Energy monitoring and its impact on individual user privacy

Michael Niedermeier[1], Nasir Ali[1], Hermann de Meer[1], Helmut Hlavacs[2], Thomas Treutner[2], L. Lefèvre[3], J.-P. Gelas[3], and Iordanis Koutsopoulos[4]

[1]University of Passau Germany, [2]University of Vienna Austria,[3]INRIA France,[4]CERTH Greece

Introduction. Resource monitoring holds a very important position within the energy efficient computing paradigm. The underlying idea is to monitor the energy consumption behavior of appliances in different scenarios to develop consumption signatures. The signatures can help to deduce detailed consumption information which can be used to improve energy savings. There are many contemporary works being done in this direction where researchers investigate resource monitoring systems and their application. Additionally, efforts are being made to utilize collected samples of monitored information in useful ways [1]. Though the overall idea is very strong, there are certain research challenges which need to be overcome before turning this vision into reality. One of the most widely discussed among these is the privacy implications of such systems [3] and usability of collected information in a constructive manner.

Detection theory through hypothesis testing is a tool that can be used to advance state-of-the-art in privacy in energy monitoring. Compromising privacy can be thought of as deducing the profile of a user from observed data. In the context of process energy data monitoring, a profile corresponds to selecting one out of a finite set of N possible processes that are ran on a physical machine. Privacy concerns are raised when a malicious entity builds a set of empirical probability mass functions (p.m.f.), one for each process. Each p.m.f. captures the statistics of instantaneously consumed energy of the process. This task can be performed easily offline by taking multiple observations from each process that is running separately on every machine.

A privacy breach exists, if, during the time a process runs on a physical machine, the malicious entity takes observations of the energy consumption level. Various factors may inherently limit the amount of observations taken. The question for the entity that seeks to compomise privacy is to identify the process that is running with good accuracy. Virtualization comes into stage to the support of privacy preservation; By appropriately mixing two or more processes (and thus p.m.f's) on a virtualized machine, the privacy is protected, in the sense that the individual processes are made indistinguishable.

This is also one of the core research questions in EuroNF SJRP SPEC[1] where we want to address this interrelation between energy consumption monitoring and their impact on user privacy.

Our primary scenario revolves around the idea of deducing activity on a certain computer based on the monitored consumption patterns. Details of our scenario are being presented in the next section along with some results depicted in the conclusion section.

[1] http://www3.net.fim.uni-passau.de/SPEC/home.html

The experiments performed during the project covered two scenarios. Firstly, it was analyzed if the startup of certain applications on a computer produced unique energy signatures, making them identifiable by their energy consumption pattern alone. This would lead to a severe privacy threat, especially in office environments, where the workstations of the employees could be spied upon by using just information already present due to the usage of smart meters. Secondly, a data center scenario was investigated, where several virtual machines (VMs) are hosted on a server (hypervisor). As a consequence of energy efficiency or load balancing concerns due to, e.g., cloud computing, the exact combination of VMs running on a specific server will most likely change over time. We present experimental results for a side-channel attack in the context of virtualized data centers, using the energy consumption of a server as a side-channel to recognize the exact combination of VMs it currently hosts. This could be valuable information for an attacker trying to track down a specific VM.This threat. This aspect can be an important threat to take into account in the design of Green Clouds architectures [2].

Experimental Details. During the first scenario, the startup energy consumption of applications was measured using an energy measurement device (ZES LMG-500 device) directly attached to the computer. The current power consumption of each tested application was measured by the device in intervals of 100ms resulting in a fine-granular energy pattern. To guarantee the soundness of the measurements, the system was booted up freshly for each test to avoid problems due to different RAM or cache usage. Before an application was measured, the system was left idle to ensure that all background services were up and running to prevent interference with the energy measurements. After several measurements of each application, the patterns were compared to see if one sample of an application matches the other samples of the same application better than those of other programs using MSE as error metric.

In the second scenario, we measured the power consumption of a server hosting 15 combinations of four virtual machines using a high precision watt meter. The VMs included a `MySQL` database VM, an `Apache` with `PHP` webserver VM, an `FTP` server VM, and a VM doing I/O operations using `bonnie++`.

To recognize the exact combination, for each combination of VMs the measured samples were split into a training set, and test set representing samples measured sequentially by an attacker. For each training set, we estimated the probability density function by using kernel estimators. When sampling N values and knowing these density functions, an attacker would then compute the log-likelihood of the N samples for each known density and choose the density with the maximum likelihood. To get reliable results, the training/test sample split was randomized and repeated 1000 times for each test sample sizes, yielding according probabilities for exactly identifying the respective VM combination (Figure 1).

We can see that for comparably small training and test set sizes, a lot of exact combinations can be recognized correctly to a high degree using a log-likelihood approach. Also, if the range of sample values of a single VM is very small then it can be detected with high probability as standalone, but with low probability if running in combination with other VMs.

Conclusion. The results of both test runs differ significantly: While in the first scenario, it is currently not possible to identify applications by their startup energy patterns, it is possible using longer energy readings, as done in the seconds scenario.

The main identification problems with individual startup energy consumption of applications are that they are indistinguishable from the consumption that background services

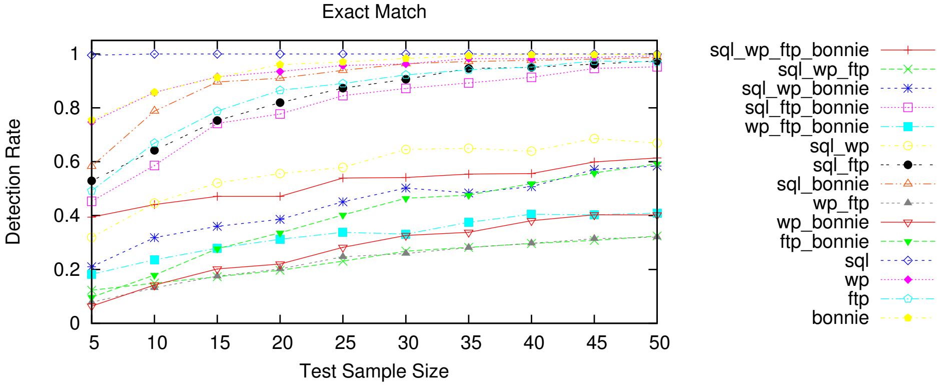

Fig. 1. Detection rates of VMs in relation to test sample size

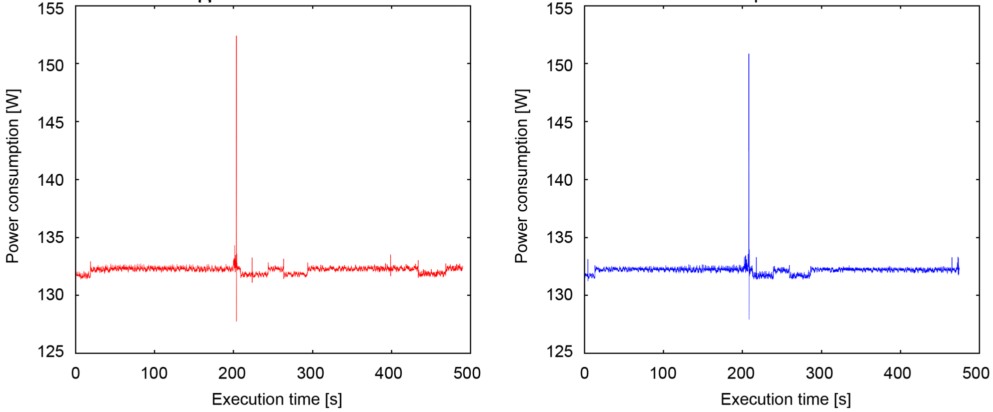

Fig. 2. Similar energy profiles of two different applications (left: disk analyzer, right: calculator)

induce. Therefore, the generation of patterns for certain programs is a time-consuming process, which is additionally complicated by the high similarity of the energy profiles of different applications (Figure 2).

As a result for the second scenario, we suggest that energy consumption data of servers must be protected carefully, as it is potentially valuable information for an attacker trying to track down a VM for further attack steps.

Acknowledgments. The research leading to the results presented in this paper has been funded by EuroNF SJRP.44 SPEC project.

References

1. Anderson, R., Fuloria, S.: On the security economics of electricity metering (2010)
2. Orgerie, A.C., Dias de Assuncao, M., Lefèvre, L.: Energy Aware Clouds, chap. "Grids, Clouds and Virtualization" - M. Cafaro and G. Aloisio editors, pp. 145–170. Springer Book (Oct 2010)
3. Quinn, E.L.: Smart metering and privacy: Existing laws and competing policies. Tech. rep., University Colorado Law School - CEES (May 2009)