Monitoring and Measurement System for Green Operation of Geographically Distributed ICT Services

Ana Carolina Riekstīns, Thomas Dandres†, Kim Khoa Nguyen*, Réjean Samson†, and Mohamed Cheriet*

∗Synchronedia - École de technologie supérieure, Université du Québec (Canada).
†CIRAIG - Polytechnique Montréal, Université de Montréal (Canada).

Abstract—Despite recent efforts and important results already achieved, the reduction of energy consumption and carbon emissions by Information and Communication Technologies is still far from the expected goals. As the annual growth in traffic is doubling every two years with more and more connections to the Internet, to be energy and carbon-aware it is paramount to implement a Monitoring and Measurement System which supports green strategies in a geographically distributed environment. Such an environment has some specific challenges that must be taken into account, such as the WAN connection, security and latency concerns. On the other hand, it also provides opportunities to reduce operational costs and emissions, improve reliability and resources management etc. This work proposes a framework which is capable of supporting green metrics and measurements, the MMSfG. The framework incorporates different models to get power and energy information and includes the method to calculate carbon equivalent emissions. This framework can be used as a support for a variety of applications which depend on energy and emissions metrics.

I. INTRODUCTION

Corporate investments in energy efficiency and reduction of greenhouse gases (GHG) emissions have increased in the Information and Communication Technologies (ICT) field. Telecommunication companies are among the ones committed to reduce emissions: Bell Canada – 50% off its 2003 GHG emissions by the end of 2020, Verizon – 50% off its carbon intensity by 2020, having 2009 as baseline [1], among others.

Despite the efforts, the amount of energy consumption and GHG emissions by ICT systems is still far from the ideal patterns. IEEE draws attention specifically to the energy required by networks: with more and more connections to the Internet, the annual growth in traffic is doubling every two years [2]. Energy consumption is not necessarily tied to carbon emissions since the energy source can be cleaner depending on where the ICT services are placed. However, for many companies, energy consumption is a key component in their GHG inventories. E.g., Verizon reports that more than 90% of its emissions are from electricity to power its networks [3].

To be energy and carbon-aware, it is paramount to have a Monitoring and Measurement System (MMS) able to support power/energy and carbon metrics, an “MMS for Green” (MMSfG). Besides the inherent challenges an MMS has to deal with [4], the MMSfG faces new ones, like accurately calculating power consumption of devices and the amount of carbon emitted for a certain volume of energy spent. Obtaining energy measurements is “challenging in terms of accuracy, timeliness, overhead, and so on and can be performed instantaneous or with a degree of aggregation, either directly on network equipment or through a model-based approach” [5].

Regarding carbon emissions, the computation of the local emission factor can be accomplished by calculating the weighted sum of the emission factors of the different sources used. This was the scenario for the development of the GreenStar Network [6], which takes the local carbon footprint of geographically distributed sites and dynamically recommends virtual machines (VM) allocation or relocation through the network aiming the lowest environmental impact. However, even when using renewables locally, the sites usually have the electricity grid as backup power. In this case, the carbon information, in general, comes from emission factors calculated as annual averages. Nonetheless, it is important to consider the variations in time (e.g. during the day, or in different seasons) and the energy matrix in the regions where the ICT equipment is located [7]. The importance of the varying emission factors was also acknowledged by [8]. The authors, however, do not give details on how to measure power/energy data from a network or the varying carbon emission factors.

In this context, this paper proposes a framework to support different green metrics and measurements, the MMSfG. The framework incorporates different models to get power and energy information and includes the method to calculate carbon emissions for different regions where the ICT equipment is located. It uses a temporally differentiated Life Cycle Assessment (LCA) model for power grid emissions, considering energy imports and exports from adjacent regions, based on the work of [9]. The remainder of this paper is organized as follows. Section II presents the previous work on monitoring and measurement, including “green” accounting. Section III presents the proposed framework, including the method to calculate emission factors varying in time and location. Section IV presents an experiment with the framework, showing the results achieved. Section V presents the concluding remarks and future work.
II. BACKGROUND

Bianzino et al. [10] and Wang et al. [11] provide evidence of a profusion of metrics related to energy and carbon assessment. There is no consensus over their usage [10]. In 2014, the Green Grid Consortium recommended a set of metrics for data centres: Data Centre energy Productivity (DCEP); Power Usage Effectiveness (PUE); Green Energy Coefficient (GEC), the amount of energy in a data centre from green sources; Energy Reuse Factor (ERF), the portion of energy that is exported for reuse outside of the data centre; and Carbon Usage Effectiveness (CUE), the total GHG emissions of a data centre divided by its ICT energy consumption [12]. These metrics, despite being recommended for data centres, could be applied to network sites or facilities.

Fiandrino et al. [13] proposed a framework of metrics related to energy efficiency and performance for cloud data centres, distinguishing communication systems from IT equipment. For energy efficiency, the authors proposed: the Communication Network Energy Efficiency (CNEE) in Watts/byte/second; the Network Power Usage Effectiveness (NPUE); and the Energy Proportionality Coefficient (EPC). Verizon and Akamai also report another type of metric, the Carbon Intensity, which is the amount of carbon emitted by the company divided by the total data transported over its networks (CO$_2$/byte) [14] [3]. All the aforementioned metrics have some common pieces: information about the network traffic, about power/energy, and about carbon equivalent emissions (CO$_2$).

A. Network Monitoring

According to [15], monitoring in networks has five subdivisions: (i) collection of raw data; (ii) representation of data into a particular format, independent of the management function; (iii) report, which aggregates data into a smaller set of management points; (iv) analysis, which extracts some high-level interpretations of the collected data; and (v) presentation, which presents data to the administrators in different formats.

The collection layer can be subdivided in active or passive, and network operators can use a combination of them for different purposes. In this sub-division, sampling techniques can be used to reduce overhead. The representation layer has the SNMP MIB as the de facto standard, but also comprises the well-known CIM and NETCONF, for instance. The reporting layer deals with reducing the traffic by suppressing redundant information or aggregating similar measurements, as well as by managing the polling frequency. The content distribution also takes privacy concerns into account. The analysis layer transforms the raw data into useful information. This layer can work with traffic analysis, traffic classification, estimation of demand, fault management etc. The presentation layer provides a visual representation of the information generated in the lower layers to the users, inferring which information is more important and presenting it in an intuitive way.

According to the Lin et al. [4], the requirements a Monitoring and Measurement System (MMS) should fulfill are: (i) integrated and cross-layer monitoring and measurement, considering a heterogeneous environment; (ii) data on-demand to support real-time decisions; (iii) support for custom metrics, since it is not feasible to previously define every metric users might need; (iv) others, such as storing historical data, near real-time processing, scalability, and elasticity [4]. Such requirements could also be applied for an MMS able to monitor power/energy and carbon data.

B. Power and Energy Information

In order to obtain power data from the infrastructure, one can get direct measurements using power meters, API requests, or management protocols like Simple Network Management Protocol (SNMP) when available. There are also model-based approaches, which monitor traffic and calculate power consumption based on power profiles, like in [16]. The instrumentation and measurement techniques are summarized in Figure 1.

Model-based approaches can also be used to measure the energy demanded by individual VMs, based on resource usage and performance monitoring counters [17]. The authors take into account the energy to powering the server, the network, the storage and the other utilities which support the operation (e.g. cooling), as well as the energy used in the migration and switch on/off events in each VM individually.

C. Emission Factors Calculation

The emission factors can be calculated for locally generated energy. For example, emission factors were used by the Green-Star Network project to decide about VM migration, following wind and sun availability in the geographically distributed data centres powered by renewables [6]. Each ICT service can have its own greenest factor which accounts for emissions from clean and dirty sources of energy.

Despite the trend of using local electricity generation from renewables, the sites still have the electricity grid at least as a backup source and electricity is an important contributor to emissions in the telecommunications industry. The electricity grid emission factors evaluation is complex, because both the electricity generation and the services demand vary significantly over time. The conventional LCA approach uses annual averages thus does not give an accurate model. Additionally, the data used to model electricity generation generally does not take into account the region where the electricity is consumed [7], an approach which is gaining importance within the current solutions on services migration across different data...
centre locations. Maurice et al. [7] developed a temporally differentiated LCA model to calculate carbon emissions related to electricity generation, described in Equation 1.

\[
\text{Carbon Footprint} (t) = \sum_{i} \text{electricity\_source}_i(\%) \times \text{electricity\_consumption}(\text{kWh}) \times \text{emission\_factor}(\text{kgCO}_2\text{e/\text{kWh}})
\] (1)

The equation is able to give the emission factor near real-time. However, it is required to calculate the emission factors for regions which do not disclose information about its own generation mix, as pointed out by [9], which will be discussed in next section.

III. PROPOSED REAL-TIME DIFFERENTIATED LCA FRAMEWORK

Regardless the network metrics to be calculated, they have some common pieces, as mentioned in Section II. Considering the requirements an MMS should fulfill, listed in Section II, such a system should be able to customize metrics according to the users’ needs. Therefore, the system should provide the ground information, while the user should define which metrics are good for his/her case. The system should also store historical data that can be used by the supported applications.

In this context, the MMSIG should not only fulfill the aforementioned requirements from [4], but also address the following: (1) be able to calculate the emission factors in real-time from the different local sources or, if using electricity from the grid, from the different regions the infrastructure is present; (2) support different power measurement techniques, model or measure-based; (3) consider different levels of aggregation, from networks down to devices and components, since the run-time decisions can be made based on the information this system provides (e.g. green routing, VM migration); (4) consider that the network is supporting different users and applications and that it may be required to report energy consumption and emission factors per user, protocol or application.

Figure 2 describes the MMSIG framework. The left part represents power and energy information gathering from the devices (including traffic information when required), while the right represents the emission factors retrieving. The calculation of GHG emissions requires both emission factors and energy information. That is why this part spans the two sides of the framework.

A. Calculation of Network Energy Consumption

Regarding the network, depending on the technology, the information can be gathered by device, component, protocol, or user. It is also important to have information about the location, which is required to calculate GHG emissions, and from indirect ICT energy consumption caused by the infrastructure which supports the equipment operation, such as cooling. In this case, the facility PUE can be used to estimate this value. The information can be represented and reported separately or aggregated by groups of users, protocol, path, sub-network or the whole network.

A simple example is obtaining information about devices and, after, putting them together to have information about the whole network. Another example is to get traffic information by protocol, such as TCP or UDP, and then calculate their flows emissions considering the geographical locations on their path. Although energy is consumed by the physical layer (equipment), each logical network layer can have emissions computed based on their traffic and characteristics. For this specific example, analytical models can also be used to estimate emissions.

The power can be obtained from direct measurements or by using model-based approaches, as described before in Figure 1. This can be accomplished with the support of Wattmeters or directly from the device’s built-in sensor. A noninvasive approach would be better for not incurring extra costs to the infrastructure [5]. When model-based approaches are used, the equipment power profile is required, as well as the equipment power state (e.g. fully operational, sleeping) and traffic information. Model-based approaches can also be used to estimate VMs energy consumption [17].

Power profiles can be fixed, almost independent from the load, or can be load-proportional, especially in modern equipment. Power profiles can be gathered from the devices if they are able to inform their own behaviour. Another alternative would be to have a power profile model repository as in [16], which contains power parameters either informed by the manufacturers, or measured beforehand using power meters.

Power state can be obtained directly from the device, using, for instance, the “Monitoring and Control MIB for Power and Energy”, which provides the device power state taking the Energy Management Framework (EMAN) into account. The EMAN framework enables the monitoring of different devices using SNMP. Another alternative is using the Green Abstraction Layer (GAL), an “interface between data and control planes for exchanging data regarding the power status of a device”, intended for SDN environments [18].

Traffic information is easily obtained with, among others, SNMP, as snmpget, or OpenFlow instructions, as ovs-ofctl -O Openflow13 dump-ports switch1. Bandwidth consumed by a switch is calculated as the sum of the received bytes counter of all ports. With the power information available, it is possible to calculate energy and standardize the output of these blocks as power or energy units. The values can be gathered for devices and their components, or aggregated to account for the whole network. One important issue at this point is the infrastructure heterogeneity: the framework must support different data gathering tools and protocols.

When it comes to getting information from equipment during runtime, the overhead is an important issue: the larger the probe rate, the higher the overhead is, incurring in low performance and high processing times. A simplistic approach which just lowers the probe rate can lower the system accuracy, depending on how frequently the traffic varies [16]. And event-triggered polling needs to be carefully defined to ensure all important events are being captured.

Besides, the collection and treatment of large amounts of
data impose various challenges, mainly with the increase in link sizes and traffic in the recent years. Even if the overhead on processing and storage is not a limiting factor, identifying the important data among huge amounts of raw data is not an easy task. Sampling is one way of reducing the overhead, and different works try to improve the accuracy of measurements, by dynamically adjusting the sampling ratios or interpolating missing values [15]. Some data and control plane abstractions specifically designed for measurement processes can also help dealing with data collection and analysis [19].

B. Calculation of Real-Time Emission Factors

Real-time emission factors can represent a significant difference compared to the traditional approach of using the annual average. For instance, Ontario (Canada) annual average is 0.10 kgCO₂e/kWh [9], and in just one summer day of 2016, the real-time emission factor varied from 0.05 to 0.15 kgCO₂e/kWh. In this province specifically, [7] also reported different profiles in the different seasons, which can also represent a significant difference in the GHG emissions when using near real-time or average factors.

When a site is partially locally powered, it is necessary to get the share of each energy source in the total consumption, as well as the emission factor of each source in CO₂e/Wh. When it comes to the electricity grid, the emission factors information for each region can be calculated from the local demand, imports and exports, as well as from the locally generated energy grid matrix. In general, such data can be obtained from the Electricity System Operators websites. For instance, Ontario’s Independent Electricity System Operator publishes data on supply, demand, and flows [20].

Unfortunately, there are regions which do not disclose this information. In this case, it is possible to combine historical information about the demand and the grid mix with the imports and exports from neighbouring regions, divided by the type of energy source. For instance, regarding the example illustrated in Figure 3. Region A which does not disclose energy generation and demand information is adjacent to the region B, C, and D, and the region C is adjacent to regions D and E. The regions B, C, D, and E disclose real time information about demand, generation, and imports/exports to other regions. It is important to note that the calculation depends on the information available and has some degree of uncertainty, which can vary depending on the types of energy sources used. To estimate real-time emission factor of A, it is necessary to:

- Obtain historical information about electricity demand in region A. If possible, classify the demand on an hourly and monthly basis, to capture the behaviour of electricity demand during the day and in different seasons of a year;
- Obtain historical information about energy generation of the grid mix in A;
- Gather flow information from B, C, and D to A and, ideally, indirectly, from D and E, connected to C. The imports can alter the region A local emission factor (e.g. more energy generated from coal in C may increase the emission factor in A, or more hydro energy generated in B may make the emission factor in A cleaner);
- Calculate the near real-time grid mix in region A.

The calculation of energy imports may consider, for simplicity, only one “level” of imports, assuming that the imports are locally consumed and not exported to another place. We refer the reader to [9] for further details on the method.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

To illustrate the information being gathered and the emission factors being calculated, we used the topology depicted in the left side of Figure 5. It represents four switches of the same institution placed in four different provinces in Canada: Quebec, Ontario, Manitoba, and New Brunswick, represented in the map in Figure 4.

Regarding network information, including power data, we started with a topology with location information. In our experimental environment, direct measurements are not available, so we used a model-based approach – which demands a power profile, the power states, and traffic information.
For simplicity, a simple linear power profile obtained from a model repository was used to get power information given the traffic load in each node. All nodes used the same power profile to allow comparing the different emission factors of the provinces. The switches are OpenFlow enabled, and, therefore, the traffic information was gathered dumping the ports  rx_bytes counter with OpenFlow 1.3. For simplicity, PUE is considered 1.2 for all nodes, a reported value for stateports.

Among the studied locations, Quebec does not disclose real-time information. Based on information from the nearby provinces and historical data, it is possible to estimate the emission factor in this province. Figure 5 depicts the different provinces emission factors in the same graph for comparison, in the right bottom\(^1\). These values were obtained during two weeks of the 2016 summer and are aligned with the annual averages reported by each province. As shown, Quebec has the smallest emission factor (as the province is mainly powered by hydroelectricity), therefore, placing a service in this province will reduce the carbon emission compared to other provinces. With the power and location information from the nodes, as well as the emission factors in different regions, it is possible to calculate the nodes and the network emissions almost in real time, as illustrated in the right upper side of Figure 5.

In our experiments, the system took, on average, eight seconds to get information from all provinces. This value can obviously vary according to the connection and the computer being used to run the experiments. This interval is not a problem since the system is refreshing the values each five minutes—they do not vary much within smaller intervals.

The framework proposed herein can support different kinds of studies and calculations. It can be used, for instance, to support the framework proposed in [14], in which average emission factors were used. Their proposal allows the provisioning of virtual data centres in a geographically distributed infrastructure taking operational costs and green SLAs violation penalties into account. By green SLAs, they mean a limit on carbon emissions generated on behalf of the client. This work can also be used to extend [8] by supporting information gathering and by providing near real-time emission factors for different locations. These values can then be used in conjunction with the Virtual Carbon Tax to optimize the profit of the system, which also takes revenues, electricity prices, PUE, operational costs and taxes into consideration.

The near real-time emission factors can also be used to estimate emissions of applications and protocols in different network layers. Energy consumed on the physical layer can be combined with the data transmission characteristics of other layers to calculate energy consumption and then be used to estimate the emissions for the different applications and protocols (e.g., estimate the emissions of TCP or UDP flows).

Using the temporally differentiated LCA approach is more accurate than the traditional LCA approach of using annual averages, as exemplified in Section III-B. As acknowledged by [9], imports and exports can also represent differences in the regional GHG emissions. For instance, in a moment in the afternoon of June 21st, 2016, the local emission factor in Ontario was 0.096 kgCO\(_2\)/kWh. However, as the province was importing from Quebec, which is mainly hydro-powered, the emission factor considering the imports went down to 0.091 kgCO\(_2\)/kWh, a difference of more than 5%, which can also represent a significant difference when considering high amounts of energy consumed.

V. CONCLUSION AND FUTURE WORK DIRECTIONS

This work proposed MMSfG, a framework to support customizable metrics, providing ground information on power/energy, traffic, and carbon equivalent emissions. Our approach includes the method from [9] to calculate the emission factors in different regions using a temporally differentiated LCA model, and considering energy flows between the different regions.

The framework can be used to support different tools which rely on power/energy and carbon information. However, the application of the framework depends heavily on information availability, both for the equipment, and for the regions in which the nodes are placed. The same applies for the information gathering for the different levels of abstraction mentioned (users, applications, protocols). Besides, despite including all life cycle phases of electricity generation, the current version of the framework considers only ICT equipment usage phase emissions.

In future, the framework can be expanded to consider other phases of the life cycle. The framework can also be expanded to inform the amount or the share of renewables being used, a value that can be established in green SLAs [14]. Another envisaged extension is to comprise near real-time electricity price as proposed by [21] or [22], who used the different electricity prices, time zones and traffic information to decide about VM migration. Information about availability and reliability of the facilities can also be incorporated, aiming at giving

\(^1\)Specifically for Manitoba, the real-time information available considers the whole region network (Midcontinent ISO). Therefore, despite the local grid mix being mainly based on hydroelectricity, the whole network mostly uses fossil fuels, and, therefore, the region emission factor is high. For New Brunswick, the local demand and flows are available online, but the energy mix used was an average gathered from recent reports. Nova Scotia and Prince Edward were not used in this experiment.
full support to prediction and decisions about migration. All this information is important to consider the trade-offs between reducing emissions, saving energy, cutting costs, and ensuring performance.

ACKNOWLEDGMENT

The authors thank NSERC and Ciena for funding the project CRDPJ 461084. This research also receives support from the Canada Research Chair, Tier 1, hold by Mohamed Cheriet.

REFERENCES


