

RESEARCH

Open Access

A decision framework for placement of applications in clouds that minimizes their carbon footprint

Marc X Makkes^{1,2*}, Arie Taal¹, Anwar Osseyran³ and Paola Grosso¹

Abstract

Cloud computing gives users much freedom on where they host their computation and storage. However the CO_2 emission of a job depends on the location and the energy efficiency of the data centers where it is run. We developed a decision framework that determines to move computation with accompanying data from a local to a *greener* remote data center for lower CO_2 emissions. The model underlying the framework accounts for the energy consumption at the local and remote sites, as well as of networks among them. We showed that the type of network connecting the two sites has a significant impact on the total CO_2 emission. Furthermore, the task's complexity is a factor in deciding when and where to move computation.

Introduction

From a user's perspective, reducing the environmental load of his computational tasks is equivalent to looking for a green data center, *i.e.* a data center with a low power usage effectiveness (PUE). Many data centers advertise their greenness as an added value for customers. A recent study [1] shows that 71% of the data centers measure the PUE and that the mean value is about 1.8. Another survey for data centers in Europe [2] came up with a higher mean value of the PUE. Some large data centers claim to have a PUE approaching the theoretical value of 1. We argue that the PUE is not the only factor to consider: the energy sources powering a data center and the network used to move the data are also important, as they determine the amount of CO_2 emitted for a given task.

We will present a framework that facilitates a user to decide where to perform a task, whether at a local data center or remotely at a clearer data center. The framework does not only take the CO_2 emission of the data centers into account, but also estimates the CO_2 emission of the transport network between them when input/output data accompanied the task. We can do this by exploiting the

relation between energy produced in kWh and CO_2 emission for different energy sources (see Equation 9). The CO_2 emission of the network of a data center is a modest part of the total CO_2 of the data center [3]. However, deciding if offloading of an individual task to an optional cleaner data center is preferable, the contribution of the network (data center LAN and transport network) can be a substantial part of the decision. This means that if the decision framework introduced in this paper will be applied to all jobs of a data center, the total CO_2 emission of both the data center and the optional cleaner data centers for offloading tasks will decrease.

The framework can make a prediction of the total CO_2 emission for different scenarios, namely software interactive computation and hot or cold data storage. For each scenario we identify the equipment required in the local and the remote data center, *e.g.*, for a computational task other equipment is used than for hot data storage. Subsequently we use models including the power consumption of the devices in use. In this paper we will focus on the computational scenario, but the interested reader can find some details of the storage scenario in [4,5].

A common aspect of all scenarios considered is the amount of data involved. The input data determines the energy cost of the data transport part, first through the LAN of the local data center, then across the core network, Internet or light path, and finally through the LAN

*Correspondence: marc.makkes@tno.nl

¹University of Amsterdam, Kruislaan 904, Amsterdam, The Netherlands

²TNO Information and Communication Technology, Eemsgolaan 3, Groningen, The Netherlands

Full list of author information is available at the end of the article

of the chosen remote data center. When output data plays a role we assume that the user is located near the local data center, so the energy cost associated with the output data is the energy cost for the local LAN versus the cost of the remote LAN and the transport network.

The equipment present in a data center, including the LAN devices, can be more realistically identified than the number of devices in a transport network to another data center. For the former we chose the same internal architecture for both data centers being compared; this allows us to purely focus on the sustainability of both. The latter instead depends not only on the type of network, Internet or light path, but also on the geographical location of both data centers. Therefore, our framework makes use of network models depending on the type of network and on the location of both endpoints to give an estimate of the minimal number of hops in the network. Furthermore, the geographical location of both endpoints determines the possible countries crossed by the shortest path transport network. These estimates make it possible to attach a CO_2 emission to the transport network. Data on the energy types used by different European countries is available. If the transport network e.g., connects a data center in the Netherlands with one in Austria, a considerable part of the shortest path network will cross Germany. So the energy cost can be divided in three contributions, according to the distance spanned in each of the countries crossed. For each country we can calculate a mean CO_2 emission based on the types of energy sources used in that country [6-11].

The rules applied to facilitate a user in his decision can also be applied by a scheduler of a data center. If a user can specify the complexity of his task, i.e., how computation time and or the amount of output data scale as a function of the input data, a scheduler can determine where to schedule the job such that the emission of the task in gr. CO_2 is minimal. In that case the user need not to know about remote data centers and their PUE's, because this knowledge resides in the scheduler's database.

Related work

There are different aspects one can focus on in the optimization process of data center infrastructure costs. We chose to concentrate on CO_2 emission costs, but there are other possible focus points such as economic costs, power utilization and infrastructure utilization. For each one of these costs there is ample existing research: namely for economic costs the work done by [12-14], for power utilization the work by [13,14] and [12,15] for infrastructure utilization.

Optimization of each of these aspects can lead to different outcomes. For example, a data center running more energy efficiently but supplied by energy produced from brown coal has a higher CO_2 emission cost than a data

center operating much less efficiently that is using hydro electric power.

In this paper we focus on CO_2 emission costs. What for us is of interest is the ever-increasing effort in modeling the power consumption of networks and data center equipment. Understanding the power consumption in more detail of networks and computer equipment and their behavior under different conditions, gives the opportunity to better predict the impact of cloud computing and storage on the environment and to develop algorithms and strategies to reduce the carbon footprint. The way we predict the energy consumption of LAN's and transport networks is based on the work of Baliga et al. [16].

We distinguish different kind of networks, LAN's, Internet and light path, each with their specific type of equipment. Our novel contribution is that we integrate and extend different models into a single decision framework for greener computing. The models used can be easily enhanced, allowing the framework to evolve if one wishes. Our main impetus for the framework presented is that not only end users but also data centers' operators and cloud service providers should think under what conditions it is better to host a job locally, or to host it elsewhere.

Energy model

When deciding to move data and the accompanied computation from a local to a remote data center we have to define an energy consumption metric that accounts for both data centers and the transport network between them. With this metric we should be able to calculate values for the following equation that indicates when movement to a remote data center is to be preferred above local processing:

$$\begin{aligned} \text{Energy cost local processing} > \text{Energy cost network} \\ &+ \text{Energy cost remote processing} \end{aligned} \quad (1)$$

where:

$$\begin{aligned} \text{Energy cost network} = &\text{Energy cost of local data center LAN} \\ &+ \text{Energy cost transport network} \\ &+ \text{Energy cost of remote data center LAN} \end{aligned} \quad (2)$$

In the following sections we will focus on two different aspects that contribute to Equation 1: how efficient a data center uses its energy, and what are the different components used in the data center and the network.

How efficiently a data center uses its energy

To rate the energy efficiency of data centers the commonly used number is the PUE. The PUE is expressed

as the ratio of the total power consumption of a data center (P_{TOT}) to the total power consumption of IT equipment like storage devices, servers, routers (P_{IT}).

$$PUE = \frac{P_{TOT}}{P_{IT}}, 1 < PUE < \infty \tag{3}$$

In the calculation of the PUE of a data center all equipment that is not considered a computing device, like pumps, air conditioners, lighting, are part of P_{TOT} only, whereas the power used by servers, storage equipment, network equipment are incorporated in both P_{IT} and P_{TOT} .

The different data center and network components used

An important conclusion of a recent study by Tucker [17] is that ‘in a global scale (data) network, the energy consumption of the switching infrastructure is larger than the energy consumption of the transport infrastructure.’ We will therefore make a distinction between optical communication systems and conventional Ethernet. We will restrict ourselves to the case where the end user is directly attached to the data center clouds/clusters via a corporate network. The user (or a scheduling application on his behalf) must decide whether the data with the accompanied computation stays at a data center or should be moved to another data center. If he decides to move data, the data will be transported over a public data network given that different data centers are mostly geographically separated. When data traverses the Internet energy consumption can be estimated by adding the contributions to the energy of switches, amplifiers, transceivers, etc. that the data traverses. At both sides, at the local and remote data center, we have the local area network (LAN) of the data center itself that connects the data storage devices and servers to the outside world, i.e., the transport network. To keep calculations simple we assume the same components are present in the LAN of any data center. Table 1 lists the typical equipment data traverses through the LAN of a data center.

Table 1 Components of a data center LAN

LAN data center
Host (network interface)
2x Switch
2x Firewall
Switch
Router

According to Table 1 we arrive (see Baliga et al. [16] Eq. 2) at the following equation for the energy consumption per bit for the LAN of a data center:

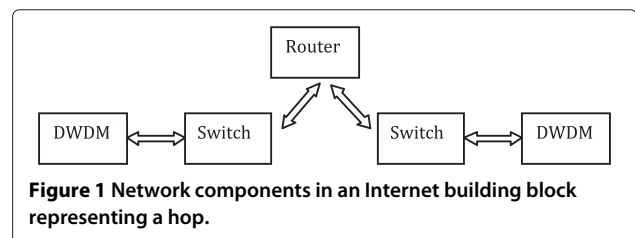
$$\hat{E}_{LAN_data_center} = \frac{PUE}{U} \cdot \left(\frac{P_{host}}{C_{host}} + 3 \frac{P_{switch}}{C_{switch}} + 2 \frac{P_{firewall}}{C_{firewall}} + \frac{P_{router}}{C_{router}} \right) [W/bit/s] \tag{4}$$

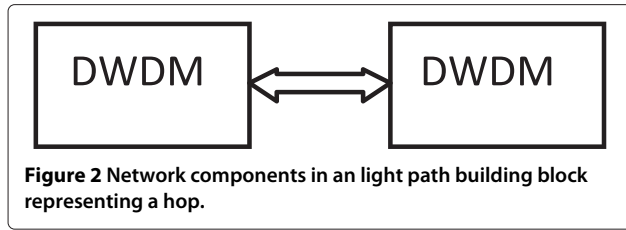
where P_{host} , P_{switch} , $P_{firewall}$, and P_{router} are the power consumed by the host computer where the data resides, Ethernet switches, firewall, and data center gateway router, respectively. The capacities of the corresponding equipment and measured in bits per second are given by C_{host} , C_{switch} , $C_{firewall}$, and C_{router} .

Here, the factor U accounts for the utilization of the network equipment, expressing the fact network equipment typically does not operate at a full utilization while still consuming 100% of the power [18], a factor we took equals to 0.5.

Data transfers across a transport network can use two different types of connections: the regular Internet and dedicated connections. The regular Internet is available to all users, while in principle dedicated connections (light paths) are more frequently encountered in scientific and corporate environments for high-end users. In both cases the data transfer can be over long or short distances, and we account for this in our model. Figure 1 and 2 show the data network building blocks we assume to be representative for Internet and light path networks.

With these building blocks we compose short and long distance network paths. Multiple Internet building blocks are connected to each other, and multiple light path building blocks are connected via a switch with each other. The entry points and exit points for any kind of data network are a switch connected to a dense wavelength division multiplexing node (DWDM). Baliga et al. [16] take a mean number of hops for each kind of network (Internet and light path), where we take the number of hops for each kind of network depending on the geographical position of both endpoints. Figures 3, 4, 5, and 6 show the example diagrams





for single hop and three hop Internet and light path networks.

We write for the processing cost of a task in Equation 1 :

$$E_{processing_data_center}(T_{processing}) = PUE_{data_center} \times P_{comp_host} \cdot T_{processing} [kWh] \quad (5)$$

where P_{comp_host} is the power consumption of a computation host in kW and $T_{processing}$ the processing time in CPU core hours. If the task is accompanied with N_{in} GByte of input data, this data will always be transferred through the LAN of the local data center. In case the task will be processed at a remote data center, this data will be once more transferred through the LAN of the local data center, subsequently the connecting transport network and the LAN of the remote data center. The transport cost of the LANs follow from Equation 4.

$$E_{LAN_data_center}(N_{in}) = \frac{PUE_{data_center}}{U} \cdot \left(\frac{P_{host}}{C_{host}} + \frac{3P_{switch}}{C_{switch}} + \frac{2P_{firewall}}{C_{firewall}} + \frac{P_{router}}{C_{router}} \right) \cdot \frac{8N_{in}}{3600} [kWh] \quad (6)$$

while the connecting transport network cost will depend on the type of network, Internet or light path, and the number of hops:

$$E_{transport_internet}(N_{in}) = \frac{PUE_{network}}{U} \cdot \left(\left(\frac{2P_{switch}}{C_{switch}} + \frac{2P_{DWDM}}{C_{DWDM}} \right) + \left(\frac{2P_{switch}}{C_{switch}} + \frac{2P_{DWDM}}{C_{DWDM}} + \frac{P_{router}}{C_{router}} \right) \times n_{hops} \right) \cdot \frac{8N_{in}}{3600} [kWh] \quad (7)$$

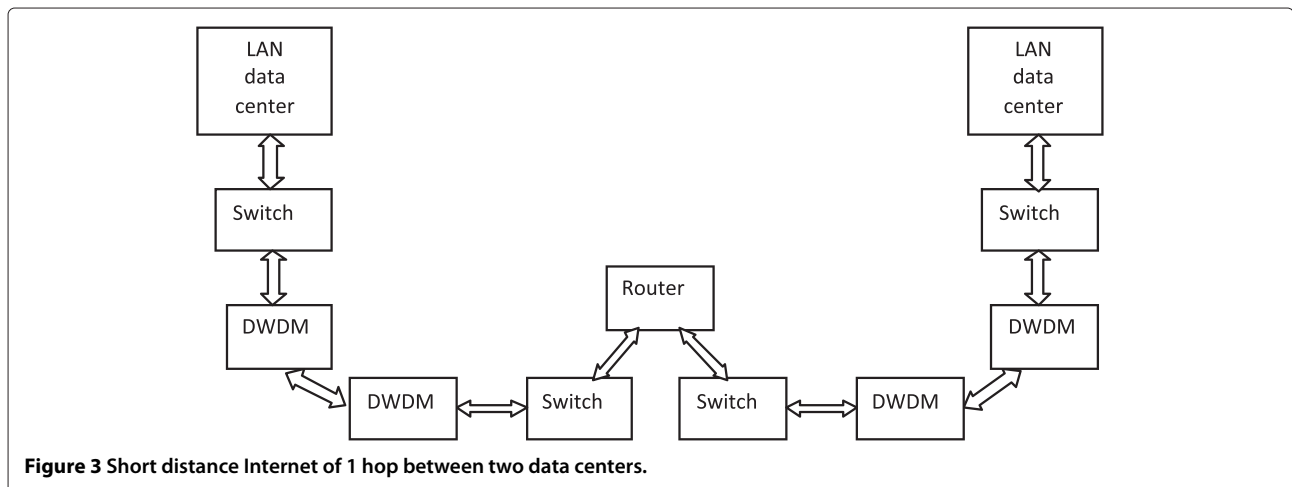
$$E_{transport_lightpath}(N_{in}) = \frac{PUE_{network}}{U} \cdot \left(\left(\frac{2P_{switch}}{C_{switch}} + \frac{2P_{DWDM}}{C_{DWDM}} \right) + \left(\frac{2P_{DWDM}}{C_{DWDM}} \right) \cdot n_{hops} + \left(\frac{P_{switch}}{C_{switch}} \right) \times (n_{hops} - 1) \right) \cdot \frac{8N_{in}}{3600} [kWh] \quad (8)$$

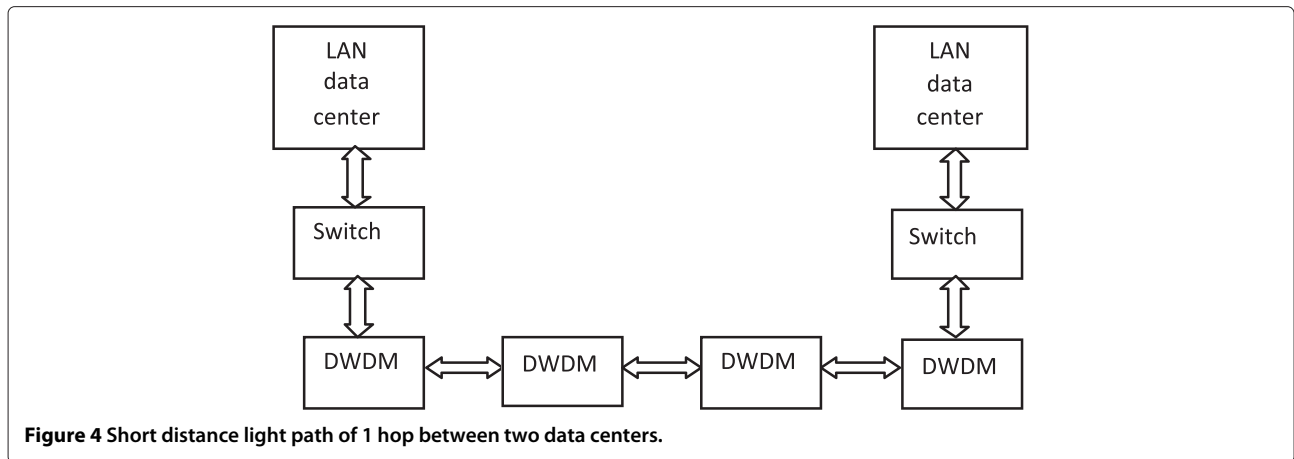
where the factor 8 accounts for the translation of bytes into bits, as the terms P/C are measured in kW/Gb/s.

In order to solve eq. 1 for the total energy consumption to move data we need values for the different equipment the data traverses. Table 2 lists the adopted values for the power per capacity (P/C) in kW/Gb/s of the devices listed in Table 1 and depicted in Figure 1 and 2. All values are taken from [16] except the value for routers which we obtained from measurements at our local data center.

Sustainability

We are interested in the sustainability aspects of the energy sources used in the data network and data centers,





and in the subsequent CO_2 emissions. One way we propose to incorporate this, is to transform energy cost in kWh to carbon emission cost effects. A kWh can be converted into grams of produced CO_2 according to the following formula

$$1kWh \sim Xgr.CO_2 \tag{9}$$

where the value of the factor X depends on the type of energy source, e.g. $X = 870$ for anthracite electricity production, and $X = 370$ for gas electricity production. In our framework values for X are compiled from different sources [6-8], leading to the values presented in Table 3.

We can now map the energy costs in kWh given by Equations 5, 6, 7 and 8 into an equivalent carbon emission cost K in terms of grams of CO_2 produced:

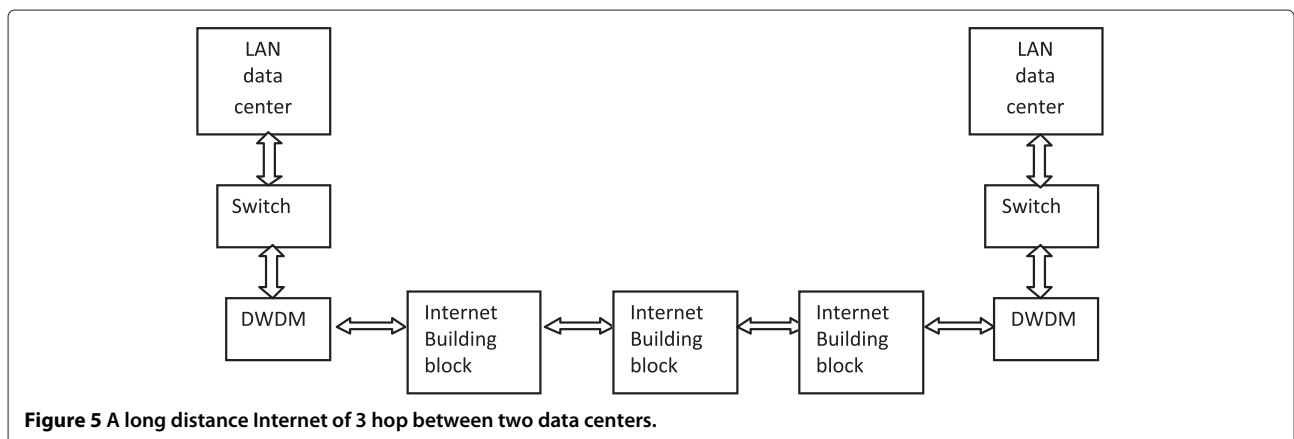
$$\begin{aligned} K_{processing_data_center}(X_{data_center}, T_{processing}) \\ = X_{data_center} \cdot E_{processing_data_center}(T_{processing}) \end{aligned} \tag{10}$$

$$\begin{aligned} K_{LAN_data_center}(X_{data_center}, N_{in}) = X_{data_center} \\ \times E_{LAN_data_center}(N_{in}) \end{aligned} \tag{11}$$

$$\begin{aligned} K_{transport_network}(X_{transport_network}, N_{in}) \\ = X_{transport_network} \cdot E_{transport_network}(N_{in}) \end{aligned} \tag{12}$$

Decision Equation 1 for transporting data with accompanied computation to another data center transformed to grams of CO_2 produced now reads:

$$\begin{aligned} K_{processing_local_dc}(X_{local_dc}, T_{processing}) \\ + K_{LAN_local_dc}(X_{local_dc}, N_{in}) \\ > 2 \cdot K_{LAN_local_dc}(X_{local_dc}, N_{in}) \\ + K_{transport_network}(X_{transport_network}, N_{in}) \\ + K_{LAN_remote_dc}(X_{remote_dc}, N_{in}) \\ + K_{processing_remote_dc}(X_{remote_dc}, T_{processing}) \end{aligned} \tag{13}$$



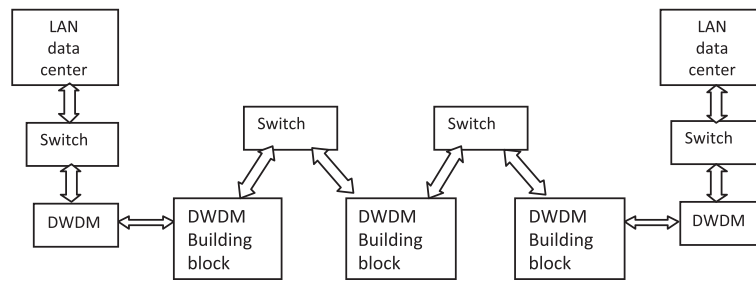


Figure 6 A long distance light path of 3 hop between two data centers.

The terms on the left of the equation describe the total emission if the computation task is performed locally, while the terms on the right site concern the emission cost if the task is offloaded to and performed at a remote data center. Left we see the contribution of the LAN for the data coming in once, while on the right we see the LAN of the local data center contributes twice, as the data needs to come in from the owner and after the decision is sent out towards the remote data center. In case we have to deal with output data from a computational task we assume that the one interested in the output data is located near the local data center, and we extend Equation 13 to:

$$\begin{aligned}
 &K_{processing_local_dc}(X_{local_dc}, T_{processing}) \\
 &+ K_{LAN_local_dc}(X_{local_dc}, N_{in}) \\
 &+ K_{LAN_local_dc}(X_{local_dc}, N_{out}) \\
 &> 2 \cdot K_{LAN_local_dc}(X_{local_dc}, N_{in}) \\
 &+ K_{transport_network}(X_{transport_network}, N_{in}) \quad (14) \\
 &+ K_{LAN_remote_dc}(X_{remote_dc}, N_{in}) \\
 &+ K_{processing_remote_dc}(X_{remote_dc}, T_{processing}) \\
 &+ K_{LAN_remote_dc}(X_{remote_dc}, N_{out}) \\
 &+ K_{transport_network}(X_{transport_dc}, N_{out})
 \end{aligned}$$

Table 2 Power per capacity for the different components in our model

Equipment	Power per capacity [kW/Gb/s]
Host data storage	0.2800
Router	0.0120
Ethernet switch	0.0230
Firewall	0.0160
DWDM terminal node	0.0034

Decision framework

Equations 13 and 14 are at the basis of our decision framework. They can be used in decision policies taken by a scheduler (section ‘Decision policies’) as well as in a web calculator available to end users (section ‘Web calculator’). A scheduler will take a decision on where to place computation based on these policies, and it will provide the user with detailed information on the CO₂ emission cost of the chosen scenario. The complexity of tasks, *i.e.*, how the computation time scales with the input data and how the output data scales with the input data, is a factor included in the decision framework too.

Decision policies

If a user submits a task and indicates the processing time and the amount of input data needed, and the amount of output data expected, a scheduler should be able to decide whether the task can be better performed locally or at another remote data center from a knowledge base. To decide whether a remote data center is a greener option

Table 3 Values for the factor X used in our framework as function of the different energy sources (in decreasing value of X)

Energy source	X value
Lignite/brown coal	950
Anthracite	870
Crude oil	640
Gas works gas	400
Natural gas	380
Nuclear power	66
Geothermal power	40
Biomass	30
Solar power	22
Hydroelectricity	15
Wind power	10

the scheduler applies Equation 14 as a decision policy, which can be written as follows:

$$\begin{aligned}
 & X_{local_dc} \cdot PUE_{local_dc} \cdot P_{comp.host_local_dc} \cdot T_{processing} + X_{local_dc} \\
 & \times PUE_{local_dc} \cdot E_{LAN_local_dc} \cdot N_{in} + X_{local_dc} \cdot PUE_{local_dc} \\
 & \times E_{LAN_local_dc} \cdot N_{out} > 2 \cdot X_{local_dc} \cdot PUE_{local_dc} \\
 & \times E_{LAN_local_dc} \cdot N_{in} + X_{network} \cdot PUE_{network} \cdot E_{network} \\
 & \times N_{in} + X_{remote_dc} \cdot PUE_{remote_dc} \cdot E_{LAN_remote_dc} \cdot N_{in} \\
 & + X_{remote_dc} \cdot PUE_{remote_dc} \cdot P_{comp.host_remote_dc} \cdot T_{processing} \\
 & + X_{remote_dc} \cdot PUE_{remote_dc} \cdot E_{LAN_remote_dc} \cdot N_{out} \\
 & + X_{network} \cdot PUE_{network} \cdot E_{network} \cdot N_{out}
 \end{aligned} \tag{15}$$

where $T_{processing}$, N_{in} , N_{out} are respectively the computation time in CPU core hours, the amount of input data and the amount of output data, both in GBytes. Furthermore $E_{LAN_local_dc}$, $E_{LAN_remote_dc}$ and $E_{network}$ are unit energy consumptions of the data center LANs and the connecting transport network, expressed in kWh/GByte. Values for X_{local_dc} , PUE_{local_dc} , X_{remote_dc} , and PUE_{remote_dc} reside in a knowledge base of the scheduler. The values $P_{comp.host_local_dc} = P_{comp.host_remote_dc} = 0.355kW$ [16] and $E_{LAN_local_dc} = E_{LAN_remote_dc} = 0.0017kWh/GByte$ (derived from Equation 6 with the adopted values for network equipment) are constants for any decision policy, whereas the value for $E_{network}$ depends on the type of network and on the number of different hops, Equations 7 and 8. In case both light path and Internet connections are possible the scheduler can try both transport networks and the number of hops for the connecting shortest path is retrieved from the knowledge base too. For reasons of simplicity we take $E_{LAN_local_dc}$ equals to $E_{LAN_remote_dc}$ and $P_{comp.host_local_dc}$ equals to $P_{comp.host_remote_dc}$. In an implementation of a scheduler, the scheduler will have knowledge of its own data center and all values concerning a remote data center will be retrieved by issuing a proposal to the scheduler of the remote data center. In that case, values for local and remote equipment maybe different.

We will illustrate a decision made with an example, where the local data center, with $PUE_{local_dc} = 1.4$, is situated in the Netherlands and is powered by electricity produced from natural gas (380 gr. CO_2/kWh). Suppose the only alternative at the disposal of the scheduler is a remote data center in Tirol, Austria, that is powered by hydroelectricity (15 gr. CO_2/kWh) and $PUE_{remote_dc} = 1.8$. Values for the connecting transport network can be prepared as knowledge to the scheduler in the following way. If the transport connection between the Netherlands and

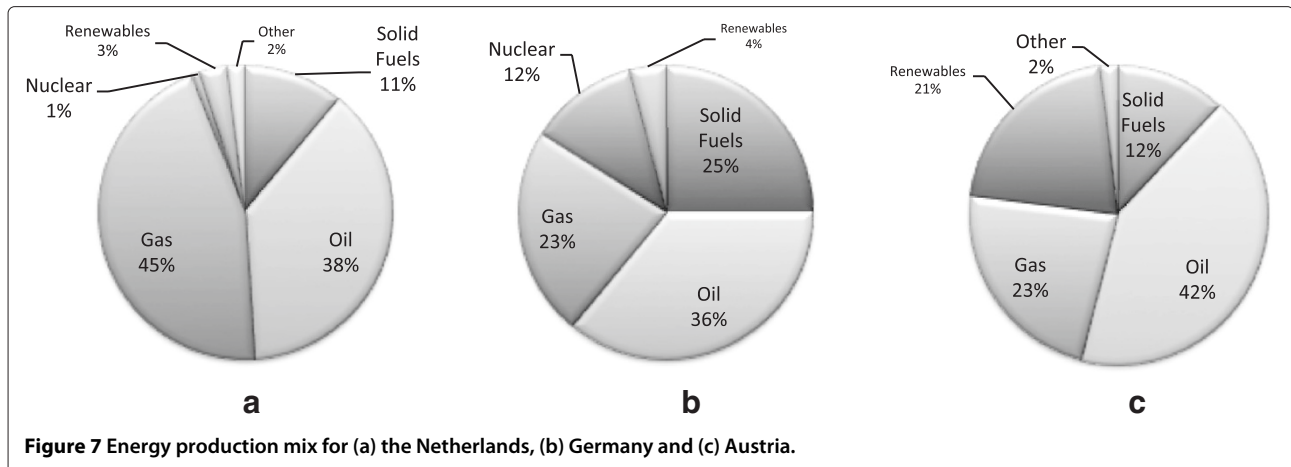
Tirol has 4 hops, then $E_{network} = 0.0014 kWh/GByte$ for an Internet connection and $E_{network} = 0.00066 kWh/GByte$ for a light path connection. For $PUE_{network}$ we use a default value of 2.2 (a value based on a recent survey [2], where we assume that more effort is put in data center equipment than in scattered network equipment), while for $X_{network}$ we use an estimate based on the shortest geographical paths between the countries and the information on the typical energy sources used in the countries crossed. In our example, the shortest path long distance network will most probably traverse the following three countries: the Netherlands, Germany and Austria. From data published by the European Commission [9-11] the energy production in the Netherlands, Germany and Austria is composed by the mixes depicted in Figure 7.

From these mixes we derive a mean value for the emission cost in gr. CO_2/kWh . For instance Germany use 36% Crude Oil (640 gr. CO_2/kWh), 25% Solid fuels (pulverized coal 870 gr. CO_2/kWh), 23% gas (380 gr. CO_2/kWh), 12% nuclear (66 gr. CO_2/kWh) and 4% renewable (30 gr. CO_2/kWh), arriving at a mean value $X_{network}$ Germany = 549 gr. CO_2/kWh . In the same way $X_{network}$ for the Netherlands = 520 gr. CO_2/kWh and $X_{network}$ for Austria = 474 gr. CO_2/kWh . The distance from say Amsterdam to Tirol is 980 km, of which 120 km in the Netherlands, 600 km in Germany, and about 260 km in Austria, or 12%, 62% and 26% respectively. So, these numbers give an estimate for the transport network $X_{network} = 0.12 \cdot 520 + 0.62 \cdot 549 + 0.26 \cdot 474 = 526$ gr. CO_2/kWh . Imagine a user submits a task needing a lot of experimental data, say $N_{in} = 10$ GByte, and producing $N_{out} = 2$ GByte of graphical data during 0.12 CPU core hours. The scheduler will respond to the user with detailed information it based its decision upon. Figure 8 shows the output the scheduler provided to the user.

In Figure 8 we see also values associated to the energy production in the country of the data center. Models used are not discussed in this paper, but can be retrieved from a report [4]. The contribution of the LAN of the local data center and of the network, occurring on the right hand side of Equation 15, due to the transport of the input data, turn out to be a considerable part of the total energy consumption. This contribution will be even higher if an Internet connection was chosen, that due to the relative high power consumption of the routers in the network path. If the user knows how the computation and its output data scale with the amount of input data, Equation 15 can be applied on a range of input data to see how the cost of the different components scale.

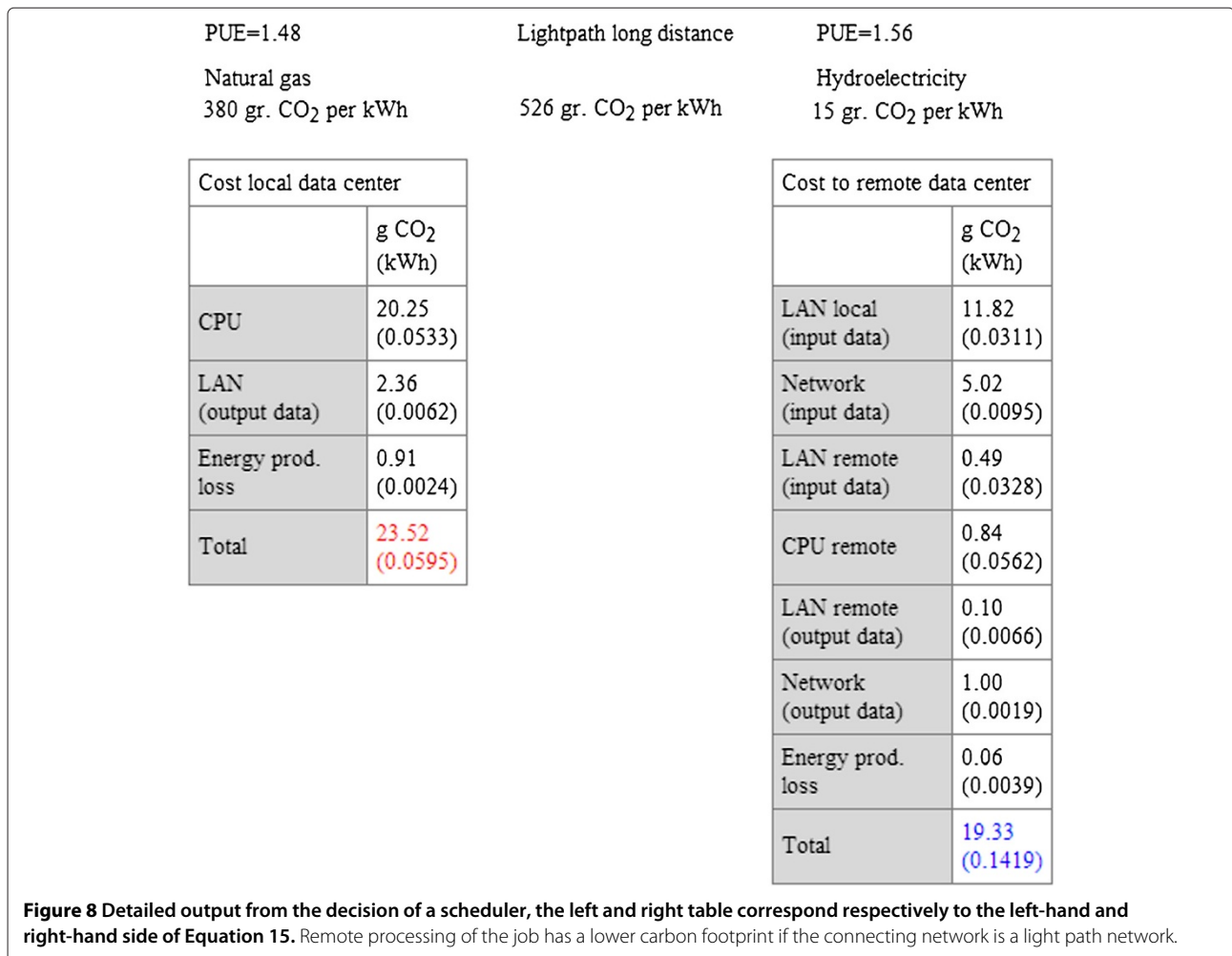
Data ranges and complexities

We introduce the complexity of a task where both the computation time and the output data scale with the input data, and define $T_{processing} = f(x)$ and $N_{out} = g(x)$ with



$x = N_{in}$. For a task with processing time and output data both scaling linearly with the input data, $O(x)$, we have $f(x) = f_1 \cdot x + f_0$ and $g(x) = g_1 \cdot x + g_0$. For a task exhibiting a processing time scaling quadratically, $O(x^2)$, and output scaling linearly, $O(x)$, we have $f(x) = f_2 \cdot x^2 + f_1 \cdot x + f_0$ and $g(x) = g_1 \cdot x + g_0$. In case the amount of input data x is

specified or expressed as a range, *i.e.*, $x \in [X_0, X_1]$, $X_0 > 0$, and the complexity of the job is specified, *i.e.*, $f(x)$ and $g(x)$ are specified, Equation 15 will decide whether local or remote processing is preferable for each $x \in [X_0, X_1]$. With these definitions we can facilitate a user or the operators of a data center in their choices of task placement with more



flexible parameters. The framework has a web calculator which allows data ranges as input for the amount of input data of a task and complexity formulas for the CPU processing time and the amount of output data as a function of the input data.

Web calculator

The web calculator [19], facilitates a user to study the output from the scheduler on submitting a task, and also to survey for which amount of input data decisions may alter. As an independent tool the user should supply all the data. Operators of a data center may use data from a knowledge base. We will introduce the web calculator according to the example used so far. Figure 9 shows the web calculator input page. The amount of input data is expressed as a range, [5, 15] GByte, and the CPU processing time exhibits a linear complexity, $O(x)$, on the amount of input data, $0.012 \cdot x$, where x refers to a value in the input range. The output data also shows a linear complexity, $0.2 \cdot x$. So we assume that computation time and amount of output data is negligible small if no input data is present ($f_0 = g_0 = 0$). For $x = 10$ GByte we have CPU time equals $0.012 \cdot 10 = 0.12$ core hours and output data equals $0.2 \cdot 10 = 2$ GByte, values used above. In case a

range is defined as input the calculator responds with a plot, Figure 10, and table output for the largest value of the range, see Figure 11.

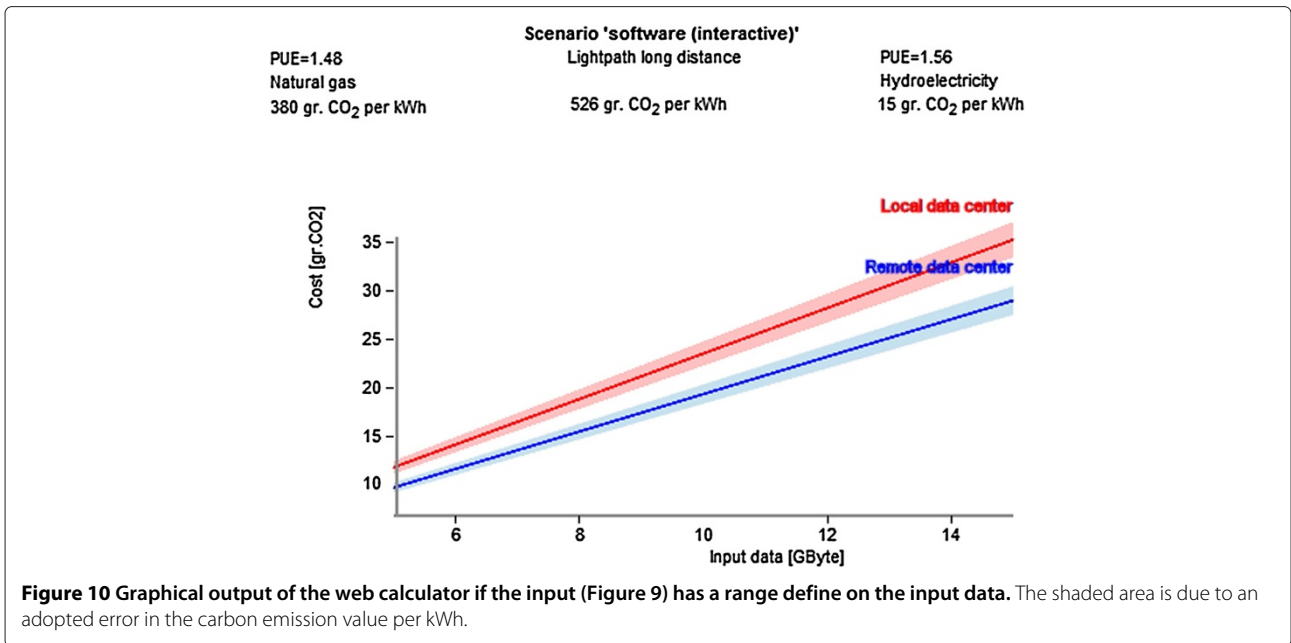
An operator might use the web calculator to study what happens if the light path long distance transport connection is not available and an Internet long distance connection is the only option. If he keeps all input the same except for the connecting transport network, and choose Internet long distance instead of light path long distance, he notices from the output, Figure 12 and 13, that the decision changes. The Internet long distance transport network spoils the greener processing advantage of the remote data center.

For quadratic behavior of the computation time it turns out that it becomes profitable to do the computation at a cleaner remote data center for even modest complexity values. This is due to the fact that the power consumption of computation nodes is relatively high. We saw that there is a difference if one compares Internet with dedicated light path connections due to the power consumption of routers in the former. This becomes clear if we transform Equation 15 into a decision boundary, i.e. substituting an equal sign for the greater sign in the formula.

The screenshot shows a web calculator interface with the following sections and fields:

- Choose a service scenario**: dropdown menu with "software (interactive)" selected.
- extra input scenario 'software (interactive)'**:
 - Input data: [5,15] GByte
 - CPU process. time: $0.016 \cdot x$ (core)hours
 - Output data: $0.2 \cdot x$ GByte
- Transport network between local and remote data center**: dropdown menu with "Lightpath long distance" selected.
- PUE**:
 - Local data center: 1.48
 - Remote data center: 1.56
 - Transport network: 1.8
- Emission energy production [gr CO₂/kWh]**:
 - Local data center:
 - X: Natural gas (380)
 - Location energy production: Local
 - Remote data center:
 - X: Hydroelectricity (15)
 - Location energy production: Local
- Transport network**:
 - X: Crude oil (640)
 - 526
- Calculate cost in gr CO₂**: button

Figure 9 Web calculator for a user or operator to decide whether a task can be greener performed at a remote data center instead of at his local data center. Input data is defined as a range, output data and CPU processing time are defined as complexity formulas on the input data range (the symbol \$0 refers to a value in the input range).

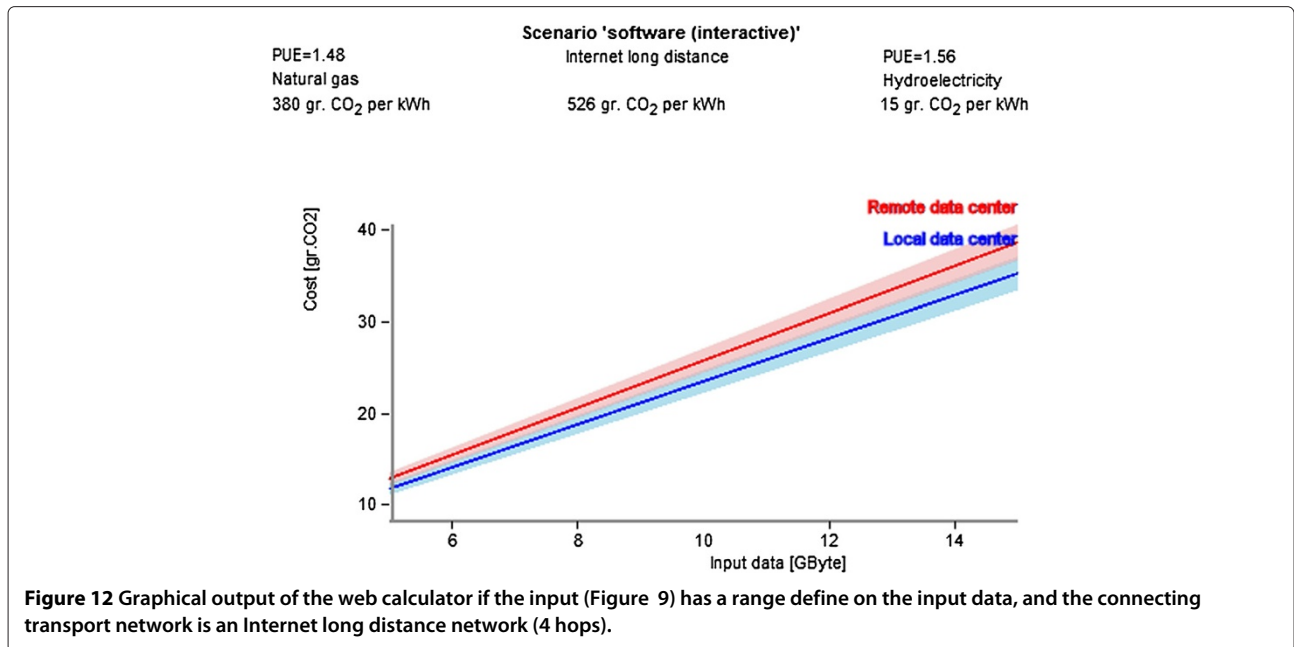


Scenario 'software (interactive)'

PUE=1.48 Natural gas 380 gr. CO ₂ per kWh	Lightpath long distance 526 gr. CO ₂ per kWh	PUE=1.56 Hydroelectricity 15 gr. CO ₂ per kWh
--	--	--

Cost local data center		Cost to remote data center	
	g CO ₂ (kWh)		g CO ₂ (kWh)
CPU	30.37 (0.0799)	LAN local (input data)	17.73 (0.0467)
LAN (output data)	3.55 (0.0093)	Network (input data)	7.52 (0.0143)
Energy prod. loss	1.37 (0.0036)	LAN remote (input data)	0.74 (0.0492)
Total	35.29 (0.0893)	CPU remote	1.26 (0.0842)
		LAN remote (output data)	0.15 (0.0098)
		Network (output data)	1.50 (0.0029)
		Energy prod. loss	0.09 (0.0058)
		Total	29.00 (0.2129)

Figure 11 Values corresponding with the maximum value of the input range [5,15] GByte for web calculator input of Figure 9.



Scenario 'software (interactive)'

Internet long distance

PUE=1.48 Natural gas 380 gr. CO ₂ per kWh	526 gr. CO ₂ per kWh	PUE=1.56 Hydroelectricity 15 gr. CO ₂ per kWh
--	---------------------------------	--

Cost local data center		Cost to remote data center	
	g CO ₂ (kWh)		g CO ₂ (kWh)
CPU	30.37 (0.0799)	LAN local (input data)	17.73 (0.0467)
LAN (output data)	3.55 (0.0093)	Network (input data)	15.60 (0.0297)
Energy prod. loss	1.37 (0.0036)	LAN remote (input data)	0.74 (0.0492)
Total	35.29 (0.0893)	CPU remote	1.26 (0.0842)
		LAN remote (output data)	0.15 (0.0098)
		Network (output data)	3.12 (0.0059)
		Energy prod. loss	0.09 (0.0058)
		Total	38.69 (0.2313)

Figure 13 Values corresponding with the maximum value of the input range [5,15] GByte for web calculator input(Figure 9), and the connecting transport network is an Internet long distance network (4 hops).

If we assume linear complexity for input and computation time, where we took $N_{out} = g_1 \cdot x$ and CPU processing time is $f_1 \cdot x$, with $x = N_{in}$, the decision boundary becomes a function of g_1 and f_1 , because x cancels out. The result is then visible in Figure 14, with two decision boundaries, $f_1 = 1.43 \cdot 10^{-2} + 4.24 \cdot 10^{-3}g_1$ for Internet and $f_1 = 9.56 \cdot 10^{-3} - 5.28 \cdot 10^{-4}g_1$ for light path. We see three regions corresponding to different choices of task location. In region 1 the task should be performed locally, independently of the type of transport network; in region 2 the task can be performed remotely provided that the connection is a light path; in region 3 the task should be done remotely for both types of transport networks. Values of the example chosen above, $f_1 = 0.012$ and $g_1 = 0.2$ give a point in region 2, a different decision for light path and Internet long distance transport network.

Discussion

As we had foreseen in the Introduction the PUE of two data centers, and even their power sources, cannot be the only guiding criteria in choosing the location of a computation or of data storage task. In case the transport network between them is powered by dirtier energy than both data centers are powered with, the contribution of the network to the total cost in gr. CO_2 for moving data can be significant. This mostly is the case if the data traverses the Internet, due to the relatively high power consumption of routers. Light path connections are preferable over Internet connections, but light path connections are dedicated connections that require a more complex setup procedure and sometimes might not be available to a user. For large input data sets and linear behavior of the computation time on the input data, it might be better to do the calculation locally, if the connecting network is Internet.

The same situation may be reversed in case the computation time shows a quadratic dependency on the input data. In that case the contribution of a dirty network becomes less prominent provided the data produced by the computation is limited and does not need to be transferred back to the user. Altogether this means that for realistic large processing, there is not one choice that can be made that is “always best” in terms of energy use and associated emissions.

Conclusions and future work

We have presented in this article a decision framework to allow users and data center operators to decide where to place an application in order to minimize the total CO_2 emitted in the process. We have shown that, if one assumes that the two data centers being considered have the same architecture and internal structure but different PUE, the network connection between them can play a significant role for the final selection of the site in which to compute or store data. Our framework depends not only on the models for the networks, which can be enhanced if one wishes, but also depends on the contents of the knowledge base it can draw upon. In the work presented here we used the energy data published by the EU and data of some European continental data centers. There are improvements we intend to include in our framework in order to obtain even more realistic carbon footprint information. For data centers that are only reachable by crossing seas, the network model should be enhanced by models of sea cables. Another aspect connected with the network topology used in the models is the knowledge of the exact numbers of hops between two locations. For this, we would like to use a detailed map of the networks for different countries. Our first step in this direction will

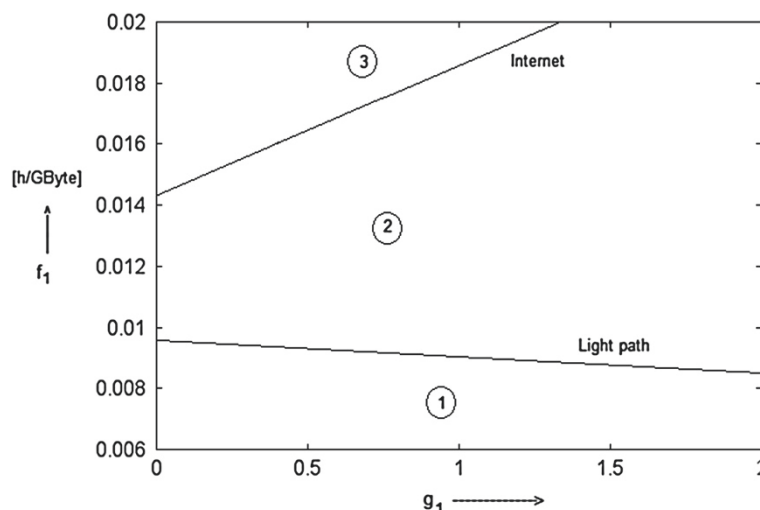


Figure 14 Decision boundaries according to Equation 15 for Internet and light path connections with 4 hops.

be to fill the knowledge base with detailed information of the transport topologies used between higher education and research data centers in the Netherlands, which are connected by the Surfnet network [20].

Abbreviations

CLF: Cooling load factor; PLF: Power load factor; PUE: Power usage effectiveness; UPS: Uninterruptible power supply; PDU: Power distribution unit; kWh: Kilo watt hour.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

AT and PG have researched the decision models and framework; AT has also developed the web portal. MM has revised the manuscript and contributed to its final version. AO has provided useful insight and feedback during the models definition. All authors read and approved the final manuscript.

Authors' information

Marc X. Makkes currently pursues a PhD degree at the University of Amsterdam and works as a researcher at TNO. His research interest include distributed computing, control and information theory. Arie Taal received his Ph.D. in nuclear science at the University of Delft in 1989. Currently, he is a part-time researcher at the University of Amsterdam in the field of network engineering. Anwar Osseyran is CEO with many years of multidisciplinary management experience in various areas of ICT including manufacturing, information management, health informatics, high performance computing, industrial automation and computer hardware and software. His current position (since 2001) is the MD of SURFSara in Amsterdam. Dr. Paola Grosso is assistant professor in the SNE group leading the activities in the field of optical networking, distributed infrastructure information modeling and GreenIT activities. She is PI in the UVA activities in GigaPort Research on Network projects to develop network models and control plane for lambda networks and topology handling. Her research interests are smart and sustainable cyber infrastructures. Therefore she investigates green ICT, provisioning and design of hybrid networks for lambda services and development of information models for hybrid multi-domain multi-layer networks. She participated in the EU NOVI project leading the information modeling workpackage and co-chaired the NML-WG (Network Markup Language Working Group). She is currently involved in the GreenClouds and Green Software projects. <http://www.science.uva.nl/~grosso>.

Acknowledgements

This research was in its initial phase sponsored by SURF and AgentschapNL during the Bits-Nets-Energy project. Further funding has been provided by the Dutch national research program COMMIT and by SURFnet via its GIGAPORT project.

Author details

¹University of Amsterdam, Kruislaan 904, Amsterdam, The Netherlands. ²TNO Information and Communication Technology, Eemsgolaan 3, Groningen, The Netherlands. ³SURFSara, Science Park 140, Amsterdam, The Netherlands.

Received: 14 July 2013 Accepted: 30 November 2013

Published: 17 December 2013

References

1. Stansberry M, Kundritzki J (2012) Uptime institute 2012 data center industry survey. [online] <http://uptimeinstitute.com>. Accessed 13 Dec 2013
2. [online] <http://www.thegreenitreview.com/2013/05/european-data-centres-are-less-energy.html>. Accessed 13 Dec 2013
3. Brown DJ, Reams C (2010) Toward energy-efficient computing. *Commun ACM* 53(3): 50–58
4. Taal A, Grosso P, Bomhof F (2013) Transporting bits or transporting energy: does it matter? [online] <http://www.surf.nl/en/knowledge-and-innovation/knowledge-base/2013/research-report-transporting-bits-or-transporting-energy-does-it-matter.html>. Accessed 13 Dec 2013
5. Taal A, Drupsteen D, Makkes M, Grosso P (2014) Storage to Energy: modeling the carbon emission of storage task offloading between data centers. In: *IEEE Consumer Communications and Networking Conference (CCNC)*. IEEE, Las Vegas
6. IEA (2011) CO2 emissions from fuel combustion—highlights. Paris. [online] <http://www.iea.org/co2highlights/co2highlights.pdf>. Accessed 13 Dec 2013
7. [online] http://en.wikipedia.org/wiki/Emission_intensity. Accessed 13 Dec 2013
8. Sovacool BK (2008) Valuing the greenhouse gas emissions from nuclear power: a critical survey. *Energy Policy* 36(8): 2950–2963
9. [online] http://ec.europa.eu/energy/energy_policy/doc/factsheets/mix/mix_nl_en.pdf. Accessed 13 Dec 2013
10. [online] http://ec.europa.eu/energy/energy_policy/doc/factsheets/mix/mix_de_en.pdf. Accessed 13 Dec 2013
11. [online] http://ec.europa.eu/energy/energy_policy/doc/factsheets/mix/mix_at_en.pdf. Accessed 13 Dec 2013
12. Rogers O, Cliff D (2012) Options, forwards and provision-point contracts in improving cloud infrastructure utilisation. *J Cloud Comput* 1: 1–22
13. Qureshi A, Weber R, Balakrishnan H, Guttaj J, Maggs B (2009) Cutting the electric bill for internet-scale systems. *ACM SIGCOMM Comput Commun Rev* 39(4): 123–134
14. Rao L, Liu X, Xie L, Liu W (2010) Minimizing electricity cost: optimization of distributed internet data centers in a multi-electricity-market environment. In: *INFOCOM, 2010 Proceedings IEEE*. IEEE, San Diego, pp 1–9
15. Kant K (2009) Data center evolution: a tutorial on state of the art, issues, and challenges. *Comput Netw* 53(17): 2939–2965
16. Baliga J, Ayre R, Hinton K, Tucker RS (2011) Green cloud computing: Balancing energy in processing, storage, and transport. *Proc IEEE* 99: 149–167
17. Tucker RS (2008) Optical Packet-Switched WDM Networks—A Cost and Energy Perspective. In: *Optical Fiber Communication Conference, Optical Society of America*. IEEE, San Diego, pp 243–252
18. Baliga J, Ayre R, Hinton K, Tucker R (2011) Energy consumption in wired and wireless access networks. *Commun Mag IEEE* 49(6): 70–77
19. A calculator for a road to cleaner computing. [Online] <http://sne.science.uva.nl/bits2energy/index.html>. Accessed 13 Dec 2013
20. [online] http://www.surfnet.nl/en/Hybride_netwerk/SURFinternet/Pages/kaart.aspx#topologie. Accessed 13 Dec 2013

doi:10.1186/2192-113X-2-21

Cite this article as: Makkes et al.: A decision framework for placement of applications in clouds that minimizes their carbon footprint. *Journal of Cloud Computing: Advances, Systems and Applications* 2013 **2**:21.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com