Measuring Energy Consumption of a Database Cluster

Volker Hudlet, Daniel Schall
AG DBIS, TU Kaiserslautern
67663 Kaiserslautern, Germany
{hudlet, schall}@cs.uni-kl.de

Abstract: Energy consumption of database servers is a growing concern for companies as it is a critical part of a data center's cost. To address the rising cost and the waste of energy, a new paradigm called *GreenIT* arose. Hardware and software developers are aiming at more energy-efficient systems. To improve the energy footprint of database servers, we developed a cluster of small-scale nodes, that can be dynamically powered dependent on the workload. This demo shows the measurement framework we set up to measure hardware components as well as an entire cluster of nodes. We'll exhibit the measurement devices for components and servers and show the system's behavior under varying workloads. Attendees will be able to adjust workloads and experience their impact on energy consumption.

1 Motivation

The growing energy consumption of data centers has become an area of research interest lately. Lots of effort is taken to improve the energy footprint of servers and to minimize overall energy consumption. Still, the amount of installed servers is constantly growing as more and more data is produced and has to be managed.

The focus on the servers' energy consumption is a rather young research field. Studies have shown that servers typically consume already more than 50 % of their peak energy when idle [THS10]. As the typical server load is between 20 to 50 % [BH07, Ran10], more energy-efficient approaches have to be developed to reduce the waste of energy of todays servers. This task addresses hardware developers as well as system designers and also the database community. Energy-proportional hardware (i.e., the energy consumed is proportional to the current load) would be highly desirable, unfortunately yet out of reach due to hardware limitations.

To avoid the limitations of a single large server in achieving energy proportionality, we are going to propose a new approach by employing commodity hardware to form an energy-proportional computing cluster. Small-scale server nodes (aka. *wimpy nodes* [AFK+09]) can be independently turned on and off, thus providing better scalability in terms of performance and power. By powering nodes based on the overall workload, the total energy consumption converges to an energy-proportional behavior. For now, however, standard server hardware components do not provide integrated probes to measure energy consumption of components or entire servers.

In this demo, we are going to introduce our idea of an energy-proportional database cluster and show how we measure power consumption. We will present an interactive front-end that controls the system's workload and illustrates how the system reacts to rising and falling loads.

2 Architecture Overview

The key ideas of the database cluster and the architecture of the measurement framework are both outlined subsequently.

Database Nodes The database cluster consists of 10 identical nodes, two of them have attached 4 hard disks, these are considered as *data nodes*, the other 8 nodes are referred to as *compute nodes*. The compute nodes are interconnected by a 1Gb/s ethernet switch (*front-end switch*). Each node has 2 GB of DRAM and one Intel Atom CPU, a rather lightweight processor with low thermal design power. The Linux-based operating system of the nodes is booted from an attached USB stick. In order to fully support dynamic node powering, we are developing our own database management system, called *WattDB*. The system is capable of reacting to changing load situations and able to adapt the cluster to the workload. By switching nodes on and off, the overall energy consumption is made proportional to the load, thus energy proportional. Each node is providing feedback about its resource utilization to a coordinator, which uses this information to re-balance the cluster by automatically powering up and down nodes.

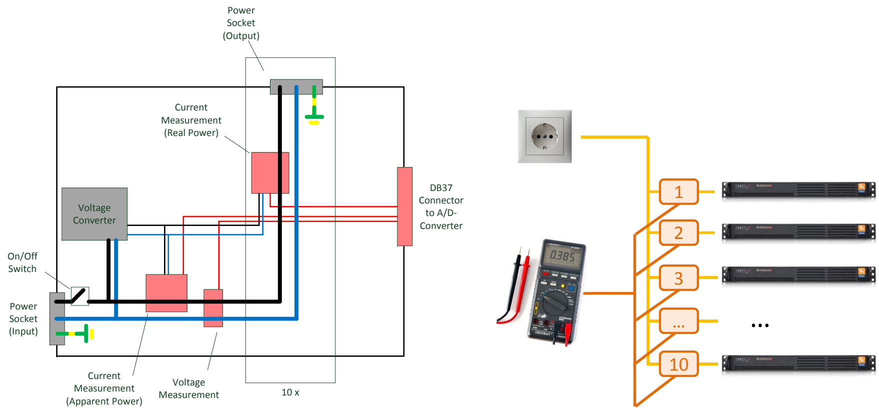
Energy Measurement We developed two devices for measuring energy consumption, one for measuring a single server down to the component level and a second device for measuring up to 10 servers in parallel.

(a) Photo (b) Schema

Figure 1: Measurement Device for a Single Node

We are able to measure the energy consumption of a single node using a custom measurement device. Figure 1 shows a picture of the front of the device and a schematic plan.

The device is connected to a standard ATX power supply and able to measure voltage and current of each power line separately. That way, we are able to track the energy consumption of the mainboard and two connected hard disks. The measurements are converted to digital signals and can be read from a connected PC. We can track the power consumption with 1% accuracy at a 10 ms resolution. To sum up all corresponding power lines, we developed an evaluation tool that is able to perform mathematical operations on the measurement, e.g., integrate multiple sources over time to capture total power consumption. Additionally, the software is able to protocol the data to file.



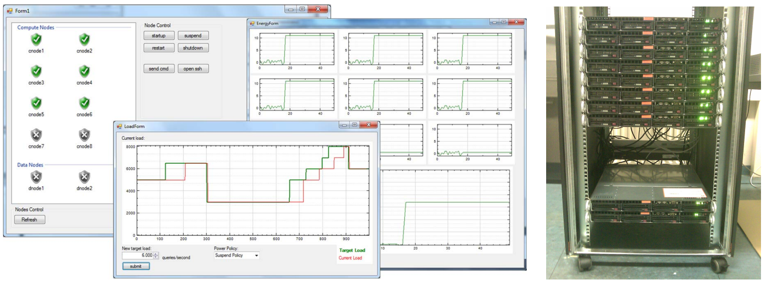
(a) Picture (b) Schema

Figure 2: Measurement Device for a Node Cluster

In order to measure the energy consumption of a whole cluster of nodes, a similar approach is taken. Measuring such a cluster requires the overall energy consumption of each node – detailed power consumption of each components is not the main focus in this scenario. Therefore, a new measurement device is needed, as depicted in Figure 2(b). As illustrated, this device is able to track the power consumption of 10 server nodes independently. To measure current, *current transducers* are used. They provide high accuracy (less than 1% deviation) and do not influence the lines under measurement. An integrated circuit is used to provide the power factor of the entire cluster. The new device can measure the total effective power as well as the apparent power. Besides the current of each server, the input voltage is needed for calculating precise power consumption. Therefore, a separated circuit is introduced to measure the voltage. All obtained values are transferred in parallel over a bus to an outlet. The values can be read out using a connected A/D converter and a standard PC. The reporting interface is identical to the single-server tool, therefore the same reporting software can be used.

3 Demo Setup

To demonstrate the capabilities of our measurement device(s), a simple workload driver is used to trigger workload on the system. Changing energy consumption based on the differ-



(a) Various GUI Forms reporting the Node States (b) Photo of the Cluster

Figure 3: Demo Setup

ent loads is reported immediately by the devices. The power consumption of each node is visualized on a GUI, additionally, the cluster's LEDs show the state of the nodes. The total power consumption is displayed in a separate graph. Figure 3(a) depicts a screenshot of the user interface for the cluster demo: The form on the left shows the various states (either *on*, *off* or *suspended*) of each of the nodes. On the right-hand side, the power consumption over time is plotted. At the front of Figure 3(a), a specific form makes the progress of the target and actual workloads visible. Both measurement devices are exhibited during the demo and can be seen in action.

The measurement device for a single server will have various storage disks attached, e.g. SSD and hard disk drives, and attendees can see the devices' power consumption while triggering different kinds of workloads. Additionally, the measurement device for a server cluster will also be displayed, attached to server nodes and showing a reporting GUI as well. The user can alter the workload of the cluster using the corresponding GUI. The power policy for the cluster can also be changed. This policy controls the way nodes are powered up and down, what sleep states the nodes should enter and how aggressive the power management is done. The effect of the user's actions is immediately visible – on the GUI as well as on the cluster itself (Figure 3(b)). Attendees can see how the system reacts to changing workloads and how the power consumption is affected.

References

[AFK+09] David G. Andersen, Jason Franklin, Michael Kaminsky, Amar Phanishayee, Lawrence Tan, and Vijay Vasudevan. FAWN: A Fast Array of Wimpy Nodes. In *SOSP*, pages 1–14, 2009.

[BH07] Luiz André Barroso and Urs Hölzle. The Case for Energy-Proportional Computing. *IEEE Computer*, 40(12):33–37, 2007.

[Ran10] Parthasarathy Ranganathan. Recipe for Efficiency: Principles of Power-Aware Computing. *Communications of the ACM*, 53(4):60–67, April 2010.

[THS10] Dimitris Tsirogiannis, Stavros Harizopoulos, and Mehul A. Shah. Analyzing the Energy Efficiency of a Database Server. In *SIGMOD Conference*, pages 231–242, 2010.