Low Carbon Virtual Private Clouds

Fereydoun Farrahi Moghaddam, Mohamed Cheriet, Kim Khoa Nguyen
Synchromedia Laboratory
Ecole de technologie superieure, Montreal
Email: ffarrahi@synchromedia.ca, mohamed.cheriet@etsmtl.ca, kimkhoa@gmail.com

***Abstract*—Data center energy efficiency and carbon footprint reduction have attracted a great deal of attention across the world for some years now, and recently more than ever. Live Virtual Machine (VM) migration is a prominent solution for achieving server consolidation in Local Area Network (LAN) environments. With the introduction of live Wide Area Network (WAN) VM migration, however, the challenge of energy efficiency extends from a single data center to a network of data centers. In this paper, intelligent live migration of VMs within a WAN is used as a reallocation tool to minimize the overall carbon footprint of the network. We provide a formulation to calculate carbon footprint and energy consumption for the whole network and its components, which will be helpful for customers of a provider of cleaner energy cloud services. Simulation results show that using the proposed Genetic Algorithm (GA)-based method for live VM migration can significantly reduce the carbon footprint of a cloud network compared to the consolidation of individual data center servers. In addition, the WAN data center consolidation results show that an optimum solution for carbon reduction is not necessarily optimal for energy consumption, and vice versa. Also, the simulation platform was tested under heavy and light VM loads, the results showing the levels of improvement in carbon reduction under different loads.**

I. Introduction

Cloud computing solutions enable small businesses to rent virtual servers as a service, instead of buying and maintaining actual servers. In this way, these businesses can focus on their goals and products rather than dealing with server maintenance, energy consumption, bandwidth, storage, etc. With increasing concerns about global warming and the increasing role of Greenhouse Gases (GhG) emissions expected in the near future, not only is the cost of services important for service providers and their customers, but the total carbon footprint of a service is of great concern to companies and governments [1]. So, it is important for cloud service providers to be able to provide their customers and their governments with measurable proof of the carbon footprint of their services, while minimizing its size in order to reduce the ultimate cost of the services. Although the goal of a cloud service provider is to reduce that cost, or arrive at a balance of cost and carbon footprint, this paper focuses only on the measurement and minimization of the total carbon footprint of services.

With the introduction of Local Area Network (LAN)-based live Virtual Machine (VM) migration [2], cloud administrators were able to move a VM from one hardware set-up to another for maintenance or energy efficiency reasons without any violation of the Service Level Agreement (SLA) with their users. In fact, many energy efficiency solutions are based on Live VM migration [3][4][5]. One of the main ideas behind these solutions is to consolidate VMs as much as possible, and reduce the amount of in-use hardware in order to save energy and reduce the carbon footprint. Because of the latency of Wide Area Network (WAN) connections and other remote access difficulties, such as the cost of remote connections and remote storage, these energy efficiency solutions were designed and tested on LAN environments. However, recent research on seamless WAN VM migration [6][7][8] proves the feasibility of Virtual Private Clouds (VPCs) [9][10]. A VPC is a uniform cloud based on a number of geographically distributed data centers which are connected through the Internet or private WAN connections.

This paper has two objectives. The first is to measure the carbon footprint of a VPC, and the second is to minimize the carbon footprint of that VPC. Considering larger WAN VM migration carbon footprints and different levels of power consumption of VMs at different data centers, these objectives include proving the feasibility of better carbon footprint reduction in VPCs compared to a number of LAN-based clouds providing the same services, as well as providing percentages of improvement in carbon footprint reduction. In this paper, a VPC is a network of data centers in different domains connected to one another through private WAN connections or via the Internet, in which a uniform cloud identity hosted and where each data center is powered by a different energy source and is situated at a different geographical location. We refer to LAN-based clouds as completely isolated data centers that are limited by their geographical location and their energy source. VPCs and LAN-based clouds provide the same services. However, in a VPC, because of the flat structure of the cloud, VMs can migrate between data centers, while in a LAN-based cloud, they can only migrate within a single data center.

In this work, the VM power measurement methods currently used [11][12] are extended with other methods in order to calculate the overall carbon footprint of the cloud. Also, existing server consolidation solutions [13] are extended with respect to carbon footprint minimization, in order to consolidate VPC data centers by moving VMs, as much as is possible, to data centers with a lower carbon footprint.

We first review the state of the art in cloud energy efficiency and VPCs. Then, we present our solution for a Low Carbon VPC (LCVPC), and build a carbon footprint model of such a cloud. Next, we examine our model on a simulation platform network and provide the results. Finally, we present our conclusions and outline future prospects.

978-0-7695-4460-1/11 $26.00 © 2011 IEEE
DOI 10.1109/CLOUD.2011.36

II. Previous work

A. Cloud energy efficiency

Two factors are limiting the development of computing systems: the cost of energy and the size of the carbon footprint [3]. While information and communications technology (ICT) usage is growing rapidly, the increasing cost of energy and the necessity to reduce GHG emissions are forcing movement towards energy efficiency in the ICT sector [14], especially in data centers. Energy efficiency there can be achieved using several methods [15]: dynamic CPU speed, energy-aware job scheduling, server consolidation [13].

All energy efficiency solutions require a measurement method to calculate the energy consumed and the GHGs emitted. Kansal *et. al.* introduced a model for VM power metering in [11]. In equation 1, the total energy consumption of a server is a combination of CPU use, memory access, and disk access.

$$E_{sys} = \alpha_{cpu}\mu_{cpu} + \alpha_{mem}\mu_{mem} + \alpha_{io}\mu_{disk} + \gamma \quad (1)$$

where E_{sys} is the energy consumption of a server; γ is the energy consumption of the server under 0% CPU, memory, and disk usage; α is the additional energy consumption of the server under 100% CPU, memory, and disk usage; and μ is the actual percentage of CPU, memory, and disk usage. The energy consumption of network cards and other elements is considered constant in γ, and it is assumed that γ does not vary greatly. We extend this formulation for our carbon footprint calculations in the modeling section, section III-A.

Among all the energy saving strategies in use in data centers, we chose dynamic VM consolidation as our tool in VPC carbon footprint optimization. Two types of controller for dynamic VM consolidation are compared by Gmach in [4]. He shows that running the trace-based workload and reactive controllers simultaneously gives a better result than running either of these controllers individually. Although combining different energy efficiency methods may result in greater carbon footprint reduction, the main objective of this paper will be to focus on the feasibility of global carbon reduction in VPCs by rearranging the services on a WAN.

In dynamic VM consolidation, VM migration is one of the elements that play a role in calculating the carbon footprint of data centers. The goal of this strategy is to move VMs, as much as is possible, from low-use servers and then turn those servers off to save energy. The cost of VM migration is considered as a contribution to the total cost function of the cloud carbon footprint [5]:

$$cost = C(Migration) + C(\#PM) + C(Utilization) \quad (2)$$

where $C(Migration)$ is the cost of VM migrations; $C(\#PM)$ is the cost of physical machine energy consumption; and $C(Utilization)$ is the cost of server use, which shows how busy the servers are.

B. Virtual Private Clouds

A Virtual Private Cloud (VPC) is a cloud identity consisting of a network of data centers connected to one another in a WAN. Recently, valuable work has been carried out on VPCs to overcome WAN-based problems, such as bandwidth limitations and high network latency.

Wood *et. al.* have used the concept of the VPC in [7] to build CloudNet, which is a prototype cloud computing platform that connects different data centers of an enterprise across the globe. CloudNet supports the WAN migration of VMs with minimal downtime over WAN connections, even with the problems of low bandwidth and high latency. They showed how CloudNet can optimize the amount of data transferred during the migration process and reduce application downtime. In contrast, other migration techniques send empty memory pages or disk blocks. In [8], we show how a Probability Density Function (PDF)-based selection of memory pages using their modification rate can minimize migration downtime. We also show that choosing to send memory pages with a low frequency of change first will prevent multiple copying of memory pages from source to destination, and works better than other methods of memory page selection. In addition, we show that a VM with a busy memory has more downtime, because it has a larger number of dirty pages. (Dirty pages are memory pages that have already been copied to the destination, but their content has been changed during the migration process. These pages need to be recopied to the destination.) We demonstrate that PDF-based selection will minimize the number of dirty pages in total and lower VM downtime.

In another work, Merwe *et. al.* used WAN VM migration to design a ubiquitous cloud computing infrastructure [6]. They studied two use case scenarios for multiple enterprise sites at geographically dispersed locations: cloudbursting, and follow-the-sun. In the follow-the-sun use case scenario, resources are relocated seamlessly to the place where they are needed most, based on the time zone. So, properly planned WAN live migration [7] will guarantee seamless migration of resources to their new locations.

In this paper, a similar concept as in [7] and [6] is used to form a flat virtual private cloud enabling the seamless movement of VMs over WAN connections from one data center to another. Unlike the follow-the-sun use case scenario in [6], where VMs follow the time of day for migration to the node nearest the active employees, we focus here on the cost function, which is set to carbon footprint reduction. This means that, in order to reduce the networks carbon footprint, the seamless migration of VMs over a WAN connection will be designed to optimize a global cost function on the whole network.

III. Our Clean Energy Efficiency Model in a VPC

Energy efficiency in VPCs is, in many ways, very similar to energy efficiency in LAN-based clouds, with a few additional features. In the latter, using live VM migration as a server consolidation tool results in lower power consumption and a reduced carbon footprint. This means that carbon footprint reduction is an immediate result of power consumption reduction. In VPCs, live WAN VM migration is used for server consolidation to reduce the power consumption and the carbon footprint as well; however, it can also be used to target power consumption and carbon footprint separately. In a WAN network with a mixture of nodes powered by both green and non-green energy sources, for example, moving a VM from a non-green powered node to a green powered one could result in a smaller carbon footprint. However, power consumption may or may not decrease, and may even increase.

WAN-based VM migration in clouds is also available for live movement of a VM from one data center to another. In this paper, a VPC manager is introduced to dynamically reallocate the position of VMs in order to reduce the overall carbon footprint of the cloud. The main idea is to create an LCVPC made up of various data centers, each with a different energy profile. The cloud manager optimizes the location of VMs in the cloud based on the availability of resources and the carbon footprint of each data center.

The data centers are situated at various geographical locations, and they have different energy sources. If a data center is powered by a renewable energy source, its carbon footprint will be small, or even zero, compared to a data center powered by non-clean energy sources. The idea is to move some VMs from non-clean powered data centers to cleaner or totally clean powered data centers, if they are available.

A total cost function will be introduced for the whole cloud, including all the data centers, based on its carbon footprint. This cost function will be optimized with heuristic algorithms to provide the best solution for the location of the VMs for a given period of time.

A. LCVPC Modeling

One of goals of this paper is to provide a formulation for measuring the carbon footprint of a VPC. There is a direct relation between power consumption and carbon footprint if the energy source is not clean. If the energy source is completely clean, the carbon footprint is zero, regardless of the amount of power used. The relation between power consumption and carbon footprint at data center d is formulated in Equation 3.

$$C_{p_d}(t) = \rho_d(t) P_d(t) \tag{3}$$

$C_{p_d}(t)$ represents the amount of carbon emitted from data center d in time t, $\rho_d(t)$ is the power-to-carbon conversion rate for data center d in time t; and $P_d(t)$ is the power consumption of data center d in time t. Every data center could be powered by different energy sources, and every energy source has its own carbon footprint [16]. As shown in Equation 4, $\rho_d(t)$ is completely dependent on the type of energy source used at each data center, and this could vary over time for those data centers with multiple energy sources, especially if the sources change over time.

$$\rho_d(t) = \frac{\sum_{source} \rho_{d_{source}}(t) P_{d_{source}}(t)}{\sum_{source} P_{d_{source}}(t)} \tag{4}$$

This equation can be rewritten as Equation 5, where $g_d(t) \in [0,1]$ represents the cleanness of all the energy sources at the data center combined, compared to the least clean energy source, where $g = 1$ represents a totally clean energy source and $g = 0$ represents a totally non-clean energy source; ρ_{max} represents the carbon-to-power conversion rate for the least clean energy source. Electricity produced from burning coal has the largest carbon footprint, according to [16].

$$\rho_d(t) = (1 - g_d(t))\rho_{max} \tag{5}$$

As formulated in Equation 6, calculation of $C_{p_d}(t)$ over all the data centers in the cloud will provide the carbon footprint of the cloud from power consumption.

$$C_p(t) = \sum_d C_{p_d}(t) = \rho_{max} \sum_d (1 - g_d(t)) P_d(t) \tag{6}$$

In a data center, $P_d(t)$, comprises the power consumed for cooling, the power processed by the Power Distribution Unit (PDU), and the power consumed by the servers:

$$P_d(t) = P_{c_d}(t) + P_{P_d}(t) + P_{s_d}(t) \tag{7}$$

Using Equation 1, $P_{s_d}(t)$ could express a summation of the power consumption of the servers:

$$\begin{aligned} P_d(t) = {} & P_{c_d}(t) + P_{P_d}(t) \\ & + \textstyle\sum_{s \in d} (\alpha_{cpu_s} \mu_{cpu_s} \\ & \quad + \alpha_{mem_s} \mu_{mem_s} + \alpha_{io_s} \mu_{disk_s} + \gamma_s) \end{aligned} \tag{8}$$

Combining Equations 6 and 8 will provide the carbon cost function for a VPC:

$$\begin{aligned} C_p(t) = {} & \rho_{max} \textstyle\sum_d (1 - g_d(t))(P_{c_d}(t) + P_{P_d}(t) \\ & + \textstyle\sum_{s \in d} (\alpha_{cpu_s} \mu_{cpu_s} \\ & \quad + \alpha_{mem_s} \mu_{mem_s} + \alpha_{io_s} \mu_{disk_s} + \gamma_s)) \end{aligned} \tag{9}$$

As mentioned in Equation 2,in order to calculate the total carbon footprint of the cloud, a migration carbon footprint should be added to the above formulation. This migration cost is roughly calculated from the additional use of source and destination servers and the in-between network routers from Equation 1 during the migration period. This cost is more for WAN VM migration, and is based on the use of more routers.

A final cost function could be produced by combining all the carbon footprint formulas, as follows:

$$\begin{aligned} C(\Delta_t) = {} & C_m(\Delta_t) + C_{DC_{on/off}}(\Delta_t) \\ & + \rho_{max} \textstyle\sum_d O_d (1 - g_d(t))(P_{c_d}(t) \\ & + P_{P_d}(t) \\ & + \textstyle\sum_{s \in d} O_s (\alpha_{cpu_s} \mu_{cpu_s} \\ & + \alpha_{mem_s} \mu_{mem_s} \\ & + \alpha_{io_s} \mu_{disk_s} + \gamma_s))\Delta_t \end{aligned} \tag{10}$$

Where O_d or O_s is a binary variable, and is equal to 1 when the data center or server is functional, and is equal to 0 if the data center or server is shut down. Δ_t is a period of

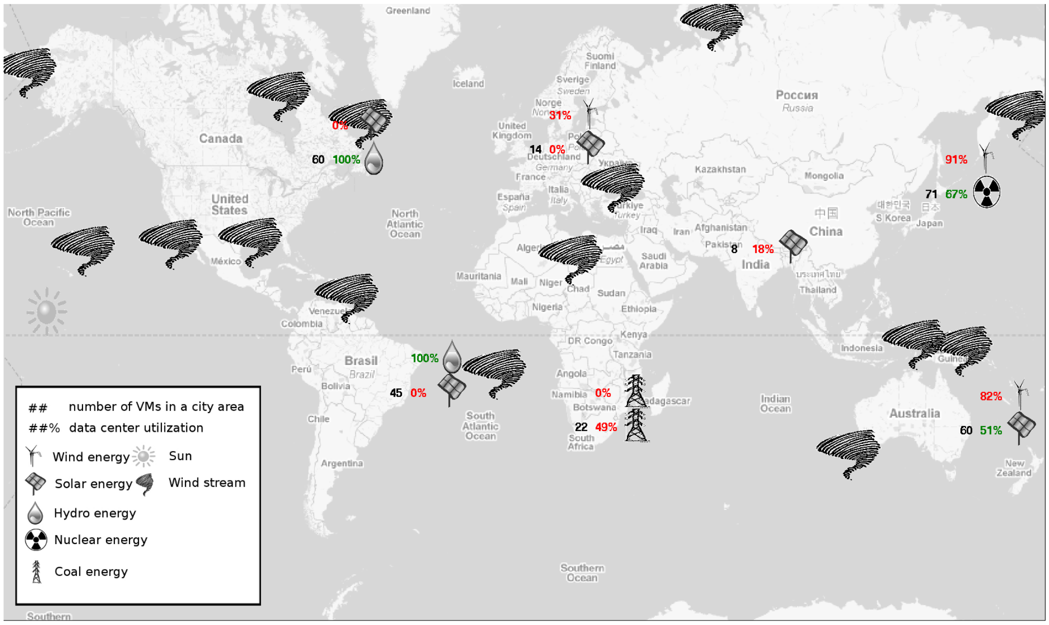

Fig. 1. Simulation platform map at hour 13. At each data center, the source(s) of energy and power use percentage(s) are provided separately.

time where power measurements are constant or with small variations. If there is no VM running at a data center or on a server, the data center or server could be shut down, in order to eliminate power consumption for cooling and overhead. There is nevertheless a modest carbon footprint generated to shut down a data center and to turn it back on, which is considered in $C_{DC_{on/off}}$. This cost function will be used for making decisions on the reallocation of VMs in our LCVPC model.

IV. Simple Use Case for the LCVPC Model

As a use case for testing the proposed cloud manager, a simulation platform is created connecting a number of data centers to form a VPC. These data centers are powered by various energy sources, both clean and non-clean. The network of data centers is tested under different test scenarios.

In the first test scenario, WAN-based data center consolidation is compared with LAN-based server consolidation. To accomplish this, a set of VMs is assigned to every data center randomly, and the carbon footprint of the whole network is calculated for a period of 24 hours. Then, a manager consolidates the VMs at each data center, in order to reduce its carbon footprint, and a network carbon footprint is calculated (LAN-based case). In the next step, the manager attempts to consolidate the VMs in the VPC. This means that the WAN migration of VMs is allowed in this step, and the carbon footprint of the network is then recalculated (WAN-based case). Finally, this network carbon footprint is compared with the carbon footprints calculated in the previous steps.

In the second test scenario, the network is tested under heavy VM load and under light VM load. In this scenario, the effect of network use percentage on carbon footprint optimization is studied, and the results are compared with the results of the first test scenario.

In the third test scenario, network carbon footprint reduction is compared with network energy consumption reduction. In this scenario, the network is first optimized for carbon footprint measurement, and energy consumption and the carbon footprint are measured. The network is then be optimized for energy consumption, and carbon footprint and energy consumption are measured and the results compared with the previous results.

To calculate the amount of energy, we modify Equation 10 and convert the carbon footprint to energy by removing the g and ρ_{max} factors and dividing the carbon parameters (C) by ρ_{max}, as follows:

$$\begin{aligned} &E(\Delta_t) = C_m(\Delta_t)/\rho_{max} + C_{DC_{on/off}}(\Delta_t)/\rho_{max} \\ &+\textstyle\sum_d O_d(P_{c_d}(t) \\ &+P_{P_d}(t) \\ &+\textstyle\sum_{s\in d} O_s(\alpha_{cpu_s}\mu_{cpu_s} \\ &+\alpha_{mem_s}\mu_{mem_s} \\ &+\alpha_{io_s}\mu_{disk_s} + \gamma_s))\Delta_t \end{aligned} \tag{11}$$

A. Simulation platform

To test different test scenarios, a platform is created to simulate a VPC and weather conditions. Seven different cities around the world have been chosen to host the WAN-based cloud data centers. Each data center is powered by a different energy source including clean and non-clean sources. Each has its own greenness factor (g), which is variable according to other environmental parameters. For example, solar power plants and wind power plants are dependent on the position of the sun and the existence of a wind stream near the plant respectively. The map of the simulation platform is shown in Figure 1. Some of the simulation platform data are real, such as geographical coordinates and sun position, and some data are randomly generated, such as wind stream movements. Each data center includes a number of servers, and each server includes a number of VMs which are assigned randomly.

The network consists of 13 data centers, for which the parameters are chosen randomly. The parameters for a typical data center are provided in Table I.

TABLE I
PARAMETERS OF A TYPICAL DATA CENTER IN A VPC

Parameter	Value
ρ_{max}	0.9 kg per kWh [16]
P_{c_d}	1 kW
P_{P_d}	0.3 kW
C_m	varied[1]
$C_{DC_{on/off}}$	0.275 kWh[2]
α_{cpu}	0.3 kW
α_{mem}	0.2 kW
α_{io}	0.3 kW
γ_s	0.3 kW

[1] 100% utilization of source and destination servers and intermediate network during migration period.
[2] 100% utilization of data center during shutting down period.

The carbon footprint of a data center network, as measured in a simulation platform with a variety of energy sources, is shown in Figure 2 for a period of 72 hours. In the figure legend, [no-opt] means no optimization has been performed, [wind] means there are wind streams in the simulation, and [no-wind] means there are no wind streams in the simulation.

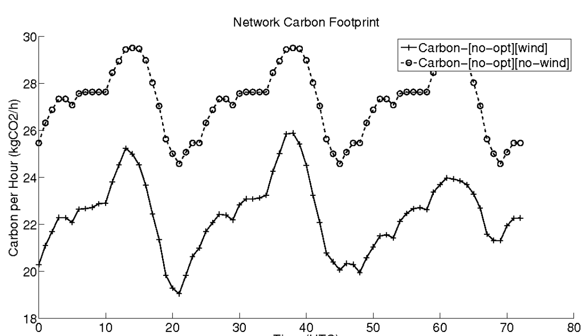

Fig. 2. Network carbon footprint

As shown in Figure 2, the carbon footprint graph is periodic when wind streams are not considered (the effect of clouds on solar power sources is not simulated in this work). When wind streams are considered, the network carbon footprint decreases and the graph is no longer periodic. The maximum carbon footprint size is reached at hour 13 (Figure 1), when the sun is in the middle of the Pacific Ocean, from where it can no longer power any of the solar sites, and there is no major wind stream near any of the wind sites.

B. Optimization

Whenever the simulation platform parameters change, the energy and carbon footprint of the network change. We are seeking to achieve the optimum network situation by reallocating some VMs over a LAN or a WAN, which is an optimization problem with the defined cost function given in Equation 10 for carbon reduction and Equation 11 for energy reduction.

Now we have two optimization problems, aimed at minimizing energy consumption and GHG emissions in the network respectively. Given N VMs running on d servers at time t, the problem of GHG optimization can be formulated as follows:

$$\begin{aligned}&Minimize\,C = C_m + C_{DC_{on/off}}\\&+\rho_{max}\sum_d O_d(1-g_d(t))(P_{c_d}(t)+P_{P_d}(t)\\&+\sum_s O_s(\gamma_s+\sum_i r_{is}(\alpha_{cpu_s}\mu_{cpuVM_i}cpu_{VM_i}/cpu_s\\&+\alpha_{mem_s}\mu_{memVM_i}cpu_{VM_i}/mem_s\\&+\alpha_{disk_s}\mu_{diskVM_i}disk_{VM_i}/disk_s)))\end{aligned} \quad (12)$$

such that:

$$\begin{aligned}&\textstyle\sum_{VM_i\in s} cpu_{VM_i} \le cpu_s\\&\textstyle\sum_{VM_i\in s} mem_{VM_i} \le mem_s\\&\textstyle\sum_{VM_i\in s} disk_{VM_i} \le disk_s\\&\textstyle\sum_{VM_i\in s} bw_{VM_i} \le bw_s\end{aligned} \quad (13)$$

$$\begin{aligned}&\mu_{cpu_s} = (\textstyle\sum_{VM_i\in s}\mu_{cpuVM_i}cpu_{VM_i})/cpu_s\\&\mu_{mem_s} = (\textstyle\sum_{VM_i\in s}\mu_{memVM_i}mem_{VM_i})/mem_s\\&\mu_{disk_s} = (\textstyle\sum_{VM_i\in s}\mu_{diskVM_i}disk_{VM_i})/disk_s\end{aligned} \quad (14)$$

and where:

$$r_{is} = 1\, where\, VM_i\, is\, on\, the\, server\, s\,,\, 0\, other\, wise \quad (15)$$

C represents the GHG emissions in the network at time t; cpu_s, mem_s, $disk_s$, and bw_s represent the total available CPU core, memory, storage, and bandwidth on server s; and cpu_{VM_i}, mem_{VM_i}, $disk_{VM_i}$, and bw_{VM_i} represent the CPU core, memory, storage, and bandwidth needed for VM_i; and μ_{cpuVM_i}, μ_{memVM_i}, and μ_{diskVM_i}represent the use percentage for CPU, memory, and storage of VM_i. The energy optimization problem can be formulated in the same way.

This is a multidimensional Knapsack problem, which is one of Karps well-known 21 NP-complete problems [17]. Since the objective function is nonlinear and there are nonlinear constraints, the optimization model is a nonlinear programming problem with binary variables. Generally, this type of problem cannot be solved by mathematical programming solvers.

We chose a Genetic Algorithm (GA) to solve the problem, because of its nature. A GA will decide which servers to consolidate, and which servers to turn off. It will also decide whether or not the entire data center in question needs to be turned off and all the VMs on its servers migrated to other data centers. The VMs on each server become variables for the GA, and the GA will optimize the carbon footprint, while considering the available memory, CPU, and storage on each server. Because of the nature of the problem, a modified mutation function is added to the GA for a faster and better result. Because we know that the optimized value can result in the removal of all the VMs from a server and their migration to other, preferably green, servers, the modified mutation function will alter the genes in such a way that the number of VMs on the servers of a few random data centers is pushed to zero. We do not focus here on the algorithmic part of the GA, but on its use. The GA parameters are tuned by running the algorithm several times with different values for those parameters. In this section, we describe the optimization tool for the network optimization problem. Following, we show the best interval on which to run the optimizer on the network.

In Figures 3 and 4, the network is optimized considering both large and small intervals. In the legend, [VPC-opt] means WAN optimization over the whole network and [Carbon-opt] means optimization for the carbon cost function.

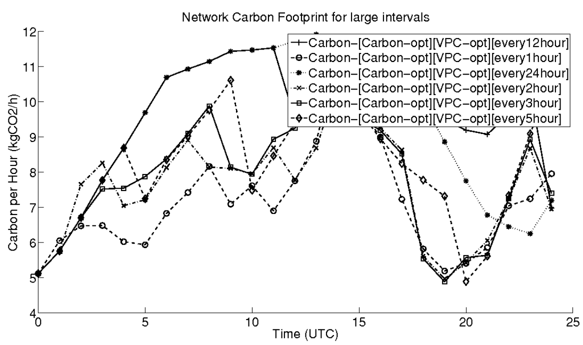

Fig. 3. Network carbon optimization under large intervals

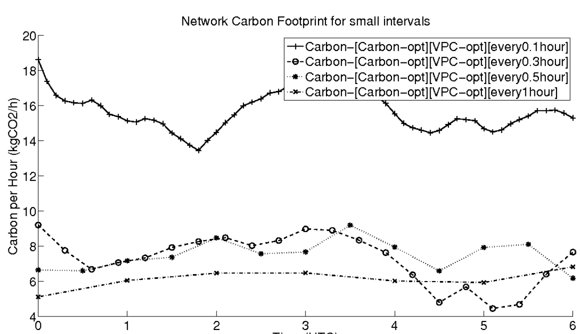

Fig. 4. Network carbon optimization under small intervals

Based on the results of using large and small intervals, we find that the optimum interval value lies between 0.5 and 2 hours. In this paper, an 1 hour interval is used for all scenario tests to run the optimization algorithm on the network.

C. Results

For the first test scenario, the network of 40 servers running on 13 data centers located in 7 different cities around the world is tested under sun movement and under the random movement of several wind streams. Three different carbon footprints were optimized and calculated for this network, and are shown in Figure 5. In the legend, [LAN-opt] means LAN optimization over each data center. The first carbon footprint corresponds to an even distribution of VMs on the servers, without any optimization. We see that this carbon footprint is at the highest level. The second carbon footprint corresponds to server consolidation within each data center, with the help of virtualization technology. This carbon footprint is much smaller than the first, as shown in Figure 5. The third carbon footprint corresponds to data center consolidation over WAN connections. Here, data centers are consolidated according to their greenness factor and their resource availability. The carbon footprint in this case is smaller than that for server consolidation.

The carbon footprint was measured for a period of 24 hours, and the results are presented in Table II under Normal VM load (47%). According to this measurement, using data center consolidation over WAN connections can lead to a 59% greater carbon footprint reduction, compared to server consolidation over LAN connections.

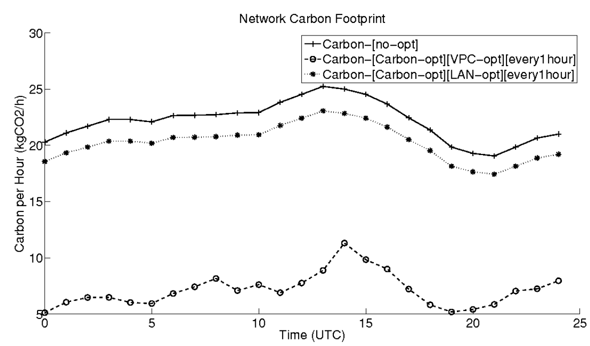

Fig. 5. Network carbon footprint for LAN-based clouds and VPCs

For the second test scenario, the network was tested under heavy load (60% CPU use) and light load (33% CPU use), as shown in Figures 6 and 7 respectively.

The total carbon footprint of the network was measured for a period of 24 hours for both test scenario 1 and test scenario 2, and the results are presented in Table II.

TABLE II
24 HOUR CARBON FOOTPRINT COMPARISON UNDER HEAVY AND LIGHT VM LOADS

Scenario	No opt CO2kg	LAN opt CO2kg	VPC opt CO2kg	LAN opt %	VPC opt %	VPC perf. %
Light	500.45	401.94	76.69	19.68	84.68	64.99
Normal	553.88	506.32	178.70	8.59	67.74	59.15
Heavy	607.32	560.29	254.71	7.74 5	8.06	50.32

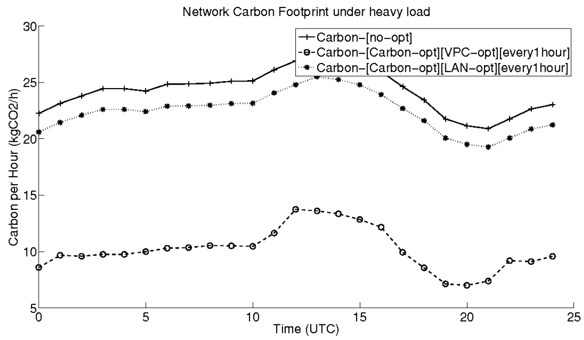

Fig. 6. Network carbon footprint under heavy VM load

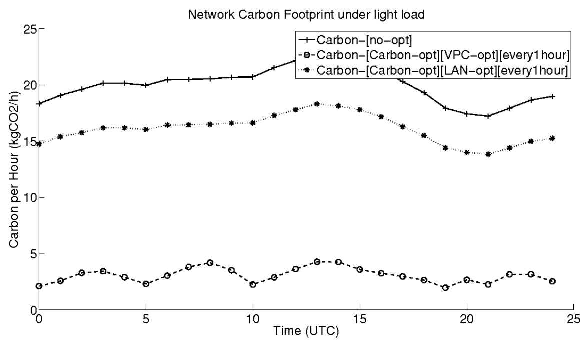

Fig. 7. Network carbon footprint under light VM load

The first three columns in the table shows the CO2 emissions in kilograms of the whole network under no optimization, and under LAN and VPC optimization. The "LAN opt %" column shows the percentage of CO2 production reduction using LAN server consolidation. The "VPC opt %" column shows the percentage of CO2 production reduction by using WAN data center consolidation. The last column, "VPC perf", shows the VPC optimization performance over LAN optimization in terms of carbon reduction percentage.

Based on the results in Table II, regardless of the VM load on the network, VPC optimization has a better performance in terms of carbon footprint reduction than LAN optimization. The best performance is achieved in cases where VPC use is less. In these cases, there are more green options available to the controller to move the VMs to the green data centers, and then shut down a larger number of non-green data centers (which includes their PDUs and cooling systems).

In the third test scenario, the network is first optimized for its carbon footprint, and second for its energy consumption. In both optimizations, carbon footprint and energy consumption are measured as shown in Figures 8 and 9 respectively. In the legend area, [Energy-opt] means that the optimization is for the energy cost function. The amount of energy consumed and the carbon footprint over 24 hours are presented in Tables III and IV respectively.

As shown in Figures 8 and 9, the energy consumption and carbon footprint of the network are lower in both energy and

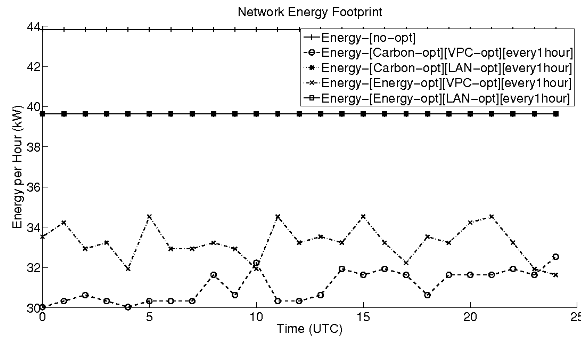

Fig. 8. Network energy measurement

carbon optimization for WAN data center consolidation. For LAN server consolidation, energy optimization is identical to carbon optimization. But, in WAN data center consolidation for the given dataset, the carbon is not optimized when the energy is optimized, and the energy is not optimized when the carbon is optimized, which is not the case at a single data center.

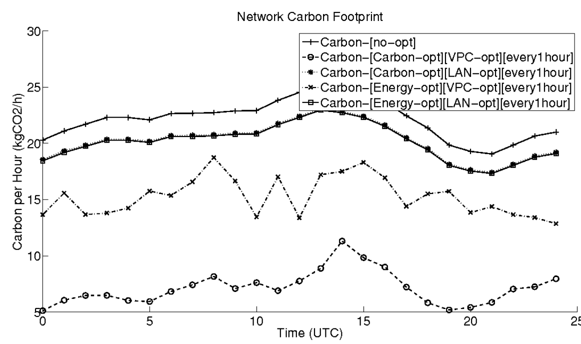

Fig. 9. Network carbon measurement

In Figure 9, there is a big difference in the network carbon footprint in the carbon and energy optimization modes, but there is not much difference in their energy consumption, as shown in Figure 8. This means that, even though the energy consumption is not optimized in VPC optimization, carbon footprint reduction is significant. In other words, a modest increase in energy consumption (4.96%) can result in a major reduction in the size of the network carbon footprint (36.64%). The figures 4.96% and 36.64% are calculated from the difference between energy and carbon performance percentage (VPC perf %) in carbon optimization and energy optimization modes.

V. Conclusion and Future Work

In this paper, a formulation for calculating the carbon footprint and energy consumption of a WAN network of data centers is presented. This formulation is used to measure the carbon footprint of a simulation platform comprising 13 data centers in seven cities at different geographical locations

TABLE III
24 HOUR CARBON FOOTPRINT COMPARISON FOR ENERGY AND CARBON OPTIMIZATION

Scenario	No opt CO2kg	LAN opt CO2kg	VPC opt CO2kg	LAN opt %	VPC opt %	VPC perf. %
Carbon opt.	553.88	506.32	178.70	8.59	67.74	59.15
Energy opt.	553.88	506.32	381.65	8.59	31.09	22.51

TABLE IV
24 HOUR ENERGY CONSUMPTION COMPARISON FOR ENERGY AND CARBON OPTIMIZATION

Scenario	No opt kWh	LAN opt kWh	VPC opt kWh	LAN opt %	VPC opt %	VPC perf. %
Carbon opt.	1,095.83	990.83	776.93	9.58	29.10	19.52
Energy opt.	1,095.83	990.83	831.23	9.58	24.15	14.56

around the world. A heuristic algorithm (a modified GA) is used to optimize the carbon footprint of the network. For tuning the optimization, different optimization intervals have been proposed to extract the best optimization interval.

The network was tested under different loads, and our results show a significant carbon footprint reduction through VPC data center consolidation compared to LAN server consolidation. Also, the network was tested under carbon and energy optimization, the results of which show that carbon reduction is not necessarily equal to energy efficiency in VPCs, which is not the case in a single data center.

For future work, different use cases can be tested under different test scenarios with different VM loads. The number of cities can be increased in order to observe the behavior of the algorithm on a very large network of data centers. Also, in this paper, randomly generated data were used for wind streams. In future, real wind stream data will be retrieved from historical weather data and used in the simulation for more realistic results. Finally, for a more realistic solar energy simulation, simulated clouds can be considered in the simulation platform.

As is concluded above, greater carbon footprint reduction can be achieved through reduced power use in a VPC. However, this may not be profitable for investors in VPCs, because of the correspondingly reduced use of their infrastructure. To address this dilemma, we will study the effect of using Low Carbon Virtual Private Clouds under carbon laws such as ”Cap and Trade”, ”Cap and Dividend”, and ”Cap and Reward,” in order to reduce the final service price.

ACKNOWLEDGMENTS

The authors thank Reza Farrahi Moghaddam for reviewing the paper and for his helpful suggestions. The authors also thank CANARIE (Canadian Network for Advanced Research in Education) for their financial support of the GreenStar Network project.

REFERENCES

[1] Lo, V., Potdar, V., *State of the Art: Carbon Emission Accountability Modeling*, Broadband, Wireless Computing, Communication and Applications (BWCCA), 2010 International Conference on , vol., no., pp.830-835, November 4-6, 2010.

[2] Christopher Clark, Keir Fraser, Steven Hand, Jacob Gorm Hansen, Eric Jul, Christian Limpach, Ian Pratt, and Andrew Warfield, *Live migration of virtual machines*, In Proceedings of the 2nd conference on Symposium on Networked Systems Design & Implementation (NSDI'05), Vol. 2. USENIX Association, Berkeley, CA, USA, 273-286, 2005.

[3] Beloglazov, Anton, Buyya, Rajkumar, Lee, Young, Choon, and Zomaya, Albert *A Taxonomy and Survey of Energy-Efficient Data centers and Cloud Computing Systems* TECHNICAL REPORT. CLOUDS-TR-2010-3. CLOUD COMPUTING AND DISTRIBUTED SYSTEMS LABORATORY. UNIVERSITY OF MELBOURNE. AUSTRALIA. JUNE 30. 2010.

[4] Daniel Gmach, Jerry Rolia, Ludmila Cherkasova, and Alfons Kemper, *Resource pool management: Reactive versus proactive or let's be friends*, Computer Networks, vol 53, issue 17, Virtualized Data centers, December 3, 2009, Pages 2905-2922.

[5] L. Liu, H. Wang, X. Liu, X. Jin, W.B. He, Q.B. Wang, and Y. Chen, *GreenCloud: a new architecture for green data center*, Proceedings of the 6th International Conference Industry Session on Autonomic Computing and Communications Industry Session, Barcelona, Spain: ACM, 2009, pp. 29-38.

[6] Van der Merwe, J., Ramakrishnan, K. K., Fairchild, M., Flavel, A., Houle, J., Lagar-Cavilla, H. A., and Mulligan, J., *Towards a ubiquitous cloud computing infrastructure*, 17th IEEE Workshop on Local and Metropolitan Area Networks (LANMAN), vol., no., pp. 1-6, May 5-7, 2010.

[7] T. Wood, K. Ramakrishnan, J. van der Merwe, and P. Shenoy, *CloudNet: A Platform for Optimized WAN Migration of Virtual Machines*, University of Massachusetts Technical Report TR-2010-002, January 2010.

[8] Farrahi Moghaddam, Fereydoun, and Cheriet, Mohamed, *Decreasing live virtual machine migration down-time using a memory page selection based on memory change PDF*, International Conference on Sensing and Control (ICNSC), pp. 355-359, April 10-12, 2010.

[9] Timothy Wood, Alexandre Gerber, K. K. Ramakrishnan, Prashant Shenoy, and Jacobus Van der Merwe, *The case for enterprise-ready virtual private clouds*, In Proceedings of the 2009 conference on Hot Topics in Cloud Computing (HotCloud09). USENIX Association, Berkeley, CA, USA, 4-4, 2009.

[10] Qian, Ling, Luo, Zhiguo, Du, Yujian, and Guo, Leitao, *Cloud Computing: An Overview*, Lecture Notes in Computer Science, Springer Berlin/Heidelberg (5931), pp. 626-631, 2009.

[11] Aman Kansal, Feng Zhao, Jie Liu, Nupur Kothari, and Arka A. Bhattacharya. *Virtual machine power metering and provisioning*, In Proceedings of the 1st ACM Symposium on Cloud Computing (SoCC10). ACM, New York, NY, USA, 39-50, 2010.

[12] D. Economou, S. Rivoire, C. Kozyrakis, and P. Ranganathan, *Full-system power analysis and modeling for server environments*, In: Workshop on Modeling Benchmarking and Simulation (MOBS), June 2006.

[13] S. Srikantaiah, A. Kansal, and F. Zhao, *Energy aware consolidation for cloud computing*, In: Workshop on Power Aware Computing and Systems (HotPower 08). San Diego, USA, December 2008.

[14] Andreas Berl, Erol Gelenbe, Marco Di Girolamo, Giovanni Giuliani, Hermann De Meer, Minh Q. Dang, and Kostas Pentikousis, *Energy-Efficient Cloud Computing* The Computer Journal, vol. 53, no. 7, (September 2010), pp. 1045-1051.

[15] Zhang, Qi, Cheng, Lu, and Boutaba, Raouf, *Cloud computing: state-of-the-art and research challenges*, Journal of Internet Services and Applications, 2010, Springer London.

[16] Lenzen, Manfred, *Current State of Development of Electricity-Generating Technologies: A Literature Review* Energies 2010, 3(3), 462-591.

[17] R. M. Karp, *Reducibility Among Combinatorial Problems*, in R. E. Miller and J. W. Thatcher (eds.), Complexity of Computer Computations. Plenum Press, New York, 1972.

[18] Zhang, Qi, Cheng, Lu, and Boutaba, Raouf, *Cloud computing: state-of-the-art and research challenges*, Journal of Internet Services and Applications, 2010, Springer London.