloud 2012	Cloud Computing	Honolulu, Hawaii	Jun. 24-29, 2012
e-Energy 2012	Future Energy Systems	Madrid, Spain	May 9-11, 2012
Sigmetrics'12	Measurement and Modeling	London, UK	Jun. 11-15, 2012
USENIX ATC'12	Computer Systems	Boston, MA	Jun. 13-15, 2012
GCC'12	Green Communications and Networks	Ottawa, Canada	Jun. 15, 2012
HPDC'12	High Perf. Distributed Computing	Delft, Netherlands	Jun. 18-22, 2012
e-Energy 2012	Future Energy Systems	Madrid, Spain	May 9-11, 2012
CCGrid 2012	Cluster, Grid and Cloud Computing	Ottawa, Canada	May 13-16, 2012
IPDPS'12	Parallel and Distributed Systems	Shanghai, China	May 21-25, 2012
IGCC'12	Green Computing	San Jose, CA	Jun. 5-8, 2012
Sigmetrics'12	Measurement and Modeling	London, UK	Jun. 11-15, 2012
USENIX ATC'12	Computer Systems	Boston, MA	Jun. 13-15, 2012
GCC'12	Green Communications and Networking	Ottawa, Canada	Jun. 15, 2012
HPDC'12	High Performance Distributed Computing	Delft, Netherlands	Jun. 18-22, 2012

Visit http://stc-sustainable-computing.ieee.net/venues for more information about these events.

To advertise a relevant venue, email Abhishek Chandra at chandra@cs.umn.edu.

To purchase a printed copy of the Register visit http://www.magcloud.com/