



UvA-DARE (Digital Academic Repository)

Transporting Bits or Transporting Energy: Does it matter? A comparison of the sustainability of local and remote computing

Taal, A.; Grosso, P.; Bomhof, F.

Publication date

2013

Document Version

Final published version

[Link to publication](#)

Citation for published version (APA):

Taal, A., Grosso, P., & Bomhof, F. (2013). *Transporting Bits or Transporting Energy: Does it matter? A comparison of the sustainability of local and remote computing*. SURF. <https://www.surf.nl/binaries/content/assets/surf/en/knowledgebase/2013/Transporting+Bits+or+Transporting+Energy+-+does+it+matter+-+May2013.pdf>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, P.O. Box 19185, 1000 GD Amsterdam, The Netherlands. You will be contacted as soon as possible.

Transporting Bits or Transporting Energy: Does it matter?

A comparison of the sustainability of local and remote computing



Colophon

Transporting Bits or Transporting Energy: Does it matter?
A comparison of the sustainability of local and remote computing

A study commissioned by NL Agency.

SURF
Graadt van Roggenweg 340
P.O. box 2290
3500 GG Utrecht
The Netherlands

T +31 (0)30 234 66 00

F +31 (0)30 233 29 60



info@surf.nl

www.surf.nl

Authors

Arie Taal - *University of Amsterdam*

Paola Grosso - *University of Amsterdam*

Freek Bomhof - *TNO*

Contributors

Josco Kester - *TNO*

The project team:

Paul Dekkers - *SURFnet*

Freek Dijkstra - *SURFsara*

Jaak Vlasveld - *Green IT Amsterdam*

Gerard van Westrienen - *SURF, project leader*

The steering group and the sounding board group:

Marga Blom - *KPN*

Maurice Bouwhuis - *SURFsara*

Walter van Dijk - *SURFnet*

Hans Gankema - *Groningen University*

Frank Hartkamp - *NL Agency*

Marco Kappe - *Vancis*

Anwar Osseyran - *SURFsara*

John Post - *Green IT Amsterdam*

Rogier Spoor - *SURFnet*

Final editing

Cecilia Willems (Gretton & Willems Translations VOF)

May 2013

Although this publication was prepared with the greatest care, the publisher, authors, and editors cannot be held liable for any errors or incomplete information, nor can any rights be derived from the content of this publication.



This document is published under Creative Commons Licence Attribution 3.0 Netherlands.

www.creativecommons.org/licenses/by/3.0/nl

SURF is the collaborative organisation for higher education institutions and research institutes aimed at breakthrough innovations in ICT (www.surf.nl/en)

This publication is online available through www.surf.nl/en/publications

Content

Foreword	4
Executive summary	5
1 Introduction	6
1.1 Energy consumption versus carbon footprint	6
2 Model of Bits-Nets component	8
2.1 Efficiently a data centre uses its energy	8
2.2 The different data centre and network components used	9
3 Sustainability	15
3.1 Decision equation	15
4 Model of Energy-Bits component	17
4.1 Losses due to energy transport	17
4.2 CO ₂ emissions associated with grid losses	19
5 Model integration	20
6 Scenarios	21
6.1 Processing (CPU-intensive) and Software (interactive)	21
6.2 Data storage	28
7 Conclusions	31
7.1 Data storage	31
7.2 Recommendations	32
7.3 Future Research	32
Acknowledgements	33
Appendix A	34
A1 Power consumption of data centre and network components	34
A2 Adopted values	35
Appendix B	36
B1 Cable losses	36
Appendix C	38
C1 Special scenario inputs	38
Appendix D	39
D1 Providing application-specific input to the calculator	39
D2 Real-life application scenarios	39
References	41

Foreword

The global economic slowdown is forcing governments and organisations to reduce costs and avoid capital investment by consolidating their ICT infrastructures. The uncertainty about the future energy supply and the rising cost of electricity are also putting data centre energy consumption in the spotlight. Enterprises are becoming more cautious about building new local data centres, adding to the attraction of alternative models in the form of outsourcing, hosting, and ultimately cloud computing services.

The move to cloud computing is one of the most dramatic ICT trends of this decade. The market for cloud computing services has continued to expand during the recession despite the decline in economic activity in most of the world. According to Cisco¹, 'global data centre traffic is projected to quadruple from 2011 to 2016, with data centre traffic specifically in the cloud forecasted to grow 6-fold during that period'. The Dutch higher education and research community has formulated a 'cloud first'² principle as its point of departure. Cloud computing models – in all their diversity across public, private, sector and hybrid clouds – will be the predominant paradigm for the next generation of ICT services. We are, however, only at the beginning, as many cloud applications have yet to be developed. New metrics and new levels of transparency are required if the impact of clouds on sustainability is to be adequately assessed.

One of the major trades-offs to be made within an enterprise or even a sector cloud strategy is how to distribute data storage and data processing across the cloud. Consolidation of data centres within the cloud makes it possible to minimize the CO₂ footprint of the ICT infrastructure by deploying the greenest data centres and minimizing energy losses due to data transport. The ideal situation would be to deploy green data centres using renewable energy and serving mainly local users. But this is not always possible: not all data centres within a cloud are equally green; sustainable energy still constitutes a small fraction of the total energy produced; and data must still be moved back and forth within the clouds. Up to now, one specific aspect of cloud computing that has been overlooked almost entirely is the role of the network on the cloud carbon footprint. What is the carbon impact of data transport across the cloud, and how does the CO₂ footprint associated with transporting data to a remote data centre using sustainable energy compare to that associated with transporting sustainable energy to a local data centre within the cloud?

This report is the result of a unique and exploratory study that looks into this issue for the first time and comes up with very interesting, albeit preliminary results. Many countries are investing heavily in consolidating their ICT infrastructures and some countries, for example Iceland, Norway, Finland and Canada (British Columbia), are positioning themselves as the place of choice for remote data centres. But what is really the most economic and sustainable solution, and would the sustainability principle of 'produce and consume locally' also apply here in the Netherlands? The answer is not only important for individuals and organisations looking for the most sustainable solution for their computing activities, but also for policymakers looking for the best location for data centres in relation to smart grids, smart cities and sustainable energy.

We are therefore very pleased that NL Agency has commissioned this study to SURF. A broad team of dedicated experts has spent five months looking into this issue. I would like to thank them for their excellent work, especially the authors of the report: Arie Taal and Paola Grosso of the University of Amsterdam, and Freek Bomhof of TNO. Valuable contributions and comments have been made by the other members of the project team: Paul Dekkers (SURFnet), Freek Dijkstra (SURFsara), Josco Kester (TNO), Jaak Vlasveld (Green IT Amsterdam) and the project leader Gerard van Westrienen (SURF). We were also very pleased with the contributions we received from the members of the steering group and the sounding board group: Marga Blom (KPN), Maurice Bouwhuis (SURFsara), Walter van Dijk (SURFnet), Hans Gankema (Groningen University), Frank Hartkamp (NL Agency), Marco Kappe (Vancis), John Post (Green IT Amsterdam) and Rogier Spoor (SURFnet).

We hope that the results of this study will give rise to further research, so that we all gain a better understanding of the sustainability effects of transporting bits versus transporting energy.

Anwar Osseyran

Chair of the Steering Group

Director of SURFsara

May 2013

¹ http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns1175/Cloud_Index_White_Paper.html

² <http://www.surf.nl/en/publicaties/Pages/IntothecloudwithSURF.aspx>

Executive summary

This document presents the results of a study of the sustainability of data management and data movement between data centres, the ultimate goal being to minimize the CO₂ footprint.

We considered two dimensions:

- the 'bit-to-energy' dimension, with data being moved to 'greener' remote data centres;
- the 'energy-to-bits' dimension, with 'greener' energy being moved to the data centre where the data resides.

We also focused on two basic questions:

- *What are the sustainability effects of data transport over the data network? How much energy is required and what is the CO₂ footprint?*
- *What are the sustainability effects of energy transport? When is it suitable to acquire green energy from elsewhere?*

We have developed basic energy models that allow us to calculate the carbon footprint under various representative data movement and data processing scenarios. Based on this model, we have been able to derive some general guidelines that can help end users and data centre operators choose the more sustainable solution to the 'bits-to-energy' and 'energy-to-bits' dilemma. We have also calculated the carbon footprint in several representative scenarios.

Our first general result is that the energy required to transport the data and the energy required for energy transport can be considerable and cannot be neglected in evaluating the overall level of sustainability in the various scenarios. Depending on the energy source, the sustainability effect of network use can have a significant impact on total CO₂ emissions.

We have observed that such application features as data processing times or required storage play a significant role in the final outcomes. Despite the need to consider the carbon footprint on case-by-case basis, we have been able to derive a set of general guidelines (see section 7.2). One important conclusion is that in many data scenarios where the local data centre can import cleaner energy from elsewhere, the best course of action is to keep the data local and perform calculations locally.

1 Introduction

Sustainability is one of the priorities of SURF and its connected higher education organisations. Such organizations should be supported in their aim of offering *greener* ICT services and applications.

The need for such support and the components and services that can provide it were investigated in an initial study carried out by TNO and commissioned by NL Agency.³ Given that storage and computing in the cloud often happen far away from the users, one could conclude that the most sustainable solution is to locate cloud data centres where renewable energy is locally produced. This intuition needs to be supported by hard figures, however. We therefore set out to answer two fundamental questions:

- What are the sustainability effects of data transport over the data network? How much energy is required and what is the CO₂ footprint?
- What are the sustainability effects of energy transport? When is it appropriate to acquire green energy from elsewhere?

We call the former the 'bits-to-nets-to-energy' case and the latter the 'energy-to-bits' case.

We believe several parties can benefit from our results:

- users who want to choose the 'greenest' solution for storing and processing their data;
- institutions that are trying to determine whether to maintain a local data centre or to offload their data to a community cloud;
- data centres in the Netherlands that need or want to specialize in areas where they can be internationally competitive from a sustainability point of view.

Ultimately, we tried to determine when it is more energy efficient to move data with accompanying computation from a local data centre to another remote data centre, rather than move 'greener' energy to the local centre where the data resides and the computing will be performed. This helped us differentiate between applications when optimizing the underlying data centre, network and energy infrastructure.

1.1 Energy consumption versus carbon footprint

When looking for the 'greenest' way to perform computationally intensive tasks, a user may have different options:

- A. Perform all computing locally, powered by locally produced energy.
- B. Perform all computing at a remote location, powered by energy that is produced sustainably at the remote location.
- C. Perform all computing locally, powered by energy that is produced sustainably at a remote location.

We can look at the *attractiveness* of each one of these scenarios from two different perspectives. One involves considering the total energy required and the other is to look at the CO₂ emissions. In essence, one can focus either on energy efficiency or sustainability, with different outcomes. Our work focuses on sustainability.

When considering the energy required, we see the following.

- In scenario A, neither data nor energy has to be transported. This solution is attractive when a lot of energy is required to either transport the data or the energy.
- Scenario B can be guided by the consideration that data processing is carried out (remotely) using sustainably generated energy, and that this energy does not have to be transported. This scenario is therefore attractive when data transport does not involve a lot of energy, and energy transport would be inefficient.
- In scenario C, the considerations are that no data has to be transported, and that the energy is generated sustainably (although at a remote location). This scenario is attractive when data transport requires a lot of energy and energy transport much less.

When these scenarios are evaluated not in terms of energy but in terms of CO₂ emissions, however, the answer to the problem changes.

In that case, scenario C would probably come out as attractive, because it does not require data transport (hence, no energy or carbon emissions due to data transport). Energy would still have to be transported, however; even if this energy were generated without any emissions, the losses due to its transport would lead to an emissions footprint because they need to be compensated by the Transmission System Operators using their local energy mixes.

In this report, we compare not only carbon emissions but also energy losses (irrespective of the associated footprint). This is because:

³ ['Cloud Computing: Grey or Green? On the energy-efficiency and sustainability of Infrastructure-as-a-Service'](#) – report for NL Agency, SURF and Green IT Amsterdam, by TNO; March 2012.

- energy, whether it is generated sustainably or not, still has to be paid for and this plays a role in the decision;
- only a limited number of renewable sources can be exploited cheaply;
- the installations that generate the electrical energy produced emissions during *their* manufacture and construction.

This is in line with the *Trias Energetica*, which states as follows.

1. Reduce energy consumption as much as possible.
2. Use renewable energy whenever possible.
3. Use (remaining) fossil fuels as efficiently as possible in order to meet energy needs.

2 Model of Bits-Nets component

The decision to move data and computation to a remote data centre depends not only on the energy consumption of the local and remote data centres but also on the energy consumption of the data network used to transport the data. We therefore need a metric for both the energy consumption of the local and remote data centres and of the transport network. Furthermore, all metrics must yield values that can be compared to one another.

When deciding to move data and the accompanying computation from a local to a remote data centre, we have to define an energy consumption metric that accounts for both of the data centres and the transport network between them. This metric should allow us to calculate values for the following equation, which indicates when movement to a remote data centre is to be preferred above local processing of the data:

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

In the event of pure data storage and consumption, with no computation being performed on the data, we can use the follow decision equation:

$$\begin{aligned} & \text{Energy cost of local storage} + \text{Energy cost of local download} > \\ & \text{Energy cost of transport network} + \\ & \text{Energy cost of remote storage} + \text{Energy cost of remote download} \end{aligned} \quad (1b)$$

It is obvious that the terms on the right side of the decision equations must be expressed in the same units. As we will show in the following section, this has never been done before and most prior research presents the two components separately and not necessarily consistently.

Similarly, when deciding to move energy and leave the data at the local data centre, we need to define a metric that permits the following decision equation:

$$\begin{aligned} & \text{Energy cost of processing with locally produced energy} > \\ & \text{Energy cost of processing with remotely produced energy} + \\ & \text{Energy cost of transporting remotely produced energy} \end{aligned} \quad (2)$$

An analogous decision equation holds for data storage and consumption.

In the following sections we will introduce the different requirements that a metric for energy consumption should meet in order to ensure that decision equations 1 and 2 are applicable:

- how efficiently a data centre uses its energy (par 2.1);
- the different data centre and network components used (par 2.2);
- how the energy is provided and how it is transported (Ch.4).

2.1 Efficiently a data centre uses its energy

The measure most commonly used to rate the energy efficiency of data centres is the power usage effectiveness (PUE). The PUE is expressed as the ratio of the total power consumption of a data centre (P_{IN}) to the total power consumption of IT equipment such as storage devices, servers and routers (P_{IT}).

$$PUE = \frac{P_{IN}}{P_{IT}} = CLF + PLF + 1, 1 < PUE < \infty \quad (3)$$

In calculating their PUE, data centres use two terms: CLF and PLF. CLF represents the cooling load factor normalized to the IT load (losses associated with chillers, pumps, air conditioners) and PLF represents the power load factor normalized to IT load (losses associated with switchgear, UPS, PDU).

A recent survey[1] by the Uptime Institute conducted in 2012 collected information about the PUEs for data centres all around the world. Figure 1 shows the distribution of PUEs collected:

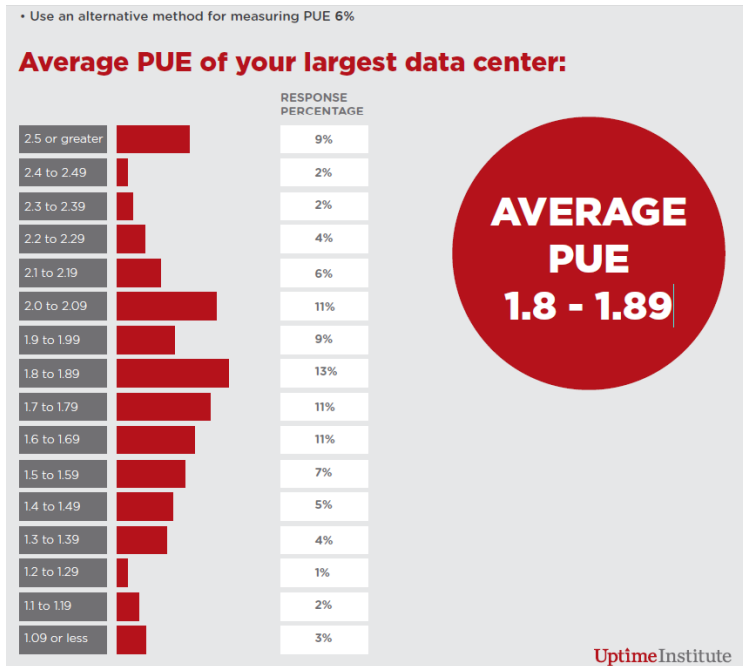


Figure 1 PUE distribution according to Uptime Institute survey.

In the same survey, it is interesting to see how the different data centres measure this PUE value. First of all, 29% of the data centres do not measure the PUE at all. Data centres can furthermore be divided into different categories, depending on how and how frequently they measure the PUE. If we want to take these differences into account, we need to introduce a measure for the trustworthiness of the PUE value that a data centre publishes. However, detailed information about how the PUE is measured and how frequently this takes place is not always available. We must therefore apply a common error value for the PUE.

If we have a measure for the costs $E_{\text{local processing}}$, $E_{\text{remote processing}}$, $E_{\text{local storage}}$ and $E_{\text{remote storage}}$ in Joules, then the PUE may be taken into account when expressing the total cost in Joules (i.e. including cooling, etc.):

$$PUE_{\text{local data centre}} \cdot E_{\text{local processing}} \quad (4a)$$

$$PUE_{\text{remote data centre}} \cdot E_{\text{remote processing}} \quad (4b)$$

$$PUE_{\text{local data centre}} \cdot E_{\text{local storage}} \quad (4c)$$

$$PUE_{\text{remote data centre}} \cdot E_{\text{remote storage}} \quad (4d)$$

2.2 The different data centre and network components used

Energy consumption in today's telecommunications infrastructure is dominated by energy in switches and routers in the metro and core networks, and by the access network.[3][4] This energy is usually expressed in *Joules per bit (J/b)*.

Figure 2 provides a quantitative impression of the energy dissipated by the different components in today's telecommunications infrastructure.[4]

Figure 2 shows that optical switches have lower energy dissipation than Ethernet switches. One important conclusion reached in a recent study by Tucker [4] 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.

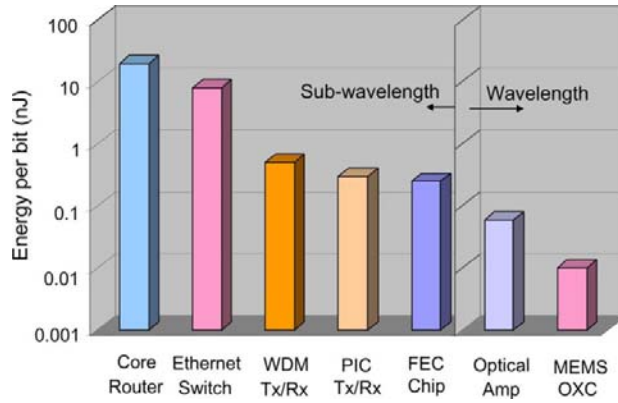


Figure 2 Energy per bit of different types of network equipment [4]

Optical communication systems consistently achieve better energy consumption at rapid rates. Tucker shows that the energy per bit for transatlantic transmission systems has fallen exponentially at an annual improvement rate of around 20%. That is around 15% for terrestrial optical transport systems. For a representative 1000-km repeated terrestrial system using 2010 generation technology, the energy per bit is 1.1nJ/b.

What kind of data networks should we consider in our Equation 1? We will restrict ourselves to a situation where the end user is connected directly to the data centre clouds/clusters via a corporate network.

The user (or a scheduling application on his/her behalf) must decide whether the data with the accompanying computation should remain at a data centre or be moved to another data centre. If he decides to move the data, it will be transported over a public data network, given that different data centres are usually geographically separate. When data traverses the Internet, it consumes energy. We can estimate how much it consumes by adding the energy contributions of the switches, amplifiers, transceivers, etc. that the bit traverses.

On both ends, at both the local and remote data centre, we have the local area network (LAN) of the data centre itself, which connects the data storage devices and servers to the outside world, i.e. the transport network. To keep calculations simple, we assume that the LAN has the same components as any data centre. Table 1 lists the typical equipment that data traverses in a data centre LAN.

Data centre LAN
Host (network interface)
2 × Switch
2 × Firewall
Switch
Router

Table 1 Components of a data centre LAN.

Table 1 brings us to the following equation for the energy consumption per bit for the data centre LAN:

$$E_{LAN_data_center} = \frac{PUE}{U} \cdot \left(\frac{P_{host}}{C_{host}} + \frac{3P_{switch}}{C_{switch}} + \frac{2P_{firewall}}{C_{firewall}} + \frac{P_{router}}{C_{router}} \right) \quad (5)$$

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 centre gateway router, respectively. C_{host} , C_{switch} , $C_{firewall}$, and C_{router} are the capacities of the corresponding equipment in bits per second. The factor U accounts for the utilization of the data network. Today's data networks typically operate at less than 50% utilization while still consuming almost 100% of maximum power (factor of 2). Cooling and other overheads are expressed by the PUE of the data centre.

Data transfers across a transport network can use two different types of connections: the regular Internet and dedicated circuits. The regular Internet is available to all users, while dedicated connections (*lightpaths*) basically are not. We include them in our model because they are offered by NRENs to academic users and for scientific applications and effectively represent a valid alternative to data transport via the Internet for our main user base. In both cases the data transfer can be over long or short distances. To account for this, we must distinguish between these different use cases. Figure 3 shows the data network building blocks we assume to be representative for Internet and light path networks.

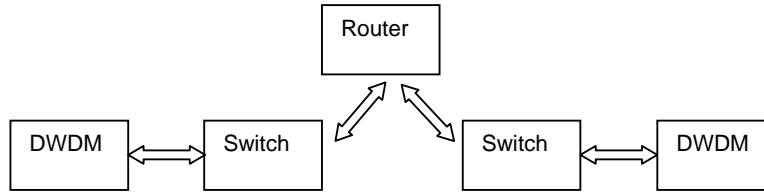


Figure 3a Network components in an Internet building block representing a hop.

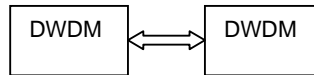


Figure 3b Network components in a lightpath building block representing a hop.

We can use these building blocks to build short- and long-distance networks. Multiple Internet building blocks are connected to one another via a switch, as are multiple lightpath building blocks. The entry and exit points for any kind of data network consist of a switch connected to a dense wavelength division multiplexing node (DWDM) (see Figure 4).

Switches are Internet switches and DWDM nodes are DWDM terminal nodes. We neglect the contribution of optical line amplifiers and regenerators in the long-distance networks as well as optical switches (OXC) (see Appendix A for our motivation).

We restrict ourselves to continental short- and long-distance data transport networks used by higher education and research institutes, where short refers to the data transport network in the Netherlands (SURFnet) and long refers to a network crossing the border of the Netherlands (GEANT) (routes via submarine cables are not considered in this study, in other words). Following Baliga et al.[3] we take a mean number of hops for each kind of network (Internet and lightpath), but additionally distinguish between short- and long-distance networks; see Figure 4a,b,c,d.

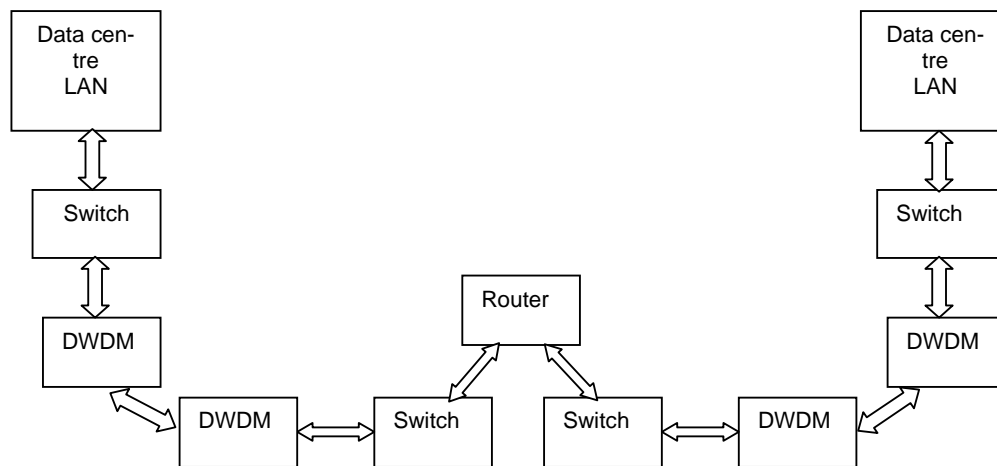


Figure 4a Short-distance Internet between two data centres.

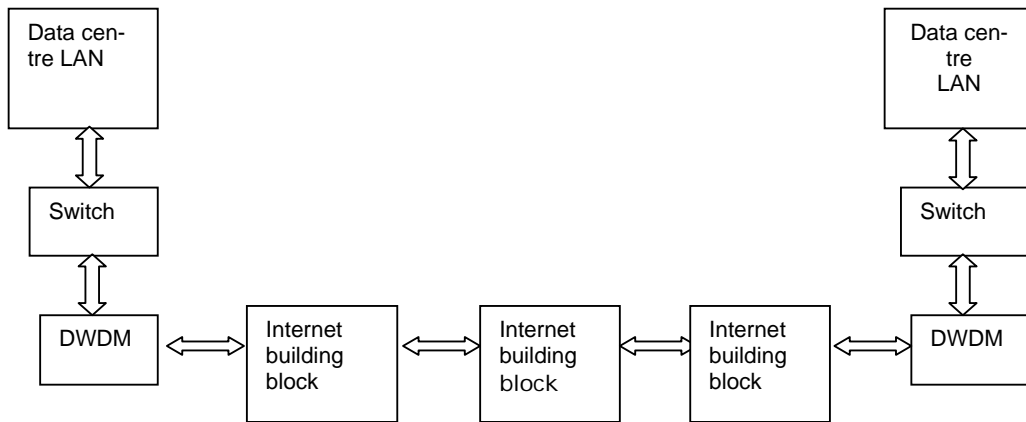


Figure 4b Long-distance Internet with 3 hops between two data centres.

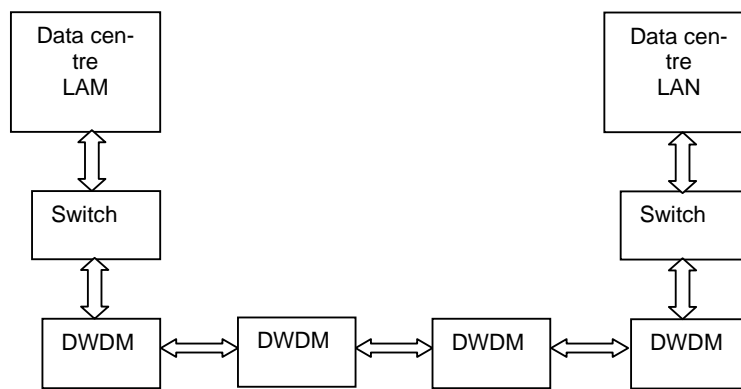


Figure 4c Short-distance lightpath between two data centres.

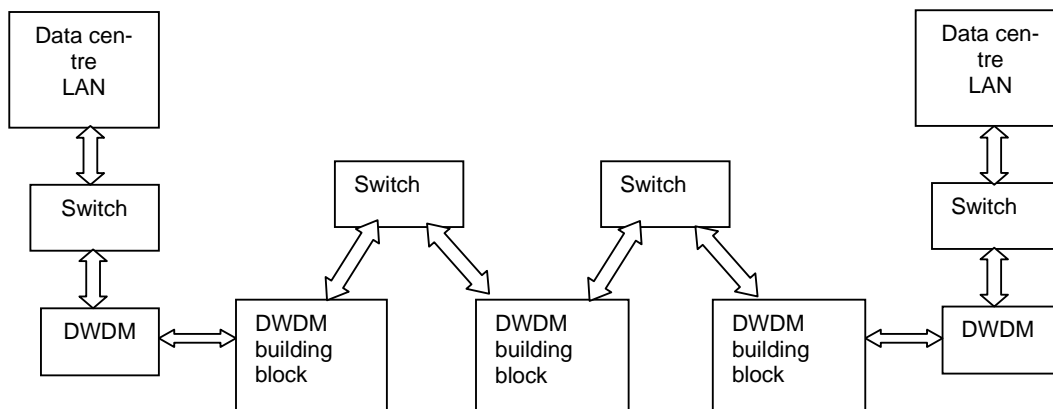


Figure 4d A long-distance lightpath with 3 hops between two data centres.

In each of the four transport cases, the correct energy needed to transport a bit contains three components:

- the energy contribution of the LAN at the local data centre;
- the energy contribution of the transport network;
- the energy contribution of the LAN at the remote data centre.

In order to apply Eq. 5 for the total energy consumption to move data, we need values for the different pieces of equipment that the data transverses. Based on the information in Appendix A and D, we will adopt the following typical values for the power per capacity (i.e. P/C) in kW/Gb/s of the devices listed in Table 1 and depicted in Figure 3:

Equipment	Power per capacity [kW/Gb/s]
Host CPU intensive	0.3550
Host data storage	0.2800
Router	0.0120
Ethernet switch	0.0230
Firewall	0.0160
DWDM terminal node	0.0034

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

Similar to a data centre LAN, we have four different equations for the energy cost of the transport network

$$\begin{aligned}
 &E_{transport-Internet-short} \\
 &E_{transport-lightpath-short} \\
 &E_{transport-Internet-long} \\
 &E_{transport-lightpath-long}
 \end{aligned}
 \tag{6}$$

With the help of Table 2 and the building blocks provided in Figure 3, we can compile an equation such as (5) for each transport network. We can also assign a PUE to a transport network.

Router energy consumption depends on the traffic load.[5] For those data networks that include routers, one could differentiate between the energy consumption during different load variations, each variation having a different value for the routers' power consumption. We will not differentiate between low traffic and peak traffic load.

In order to calculate the energy used for (hot) storage, we have defined a storage array network as depicted in Figure 5.

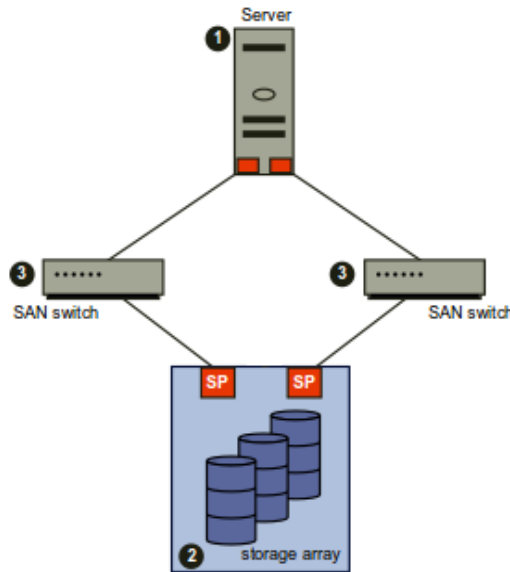


Figure 5 Storage Array Network.

The storage array network consists of a content server (1) and a storage array (2). Switches allow the server to be connected to different storage arrays. For hot data storage, the storage array may consist of a RAID system (Redundant Array of Independent Disks); for cold data, it may consist of a MAID system (Massive Array of Idle Disks).

The energy consumption of storing N GByte of data depends on the capacity of the disks in the storage array. Suppose the storage capacity of a single disk is S_{disk} GByte. Storing N GByte of data would then require $\lceil \frac{2N}{S_{disk}} \rceil$ disks, as there is a redundancy of 2 to store 1 GByte (i.e. 2 GByte of storage capacity is used).

The energy needed to keep N GByte of data for RT (*retention time*) hours in storage is:

$$E_{storage\ hot}(N, RT) = PUE \cdot f \cdot \left\lceil \frac{2N}{S_{disk}} \right\rceil \cdot P_{disk\ active} \cdot RT \text{ [kWh]} \quad (7)$$

$$E_{storage\ cold}(N, RT) = PUE \cdot f \cdot \left\lceil \frac{2N}{S_{disk}} \right\rceil \cdot P_{disk\ standby} \cdot RT \text{ [kWh]} \quad (8)$$

where:

- $P_{disk\ active}$ and $P_{disk\ standby}$ are the power consumption of a single disk in kW for a disk in active mode (hot storage) and a disk in standby mode (cold storage), respectively;
- f is the ratio between the real physical disks actually in use and the requested virtual disks. This value is always higher or equal to 1, which means that the actual energy needed for storage is higher. For simplicity's sake, we assume that $f=1$ in our calculations.

To account for the write and read activity of the data, which we take to be equal for hot and cold data, we derive, for the energy required:

$$E_{write}(N) = E_{read}(N) = \frac{PUE}{U} \cdot \frac{8N}{3600} \cdot \left(\frac{P_{content\ server}}{C_{content\ server}} + \frac{P_{switch}}{C_{switch}} \right) + PUE \cdot \frac{8N}{3600} \cdot \left(\frac{2 \cdot P_{disk\ active}}{C_{disk\ active}} \right) \text{ [kWh]} \quad (9)$$

where the factor 8 accounts for the number of bits per byte, as C is expressed in Gb/s.

For a RAID system, we assume 2.5" disks with a typical active power consumption of about 12 W per disk, so $P_{disk\ active} = 0.012$ kW and S_{disk} is 200 GBytes. For a 2.5" disk in a MAID system, we assume that only the electronics are on but that the disks are no longer spinning and the heads are unloaded, so $P_{disk\ standby} = 0.0002$ kW.

For the read-write capacity of a 2.5" disk in both storage systems, we assume 1 Gb/s.

3 Sustainability

We are interested in the sustainability aspects of the energy sources used in the data network and data centres, and in the subsequent CO₂ emissions. One way to consider this is to transpose energy costs in kWh into carbon emission cost effects. A kWh can be converted into grams of produced CO₂ [7] according to the following formula:

$$Q = 1KWh \sim X \text{ gr CO}_2 \quad (10)$$

The X values depend on the type of energy source, e.g. X = 870 for anthracite electricity production, and X = 370 for gas electricity production. Values for X are taken from different sources.[8,9,10]

3.1 Decision equation

If we know the amount of (input) data N [GByte] that will be transported through the data network connecting both data centres, we can transpose the energy cost in kWh into an equivalent carbon emission cost in terms of grams of CO₂ produced:

$$K_{LAN \text{ data center}} = X_{data \text{ center}} \frac{8N}{3600} E_{LAN \text{ data center}} \quad (11a)$$

$$K_{transport \text{ network}} = X_{transport \text{ network}} \frac{8N}{3600} E_{transport \text{ network}} \quad (11b)$$

with $E_{transport \text{ network}}$ in kWh/Gb (the factor 8 accounts for the translation of bytes into bits).

3.1.1 Decision equation for software (interactive)

We can transform Equation 1 into a decision equation for transporting data with accompanying computation to another data centre:

$$K_{processing \text{ local data center}} > K_{LAN \text{ local data center}} + K_{transport \text{ network}} + K_{LAN \text{ remote data center}} + K_{processing \text{ remote data center}} \quad (12)$$

If we are dealing with the output data of a computational task, we can assume that the party interested in the output data is located near the local data centre. Equation 12 becomes:

$$K_{processing \text{ local data center}} + K_{LAN \text{ local data center}}^* > K_{LAN \text{ local data center}} + K_{transport \text{ network}} + K_{LAN \text{ remote data center}} + K_{processing \text{ remote data center}} + K_{LAN \text{ remote data center}}^* + K_{transport \text{ network}}^* \quad (13)$$

where K^* is the contribution due to the output data.

3.1.2 Decision equation for data storage

To decide whether the movement of hot/cold data to another data centre is energetically preferable, a more complex decision is needed.

Suppose N GByte of data are present in the local data centre for RT (retention time) hours. Hot data that is stored will generally be used by users, so we assume that N^* GByte will be downloaded by users during the retention time.

The carbon emission cost for hot data in the local data centre becomes:

$$K_{data \text{ storage local data center}} = X_{local \text{ data center}} (E_{write \text{ local data center}}(N) + E_{storage \text{ hot local data center}}(N, RT) + E_{download \text{ local data center}}(N^*)) \quad (14)$$

where $E_{write \text{ local data center}}(N)$ is produced by Eq. 9, $E_{storage \text{ hot local data center}}(N, RT)$ by Eq. 7 and where $E_{download \text{ local data center}}(N^*)$ is a combination of Eq. 9 and Eq. 5 :

$$E_{download \text{ local data center}}(N^*) = E_{read \text{ local data center}}(N^*) + \frac{8N}{3600} E_{LAN \text{ local data center}}^4 \quad (15)$$

For cold data, an equation analogous to Eq. 14 applies, with N^* equal to zero.

⁴ According to Eq. 5, E_{LAN} contains a P_{host} , which term is already accounted for by $P_{content \text{ server}}$ in Eq. 9. Consequently, the term P_{host} should be skipped in Eq.5.

If the hot data is stored in a remote data centre, the energy cost is the sum of a number of terms:

- $E_{write\ local\ data\ centre}(N)$ as the data first resides at the local data centre
- $E_{storage\ hot\ local\ data\ centre}(N, RT^t)$, with RT^t being the time data is at the local data centre before it is moved, we will neglect this term
- $E_{read\ local\ data\ centre}(N)$
- $E_{LAN\ local\ data\ centre}(N)$
- $E_{transport\ network}(N)$
- $E_{LAN\ remote\ data\ centre}(N)$
- $E_{write\ remote\ data\ centre}(N)$
- $E_{storage\ hot\ remote\ data\ centre}(N, RT)$
- $E_{download\ remote\ data\ centre}(N^*)$

The last term, $E_{download\ remote\ data\ centre}(N^*)$, also contains the energy cost of transporting N^* GByte over the transport network. Each term is multiplied by the appropriate value of X , i.e. the local data centre terms, the transport network terms and the remote data centre terms.

4 Model of Energy-Bits component

4.1 Losses due to energy transport

In the 'energy-to-bits' scenario, storage and computation are kept local. They are powered by renewable energy imported from abroad. To calculate the energy use and the CO₂ emissions for this scenario, we need to know not only the CO₂ emissions for remote renewable generation but also the energy losses for energy transport and its associated CO₂ emissions.

If energy is imported from abroad, we distinguish between two modes of transport to the data centre's home country: over land, through EHV AC (Extra High Voltage Alternating Current) overhead lines; or by sea, through an HVDC (High Voltage Direct Current) submarine cable. The import of hydro-electricity from the Tyrolean region of Austria is an example of power being transported over land through EHV AC overhead lines. The import of hydro-electricity from southern Norway is an example of power being transported through an HVDC submarine cable.

Appendix B gives a derivation for estimating the order of magnitude of the total energy losses for energy transport in three different cases:

1) *import of hydro-electricity from Tyrol in Austria:*

$$L_{tot_remote_AT} = 0.156 E_{DC}$$

or 15% of the energy consumption of the data centre.

2) *import of hydro-electricity from southern Norway:*

$$L_{tot_remote_NO} = 0.107 E_{DC}$$

or 11% of the energy consumption of the data centre.

3) *use of renewable energy generated in the Netherlands:*

$$L_{tot_NL} = 0.04 E_{DC}$$

or 4% of the energy consumption of the data centre.

Taking these energy losses into account, Equations 12 and 13 can be extended by the appropriate losses, with

$$E_{DC} = E_{processing\ local\ data\ center} + E_{LAN\ local\ data\ center}^* \quad (14)$$

for the local data centre, and

$$E_{DC} = E_{LAN\ remote\ data\ center} + E_{processing\ remote\ data\ center} + E_{LAN\ remote\ data\ center}^* \quad (15)$$

for the remote data centre.

It is difficult to calculate an exact value for the size of these energy losses. This is because the energy loss for a particular additional power flow across a connection depends on the size and direction of the existing power flow on that connection. The resistive energy loss caused by a particular amount of additional power flow is proportional to the size of the existing power flow.

Furthermore, the value of the energy loss can become negative. This is the case when the energy transport for the data centre is in the opposite direction of the existing power flow. The total amount of power transported is then reduced, and so is the amount of energy lost. The combined effect of these two phenomena is shown in Figure 6 below. (We can see that the energy loss L becomes negative if the existing power flow P_{ex} is the opposite of the power transported.) This situation can occur at certain times of day or on parts of the energy transport flow route.

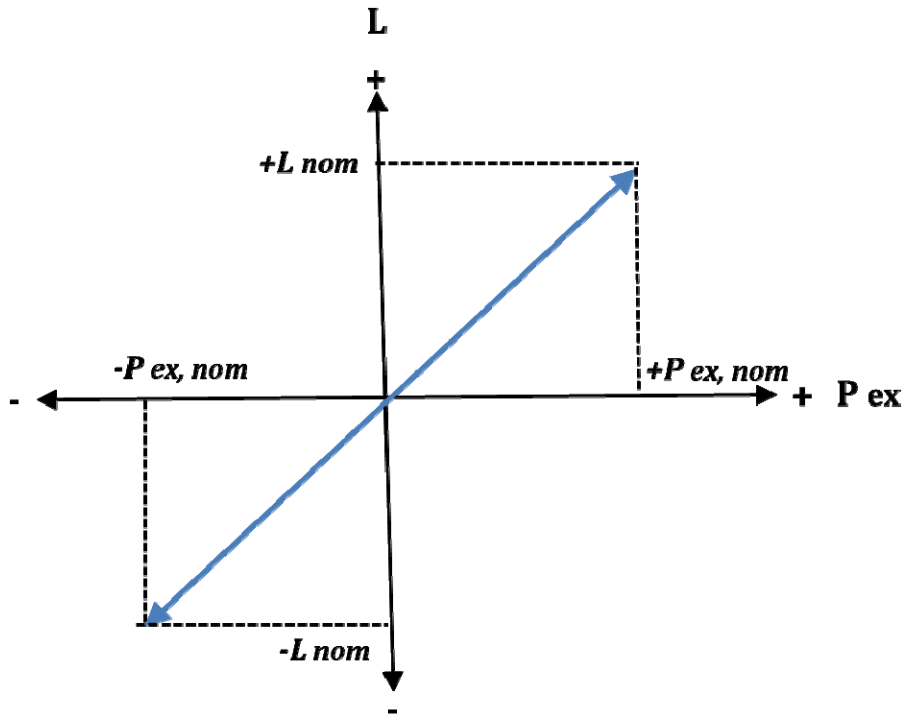


Figure 6 Energy losses L for an additional amount of power transported on a connection, as a function of the existing power flow P_{ex} on that connection.

In the case of a meshed network, there are several 'parallel' routes connecting any two buses in the network and the power flow will distribute itself over these 'parallel' routes. This makes it even more complicated to calculate the power losses caused by any additional energy transport between two buses.⁵ TSO TenneT was therefore unable to supply generic key figures for energy losses due to energy transport as a function of the distance. Before estimating the order of magnitude for this value, we will first discuss the case of the HVDC connection.

Fortunately, it is much more straightforward to estimate the energy losses for the NorNed HVDC connection. In the NorNed case, the network is not meshed. The NorNed cable is usually operated at a constant load (i.e. maximum capacity). Furthermore, in 2011 the NorNed cable was operated with power flowing from Norway to the Netherlands around 95% of the time. In such a situation, the energy losses for energy transport are constant and can easily be measured. The measurements showed an energy loss of 4% of the power transported. This includes the energy losses for AC/DC conversion and DC/AC conversion (separate from the power transported) and ohmic energy losses in the HVDC cable (losses that increase linearly with the amount of power transported).

The 4% energy loss figure for the NorNed HVDC connection (a distance of 580 km) can be used as a reference for estimating the order of magnitude of the losses for transport via overhead lines over land. One of the advantages of an HVDC connection is that its grid loss across longer distances is lower than the grid loss of an AC connection with the same power. The breakeven point for grid loss is several hundred kilometres. Given the 800 kilometre distance between Tyrol in Austria and the border of the Dutch grid, the energy loss for this connection can be estimated at 8% (twice the amount of the NorNed HVDC connection). It should be said that this is a very rough figure. It could be on the low side if the difference in energy loss between HVDC and AC connections has been underestimated.⁶ On the other hand, if the power flow for the data centre is opposite to the existing power flow along much of the route between Austria and the Netherlands,⁷ this estimate could be too high.

⁵ In theory it is possible to calculate the transport losses in the grid. But that requires a full network calculation for all nodes in the whole area involved. That is far beyond the scope of this study. The calculation would also require us to make many assumptions, e.g. about the generation profiles of all generators in the wider grid area and the consumption profiles of all loads. (This mechanism behind the transport losses in an international situation was confirmed by S. Ongkiehong of NL Agency.)

⁶ No references were found for this in the literature.

⁷ This is quite likely in the current operating conditions of the German power grid. Nowadays, electric power is transported from renewable and non-renewable generators in northern Germany to load centers in southern Germany for many hours of the year.

4.2 CO₂ emissions associated with grid losses

We have shown that the grid loss for energy transport is not totally negligible. But what does this mean for the CO₂ emissions? What level of CO₂ emission is associated with these grid losses?

To estimate this, we should bear in mind that it is the grid operator's responsibility to buy the energy needed to compensate for the grid losses in its grid area. This power needs to be generated locally for many applications, so the grid operator is not totally free to choose just any remote generator. Furthermore, no grid operators have a stated policy of prioritizing a particular type of energy (e.g. renewable generators). This means that the best estimate for the CO₂ emission factor of the energy losses involved in energy transport is the average CO₂ emission factor of the country involved. For continental Europe, we can use the emission factor of coal. For Norway, we can use the emission factor of hydro-electricity. The Netherlands will account for a smaller share of the energy losses for the NorNed cable, and Norway for a larger share. A practical approach might be to set the emission factor for the cable's energy transport losses at a value corresponding with 50% hydro-electricity and 50% coal-fired generation.⁸

⁸ Coal-fired power plants in fact generate less than 50%. This compensates for the low emission factor estimate in Norway.

5 Model integration

The decision equations discussed above have been integrated into a website calculator: <http://sne.science.uva.nl/bits2energy/index.html>

Users must answer a number of questions, as shown in Figure 7. First, they must choose a scenario (see next section) and supply additional input, depending on the scenario chosen. For the scenario 'software (interactive)', they must supply the amount of input data, the CPU processing time, and the amount of output data. Appendix D discusses how the scenarios and the calculator input relate to real-life applications.

The screenshot shows a web calculator interface with several sections:

- Choose a service scenario**: A dropdown menu set to "software (interactive)".
- extra input scenario 'software (interactive)'**:
 - Input data: [1,60] GByte
 - Output data: 9.1 GByte
 - CPU process. time: 3.7 (core)hours
- Transport network between local and remote data center**: A dropdown menu set to "Lightpath long distance".
- PUE**:
 - Local data center: 1.48
 - Remote data center: 1.56
 - Transport network: 2
- Emission energy production [gr CO₂/kWh]**:
 - Local data center**: X: Natural gas (380) 340
 - Remote data center**: X: Hydroelectricity (15) 20
 - Location energy production**: Local
 - Transport network**: X: Crude oil (640) 520

Figure 7 Input fields for web calculator.

Next, users must have an idea of how far the remote data centre is situated from the local data centre, as they have to choose the type of transport network (short-distance for SURFnet and long-distance for GEANT).

Third, users have to input the PUE of the local and remote data centres. A default PUE of 2 is presented for the transport network.

Finally, users must indicate which energy production sources the data centres and the transport network use. X-values for different forms of energy production are presented in a drop-down menu and may be changed by users. The data centre's energy production location is also taken into account. If local energy production is chosen, the estimate makes allowance for energy transport losses within the country and distribution losses, as shown in Appendix B.

The additional scenario inputs allow a range as input, e.g. '[1,60]' which indicates that input values vary between 1 and 60. If one of the scenario inputs contains a range, the calculator responds with a plot (see Figure 8).

Range data is presented as a coloured band due to a fixed error of 10% in the X-values (see Figure 17 for the spread in different X-values). Blue represents the data centre that has the lowest carbon emission cost for the maximum value of the range. A second calculation produces a table showing the values for the maximum of the range (see Figure 8), which is also how a result is presented in the absence of a range.

In the model outcomes presented below, please note that energy transport losses have been accounted for using the footprint that reflects the generation of that energy. In other words: sustainably generated energy (with a 'zero' footprint) will be lost during transport, amounting to about 5-17% depending on the transport mode. These losses are compensated for, but NOT by using renewable energy sources. The model could be refined to account for this.

6 Scenarios

Three different scenarios have been defined:

- processing (CPU-intensive);
- software (interactive);
- data storage.

Processing (CPU-intensive) is any scenario in which a calculation is performed on data. The calculation is performed on a computation server that also contains the input data for the computation. Two input parameters characterize this scenario: the amount of data involved and the CPU processing time in CPU core hours. It is assumed that the same type of computation server is present at both data centres, local and remote, with respect to energy consumption.

The second scenario, Software (interactive), is an extension of the first. It has an additional input parameter: the amount of output data generated by processing. Here we assume that the party interested in the output data is located near the local data centre, so the output data, like the input data, traverses the transport network between both data centres.

Data storage is the subject of the third scenario. In this scenario we make a distinction between hot and cold data. Hot data resides on content servers and is directly accessible, whereas cold data resides on content servers in low power mode or on tape or other offline data storage devices, and requires preparation to be made accessible.

6.1 Processing (CPU-intensive) and Software (interactive)

6.1.1 Software (interactive), large data sets

A typical representative of this scenario is when performing calculations on a large amount of data generated in large-scale experiments (high-energy physics, astronomy). Suppose the experimental data (Input data) varies between 1 and 60 GByte. The CPU processing time and the amount of output data both show a linear dependency on the input data (see Appendix C for the applicable function inputs).

The users of the data perform their calculations in a local data centre in the Netherlands with a PUE=1.48. That local data centre runs on energy produced by natural gas (where we changed the default value of 380 to 340 gr. CO₂/kWh). The calculations could also be performed at a remote data centre (PUE=1.56) that runs on cleaner energy, e.g. a data centre in Norway that makes use of hydro-electricity (with a higher value – 20 gr. CO₂/kWh – than the default value). We assume that both data centres obtain their energy from power plants in their home country. As the local and remote data centres are far apart, a long-distance network should be chosen. In this scenario, the long-distance Internet and lightpath variants have 3 hops (see Figure 4b,d for an example).

Whether we should prefer the remote data centre above the local one depends on the type of transport network and on the type of energy on which the network equipment runs. We assume the transport network has a PUE of 2.0 and the energy type is a mix of 10% natural gas (Netherlands) at X=340 gr. CO₂/kWh, 10% hydro-electricity (Norway) at X=20 gr. CO₂/kWh, and – for the remaining 80% of the network (North West Europe) – roughly 600⁹ gr. CO₂/kWh. This gives us a total value of about 520 gr. CO₂/kWh for the transport network between the local and remote data centres. Figure 8 shows the result when plotted for slow linear growth of the CPU processing time (see Appendix C)

⁹ This value of 600 gr. CO₂/kWh is a mix of values derived from http://ec.europa.eu/energy/energy_policy/doc/factsheets/mix/mix_de_en.pdf, where it can be seen that Germany uses about 27% solid fuel (pulverised coal), 28% crude oil, 25% gas and 12% nuclear energy for its energy production. We use this as a representative value for North West Europe.

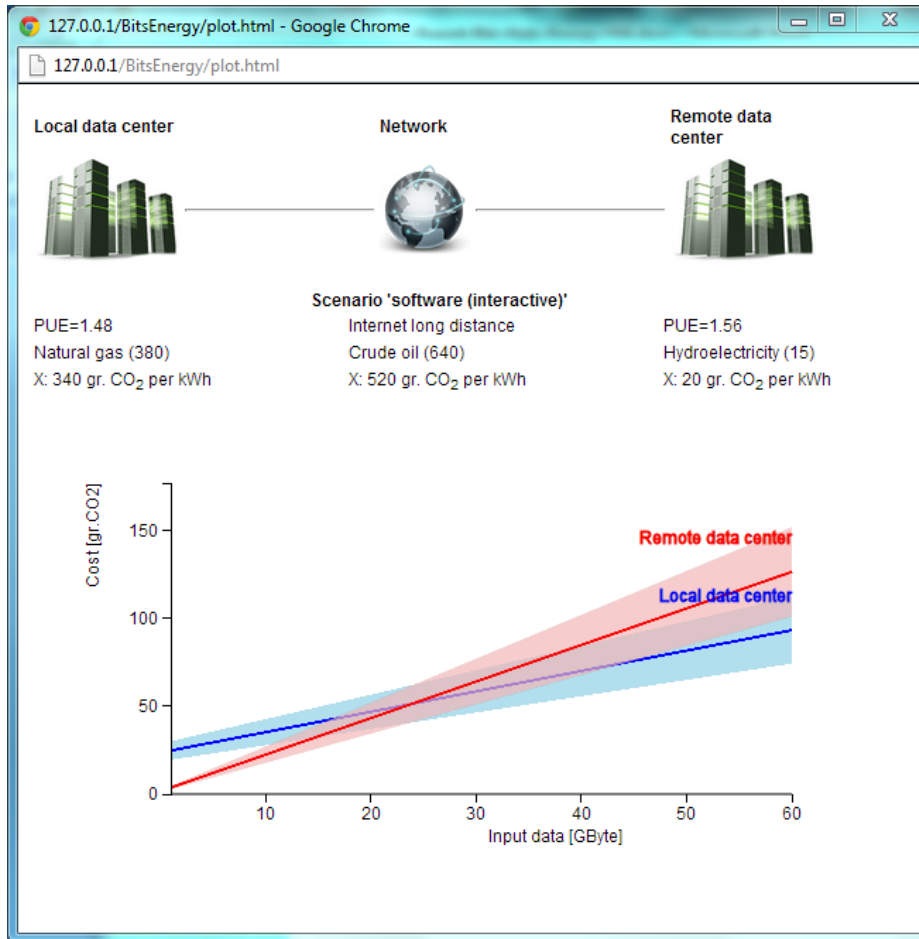


Figure 8 Output of scenario 'software (interactive)' with large data sets and long-distance data transport on the Internet. Input data in the range [1,60] GByte, linear dependency of output data, and slow linear growth of CPU processing time, as indicated in Appendix C.

It follows that it is preferable to do the calculation in the local data centre for input data larger than 25 Gbyte owing to the 'dirty' Internet transport network.

If we opt for another long-distance network, i.e. a long-distance lightpath network but with the same dirty energy, the situation changes as shown in Figure 9.

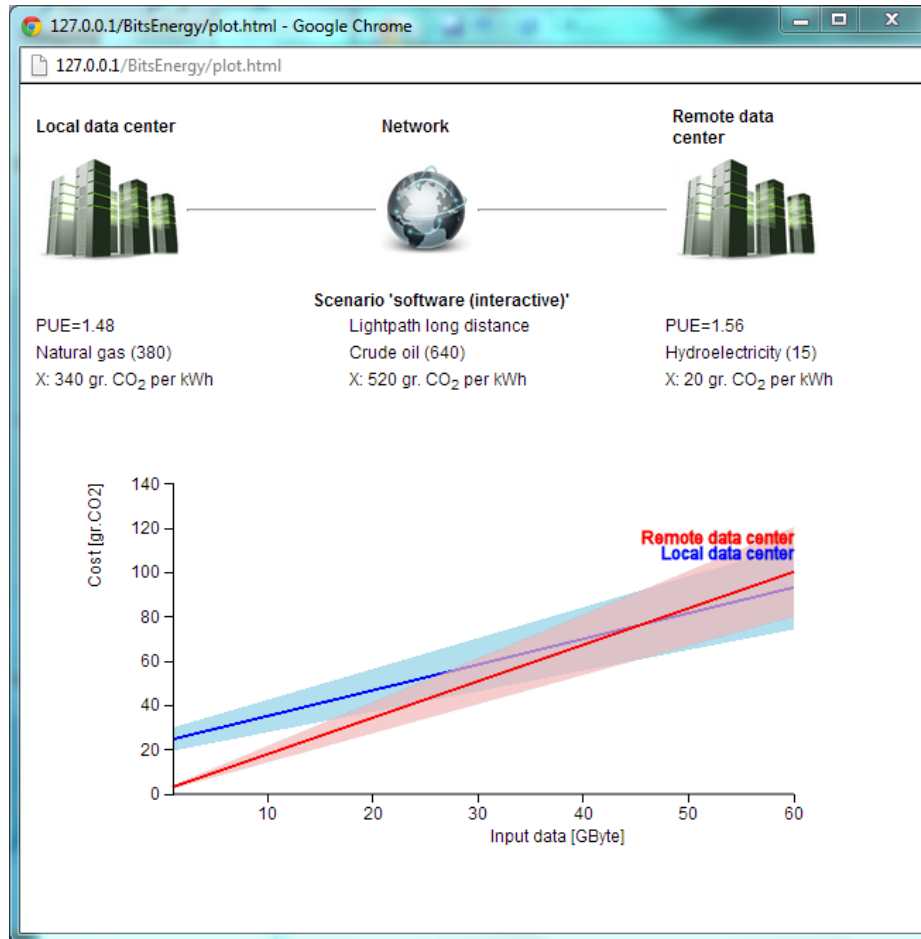


Figure 9 Output of scenario 'software (interactive)' with large data sets and long-distance lightpath data transport. Input data in the range [1,60] GByte, linear dependency of output data, and a slow linear growth of the CPU processing time as indicated in Appendix C.

If a dirty long-distance lightpath network is possible, we could opt for the remote data centre for a larger range of input data (<45 GByte). The difference is due to the fact that a long-distance Internet network has routers, devices with a relatively high energy consumption. Looking at the tabular output, Figure 10, for the long-distance Internet solution (Figure 8), the different contributions to the total carbon emission costs in gr. CO₂ for the maximum amount of the input data are:

Local processing		Remote processing	
•89.7%	• Computation	•31.0%	• LAN local (input data)
•6.3%	• LAN local (output data)	•54.2%	• Transport network (input data)
•4.0%	• Energy production loss	•1.9%	• LAN data centre (input data)
		•4.1%	• Computation
		•0.3%	• LAN data centre (output data)
		•8.2%	• Transport network (output data)
		•0.3%	• Energy production loss

Table 3 Relative carbon emission contributions for the maximum value of input data as depicted in Figure 8 and listed in Figure 10.

The CPU processing energy consumption in kWh for the local and remote data centres is 0.2464 and 0.2597 kWh respectively, and the ratio of these values equals the ratio of the PUEs, 1.48 and 1.56 respectively.

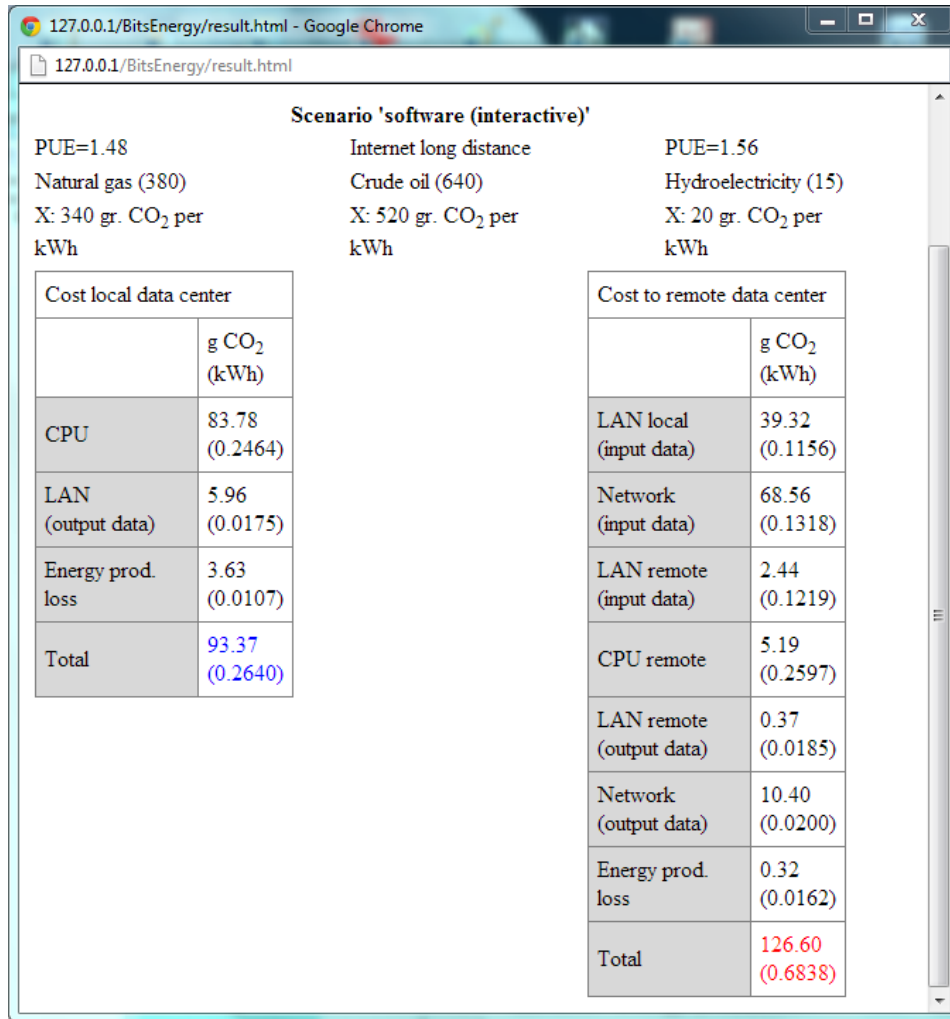


Figure 10 Output table for the maximum value of input data as depicted in Figure 8.

Importing clean energy from Norway would produce major advantages. The calculator allows two identical data centres to be compared, except for the energy production.¹⁰ The blue line of the 'remote' data centre (which now functions as the local data centre) is powered by energy imported via submarine cable from Norway. Figures 11 and 12 show the output for this comparison. The curve for the local data centre in Figure 11 is the same as the one in Figure 8; the blue curve represents the local data centre when hydro-electricity is imported from Norway.

¹⁰ Input as shown in Figure 7 should be changed such that both PUEs are equal (1.48) and the X-value for the transport network should be set to 0. Not only is the contribution of the input data through the transport network then set to zero, but so is the contribution of the input data through both data center LANs (cancelling out transport between the data centres).

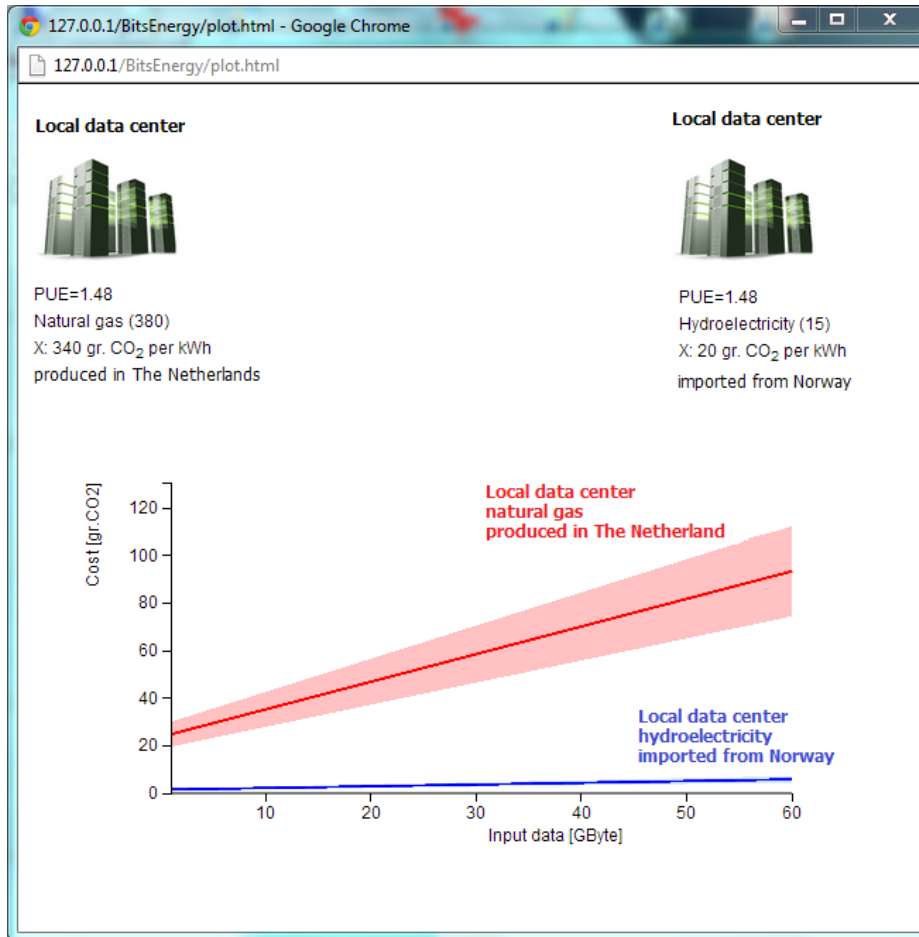


Figure 11 Comparison of two situations: local data centre powered by natural gas, delivered by a power plant in the Netherlands versus the same local data centre importing cleaner energy, hydro-electricity, from Norway (blue curve). This is for the scenario ‘software (interactive), large data sets’, with the input data being in the range [1,60] GByte, and with linear dependency of the output data and slow linear growth of the CPU processing time as indicated in Appendix C.

It follows from Figure 12 that CPU and LAN energy consumption in kWh is the same in both situations, as it should be. If the energy is imported via a submarine cable from Norway, the energy loss (see Appendix B) increases from 0.0107 kWh to 0.0282 kWh and the total energy cost increases from 0.2640 kWh to 0.2922 kWh. But because the energy imported from Norway, hydro-electricity, is much cleaner, the total cost in gr. CO₂ is much lower.

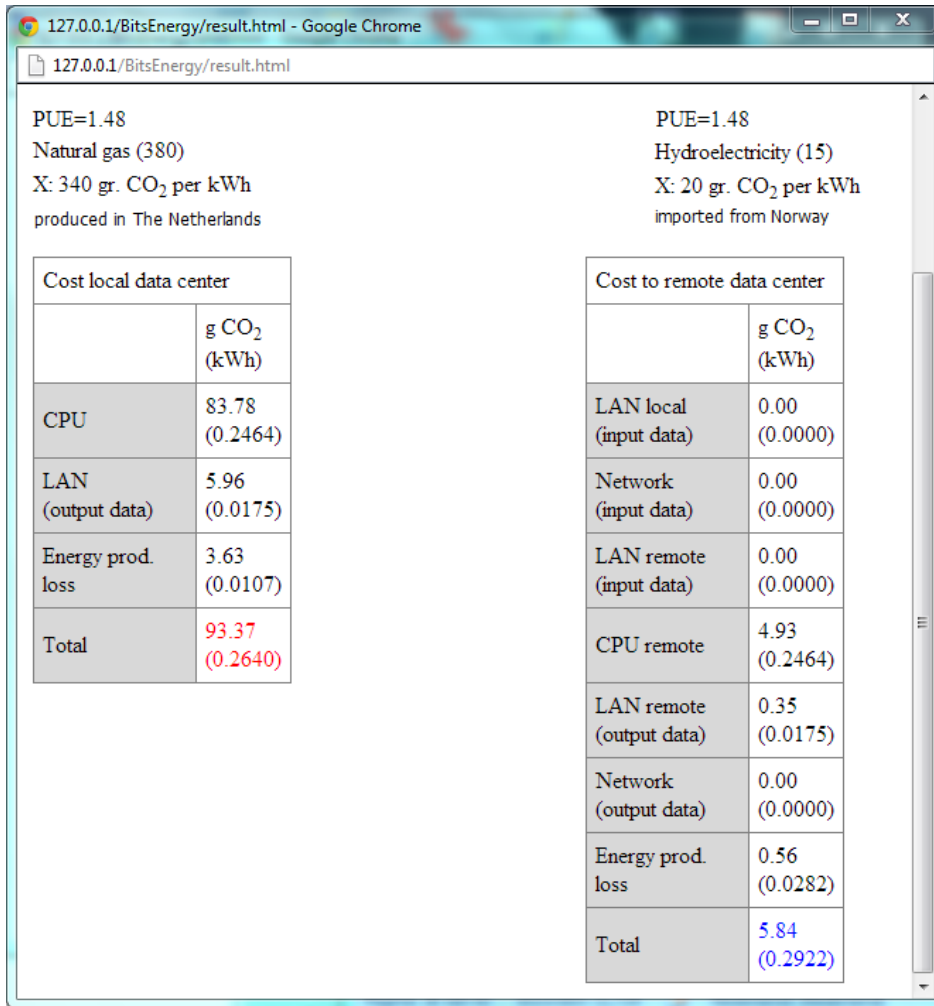


Figure 12 Output table for the maximum range values in Figure 11.

If the CPU processing time shows a faster linear growth (see Appendix C), the situation may change, as Figure 13 shows.

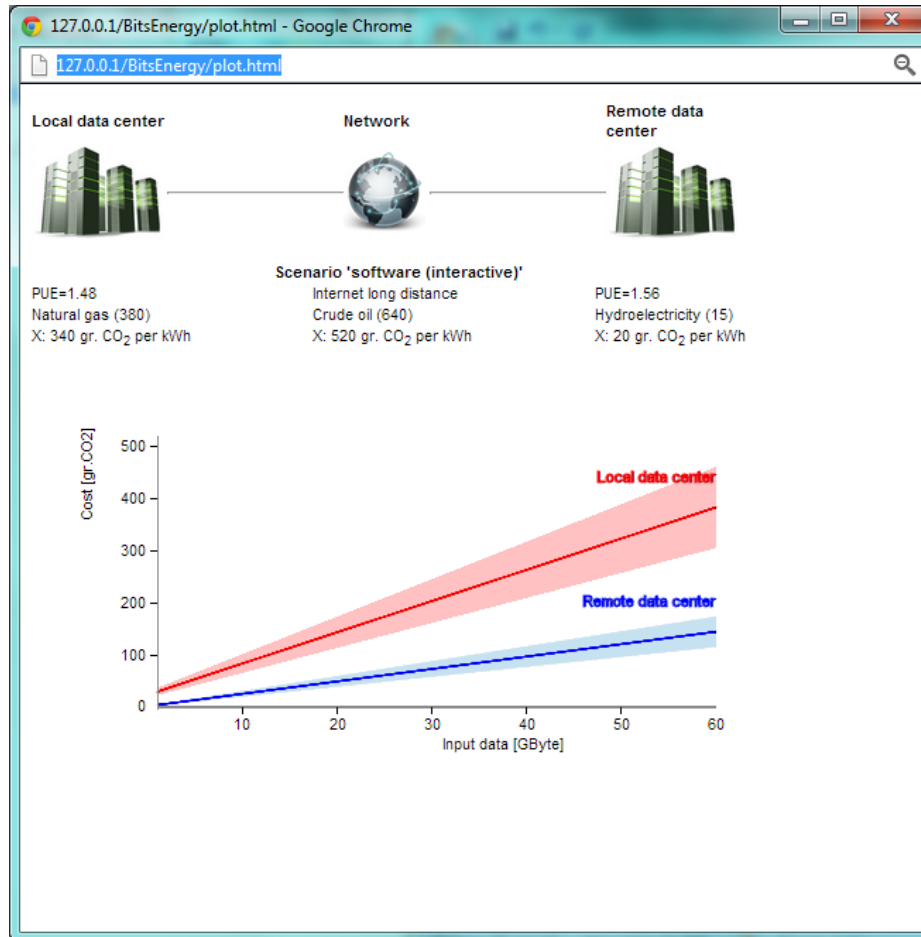


Figure 13 Output of scenario 'software (interactive), large data sets', with long-distance Internet data transport. Input data in the range [1,60] GByte, linear dependency of output data, and fast linear growth of the CPU processing time as indicated in Appendix C.

So far the decision problem has been a linear decision problem, with all the terms in the decision equation being linear. In this case, the network and CPU processing time contributions are also both linear as to the amount of data involved, and, as Figures 8, 9 and 13 indicate, the network may be of major importance. In fact, in some real-life scenarios, the transport network may be an obstacle to remote computation.

If the processing time reveals a quadratic dependency on the input data, i.e. an $O(n^2)$ problem, the decision problem becomes a quadratic one and the outcome can be more surprising. Consider the above example with the same input, except for a quadratic dependency of the CPU processing time (see Appendix C). Figure 14 shows the result of the web calculator for the adjusted example.

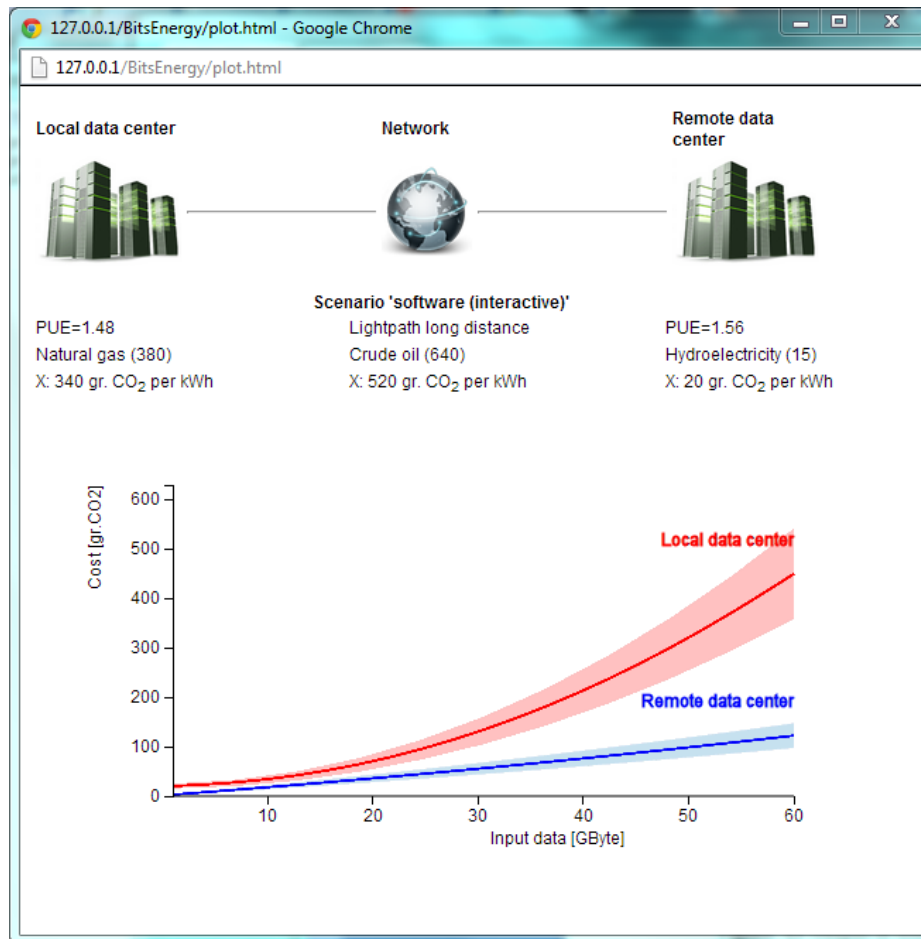


Figure 14 Output of scenario 'software (interactive), large data sets', with long-distance Internet data transport. Input data in the range [1,60] GByte, linear dependency of output data, and quadratic dependency of the CPU processing time as indicated in Appendix C.

The best option is to move the data and the accompanying calculation to a remote data centre for the whole range of input data. Due to the quadratic behaviour of the computation, it will be more costly to opt for local processing in the case of large data sets. In this instance, the network contribution, which is still linearly dependent on the amount of input data, will increase more slowly than the CPU processing contribution.

The software (interactive) scenarios presented in this section lead us to conclude that acquiring greener energy from abroad will have a favourable impact when the network is powered by dirty energy sources.

6.2 Data storage

Using the same settings for data centres and transport network, we now look at hot and cold data storage.

6.2.1 Hot data

The data to be stored is [1,200] GByte and the retention time is taken to be 144 hours or 6 days. Figure 15 shows the carbon emission cost for storage at a local data centre versus storage at a remote data centre; the jump in the curve is due to the ceiling term in Eq. 7 (it takes 2 disks with storage capacity of 200 GByte to store more than 100 GByte).

We can conclude that hot data should preferably be stored at a remote and 'clean' data centre. This might change if the data is accessed frequently, in which case network contribution starts to play an important role.

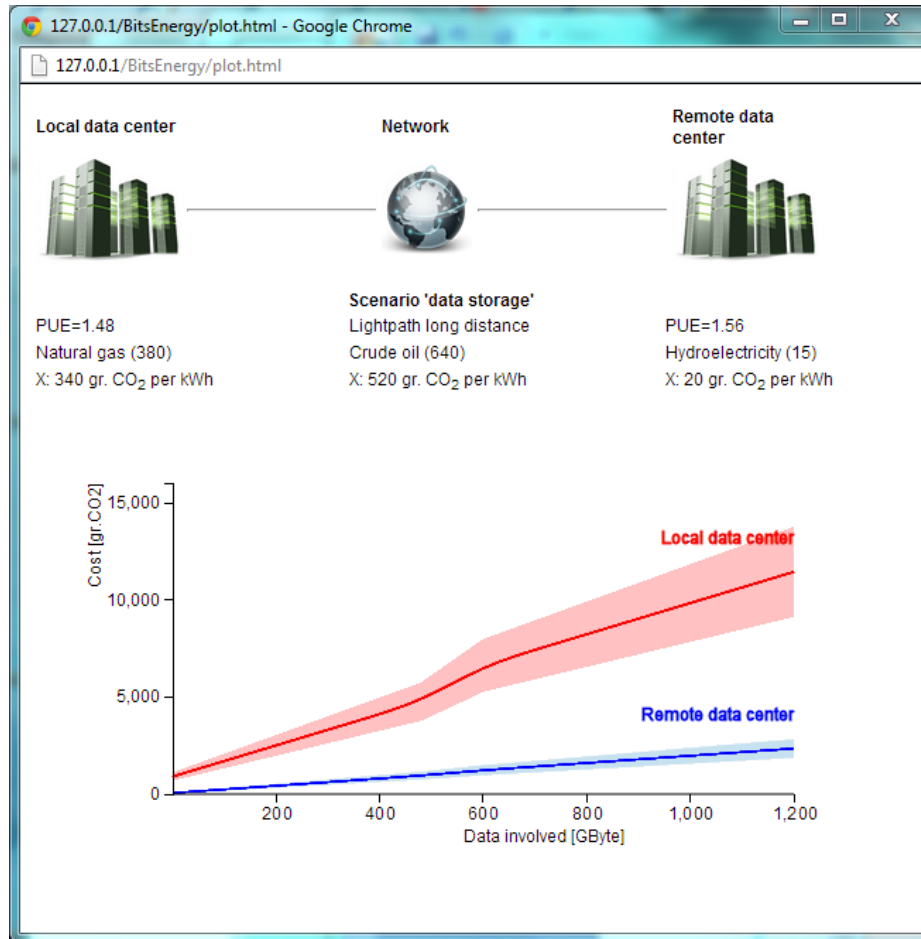


Figure 15 Output of scenario 'data storage', with hot data in the range [1,200] GByte, a retention time of 6 days and long-distance lightpath data transport.

6.2.2 Cold data

If the data is cold, the situation changes as depicted in Figure 16. The storage cost is much lower than the transport cost. We can conclude that with a 'dirty' transport network between two data centres, it is better to keep cold data at the local data centre (depending on the amount of data and the retention time).

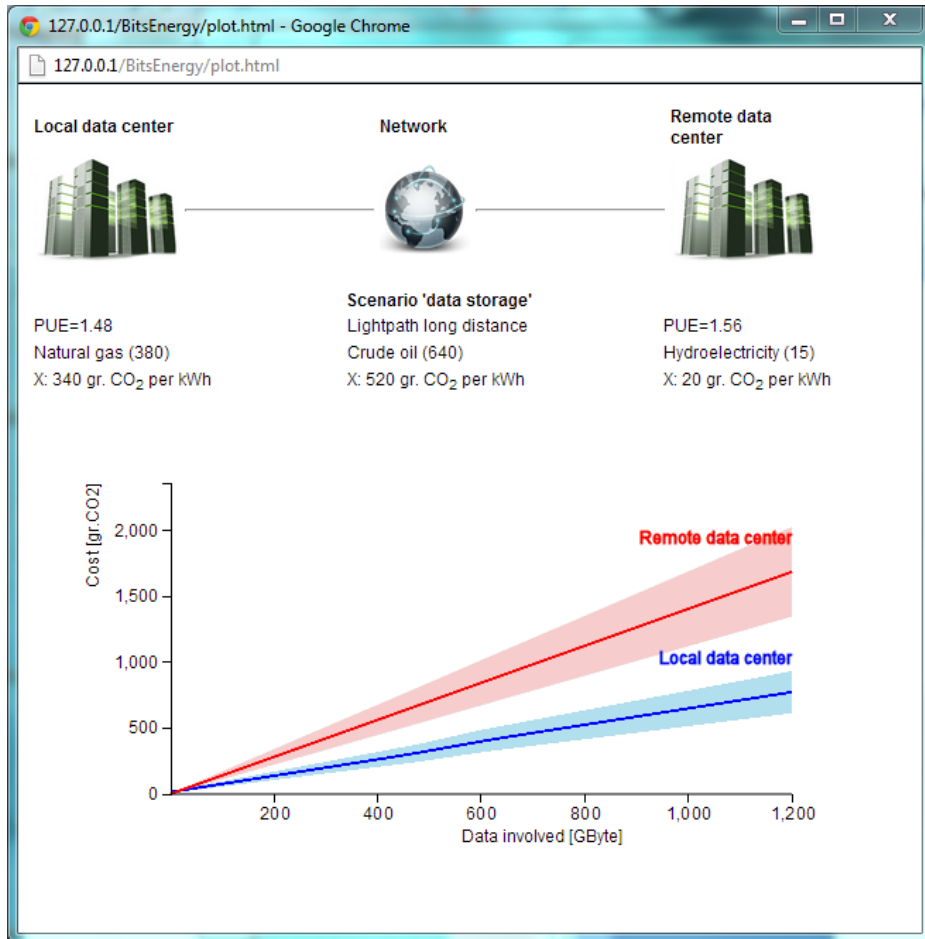


Figure 16 Output of scenario 'data storage', with cold data in the range [1,200] GByte, a retention time of 6 days and long-distance Internet data transport.

7 Conclusions

7.1 Data storage

From the above outcomes of some representative examples, a few general features become clear. Figure 17 shows that there is a large gap between the cleanest of dirty energy production, i.e. natural gas, and the dirtiest of clean energy production, i.e. solar energy. Since we are looking for situations with low gr. CO₂ emissions, using data centres that are powered with cleaner energy is advantageous, but this positive effect may be negated when the transport network equipment is powered by 'dirty' energy.

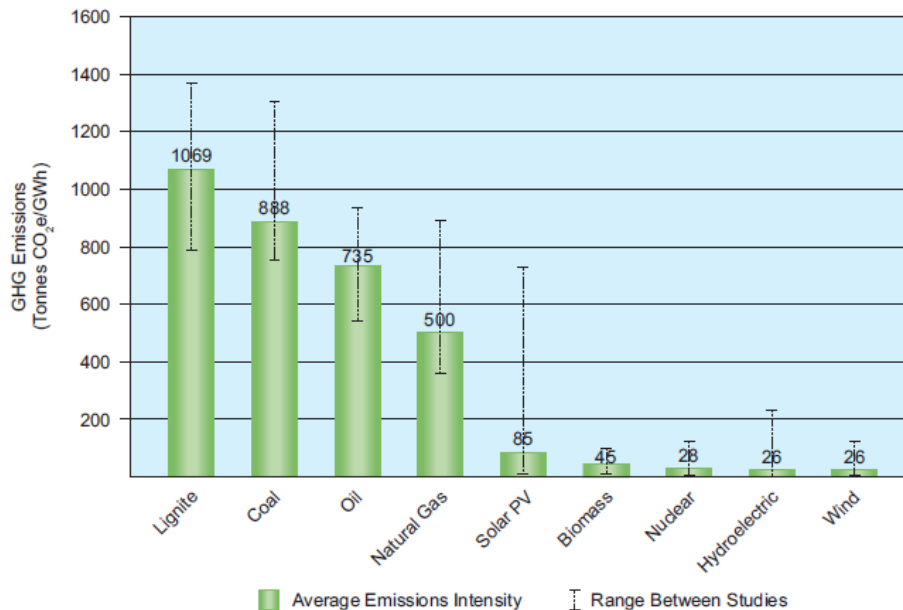


Figure 17 Taken from reference [11]

If the transport network is powered by dirtier energy than both of the data centres, the contribution of the network to the total cost of moving data in gr. CO₂ can be significant. This is mostly the case if the data traverses the Internet, due to the relatively high power consumption of routers. Lightpath connections are to be preferred above Internet connections, but lightpath connections are dedicated connections that require a more complex set-up procedure and sometimes might not be available to a user. For large input data sets and linear behaviour of the computation time on the input data, it might be better to perform the calculation locally, if the connecting network is the Internet. The same situation may be reversed if the computation time shows a quadratic dependency on the input data. In that case, the contribution of a dirty network is a less prominent factor.

Importing 'clean' energy from elsewhere leads to a large reduction in emissions. Emissions can be reduced further if the 'clean' energy can be produced locally, because importing green energy gives rise to losses, which need to be compensated and thus produce emissions; this is because compensation takes place in the transport network using an energy mix that is usually *not* 'green', so that the carbon footprint of the transported green energy increases.

Referring to the original research questions:

- *What are the sustainability effects of data transport over the data network? How much energy is required and what is the CO₂ footprint?*
- *What are the sustainability effects of energy transport? When is it suitable to acquire green energy from elsewhere?*

we can formulate the following conclusions.

- The energy involved in data transport can be considerable, compared to the energy involved in data processing.
- The energy involved in energy transport when acquiring (green) energy from distant locations can be considerable as well and should not be ignored.

Additionally, we can conclude the following:

- Using locally generated green energy has the largest (positive) sustainability effects.
- The energy sources and the energy required in data transport also have a non-negligible effect.
- The characteristics of the application, such as the amount of data transported and the amount of processing and storage required, have a significant effect on total CO₂ emissions.
- Acquiring green energy from abroad will generally be more positive than transporting data to a data centre abroad if the network transport components have a large carbon footprint.

Taken altogether, this means that for realistically large processing jobs (in the academic world), no one choice is 'always best' in terms of energy use and the associated emissions.

The next section outlines the choices that can be made in specific situations.

7.2 Recommendations

The above discussion has led us to develop the following general guidelines, in which we assume that the energy powering the local data centre is less clean than that powering a remote candidate data centre to which the data and the accompanying computation has been moved. When we refer to cleaner energy, we mean energy as clean as the energy of the remote candidate data centre. When we refer to dirtier energy, we mean energy with a higher gr. CO₂ emission than the energy powering the local data centre.

Scenario	Decision
Data scenario where the local data centre can import cleaner energy from elsewhere.	Keep data local and perform calculations locally.
Software (interactive) scenarios where the computation time depends linearly on the input data.	Be aware that if the network is long-distance, and the network is powered by dirtier energy, it might be preferable to keep the data local. This is truer for smaller amounts of input data on long-distance Internet paths than on long-distance light-paths.
Software (interactive) scenarios where the computation time depends on the data in a more than linear sense.	The type of transport network and energy determine above which amount of data this is preferable in order to obtain the minimal carbon footprint. The footprint of the remote data centre is also an important factor, given the quadratic behaviour.
Data storage scenario with hot data.	In general the carbon emission cost of the transport network is much lower than the storage cost at the local data centre. Moving hot data to a cleaner remote data centre is therefore preferable in terms of obtaining the minimal carbon footprint. This might change if the data is used/accessed frequently.
Data storage scenario with cold data.	In general the carbon emission cost of the transport network is much higher than the storage cost at the local data centre. Keeping cold data at the local centre is therefore preferable.

7.3 Future Research

Our study has yielded some general guidelines for data transport with accompanying computation to a remote data centre. We have been able to show under which conditions the transport network influences the decision. We must, however, consider that our work rests on some rough approximations and estimates that require further investigation. Nevertheless, we identified three points of interest:

- **Variable control.** Analysis showed that different combinations of variables play an important role in the outcome. Computation time and data centre energy resources and PUE, amount of data to be transported and PUE, and the energy resources of the network are all sets of variables that can play a significant role when deciding whether or not to move data to a remote data centre.
- **Energy database for networks.** In our model we use monitored and measured power consumption values from network equipment in real-time situations, and not only data sheet values. Despite all this, one major drawback of our analysis is that the actual number of network devices present in a transport network is very much dependent on the geographical position of both the local and remote data centre. The general guidelines presented here would be more realistic if we had a geographical database describing the footprint of the SURFnet and GEANT networks. Such a database should ideally contain information

from the energy companies that deliver energy to the SURFnet and GEANT network about what types of energy they use, and especially the amount of sustainable energy that they make use of. Having such data available, the calculator would allow users to express only what they want and it would suggest possible cleaner solutions from which users could choose. In essence, users could choose between different routes created dynamically via optical networks to maximize the sustainability effect of their data transport

- **Selection of energy provider.** Locally produced clean energy is the most preferable option, and then importing clean energy from abroad. Since energy companies control energy imports, we recommend selecting an energy company that buys clean energy and opting for that kind of energy to power data centre equipment.

Acknowledgements

The authors wish to thank the steering group and the advisory committee for all their feedback while conducting their research.

The authors are also grateful to NL Agency for commissioning this study.

Appendix A

A1 Power consumption of data centre and network components

Different sources are available that provide values for the power consumption of data centre and public network components, such as publications (including vendor data sheets and measurements).

A 1.1 Publications

Baliga et al. [3] present some values for the power consumption of data centre and public network components; see Tables A1 and A2. These tables give an indication of the typical power consumption of devices involved in data transport.

	Equipment	Capacity	Power Consumption
Storage	HP 8100 EVA	604.8 Tb [37]	4.9 kW [37]
Content Server	HP DL380 G5	800 Mb/s [38]	225 W [39]
Computation Server	HP DL380 G5	-	355 W [39]
LAN	Cisco 6509	320 Gb/s [40]	3.8 kW [40]
Gateway Router	Juniper MX-960	660 Gb/s [41]	5.1 kW [41]

Table A1 Power consumption of data centre equipment.

	Equipment	Capacity	Power Consumption
Ethernet Switch (Small)	Cisco 4503	64 Gb/s [40]	474 W [40]
Ethernet Switch	Cisco 6509	160 Gb/s [40]	3.8 kW [40]
BNG	Juniper E320	60 Gb/s [41]	3.3 kW [41]
Provider Edge	Cisco 12816	160 Gb/s [40]	4.21 kW [40]
Core router	Cisco CRS-1	640 Gb/s [40]	10.9 kW [40]
WDM (800 km)	Fujitsu 7700	40 Gb/s [45]	136 W/channel [21]

Table A2 Power consumption of public network components.

In our model we consider DWDM terminal nodes only and ignore optical line amplifiers, regenerators, and optical switches such as OXCs.

In DWDM terminal nodes, the transponder is the power consumer; a typical value for a transponder is 34.5 W @ 10 Gb/s = 3.45 W/Gb/s, which value equals the value given in Table A2: 136 W / 40 Gb/s = 3.4 W/Gb/s. Optical line amplifiers (EDFAs) show values of 200 W for 40 wavelengths @ 10 Gb/s, or 0.5 W/Gb/s. For regenerators, a typical value is 50 W for 1 wavelength @ 10 Gb/s, or 0.5 W/Gb/s. Optical switches operate at even lower values: 0.5 W for 1 wavelength @ 10 Gb/s, or 0.05 W/Gb/s.

A 1.2 Measurements

We compare the number for routers from Table A1 and A2 with data on power consumption collected from routers in use at SARA.

A 1.2.1 Routers

Routers from different vendors were monitored for their power consumption. These were Juniper MX960 routers, and Cisco Catalyst 6500 routers.

Juniper MX960 routers

These routers are equipped with 4 Power Entry Modules of max. 4100 W each. Output via JUNOS is:

PEM 0: DC output: 114 W

PEM 1: DC output: 399 W

PEM 2: DC output: 228 W
 PEM 3: DC output: 285 W
 The PEMs in this case consume 1026 W in total.

For another JUNIPER router, we get the following data:

PEM 0: DC output: 171 W
 PEM 1: DC output: 570 W
 PEM 2: DC output: 171 W
 PEM 3: DC output: 741 W
 In this case, the PEMs consume 1653 W in total.

The JUNIPER router of 1026 W has 8 interfaces of 10 Gb/s each and 20 interfaces of 1 Gb/s each, whereas the second one has 4 interfaces of 10 Gb/s and 40 interfaces of 1 Gb/s.

If all the interfaces are active, a typical value for the capacity of these routers is: $8 \cdot 10 + 20 \cdot 1 = 100$ Gb/s and $4 \cdot 10 + 40 \cdot 1 = 80$ Gb/s, respectively.

Combined with the power consumption, we arrive at values of 1026 W / 100 Gb/s and 1653 W / 80 Gb/s, or 10.26 W/Gb/s and 20.66 W/Gb/s. Given that the capacity of the second router is double that of the first, these values are not far off from the one given in Table A1, 5.1 kW/660 Gb/s = 7.7 W/ Gb/s

The Cisco Catalyst 6500 router

An enhanced 9-slot Chassis System with 8 modules produces the following data:

module 1 power consumption: 325.50 Watts (7.75 Amps @ 42V)
 module 2 power consumption: 325.50 Watts (7.75 Amps @ 42V)
 module 3 power consumption: 325.50 Watts (7.75 Amps @ 42V)
 module 4 power consumption: 325.50 Watts (7.75 Amps @ 42V)
 module 5 power consumption: 282.24 Watts (6.72 Amps @ 42V)
 module 7 power consumption: 325.50 Watts (7.75 Amps @ 42V)
 module 8 power consumption: 325.50 Watts (7.75 Amps @ 42V)

Total power consumption is 2235.4 W. A forwarding engine in each module delivers 40 Gb/s. As not all ports were connected or enabled, we arrive at $2235 \text{ W} / 7 \cdot 40 \text{ Gb/s} = 7.9 \text{ W/Gb/s}$ for this router.

A2 Adopted values

Following on from the above discussion, we have adopted the values below for the equipment, with the values for a router based on our own measurements and the other values being taken from Tables A1 and A2.

Equipment	Power per capacity [kW/Gb/s]
Host CPU-intensive	0.355
Host data storage	0.280
Router	0.012 (= (10.26 + 20.66 + 7.9)/3 W/Gb/s)
Ethernet switch (small)	0.007
Ethernet switch	0.023
DWDM terminal node	0.0034
Firewall	0.016

Table A3 Adopted values for the power per capacity of different devices used in our model

For the firewall, we adopted the value: $650 \text{ W} / 40 \text{ Gb/s} = 0.016 \text{ kW/Gb/s}$.

Appendix B

B1 Cable losses

The energy losses that are associated with energy transport are defined as follows:

- f_{T2C} [-] is the fraction lost for the transport from the generator abroad up to the border of the home country;
- f_{TIC} [-] is the fraction lost in the transmission grid of the data centre's home country, from its border up to the transmission grid node to which the data centre distribution grid is connected;
- f_{DIC} [-] is the fraction lost due to distribution from the transmission grid node up to the point of common coupling of the data centre.

The losses for energy transport can then be calculated using the following formulas (where E_{prod} is the amount of electric energy generated):

Transport loss from the generator to the data centre's home country (E_{T2C}):

$$E_{T2C} = E_{prod} * f_{T2C} \quad (B1)$$

Transport loss within the data centre's home country (E_{TIC}):

$$E_{TIC} = E_{prod} * (1 - f_{T2C}) * f_{TIC} \quad (B2)$$

Distribution loss within the data centre's home country (E_{DIC}):

$$E_{DIC} = E_{prod} * (1 - f_{T2C}) * (1 - f_{TIC}) * f_{DIC} \quad (B3)$$

The energy arriving at the data centre (E_{DC}):

$$E_{DC} = E_{prod} * (1 - f_{T2C}) * (1 - f_{TIC}) * (1 - f_{DIC}) \quad (B4)$$

The total transport loss before arrival at data centre ($L_{tot, remote}$) is:

$$\begin{aligned} L_{tot, remote} &= E_{prod} * f_{T2C} + E_{prod} * (1 - f_{T2C}) * f_{TIC} + E_{prod} * (1 - f_{T2C}) * (1 - f_{TIC}) * f_{DIC} \\ &= E_{prod} * \{ f_{T2C} + (1 - f_{T2C}) * f_{TIC} + (1 - f_{T2C}) * (1 - f_{TIC}) * f_{DIC} \} \end{aligned} \quad (B5)$$

From (B4), the energy generated E_{prod} can be expressed in terms of the energy consumption at the data centre E_{DC} :

$$E_{prod} = E_{DC} / \{ (1 - f_{T2C}) * (1 - f_{TIC}) * (1 - f_{DIC}) \} \quad (B6)$$

This produces for the total loss $L_{tot, remote}$:

$$L_{tot, remote} = E_{DC} * \{ f_{T2C} + (1 - f_{T2C}) * f_{TIC} + (1 - f_{T2C}) * (1 - f_{TIC}) * f_{DIC} \} / \{ (1 - f_{T2C}) * (1 - f_{TIC}) * (1 - f_{DIC}) \} \quad (B7)$$

If the energy is generated in the data centre's home country, the losses are:

$$L_{tot, home} = E_{prod} * f_{TIC} + E_{prod} * (1 - f_{TIC}) * f_{DIC} = E_{prod} * \{ f_{TIC} + (1 - f_{TIC}) * f_{DIC} \} \quad (B8)$$

Using (B4), the energy generated E_{prod} can be expressed in terms of the energy consumption at the data centre E_{DC} :

$$E_{prod} = E_{DC} / (1 - f_{TIC}) * (1 - f_{DIC}) \quad (B9)$$

This produces for the total loss $L_{tot, home}$:

$$L_{tot, home} = E_{DC} * \{ f_{TIC} + (1 - f_{TIC}) * f_{DIC} \} / \{ (1 - f_{TIC}) * (1 - f_{DIC}) \} \quad (B10)$$

We have adopted the following values for the loss fractions:

$$f_{DIC} = 3.4\%^{11}$$

$$f_{TIC} = 0.5\%^{12}$$

$$f_{T2C, NO} = 6\% \text{ for the submarine cable to Norway (580 km)}$$

$$f_{T2C, AT} = 10\% \text{ for the overhead line to Tyrol, Austria (800 km).}$$

With these numbers, the losses become:

$$L_{tot, remote, AT} = 0.156 E_{DC}$$

or 15% of the data centre's energy consumption.

¹¹ Typical value for grid losses in the distribution grid of Dutch distribution grid operator Liander.

¹² Typical value for grid losses in the Dutch transmission grid as supplied by TenneT.

$L_{tot_remote_NO} = 0.107 E_{DC}$
or 11% of the data centre's energy consumption.

$L_{tot_NL} = 0.04 E_{DC}$
or 4% of the data centre's energy consumption.

Appendix C

C1 Special scenario inputs

Scenario inputs may consist of a range. For instance, Input Data = [1,80] GByte means that the amount of input data varies between 1 and 80 GByte.

If the input data varies, CPU processing time will in general also vary, and so will the amount of output data. To facilitate such a dependency, dependent inputs may be expressed as a function of the range values.

If the CPU processing time shows a linear dependency on the amount of input data, this can be expressed, for example by the following function input for the CPU processing time: $f(\$0)=0.009*\$0+0.20$, where $\$0$ is a placeholder for a value in the first range given (here, the input data range). At the moment, the calculator allows for one range, but more than one function input is possible.

The following input examples are used in the scenario 'software (interactive) large data sets' (5.1.1.) for any range of input data:

CPU processing time (slow linear growth)	= $f(\$0)=0.009*\$0+0.20$
CPU processing time (fast linear growth)	= $f(\$0)=0.05*\$0+0.20$
CPU processing time (quadratic)	= $f(\$0)=0.001*\$0*\$0+0.17$
Output Data	= $f(\$0)=0.15*\$0+0.1$

Appendix D

D1 Providing application-specific input to the calculator

The calculator deals with raw input numbers: it asks the user of the calculator to provide information, and turns this into numbers on how much energy the collective servers and other equipment in a data centre are using, and how much energy a transport network could be using as a result.

Basically, the user of the application has to model

- a) application-specific aspects:
 - *What does a specific application look like from a data processing, data storage and data transport perspective?*
- b) infrastructural aspects:
 - *What characterizes the energy production facilities that power the local data centre, the remote data centre and the transport network?*
 - *What characterizes the data transport network between the local and the remote data centre?*

The ideal future calculator would have a database with all current infrastructural data, so that the calculator's user can start his investigation by merely entering the location of the local and remote data centres.

Entering the application-specific aspects can also be simplified. Currently, the user now has to know the following application input data:

- a) How long does the CPU calculate?
- b) How much data is being stored, retrieved or processed?
- c) In what type of medium is the data being stored?

Ideally, the user would not enter these numbers but only the application that he would like to investigate. The calculator would then deduce the three variables of the processing task, ideally combined with an estimation of the typical energy profiles of the hardware that is running at this data centre.

Since the calculator has not reached that level of maturity yet, however, we still need to input the application input data manually. This appendix explains what kind of application can be input into the calculator. It also describes how different applications can have differing impacts on the amount of energy a CPU or a hard drive uses in relation to idle time and a server's total power usage. We hope that putting things into perspective will help calculator users understand the need to diversify their cloud strategies based on application-specific needs.

D2 Real-life application scenarios

Below we describe three scenarios run by the Software Energy Efficiency Lab (SEFLab), which was set up and operated by Amsterdam University of Applied Sciences and the Software Improvement Group in order to analyse the energy footprint of software applications.¹³ The lab is equipped to perform measurements on components of server hardware and to link these measurements to software running on this server. The SEFLab has run three application scenarios on a Dell PowerEdge SC1425 Server running Windows, and analysed the load of the server and the load on the components of this server.

This allowed us to survey the energy usage characteristics of processors, memory, hard disks and other components when a server is asked to transfer data, to perform calculation-heavy operations, and to provide a remote application to an end user. The measurements are compared to the idle state of the server, and therefore show the additional power consumption resulting from a request for service.

In the 'idle' situation, the system uses 150.4 W of power, meaning that when idling for one hour the server would consume approximately 0.15 kWh. The breakdown over the components is as follows:

Server load	CPU 1	CPU 2	MB	MEM	HDD1	HDD2	FANS
150.4	23.4	23.0	26.6	24.5	9.18	2.69	3.32

Table D1: idle power usage (in W)

¹³ The SEFLab collects such footprints in a database for further analysis in order to support the design of software that will lead to a smaller energy footprint. For more information, see: <http://www.hva.nl/kenniscentrum-dt/labs-ateliers/seflab/>

Scenario 1: data retrieval

The first scenario inspected was the transfer of data. A file was downloaded from the server using the FTP protocol. The throughput was 7.8 Mbit/s, resulting in a download of a 1 GB file in 131 seconds.

What is striking is that the server load increased by only 4% compared to the idle situation, mainly because CPU power consumption increased (11%). This is not significant, as the FTP server was running additional logging functionality.

Server load	CPU 1	CPU 2	MB	MEM	HDD1	HDD2	FANS
155.9	26.2	24.3	26.7	25.3	9.14	2.69	3.32

Table D2: data transfer power usage (in W)

Scenario 2: interactive software

The second scenario consisted of running an application for the remote control of a desktop (Ultr@VNC), with a video file running on the server that was being viewed at a connected client computer. This application is in the category interactive software because it requires the server to perform calculations and to transfer data to and/or from a client.

To start with, watching a 1 hour movie would result in the transfer of 560 MB of data from the server to the client (and 15 MB of data from the client to the server). The power consumption during this hour of playback would be 0.20 kWh.

Playing this movie for the client increases the server's power consumption by 35%. This is mainly the result of an increase in CPU power consumption by 91%, although the memory banks have also increased their power consumption in these measurements, by 7%.

Server load	CPU 1	CPU 2	MB	MEM	HDD1	HDD2	FANS
204.1	47.7	42.9	26.9	26.7	8.83	2.64	3.26

Table D3: remote video playback power usage (in W)

Scenario 3: CPU-intensive processing

The third scenario is an illustration of CPU-intensive processing: the re-encoding of a video file in a different format. By re-encoding an MP4 file to DivX, we simulate a CPU-intensive task.

Re-encoding results in a 40% increase in power consumption as compared to the idle state, and a 110% increase in power consumption resulting from the CPUs.

Processing a 1 GB file would take about 3 hours, during which time the server would use a total of 0.65 kWh according to the SEFLab measurements. These correspond neatly with the calculations resulting from the Bits-Nets-Energy calculator, which estimate the power consumption at 0.675 kWh.

Server load	CPU 1	CPU 2	MB	MEM	HDD1	HDD2	FANS
209.2	52.1	46.1	26.8	24.7	9.24	2.68	3.33

Table D4: processing: re-encoding a video file (in W)

References

- [1] M. Stansberry and J. Kudritzki 'Uptime institute 2012 data center industry survey', <http://uptimeinstitute.com>
- [2] http://www.thegreengrid.org/events/technical_forum/Day_1.Track_1._Rating_Systems_for_Data_Centers_P.pdf
- [3] J. Baliga, R. W.A. Ayre, K. Hinton, W. and R. S. Tucker, 'Green Cloud Computing: Balancing Energy in Processing, Storage, and Transport', Proceedings of the IEEE, vol. 99, no. 1, pp. 149–167, Jan. 2011.
- [4] R. S. Tucker, 'Optical packet-switched WDM networks—A cost and energy perspective,' presented at the Opt. Fiber Commun. Conf. Expo., San Diego, CA, 2008.
- [5] S. Ricciardi, D. Careglio, F. Palmieri, U. Fiore, G. Santos-Boada, and J. Solé-Pareta, 'Energy-oriented models for WDM networks', Proc. of BROADNETS 2010, LNICST 66, pp. 534-547, 2012.
- [6] R. Basmdjian, N. Ali, F. Niedermeier, H. de Meer, G. Giuliani, 'A methodology to predict the power consumption of servers in data centres', Proceedings of the 2nd International Conference on Energy-Efficient Computing and Networking (e_energy '11), pp. 1-10, (2011).
- [7] Z. Hasan, H. Boostanimehr and V. K. Bhargava 'Green Cellular Networks: a survey, some research issues and challenges', In Communications Surveys & Tutorials, IEEE (2011).
- [8] International Energy Agency (IEA), 'CO₂ emissions from fuel combustion: Highlights, Report, 2009'; p. 14.
- [9] B. K. Sovacool, 'Valuing the greenhouse gas emissions from nuclear power: A critical survey', Energy Policy 36, pp. 2940-2953, (2008).
- [10] http://en.wikipedia.org/wiki/Emission_intensity.
- [11] WNA report, 'Comparison of Lifecycle Greenhouse Gas Emissions of Various Electricity', http://www.worldnuclear.org/uploadedFiles/org/WNA/Publications/Working_Group_Reports/comparison_of_lifecycle.pdf