

The Environmental Footprint of Data Centers: The Influence of Server Renewal Rates on the Overall Footprint.

Willem Vereecken¹, Ward Vanheddeghem¹, Didier Colle¹,
Mario Pickavet¹, Bart Dhoedt¹ and Piet Demeester¹

¹ Internet Based Communication Networks and Systems,
Ghent University, IBBT
Gaston Crommenlaan 8 bus 201
B-9050 Gent – Belgium

{Willem.Vereecken, Ward.Vanheddeghem, Didier.Colle,
Mario.Pickavet, Bart.Dhoedt, Piet.Demeester}@intec.ugent.be

Abstract. The environmental footprint of ICT is rising. Data centers are a key contributor to this footprint. In this paper we investigate the influence of the renewal rate of servers on the footprint of the data center. We take into account both the use phase power consumption as well as the contributions of the other life cycle stages. Based on this we construct an analytical model. From the results, we demonstrate that in a scenario where the data center needs to keep up with the increasing processing capacity of the servers, the footprint increases annually and keeping the servers in operation as long as possible is necessary. However, when the capacity remains constant, the footprint is decreasing and an optimal renewal rate is obtained.

Keywords: Green IT, Carbon Footprint, Data Center, Life Cycle Assessment

1 Introduction

In the past years, it has become clear that increasing carbon emissions are a global challenge. In every sector of the economy, initiatives are being taken, and endorsed by governments to reduce carbon footprint. Also in ICT these challenges need to be tackled. Studies [1,2] have shown the power consumption of ICT is growing even faster than the world's global power consumption, thus being responsible for an increasing fraction of this global power consumption. Data centers are currently responsible for about 1/6th of the ICT footprint. It has already been demonstrated that also their power consumption is increasing. [3,4]

On the other hand, it is estimated that ICT can play an important role in the reduction of the global carbon footprint. Through dematerialization of streams, ICT services enable people to massively reduce their carbon footprint while still fulfilling their needs. In the SMART 2020 report [2], it is estimated that ICT can reduce up to 5 times its own footprint. It is important to note that these claims will only be achieved when a certain adoption of these technologies is obtained.

Hence, it is essential that the ICT sector can keep its own carbon footprint under control, and even reduce it in the near future in order to be able to support the claims of the carbon footprint reduction capabilities of ICT. The sector itself realizes this and has taken several initiatives to tackle this issue. In the work that is being performed in these initiatives (e.g. [5,6]), however, the focus is mainly on energy consumption of equipment. However, in order to estimate the full impact of a technology, we need to take into account the full impact of a product's life cycle (i.e. from material extraction until disposal of the product). As we will demonstrate in this paper, if the life cycle impact is not taken into account, wrong conclusions might be drawn and efforts to reduce the carbon footprint could lead to an increase of carbon footprint. Vice versa, decisions that seem to increase the carbon footprint could actually lead to overall reductions thanks to other life cycle stages that are more advantageous.

In this work, we will focus on the carbon footprint of data centers. In section 2, we will investigate the power consumption of the ICT equipment in a data center. Next, we will relate this power consumption to full life cycle impact of the ICT equipment. In section 3 we identify values of the key parameters in the model. Then, in section 4, we compare the influence of the different factors on the overall datacenter power footprint. Finally, in section 5 we analyze these results and draw the main conclusions.

2 Modeling the Footprint of the Data Center

We construct a model that describes the footprint of a data center. We analyze the server power consumption and the related footprint. For this, we consider two scenarios. In the '*constant number of servers*' scenario, we assume that the number of servers in a data center remains constant, and every removed server is replaced by a new server. In the '*constant data center capacity*' scenario, we assume the processing capacity of the data center remains constant and that, with increasing server capacity, a higher number of old servers is replaced by a lower number of new servers.

2.1 Server Power Consumption

We consider a data center. In this data center a number of servers are present. Every year, new servers are brought into the data center. In this model, we assume this happens at the beginning of the year. At the same time, old servers are removed from the data center. We assume that a server is used for n years before being removed. This means, in a given year y , there are servers present from year $y, y-1, \dots, y-n+1$. We denote the number of servers added in year y as N_y . When we denote the average power consumption of a server purchased in year y as P_y^s we get for the total server power consumption in the data center:

$$P_y^{dc} = \sum_{i=y-n+1}^y N_i P_i^s. \quad (1)$$

We assume an exponential growth rate for the server power consumption. We denote the growth factor as β . This implies:

$$P_{y+i}^s = P_y^s \beta^i. \quad (2)$$

Constant number of servers – First, let's assume that every old server gets replaced by a new server. This means that N_y is constant. Since every year we replace N_y servers which remain in the data center for n years, the total number of servers in the data center is $n \times N_y$, which we denote as N . With (2), we get for the power consumption:

$$P_y^{dc} = \frac{N}{n} P_y^s \sum_{i=y}^{y-n+1} \beta^i. \quad (3)$$

Or, with the formula for geometric series:

$$P_y^{dc} = \frac{N}{n} P_y^s \frac{1-(1/\beta)^n}{1-(1/\beta)}. \quad (4)$$

Constant data center capacity – In the previous case we assumed the number of servers in the data center to remain constant. However, this does not take into account the fact that server capacity is increasing. From Moore's law, we know that the processing capacity of a server doubles every 18 months. This means, if we replace servers, that we can replace them by a smaller number of servers if we want the processing capacity of our data center to remain the same.

Lets denote the processing capacity of a server as C_y^s . Again, we assume an exponential growth for this capacity:

$$C_{y+i}^s = C_y^s \gamma^i. \quad (5)$$

The total capacity present in the data center is C^{tot} . Since this remains constant, and each year we replace a fraction $\frac{1}{n}$, we get:

$$\frac{C^{tot}}{n} = N_y C_y^s. \quad (6)$$

This needs to be valid every year. Hence:

$$N_{y+i} C_{y+i}^s = N_y C_y^s. \quad (7)$$

or, with (5):

$$N_{y+i} = N_y (1/\gamma)^i . \quad (8)$$

Similar to the calculation of (4), we get:

$$P_y^{dc} = N_y P_y^s \frac{1-(\gamma/\beta)^n}{1-(\gamma/\beta)} . \quad (9)$$

In this formula, we can no longer simplify by eliminating N_y .

2.2 Server Footprint

After determining the power consumption of the server, we need to calculate the carbon emissions associated to this power consumption. First we need to multiply the power consumption (in W) with a factor of 8.766 $\left(= \frac{365 \times 24}{1000} \right)$ to get the yearly electrical energy consumption (in kWh). Second, we need to account for the carbon emissions associated to this energy consumption. These carbon emissions are expressed as carbon emission intensity I , i.e. the mass of CO₂ (in g) emitted per used kWh. This emission intensity is dependent on the production technology (e.g. based on oil, gas, charcoal, etc.).

Next to the power consumption, we need to take into account the full life cycle of the server. Life cycle assessment is a field in which the environmental impact of a product or service is measured taking into account the material extraction, production, transportation, use and disposal. Only the use phase impact is determined by the time the product is operational. Thus, we can model the other life cycle impacts as a single emission at the moment the server is purchased. We denote the non-use-phase impact of a server as L . Every year N_y servers with this footprint are purchased.

Based on the above, we get for the footprint (in kg CO₂) in year y :

$$F_y = N_y L + 8.766 I P_y^{dc} . \quad (10)$$

with P_y^{dc} as denoted by either (4) or (9).

2.3 Power Usage Effectiveness

Next to the power consumption of the servers, other equipment is present in a data center as well. This equipment also consumes power. It is on the one hand other ICT equipment such as switches, storage networks, etc. and on the other hand equipment used for cooling, uninterruptable power supplies, lighting etc. The latter overhead is considered proportional to the ICT equipment power consumption and is expressed by

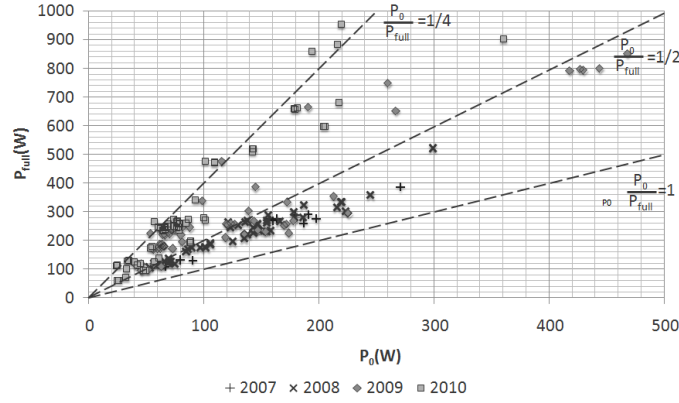


Fig. 1. Server Power Consumption P_0 and P_{full} for Different Years.

Power Usage Effectiveness (PUE). By multiplying the ICT power consumption with the PUE one obtains the total data center power consumption. The PUE is considered to be approximately 2.[7]

In many studies, the PUE is also considered in the modeling of the data center. This in itself is valid. However, in this paper we want to incorporate the life cycle impact of the servers. If we would incorporate the PUE, this would require accounting the life cycle impact of the other equipment as well. As we consider these impacts out of scope for the effects we wish to model, we do not account for PUE.

3 Parameter Value Estimation

After constructing the model, we identify the value of the different parameters in order to be able to draw conclusions.

First, we need to understand how the *server power consumption* evolves. We used the server power consumption measurements submitted to SPEC using the SPEC Power benchmark [8]. In these measurements two values are important. The idle server power consumption P_0 and the full load server power consumption P_{full} . With linear interpolation the average server consumption can be determined in function of the server CPU load λ :

$$P^s = (1 - \lambda)P_0 + \lambda P_{full} . \quad (11)$$

In Fig. 1. we have displayed P_0 and P_{full} for the data collected in different years. We see that the fraction $\frac{P_0}{P_{full}}$ is decreasing and thus the server power consumption becomes more and more dependent on the load. This means the server power consumption will have a different growth rate depending on the assumed load. We

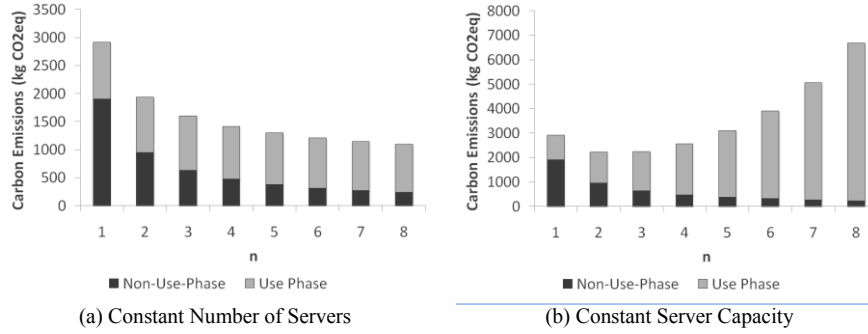


Fig. 2. Normalized Data Center Footprint in Function of the Server Renewal Rate n in 2011.

assume an average load on the servers of $\lambda=80\%$, which means that the server capacity in the data center is well used.

Based on this data we performed a linear regression to determine the growth factor of the server power consumption. Note that in the benchmarking results in [8], next to the year of submission, the month is given as well. This allows for a more fine-grained analysis. As a result, we get for the server power consumption in 2007 a value of 187 W. The growth rate of the power consumption is 5.2% p.a. (i.e. $\beta=1.052$).

The dataset we used to perform the linear regression has the advantage of being a consistent dataset with 217 data points. On the other hand, the servers submitted in this dataset will probably be better performing in terms of power consumption than the average server as the submission of the results implies a consideration for power consumption aspects during the design of the machines.

Second, we model the growth of the *capacity* of servers. Moore's law states that the number of transistors on a chip doubles every 18 months. In the past, this increase was enabled by ever increasing the clock frequency of the CPU. Currently, the trend is to provide multiple cores per server. In this work, we assume Moore's law also describes the increasing processing capabilities of servers. This implies $\gamma=1.59$ ($=2^{2/3}$).

The value for the *carbon emission intensity* is based on either production technology of the energy or either the weighted average of the technologies used in the energy mix for a certain location. In [9] we can find values for the several energy production technologies as well as the regional averages. In this paper, we assume the world average emission intensity of $I=504$ gCO₂/kWh.

Finally, we need to determine the non-use-phase footprint of the server. In [10] an estimation is made expressed in Mega-Joules. Again using the world average emission intensity we get $L = 1903$ kgCO₂ per server. Note that the operating model of the considered server in [10] is significantly different from the assumptions made in this paper.

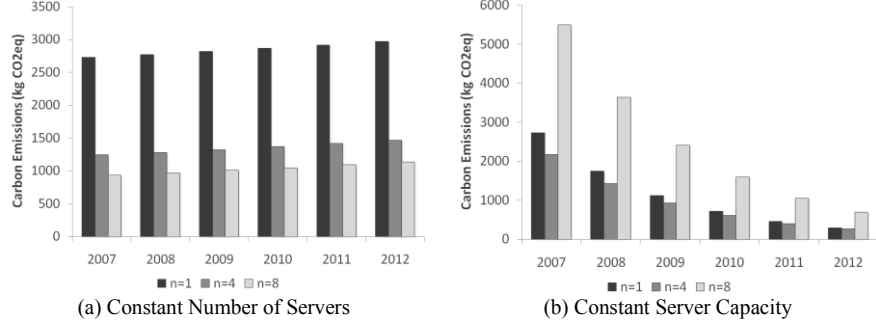


Fig. 3. Data Center Footprint Evolution in Function of the Server Renewal Rate n .

4 Results

We evaluate the constant server and the constant capacity scenario for a data center in 2011. For the constant server scenario, we normalize the footprint to the footprint of 1 server. This implies dividing (4) by N .

In the constant capacity scenario, we normalize to the capacity of one server in 2011. This capacity in itself remains undefined but it allows us to make quantitative comparisons. This assumption implies substituting N_y by $1/n$ in (9). Note that in these assumptions, for $\gamma=1$, formulas (4) and (9) are equal.

In Fig. 2, we have displayed the footprint for the considered data center in function of the renewal rate n . In Fig. 2(a) we consider the scenario with a constant number of servers. The longer the renewal rate, the lower the footprint becomes. This is because the non-use-phase footprint decreases and a direct consequence of the lower number of replaced servers. Additionally, since the server power increases yearly, the number of servers that is replaced is lower, and the overall power consumption decreases.

In Fig. 2b, we have displayed the footprint for the considered data center in the constant capacity scenario. Now, there is an optimum for $n=2$. Because the server capacity is increasing, every year we need to install less servers. Since the capacity is increasing faster than the power consumption ($\gamma/\beta > 1$), the large amount of old servers outweighs their lower power consumption on the long term. This means, in this scenario, it is best to handle a replacement period of 2 years in order for the carbon emissions to be minimal.

This analysis could lead to the conclusion that it is better to replace a larger number of servers. When we compare the values for $n=8$, one notices in the constant server scenario the footprint is a lot lower than in the constant capacity scenario. This is a wrong conclusion, though, and originates in the normalization to the server capacity from 2011. In Fig. 3, we represent both cases on a longer term. We now normalize for the year 2007 in the same way as described before and we look at the evolution of the total footprint until 2012. In the constant number of servers scenario the footprint increases yearly, due to the increasing server power consumption. In the constant capacity scenario, the footprint decreases due to the reduced equipment requirement.

5 Conclusions

We evaluated the evolution of data center power consumption. We analyzed two scenarios. One in which we assume every old server needs to be replaced by new server and one where the processing capacity of the data center remains constant.

If during the replacements, the number of servers remains constant, we see that over the years the footprint of the data center increases. In this case it is essential to extend the lifetime of the servers as long as possible.

In the constant capacity case, there is a yearly decrease of the footprint. In this case, we also observe there is an optimal replacement period of about two years.

The driver for the capacity is the software running on the servers. The fact that data center power consumption is still increasing indicates that software is increasingly demanding capacity. As long as this situation persists, it is key to keep servers in operation as long as possible in order to reduce the full life cycle cost of the servers.

It is however important to strive for the scenario in which the software allows for the processing capacity to remain constant. In this case the power consumption of data centers will reduce. At this time however, we need to carefully evaluate the optimal renewal rate for data centers.

Acknowledgments. The work described in this paper was carried out with the support of the TREND project ("Towards Real Energy-Efficient Network Design"), a Network of Excellence funded by the EU's Seventh Framework Program.

References

1. Pickavet, M., Vereecken, W., Demeyer, S., Audenaert, P., Vermeulen, B., Develder, C., Colle, D., Dhoedt, B., Demeester, P.: Worldwide energy needs for ICT: The rise of power-aware networking. Advanced Networks and Telecommunication Systems, 2008. ANTS '08. 2nd International Symposium on , pp.1-3, (2008)
2. Webb, M.: SMART 2020: Enabling the low carbon economy in the information age. The Climate Group (2008)
3. Koomey, J.: Estimating Total Power Consumption by Servers in the U.S. and the World. Oakland, CA: Analytics Press, 2007.
4. U.S. Environmental Protection Agency: Report to Congress on Server and Data Center Energy Efficiency, Public Law 109-431 (2008)
5. Lefevre, L.; Orgerie, A.-C.: Designing and evaluating an energy efficient Cloud. The Journal of Supercomputing, vol. 51, pp. 352-373 (2010)
6. Vereecken, W., Deboosere, L., Simoens, P., Vermeulen, B., Colle, D., Develder, C., Pickavet, M., Dhoedt, B., Demeester, P.: Power efficiency of thin clients. European Transactions on Telecommunications, vol. 26 issue 6, pp. 479-490, (2010)
7. The Green Grid: Green Grid Metrics: Designing Datacenter Power Efficiency. (2007)
8. SPECpower_ssj2008 Results, http://www.spec.org/power_ssj2008/results/
9. IEA, CO2 Emissions from Fuel Combustion - highlights, <http://www.iea.org/co2highlights/> (2010)
10. Hannemann, C.R.; Carey, V.P.; Shah, A.J.; Patel, C.: Lifetime exergy consumption of an enterprise server. Electronics and the Environment, 2008. ISEE 2008. IEEE International Symposium on , vol., no., pp.1-5, (2008)