LUND UNIVERSITY

# A forecast of the Cloud

## – An investigation of the energy use from one of the fastest growing phenomena of the IT sector - the Cloud

*Ingela Oscarsson*

Master Thesis 2014
Environmental and Energy Systems Studies
Department of Technology and Society
Lund University

LUNDS UNIVERSITET
Lunds Tekniska Högskola

# A forecast of the Cloud

## An investigation of the energy use from one of the fastest growing phenomena of the IT sector – the Cloud

Ingela Oscarsson

Master's Thesis

September 2014

ISRN LUTFD2/TFEM--14/5087--SE + (1-72)

Dokumentutgivare, Dokumentet kan erhållas från

LUNDS TEKNISKA HÖGSKOLA
vid Lunds universitet
Institutionen för teknik och samhälle
Miljö- och energisystem
Box 118
221 00 Lund
Telefon: 046-222 00 00
Telefax: 046-222 86 44

Dokumentnamn

Examensarbete

Utgivningsdatum

September 2014

Författare

Ingela Oscarsson

Dokumenttitel och undertitel

En prognos för Molnet – En undersökning av energianvändningen från ett av de snabbast växande fenomen på IT-marknaden - Molnet

Sammandrag

IT ses ofta som en del av lösningen för att uppnå ett hållbart samhälle genom till exempel minskat resande, optimering av industri- och jordbruksprocesser, intelligenta elmätare och smarta hem. Något man sällan reflekterar över är att IT-branschen själv också bidrar till elanvändningen. Ett nytt fenomen inom IT-världen är molnet som ger möjlighet till som det verkar outtömliga resurser i form av lagrings- och beräkningskapacitet, konstant uppkoppling och snabb överföring. Det finns många definitioner av molnet men om man ser på det materiellt består det av datahallar i olika storlek samt fasta och trådlösa nätverk som drar el dygnet runt. Om användningen av molntjänster ökar – hur mycket kommer elanvändningen öka och med den också den globala uppvärmningen? I denna studie kommer molnet definieras, materialiseras och kvantifieras för att kunna bedöma dess elbehov idag och i framtiden. Lagar och regler för energieffektivisering kommer undersökas och framtida prognoser tas fram genom tillväxtmodeller.

De huvudsakliga resultaten är:

- Det finns inga lagar för hur energieffektiva datahallar måste vara, även om det görs en del inom området på frivillig basis och företag tar på sig egna miljömål. Europeiska unionen inkluderade vissa delar av servrar i ekodesigndirektivet år 2014 vilket visar på att problemet har börjat tas upp.
- Användningen av molnet kommer öka explosionsartat i framtiden och det finns stor potential för energieffektivisering när det gäller lagring, bearbetning och överföring av data. Beroende på hur mycket som energieffektiviseras kan molnet komma att konsumera mellan 5 000 och 10 000 TWh år 2040. Detta kan jämföras med hela IT-branschen som 2010 drog mellan 700 och 1 000 TWh. Om man jämför molnet och traditionell IT är molnet oftast mer energieffektivt bland annat därför att resurser förbrukas efter behov och servrar utnyttjas optimalt. Det finns alltså ännu större potential för energieffektivisering om hela IT-sektorn inkluderas.
- Om inte energieffektivisering sker alls kommer molnets energiförbrukning öka bortom greppbara magnituder. Det finns dock också studier som pekar på att den totala energikonsumtionen kan minska sett från idag, även om användningen av molnet ökar, då teknik med mycket bättre energiprestanda håller på att utvecklas.
- Då dagens lagstiftning inte täcker in energieffektivisering ligger ett stort ansvar på företag att göra detta på frivillig basis, vilket till viss del motiveras av att de sparar pengar genom att energieffektivisera. Det är dock mycket viktigt att denna energiåtgång uppmärksammas och att den inte tillåts skena iväg i framtiden.

Nyckelord

Molnet, IT, Informationsteknik, elförbrukning, effektivitet,

| Sidomfång | Språk | ISRN |
|---|---|---|
| 74 | Engelska | ISRN LUTFD2/TFEM--14/5087--SE + (1-72) |

Organisation, The document can be obtained through
LUND UNIVERSITY
Department of Technology and Society
Environmental and Energy Systems Studies
Box 118
SE - 221 00 Lund, Sweden
Telephone: int+46 46-222 00 00
Telefax: int+46 46-222 86 44

Type of document
Master thesis

Date of issue
September 2014

Author
Ingela Oscarsson

Title and subtitle
A forecast of the Cloud. An investigation of the energy use from one of the fastest growing phenomena of the IT sector – the Cloud

Abstract

Information and communication technology (ICT) is often seen as part of the solution for a sustainable society, for example through reduced travel, optimization of industrial and agricultural processes, smart meters and smart homes. However, something usually left unconsidered is the electricity consumption of ICT itself. A new phenomenon of the ICT industry is the Cloud that enables seemingly inexhaustible resources in terms of storage and computing capacity, constant connection and fast transfer. There are many definitions of cloud but if looked at from a material point of view it consists of data centers in different sizes as well as wired and wireless networks that consumes electricity. If usage of cloud services increase - how much will electricity consumption and with it global warming increase? In this study the cloud is defined, materialized and quantified in order to estimate its electricity demand today and in the future. Laws and regulations for energy efficiency will be examined and future forecasts are created with the use of growth models. The main results are:

- There are no regulations of how energy efficient data centers must be, even though some companies set their own environmental goals and voluntary projects are carried out. The European Union included some parts of servers in the Ecodesign Directive in 2014, which shows that the problem has begun to be addressed.
- The usage of the cloud will increase dramatically in the future and there is great potential to improve energy efficiency in terms of storage, processing and transmission of data. Depending on how energy efficient the cloud will be it can consume between 5 000 and 10 000 TWh in 2040. This can be compared to the entire ICT industry which consumed between 700 and 1 000 TWh in 2010. If the cloud is compared with traditional IT it is usually more energy efficient as resources are pooled and used when needed and servers are utilized optimally. Therefore there is even greater potential for improving energy efficiency if the entire ICT sector is included.
- If there are no energy efficiency improvements at all the cloud’s energy consumption will increase beyond graspable magnitudes. However, there are also studies that indicate that the total energy consumption can decrease in the future, even though the use of the cloud increases, due to new efficient technologies currently under development.
- As current regulations do not include energy efficiency of data centers, a huge responsibility is placed at companies to do this on a voluntary basis. The companies do however have a self-interest to improve energy efficiency as it saves them money - but is it enough? It is very important that this energy consumption is recognized and is not allowed to increase out of control in the future.

Keywords

The Cloud, ICT, information and communication technology, electricity consumption, efficiency,

| Number of pages | Language | ISRN |
|---|---|---|
| 74 | English | ISRN LUTFD2/TFEM--14/5087--SE + (1-72) |

# Preface

This work is written as a master thesis for the program Environmental Engineering at the Faculty of Engineering at Lund University and Pål Börjesson, professor at the faculty, is examiner.

When writing this thesis I have found myself in many interesting conversations as the topic concerns everyone. Those conversations have helped me stay motivated and provided me with useful input. I want to thank all that has taken interest in my thesis. Especially I want to thank my three thesis supervisors who all have inspired me: Johan Holmqvist, dr., Senior Environmental Specialist at Sony Mobile Communications, who has given me many interesting reports to look into and explained how companies take their decisions. Max Åhman, dr. and researcher at LTH, who continuously has given me great feedback. Lars J. Nilsson, professor at LTH, who has provided me with many reports and contacts.

Last but not least I want to extend a huge thanks to my brother Jonas who has been my anchor this year by reading my drafts and by giving me knowledge about a subject where I have found myself hopelessly lost several times.

# TABLE OF CONTENTS

# 1 INTRODUCTION

## 1.1 BACKGROUND

The first digital computers saw the light of day during the 40s (Koomey, et al., 2011 b). The computers were huge and demanded several people to maintain them. In the figure below a U.S army picture of the pioneer computer ENIAC is shown. Today it is hard to imagine a world without computers and internet, as they constitute an essential part of our lives. The famous quote when Thomas J. Watson states that there is a market for about five computers in the world seem long gone. Today you can reach a person in the other end of the world whenever you want, using a device you can easily carry with you. Information in the form of bytes are created, distributed, transformed, stored and received all over the world. In developed as well as in developing countries more and more people connect to the internet with more devices per person each day. 50 billion devices will be connected to the internet in 2020 - this is more than six times the predicted world population of that same year (IVA, 2013).

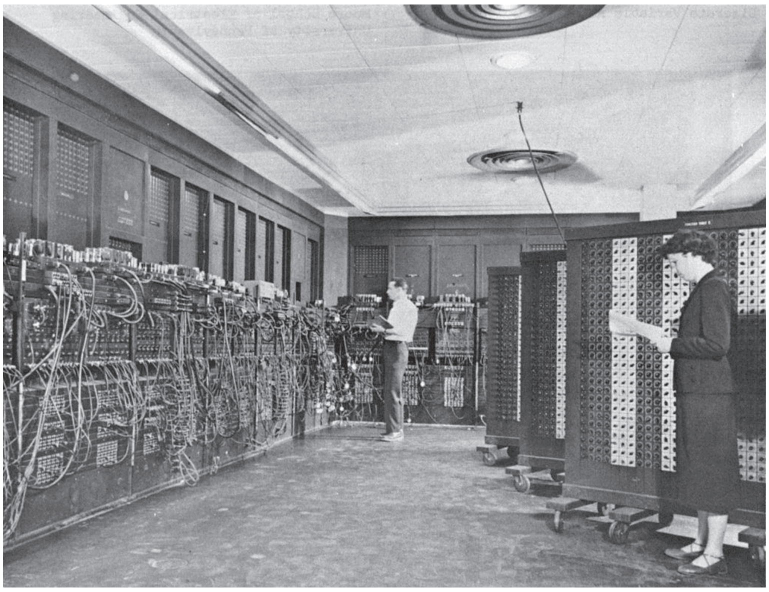
Figure 1. Shows the digital computer ENIAC (U.S Army photo, 1947).

The amount of electricity dense data centers are also increasing, as storage and computing demand grow. A special form of IT services, so called Cloud services, have gained ground. The term cloud can be interpreted in many ways, as there are many definitions and descriptions of it. One way to describe the Cloud is that it offers a measured and scalable IT service according to demand. One of the pioneers to offer cloud services was Amazon Web Service. The book selling company had most customers during Christmas and had thus to dimension its servers according to this peak in demand when a lot of people ordered books online. During the rest of the year the servers were over dimensioned and the launch of cloud service sprung from this fact: in order to save money and actually use this capacity Amazon Web Services started in 2002 to lease computing and storing resources to other companies which had peaks in other parts of the year than December (Rhoton, 2011).

However, knowledge about the environmental impact from the cloud is scarce and there are both advantages and disadvantages concerning the cloud. This thesis is an investigation of the energy use of the cloud today and in the future. It also summarizes and explains terminology concerning the cloud.

## 1.2 PURPOSE

This study is aiming at clarifying the concept “the Cloud” and the energy use related to the service. The main question to be answered is: “What is the energy demand for cloud services now and in the future?”. In order to answer this question some sub-questions are raised:

- What is the definition of the cloud?
- How will cloud services evolve in the future?
- Which incentives exist for lowering the energy demand concerning the cloud?

## 1.3 METHOD

An extensive literature study is conducted based on books, reports, articles and journals. Interviews with IT companies and key figures within the studied topic is performed. Scenarios based on knowledge gained from the interviews and the literature study is elaborated and the cloud's future electricity consumption is forecasted through complementary top-down and bottom-up approaches. A system perspective is used throughout this thesis.

The structure of this thesis is as follows:

1. The definition of the cloud is investigated by comparing the descriptions and definitions in circulation.
2. The cloud is materialized and a general description of which devices are needed for it to work is be presented.
3. The sources of the electricity powering the cloud are covered.
4. The energy use and other quantifying terms are be studied for ICT and the cloud.
5. The remaining sections concentrate on future predictions and projections before the reports final analysis, discussion and conclusions are presented.
6. In the appendix the interviews of some IT companies are summarized as well as detailed assumptions and calculations.

## 1.4 SCOPE

This report is not a full-scale life cycle assessment even if a life cycle approach is used when determining the energy demand. It is only the operational phase of the life cycle that is studied in this report. The energy use from the cloud and ICT will be studied, but further emissions will not be included in this report apart from mentioning some greenhouse gas emissions from ICT in order to relate the energy use to climate change. The study has no geographical border as cloud services are global. This paper does not describe technical details but instead concentrate on getting an overview of the whole system.

# 2 DEFINING THE CLOUD

In this chapter the definition of the cloud, or cloud computing, will be discussed. There are several different definitions in circulation, and it is hard to determine if there is one "correct" definition. It might be more beneficial to understand the different views of cloud computing to enable comparison. In the report *A Break in the Clouds: Towards a Cloud Definition* by Vaquero et al. (2009) over twenty different definitions of the cloud are presented. A similar list, covering definitions and explanations stated after 2009, can be found in appendix Table 15.

Yet there are some similarities in the majority of definitions, which also Vaquero et al. found out. Vaquero et al. identified the common **attributes**; virtualization, pay-per-user model and scalability. Other reports further divide cloud computing into **service models** and **deployment types**. All of these three categories of the cloud will be explained in this chapter. Definitions diverging from these categories will also be mentioned in section 2.4 in order to get a comprehensive view of the definition of the cloud. In Figure 2 below the cloud is summarized in a mind map.

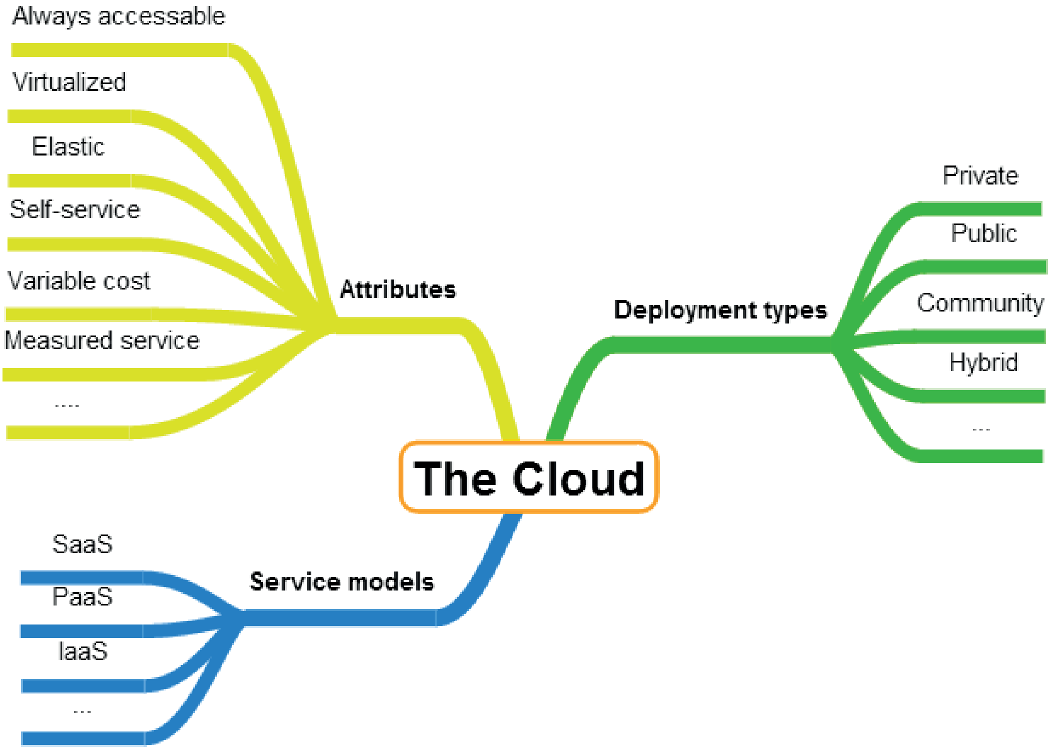

Figure 2. A mind map of the cloud. The dots indicate that there can be more terms describing the cloud. SaaS stands for Software as a Service, PaaS stands for Platform as a Service and IaaS stands for Infrastructure as a Service. Sources: (Mell & Grance, 2011; European Commission, 2010; Rhoton, 2011; Vaquero, et al., 2009)

### The common thread

In this paper the cloud's energy demand is to be examined. Exactly which service model or deployment type used is therefore not of great importance, apart from determining if a service can be identified as cloud service or not. The majority of definitions emanate from a user perspective - they are not defined depending on which hardware is used but instead of how the user experiences the service. This point of view makes it more complicated to separate cloud hardware from traditional IT hardware and the whole ICT sector. The next chapter will go deeper into this issue and try to materialize the cloud.

## The metaphor cloud


Figure 3. A picture of a natural cloud. (Jastremski, M., 2004)

The clouds appearing in nature can be seen from everywhere and by everyone on the ground if you choose to look up. Yet they are very distant. They are ever changing in size and appearance. It is no big surprise as to why cloud computing uses this metaphor. Many of the characteristics, or attributes, of the digital cloud can be applied to the natural cloud as well; elasticity, always present and always connectable - or in this case, visible.

## Presentation of reports using the three categories when defining the cloud

A newly published book discussing the cloud, *The basics of cloud computing* by Rountree and Castrillo (2014), states that NIST's definition of cloud computing is the most widely used. NIST is the National Institute of Standard and Technology in the United States. In a seven pages long report the institute explains their definition, dividing it into five essential attributes, three service models and four deployment models. This definition will be studied closely as it provides explanations of the three categories. The summarizing definition according to NIST is as follows:

*"Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g. networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction."* - (Mell & Grance, 2011, p. 2)

Another report explaining the cloud's features is found in the European Commission's report *The Future of Cloud Computing – Opportunities for European Cloud Computing Beyond 2010* published in January 2010. The report describes how cloud computing should evolve in order to benefit European society the most and which possibilities and challenges arise when trying to evolve the service. This future perspective is reflected in the definition in the terms describing how the service should operate. Similar to NIST's definition, service and deployment models and attributes are defined in the European Commission's report as well. The summarizing definition of the cloud, according to the European Commission is:

*"… a 'cloud' is an elastic execution environment of resources involving multiple stakeholders and providing a metered service at multiple granularities for a specified level of quality (of service). "* - (European Commission, 2010, p. 8)

These two reports and a book by John Rhoton named *Cloud Computing Explained* from 2011 constitute the core background material when explaining the categories of the cloud. Also some of the definitions in Table 15 in Appendix are used.

## 2.1 ATTRIBUTES

In the table below some attributes, or characteristics, of the cloud are collected. These seem to be the foundation of the cloud - what makes the cloud what it is and describes how it works. Rhoton (2011) states that the definitions of the cloud that appear in different reports do not directly collide but focus on different parts of the cloud's features. He continues to explain that the easiest way to understand the cloud is to understand the attributes it is built upon. A service having one of the characteristics should not be considered a cloud service, but the more characteristics that applies, the higher the probability that the service is a cloud service in the public sense according to Rhoton.

Table 1. Shows different attributes of the cloud.

| Attribute | Explanation | Source |
|---|---|---|
| *Elasticity/ scalability* | The computing resources are adapting to demand in real-time. If one tenant[1] demands more resources, these can be provided rapidly by using more servers and thus scale up the resources. When the demand decreases the capacity is liberated, matching the current demand. The resources seem to be unlimited to the tenant. | (Mell & Grance, 2011; European Commission, 2010; Vaquero, et al., 2009; Rhoton, 2011) |
| *Variable cost and measured service* | The service is measured to enable variable cost for the tenant. The measured resources can for example be storage, bandwidth, computing or number of user accounts. The cost for the tenants is only operational, dependent on current use of computing resources. | (Mell & Grance, 2011; European Commission, 2010; Vaquero, et al., 2009; Rhoton, 2011) |
| *Virtualization, multi-tenancy and resource pooling* | One server can be used by multiple tenants with the assistance of virtualization; the creation of multiple virtual operating systems. The tenants do not need to know exactly where their data is stored (which server, which data center or sometimes which country), enabling the cloud provider[2] to choose place according to access to resources. | (Mell & Grance, 2011; European Commission, 2010; Vaquero, et al., 2009; Rhoton, 2011; Accenture and WSP, 2010; Koshy, et al., 2012) |
| *Accessible from anywhere and anytime* | When having access to internet, the cloud is just a click away. The cloud can be accessed via computers, lap tops, smart phones, tablets or other devices with internet access. | (Mell & Grance, 2011; European Commission, 2010; Rhoton, 2011) |
| *Self-service and minimized time to market* | The process from an idea to launch can go fast if needed. There is no requirement of human contact when ordering a new service – this quickens the process of developing new services. | (Mell & Grance, 2011; European Commission, 2010; Rhoton, 2011) |

---

[1] **Tenants** are companies, organizations, authorities or persons that rent cloud services.

[2] The **cloud provider** manages and owns the data center(s) from where the cloud service is delivered (infrastructure). Cloud providers can also rent out software and platform services based on such infrastructures - this is also called vendors or adopters.

The attributes presented in Table 1 are the most common appearing in the studied reports, but there exist even more attributes describing the cloud. The European Commission's report (2010) describes moreover how the cloud should work in order to gain the society the most.

According to the European Commission it is important for cloud services to be cost efficient, the benefits for the tenant must exceed the disadvantages. The tenant should also get access to an easy understandable environment; the complexity of the cloud service should be hidden from the tenant. The tenant can then focus on creating applications instead of dealing with technical problems. The cloud provider must be aware of where the data is stored and keep all duplicates up to date. The service must also be reliable and available, meaning that no data should be lost or connections disabled and the data should be stored at multiple places, enabling easy and secure access independent of the location of the user[3]. Furthermore the service must have a short response time when scaling up and down in order for the elasticity to work properly. (European Commission, 2010)

## 2.2 SERVICE MODELS

In this section three service models will be presented, along with examples of cloud providers offering the services. These three service models seem to be generally accepted as both NIST, the European commission, Rhoton and some of the authors in Table 15 uses them.

### Infrastructure as a Service (IaaS)

When the cloud provider owns and manages the physical servers and the tenant is offered storage, network and processing power it is called Infrastructure as a Service, IaaS. IaaS are sometimes called Hardware as a Service as well. The tenant can have access to network configurations as well, such as firewalls. The software and the operating system is up to the tenant to decide (Mell & Grance, 2011). The European Commission (2010) divides IaaS into two types; computing and data & storage.

Amazon Web Service is the dominating Infrastructure as a Service on the market today according to Rhoton (2011). Another example of IaaS is vCloud Express (Rhoton, 2011). An examples of data and storage provider are Amazon S3 and examples of compute providers are Amazon EC2, Zimory, and Elastichosts (European Commission, 2010).

### Platform as a Service (PaaS)

The cloud provider offers a platform for the tenant to build its cloud on (European Commission, 2010; Mell & Grance, 2011). Some programming is needed, using the program languages, tools and services supported by the provider. The tenant does not control the underlying features such as servers, network and operating system (Mell & Grance, 2011). The created applications cannot be transferred to another cloud provider today, as the API's[4] are specific for each provider, but efforts in achieving uniform models are in progress (European Commission, 2010).

Google App Engine, Windows Azure, Salesforce.com (Force.com) are some examples of PaaS (Rhoton, 2011; European Commission, 2010). Another example is AWS Elastic Beanstalk (Amazon Web Service, 2014).

### Software as a Service (SaaS):

The provider offers an application that can be used by the tenant, i.e. no programming is needed from the tenant (Mell & Grance, 2011; European Commission, 2010). The tenant does not control the application's underlying features, such as servers, network and operating system. The exception can

---

[3] The **user** or **consumer** is a person who uses the cloud-based website or service.

[4] Application Programming Interface, see glossary.

be some basic settings acquired by the application. A web browser or similar basic tools are used to connect to SaaS. One example of SaaS is cloud-based email (Mell & Grance, 2011).

Software as a Service is the most common and oldest cloud service model according to Rountree and Castillo (2013), which can also be seen in the variety of companies offering the service. One example of SaaS is online word processing applications (Vaquero, et al., 2009). Google Apps, Apple, Cisco WebEx Collaboration Cloud, Microsoft Online, Salesforce.com and Workday are further examples off SaaS according to Rhoton (2011). The European Commission (2010) also presents Google Docs, today called Google Drive, and SAP Business by Design as SaaS.

## 2.3 DEPLOYMENT TYPES

Deployment models describe who has access to the cloud and what the purpose with it is. There are four deployment types in NIST's definition and five in the European Commission's definition. The special purpose cloud, which is unique for the European Commission, is also described as a usual cloud but designed for managing dedicated functionalities. Therefore, as the special purpose cloud could fit into the other deployment types, only the other four types are presented in this section.

### Private cloud

Private clouds can both exist at or apart from a company's location. The cloud's hardware is owned and operated by the company using it or by a third party, or both. The cloud is provisioned to satisfy one company's need, but can be used by many consumers. (Mell & Grance, 2011; European Commission, 2010). To the user, the private cloud works like a SaaS according to the European Commission (2010), an example is the internet trading site ebay.

### Public cloud

When a cloud provider offer cloud services to (several) other enterprises it is called a public cloud. The public cloud can, as the name reveals, be accessed by the general public. It can be owned and handled by a company, organization, authority or a mixture of them. The hardware is located where the cloud provider is seated. (Mell & Grance, 2011) Examples of public clouds are Amazon Web Service, Google App Engine and Windows Azure according to the European Commission (2010).

### Community cloud

The description of community clouds differ some between the European Commission and NIST, the former interpreting it as companies sharing computing resources with each other and the latter as organizations or companies with a common concern. These two definitions do not need to disagree, but they do not say the same thing either. The European Commission, though, stresses that community clouds are not common yet. NIST describes that a community cloud is used by organizations with a mutual concern, but is closed to the public. The concern can for example be a common policy or a common mission. The cloud's ownership is not of great importance for the definition nor is the location of the infrastructure. (Mell & Grance, 2011)

The European Commission states that community clouds are not common at the moment but are expected to increase. There are private community clouds that consist of small enterprises that join their server capacity. This increases their server efficiency by applying the cloud idea; sharing resources between the enterprises according to demand. There are also public community clouds which consist of several public clouds coming together, sharing resources. (European Commission, 2010)

### Hybrid cloud

A hybrid cloud is collaboration between private, community or public clouds. When needed, the data capacity can be shared between the different clouds, enabling more storage, network and processing capacity (Mell & Grance, 2011). The hybrid cloud enables the outsourcing company to maintain control over the infrastructure and data, ensuring high security and full control (European Commission, 2010). The resources that the outsourcing company offers to other tenants work as a public cloud. There are few hybrid clouds in use at the moment, but more are expected to appear according to the European Commission (2010).

## 2.4 OTHER DESCRIPTIONS OF THE CLOUD

Many reports about the cloud have been studied during the progress of this paper and some of them do not define the cloud at all before making calculations and statements of the subject. Greenpeace (2011) states that sometimes internet is interpreted as the cloud and acknowledges that there are different meanings of the term *cloud*. The definition of the cloud depends on what purpose you want to achieve when writing a report according to Kihl (2014), associate professor at Lund University. The borders of the cloud are unclear. Sometimes end user devices can be included in the definition, seen from an energy consumption view, as well as data centers and network connections. Sometimes it is preferable to only define big data centers as the cloud (Kihl, 2014).

Internet is a worldwide computing network which consists of regional as well as local networks that are connected to each other (NE, 2014 c). If the cloud was to be described as a synonym to internet the term would be very wide and all IP[5] data transfer occurring would be the cloud. Many of the definitions in Table 15 refers to the cloud as an external server space or a service being delivered across internet. These definition do however not include private, hybrid and community clouds that actually can be owned and managed by the companies using it and thus are not necessary delivered across internet. One explanation as to why the definitions differ is the fast development of cloud services – new definitions are required as the service evolve. It is like trying to embed a river, sometime it will flood and a new, even bigger, reservoir have to be built.

In Greenpeace's report *How clean is your cloud* from 2012 big data centers are considered as the foundation of the cloud. Companies like Facebook, Apple, Microsoft, Amazon, Spotify and IBM, among others are part of the cloud according to Greenpeace.

There is a strong correlation between the cloud and ICT as the cloud is a part of ICT, but some differences can be noted as well. ICT stands for Information and Communication Technologies and it is a much broader subject than the cloud (Kihl, 2014). Some sources though, refer to ICT as the cloud (Mills, 2013). The Swedish dictionary Nationalencyklopedin, NE, describes ICT as "a generic term for all the technical possibilities that has been created through progress within the computer technique and telecommunications" (NE, 2014 b).

In section 2.3 some cloud providers were presented, but these are only a fraction of all cloud providers operating. There are many other companies offering cloud services. Facebook can be taken as an example of a cloud service although all of the attributes in Table 1 however do not correspond to Facebook. The users of Facebook do not pay for the service, but instead Facebook gets its income from advertising, but according to Rhoton not all of the attributes have to match the service. The more attributes that matches a service the higher the probability is that the service is part of the cloud in the general public's point of view.

In the list of references for the European Commission's report, the NIST definition of the cloud can be found, which can explain some of the similarities of the definitions. Large parts of this chapter builds upon the NIST's definition and Rountree and Castrillo (2014) states that it is the most widely used. The NIST definition might though be the most widely used from a user perspective – but when determining the energy demand it might be more manageable to describe the cloud as big data centers and networks as Greenpeace did.

---

[5] Internet Protocol

# 3 MATERIALIZING THE CLOUD

In the previous chapter the definition of the cloud was studied and it turned out to be user oriented. The definition does not say anything about the energy use of the cloud or which hardware that is used. In order to later be able to estimate the energy use the hardware required to run the cloud must be investigated. This will be done in this chapter. In the following text the different parts of the cloud will be presented. It should be noted that this is not a complete list of all equipment of the cloud but an overview. In Figure 4 below ICT is divided into data storage & processing, data transfer and end user devices. As has been earlier mentioned, there is a strong correlation between ICT and the cloud and more exactly which part of ICT that the cloud stand for will be discussed in each section, but as can be seen in the figure end user devices might not be included in the cloud.

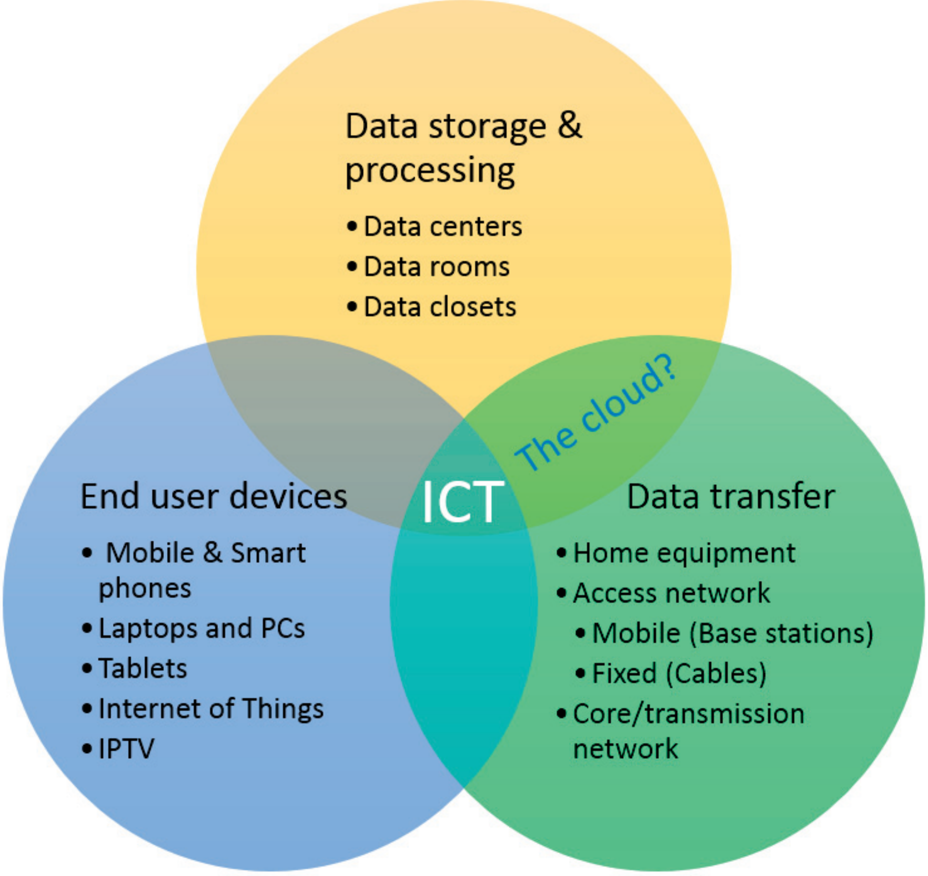

Figure 4. ICT and the cloud materialized. The question mark stresses the fact that the cloud can be defined in many ways.

## 3.1 DATA STORAGE, LOCAL NETWORK AND COMPUTATION

The core of the cloud is the data centers, they enable storage and computation. There are industrial-sized data centers as well as small data rooms and closets.

### Data centers

Data centers are clusters of servers where processing and storing of data take place. The servers are located in racks and the racks are placed in rows. A traditional data center usually runs multiple small applications on servers isolated and protected from each other. Cloud data centers, on the other hand, can be said to run as one huge computer (Barroso & Hölzle, 2009). Each server communicating with the other to enable fast services. Cloud data centers are usually owned by a single organization, often run huge applications and have similar hardware configuration. The size of a data center varies; some are so large that they can be seen from space according to Greenpeace (2012). Data centers are often described as the 21th centuries industries.

Data is often stored in multiple servers and in several data centers in order to secure the information in case of hardware failure and to reduce the transfer time to clients. A website, for example, is preferably hosted at multiple places in order to avoid data transfer all over the globe each time a customer visits the site (Barroso & Hölzle, 2009). A usual data center consist of servers, switches, lighting, Power Distribution Unit (PDU), Uninterruptable Powering Supply (UPS), cooling system and ventilation (Barroso & Hölzle, 2009). The cooling system and ventilation are essential as the servers have an optimal working temperature. All the electricity used by the servers eventually becomes heat which must be disposed of in order for the servers to have a long life time and work optimally (Sjöström, 2013). Cooling potential and a reliable grid are two subjects to consider when choosing location site for a data center.

There are also a lot of network traffic occurring within a data center, which will be further described in section 5.1, and the servers are often specialized on one task; for example storing files, hosting websites or processing data.

### Server rooms and closets

There are server closets and server rooms at many enterprises enabling storage of data and hosting websites. These rooms or closets have often been proven not to be very energy efficient as they are usually over-dimensioned in order to manage the max load that sometimes occur. They could be part of a community cloud, as have been discussed in the previous chapter, but these cloud are not very common yet. The enterprises using server rooms and server closets are however examples of possible customers to cloud providers if they choose to move into the cloud. There are many advantages with moving into the cloud but also some disadvantages, as will be described later in chapter 6.1.

### Mobile computing during night

Sony is developing a new technique in which cellphones are used for computing during night time. There are powerful processors in smart mobile phones today and if all of them are connected during night very powerful computing resources can be achieved (Holmqvist, 2014). This can also be part of the cloud - as well as data centers offers computing resources, also the cell phone can be used as such. This will however not be studied in this thesis.

## 3.2 DATA TRANSFER

The transfer of data involves several devices depending on origin and destination of the data package and choice of path. Telia Sonera and Ericsson made a life cycle assessment of ICT networks in 2010, analyzing the parts of their ICT network in Sweden (Malmodin, et al., 2011). This study will be used to identify the physical parts of data transfer.

### Home equipment

There are modems, routers, switches and gateways to enable communication in homes and offices (Malmodin, et al., 2011). Modems send and receive signals over, for example, the telephone line at higher frequency than normal analog calls, enabling digital communication. Routers navigate digital packages to the right address within a local network or to the internet. All equipment must use the same language in order for the transfer to work, and thus some rules applies when sending digital packages. The most common rule, or language, is the internet protocol, IP. Each end user device has an IP address enabling digital packages to be received and enabling routers to know where to send each package. Network switches link computers together, forming local networks. There are several forms of gateways; the default gateway is a router from where the internet is distributed to the local network. Gateway, in terms of communication, translate between "languages" when different protocols are used. Also set-top boxes can be regarded as IPTV is spreading.

### Access network

The access network connects the home network to the backbone, i.e. transmission network. There are both wireless as well as wired solutions for this. The digital transfer from a cellphone uses the wireless mobile communication network, base stations receive and transmit signals from mobile phones and transfer them. There are also subscriber databases and mobile switches. There are several mobile connection generations, each enabling faster transfer rate than the one before. At the moment the fourth generation, 4G, mobile network is expanding. Also the coverage of the other generations 2G and 3G are increasing.

The wired access network is using the telephone line, the electric grid or lines installed particularly for digital transfer (NE, 2014 d). The wired transfer is sometimes called broadband, depending on how fast the digital transfer is. There are several types of broadband or Digital Subscriber Lines, such as ADSL, VDSL and RDSL. The fastest transfer is obtained using optical fibers, which can be used for internet access, telephone and TV at the same time.

### Transmission network

Routers are used in the backbone to navigate digital packages to the receiver which, for example, can be a data center, a smart phone or a personal computer. A package travelling over the internet does seldom exceed 20 hops (Koshy, et al., 2012). One hop is the passage between two routers. The backbone in, for example Sweden, consists mainly of optical cables enabling transport of more than 1 Tbit/s (NE, 2014 d). Copper cables and radio links are also used to transmit data at some places, but are not as fast as optical cables.

## 3.3 END USER DEVICES

In the NIST definitions it is stated that the cloud should be accessible from any device with internet access. There are many devices with internet access on the market and thus the list of user devices can be long.

### Smart phones, tablets, laptops and computers

Historically only computers were used to send email and access websites but now both tablets and smart phones can do this as well. Services such as Skype also enables voice calls using the computer. The borders between the fields of application of smart phones, computers and tablets are not clear as they can perform the same services (NE, 2014 e). They are however all examples of end user devices that can connect to the cloud.

### Mobile phones

Mobile phones, in this paper regarded as regular cell phones and not smart phones, are often included when describing ICT but cannot be considered part of the cloud as little data transfer occur during usage.

### TVs and Digital Boxes

There are TVs and digital boxes which are connected to the internet to provide on demand streaming of video, VoD.

### Consoles

Online gaming is enabled with consoles like Xbox, PlayStation and Wii. The consoles are connected to the internet constantly receiving and transmitting data.

### Internet of Things

Devices in our homes and companies which are connected to the internet and send and receive data is part of Internet of Things, IoT. Fridges that regularly report temperature and electricity meters that report power usage are examples of IoT. In the future a household's things might make hundreds of status updates per day. Internet of Things is sometimes also called machines to machines, M2M. (Vermesan, et al., 2011)

## 3.4 THE DIFFERENCE BETWEEN CLOUD SERVICES AND TRADITIONAL IT SERVICES

It hard to draw a clear line between cloud services and traditional IT and the whole ICT sector. In this section it will be investigated if the definition of the cloud can help separate the cloud from traditional IT services when seen from the hardware that is needed. Note that a single data center most probably know how many of their servers is part of the cloud, but the difficulties arise when trying to estimate cloud servers on global scale.

Traditional IT can also be available and accessible from anywhere – just as the attributes describing the cloud. For example a server providing a companies' web pages to the public. One attribute that can be seen as unique for the cloud is the scalability. The cloud is adopting its resources according to demand. In non-cloud data centers servers might not be used optimally, working at low percentage of their full performance. They are instead dimensioned for their max load that might not happen so often. Studies have concluded that small traditional data centers can save as much as 90 % of electricity use by moving into the cloud, this will be described in section 6.1.

Tieto, a company offering cloud services among other things, declare that about 65 % of their servers are virtualized, even though only about 20 % of their servers are used for cloud services (Landström, 2014), see Appendix 13.4.3. Virtualization are thus used in traditional IT services as well.

It is easy to think that effective data centers are most probably cloud data centers, but this does not have to be true. The hardware can be very effective but the virtualization and the performance of the software might be bad. The efficiency measurement PUE[6] does not take software efficiency into account. However some studies have concluded that cloud data centers usually have lower PUE values than traditional IT (Verdantix, 2011).

Small data centers are also not naturally traditional IT, as they can be part of a community cloud - sharing resources with each other and thus be very effective. But as these are not very common at the moment data rooms and data closets can be regarded as traditional IT.

Apart from determining if a data center is part of the cloud, there are also data transfer and end user devices to consider. Without the data transfer the cloud would not exist and therefore routers, switches, cell towers, optic cables etc. should also be part of the cloud. But all servers are not part of the cloud and therefore all data transfer is neither. The tricky part is to determine how much of today's data transfer that can be linked to the cloud.

As can be seen in this chapter the cloud hardware is very complex and many types of equipment can be regarded as part of the cloud. End user devices are important when using the cloud – but one can think that these devices would be used regardless if the owner uses the cloud or not. These devices might be the link to the cloud but not *the cloud*. Some estimations of energy use from end user devices will however will be presented along with rest of the ICT sector to give the reader a comprehensive view in chapter 5.

---

[6] Power Usage Effectiveness

# 4 POWERING THE CLOUD

Increasing the efficiency of data centers and networks is one way to reduce the environmental impact of the cloud but there are also other factors to take into account. The carbon usage efficiency, CUE, describes how much carbon dioxide equivalents that is released to the atmosphere when using 1 kWh electricity in a data center. This efficiency is dependent on the source of the electricity, a wind power plant releases less carbon dioxide than a coal-fired plant for example. One organization which investigate data centers' electricity sources is Greenpeace. The latest report was published in April 2014, and it presents among other things data centers' use of renewable resources and their environmental goals. This chapter does not include data transfer and end user devices but note that the emissions from these sectors are also dependent on electricity source and the carbon emission factors can be applied to them as well. The placement of data centers also decide where the data traffic will be abundant and thus the electricity source of the country is important.

## 4.1 GLOBAL TRENDS

In Greenpeace latest report *Click Green* from 2014 three global companies were prominent in their environmental work; Google, Apple and Facebook. All three companies are working towards 100 % renewable energy sources for their data centers. Greenpeace states that these companies are pushing electricity companies to invest in renewable resources in order to fill their clients' need. Apple and Google are also investing in building their own new renewable energy facilities. In Greenpeace's report from 2012, *How Clean is your Cloud?,* Apple's share of renewable resources were 15 % and the company was criticized for its environmental work, but in the report published in 2014 this percentage has increased to 100 % as the company has made efforts improving their environmental status. Microsoft and Yahoo have also started to work towards more environmental friendly data centers. Yahoo invests in renewables and Microsoft has signed multiple-year contracts with renewable energy suppliers (Greenpeace, 2012).

Other companies, however, still have a long way to go when it comes to renewable sources. Amazon Web Services and Twitter are pointed out as non-transparent and have a small share of renewable resources to their data centers (Greenpeace, 2014). Overall Greenpeace conclude that the IT sector is focusing on rapid growth instead of renewable electricity sources.

Taylor (2014) answered the question if he believed that the majority of data centers worldwide will use renewable energy sources in the future. He described that he believed that more than 50 % ought to be driven by renewable energy 20 years from now, but it might take 50 years to reach over 80 % of data centers driven by renewable energy. Taylor describes that large energy storage is needed to deliver renewable energy when the sun does not shine and the wind does not blow. Other critical factors are available land for wind farms and solar energy and how fast more data centers will be built and capital costs for these investments.

Data centers have a backup system in case of power failure, the so called uninterruptable powering supply (UPS). The UPS often consist of diesel generators that can provide the whole facility with electricity. The emission when using diesel generators are large and some data center have started to use natural gas in fuel cells instead as UPS (Greenpeace, 2014).

## 4.2 GREENHOUSE GAS EMISSION FACTORS

In Table 2 some emission factors of greenhouse gases (GHG) measured in carbon dioxide equivalents per KWh ($CO_2$ eq/ kWh), are presented based on regions and electricity production methods. The values range from 3.3 g $CO_2$ eq/kWh to 1005 g $CO_2$ eq/kWh. The values presented in the table are generated electricity, if distributed electricity values would have been used the values would have been slightly higher; for Forsmark Nuclear Power the value is increased from 3.3 to 4.8 g $CO_2$ eq/kWh for example. GHG emission factor is often called carbon emission factor (CEF) as well.

Table 2. Showing Greenhouse Gas Emission Factors of different regions and production methods.
* are presented as $CO_2$ –emissions, not equivalents.
**are presented as $CO_2$ –emissions, not equivalents, and do only account for fossil fuel emissions, year 2011.

| *Electricity production* | Emission (g $CO_2$ eq/kWh) | Source |
|---|---|---|
| | | |
| **Source** | | |
| *Forsmark Nuclear power* | 3,3 | (Vattenfall, 2013 a, p. 5) |
| *Nordic Hydro power* | 8,6 | (Vattenfall, 2011, p. 4) |
| *Nordic Wind power* | 14 | (Vattenfall, 2013 b, p. 5) |
| *Solar power (Photovoltaic)* | 46 | (IPCC, 2011, p. 982) |
| *Natural gas* | 400* | (IEA, 2013, p. 43) |
| *Black coal* | 860* | (IEA, 2013, p. 43) |
| *Brown Coal* | 1005* | (IEA, 2013, p. 43) |
| | | |
| **Geographical** | | |
| *World mean* | 600 | Used in (GeSI, 2012, p. 207), provided by IEA |
| *World mean* | 536** | (IEA, 2013, p. 110) |
| *European Union* | 352** | (IEA, 2013, p. 110) |
| *Asia* | 707** | (IEA, 2013, p. 110) |
| *Africa* | 596** | (IEA, 2013, p. 110) |
| *Non-OECD*[7] | 633** | (IEA, 2013, p. 110) |
| *OECD* | 434** | (IEA, 2013, p. 110) |

As can be seen in Table 2 above the greenhouse gas emission factors differs a lot, using for example brown coal as electricity source emits more than 100 times of carbon dioxide equivalents compared to Nordic hydro power.

## 4.3 LOCATION OF DATA CENTERS

### Data center location

The majority of data centers are located in North America and Western Europe today, more than one thousand data centers using power more than 5 MW are located in the two areas. (Data Center Map, 2014; Länsstyrelsen Norrbotten, 2014). Set in relation to the rest of the world Asia has slightly less than 400 data centers and Africa and South America has about 80 data centers together (Länsstyrelsen Norrbotten, 2014). Greenpeace (2014) note that a large part of the future growth of data centers will occur in China and that it is important that IT companies take an active part to support renewable energy sources in this country. Today the electricity in China is mainly produced from coal and natural gas. Another study reports that there will be built an equal number of data centers in Europe, North America and Asia, while smaller number will be built in Africa and South America by 2020 (Länsstyrelsen Norrbotten, 2014). There are diverging opinions of where new data centers will be built in the future.

---

[7] Organization for Economic Co-operation and Development

## Data center risk index

When building a new data center there are many factors to take into consideration. Business Sweden, which is the Swedish invest and trade council, have developed an index which estimate the risk to build data center in a certain country. The lower the risk the better the conditions are for investing in that country. The index has been compiled by the help of data center experts and interviews with data center owners and users. The most important parameters for the index is energy price, international bandwidth speed and ease of doing business, see Figure 5 below. Business Sweden stresses that this is a guide for data center investors, often the business opportunities play a bigger role when deciding location, but the index can help to create awareness of the risks and prepare for them. The top three countries in the index for 2013 was US, UK and Sweden. (Business Sweden, 2013)

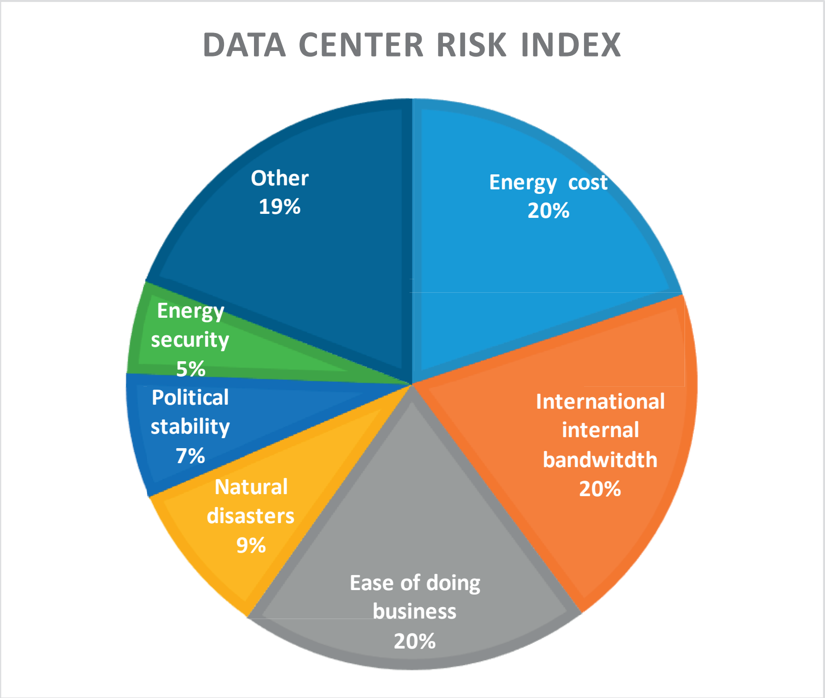

Figure 5. Importance of the different parameters in the Data Center Risk Index. Number from (Business Sweden, 2013)

# 5 QUANTIFYING THE CLOUD

Quantifying the energy demand of the cloud is a complex task as the most common definitions are defined from a user perspective and not from which hardware that is used. However, the operations the cloud performs can be divided into storage, processing and transfer of bytes. A functional unit can be defined as *number of bytes managed by the cloud during one year*. The term manage include stored bytes, transferred bytes from, to and within a data center and processed bytes. The functional unit thus includes storage, networks and processing in data centers as well as the transfer to and from the end user device and other data centers. The end user devices themselves are not included. In equation (1) below the functional unit is used to illustrate the clouds energy demand during one year.

$$E_{cloud} = E_s + E_p + E_t = (n_s \cdot B_s) + (n_p \cdot B_p) + (n_t \cdot B_t) \qquad (1)$$

Where
$E =$ Energy use per year
$n =$Number of bytes per year
$B =$Energy use per byte per year

The subscriptions describe which type of operations that occur:
$s =$ Storage
$p =$Processing
$t =$Transfer

If the greenhouse gas emissions are to be included, the energy demand is multiplied with greenhouse gas emission factors which were explained in the previous chapter. In this chapter numbers that can fit into equation (1) will be presented, but these numbers often represent the ICT sector and not the cloud alone. The lack of data is one reason as to why it is problematic to estimate the energy use of the cloud. The cloud is seldom separated from ICT. Malmodin et al. (2014) made a life cycle assessment of ICT in 2014 and also points out that it is hard to estimate the energy use per transmitted data as it is difficult to allocate energy use of different services and such estimations should be handled with caution as there are very general.

Estimates of energy use and greenhouse gas emissions of ICT will be presented in this chapter as well to get an overview of the whole system. Some forecasts will be mentioned as well but there will be more focus on forecasts and trends in chapter 0. The following headings are divided into data centers, data transfer, end user devices, ICT and the cloud. But first the unit byte will be explained:

## How much is a byte?

Table 3 below presents some examples of how much information different amounts of bytes can contain.

Table 3. Explaining how much a byte, with various prefixes, is. Sources: (Short, et al., 2011) and whatsabyte.com

| *Amount* | *Bytes* | *Approximate relation* |
|---|---|---|
| *Byte* | 1 | One character |
| *Kilobyte* | 1000 or 1024 | A paragraph of text |
| *Megabyte* | $1000^2$ or $1024^2$ | A small book or a small photo |
| *Gigabyte* | $1000^3$ or $1024^3$ | One hour high definition video or nine meters of books on a shelf |
| *Terabyte* | $1000^4$ or $1024^4$ | The largest consumer hard drive in 2008 |
| *Petabyte* | $1000^5$ or $1024^5$ | 500 billion pages of text |
| *Exabyte* | $1000^6$ or $1024^6$ | - |
| *Zetabyte* | $1000^7$ or $1024^7$ | - |

There are different ways to translate for example one megabyte to bytes - according to the binary notation one megabyte is $1024^2$ bytes while the decimal notation states $1000^2$ bytes. In glossary, chapter 12, this is further explained. As to if the different sources have used the binary notation or the decimal notation are often not stated, but one can have in mind that one zetabytes is 18 % larger using the binary notation compared to the decimal notation. In chapter 8, where future estimations of the cloud energy demand is projected, the binary system is used.

When transferred and stored, bytes are often compressed in order for the transfer to go faster and to be able to store more. Table 3 present uncompressed bytes.

## 5.1 DATA CENTERS

This section describes data centers storing capacity, computing capacity, local network transmission, energy efficiency measurements and energy demand.

### *5.1.1 STORAGE, NETWORK AND PROCESSING CAPACITY*

The International Data Corporation, IDC, offers advising services, hosts events and provides market intelligence for actors of the ICT market. Among other things IDC present statistics of for example ICT equipment shipped each year. Unfortunately these statistics are not for free and therefore second-hand sources are used below, but many of them use numbers from IDC.

#### Storage

In 2011 1.8 zetabytes of data was created or replicated according Jay Taylor from the IT Business Unit at Schneider Electric (2014). He based his statement on an IDC Digital Universal study from 2011 (Gantz & Reinsel, 2011). The same study states that this number will increase nine-fold in five years and increase 50 times until 2020. The majority of the created information will be metadata, i.e. information about information. The actual available storage of the world servers will be nearly 8 zetabytes in 2015 according to the IDC study, this can be compared to around 1 zetabyte in 2010.

The number of installed servers in the world was 32.7 million in 2010, according to Koomey (2011). He based his number on statistics provided by IDC as well. For 2014 the forecasted number of server which will be delivered are estimated to 9.2 million (IDC, 2014).

#### Processing

In 2008 9.57 zetabytes of information was processed by enterprise servers (Short, et al., 2011). Short et al. included all information delivered as input to the servers as well as all information derived from the servers, i.e. output. This means that an email might be counted multiple times as it flows though several servers. The information do not have to be new. Short et al. calculated total processed information by first classifying servers into processing type, assuming different flows according to type. Examples are database, web, mail and application servers. The flows were then multiplied with number of servers and a load factor according to type. The number of servers in the world are based on numbers from IDC and Gartner. Short et al. points out that their result differ from other studies as they have much bigger numbers. This is due to their broad definition of information and that they included servers from some big companies, such as Google and Microsoft, which normally is not part of the statistics from IDC.

#### Network

There are huge amounts of data that is transferred within data centers, exceeding the data transfer via internet many times. In 2012 the traffic within data centers were almost 2 zetabytes and are expected to reach 5.9 zetabytes in 2017 (Cisco, 2013 b).

### 5.1.2 EFFICIENCY MEASUREMENTS

A taskforce consisting of inter alia members of the European Commission Joint Research Centre, Japans Ministry of Economy, Trade and Industry, the U.S Star program and the Green Grid compiled four measurements of data center efficiency and carbon emissions. The purpose was to create unified measurements all over the globe as the expansion and thus significance of data centers energy use increase. (TGG, 2014) The four measurements are described in the following text.

#### Power Usage Effectiveness

Power Usage Effectiveness, PUE, is a common measure of data center efficiency, developed by the Green Grid. PUE describes a ratio of how much of the utilized energy that goes to the actual servers and not to cooling, lights and peripherals in a data center. The equation, (2), is as follows:

$$PUE = \frac{E_{total}}{E_{IT\ equipment}} \qquad (2)$$

If all energy goes to the IT equipment, the ideal value 1 is achieved. Very efficient data centers can have values slightly under 1.1 (Koomey, 2011 a). The average value of PUE, based on numbers from 2009, is between 1.8 and 1.89 (Yuventi & Mehdizadeh, 2013). The taskforce, described in the beginning of this chapter, says that the energy should be measured during one year. There have been some criticisms to PUE as it is a dimensionless unit that does not say anything about the total energy consumption or the origin of electricity. Some data centers wrongly advertise low PUE values as proof of sustainability according to Greenpeace (2011).

In the report *Cloud Computing – the IT solution for 21st century* PUE is estimated for traditional IT, private and public cloud from 2011 to 2020, see Table 4. The report used expert interviews to estimate PUE.

Table 4. Estimations of PUE from 2011 to 2020 dependent on deployment type. Values from (Verdantix, 2011, p. 23)

| *Deployment type* | PUE 2011 | PUE 2020 |
|---|---|---|
| *Private Cloud* | 1.8 | 1.5 |
| *Public Cloud* | 1.5 | 1.25 |
| *Traditional IT* | 2 | 1.65 |

#### Carbon Usage Effectiveness

Carbon Usage Effectiveness, CUE, is an extension of PUE that takes emissions of carbon dioxide equivalents per kWh into account. Thus the measurement includes all GHG, not only carbon dioxide, even though the name can be misleading. If the energy source to a data center is based on fossil fuels the value increases compared to if renewable electricity had been used. The ideal value for CUE is zero, indicating no carbon emissions at all from the data center. The formula is developed by the Green Grid, see equation (3) below: (TGG, 2010)

$$CUE = PUE \cdot CEF = \frac{P_{total}}{P_{IT\ equipment}} \cdot \frac{emitted\ CO_2\ eq\ (kg\ CO_2)}{per\ unit\ of\ electricity\ (kWh)} \qquad (3)$$

The value of CEF can be provided by the electricity provider or from reports covering the subject (TGG, 2010), in section 4.2 some CEFs where presented for different regions and energy sources. CUE is not adopted by many data centers yet (Greenpeace, 2012).

#### Green Energy Coefficient

The percentage green energy used by the data center is described by the Green energy coefficient, GEC. Green energy is energy produced from renewable resources as wind, waves, hydropower and the sun. To calculate GEC the amount green energy used at the facility is divided by the total energy use, see equation (4) (TGG, 2014). The ideal value is 1 - stating that all energy used are green.

$$GEC = \frac{E_{Green}}{E_{Total}} \tag{4}$$

### Energy reuse factor (ERF)

The internal reuse of energy is covered by the PUE factor, as the efficiency increases when more energy is reused within the data center. For example if an adjacent office is heated by the excess heat from the servers, no other energy is needed for heating the office and the PUE gets a lower value than if the office should have been heated by external sources. The energy reuse factor, EFR, covers reuse of energy outside the data center, for example delivering heat to a nearby city's district heating system. ERF is calculated as follows, equation (5): (TGG, 2014)

$$ERF = \frac{E_{Reuse\ outside\ data\ center}}{E_{Total}} \tag{5}$$

### 5.1.3 ENERGY USE

Jonathan G. Koomey, professor at Stanford University, is one of the few people who has made calculations of the electricity use of data centers worldwide. There are two reports covering data center electricity use published by Koomey, one in 2006 and one in 2011. Koomey's estimation of data center electricity consumption is for example used in GeSI's[8] report *SMARTer 2020* from 2012, when estimating the carbon dioxide emissions from data centers worldwide. Also Malmodin et al. (2010) uses Koomey's estimations. Greenpeace uses the *SMART 2020* and *SMARTer 2020* reports in their calculations. First hand estimates are thus not abundant.

In the report *Growth in data center electricity use 2005 to 2010* published in 2011 Koomey states that approximately 1.3 % if the worlds produced electricity was used by data center during 2010, 237 TWh. Within the United States this percentage was larger, 2 % of the electricity was used by data centers. Koomey base his estimations of installed base of servers from IDC reports and multiplies this with the amount electricity each server demands per year. Koomey includes server closets and servers rooms as well as data centers in his calculations. For network communication in the data centers he adds 15 % of server electricity and for storage he adds 24 % of server electricity. At last Koomey adds the electricity use of fans, coolers, light, other peripheral and efficiency losses, using PUE values. For more detailed explanation of Koomey's assumptions and calculations, see Appendix 13.2.

Koomey also made some calculations of Googles portion of the worldwide data center electricity use. His calculations showed that the company accounted for less than one percentage of the worldwide data center electricity use. Google itself has made some research of how much electricity one query in their search engine accounts for and arrived at 1 kJ, or 0.3 Wh (Google Official Blog, 2009)

The US Environmental Protection Agency made an estimate of US data center electricity use in 2006. The agency arrived at 61 TWh, representing 1.5 % of the total electricity consumption in the US. About 23 TWh or 38 % of the electricity usage derived from enterprise data centers, which also are the kind of data centers that are growing fastest. The agency further made a forecast of US data center electricity demand in 2011, arriving at 100 TWh. (EPA, 2007)

---

[8] Global eSustainability Initiative

## 5.2 DATA TRANSFER

Data transfer can be divided into wireless, or mobile, and wired transfer. In the report *SMARTer 2020* GeSI (2012) presented the number of subscriptions and electricity use per subscription for networks in 2011, see Table 5 below. The total electricity use is calculated by multiplying the first two columns. See Appendix section 13.3.1 for sources and assumptions.

Table 5. Showing network electricity use worldwide during 2011 (GeSI, 2012).

| ***Network*** | **Subscriptions (billion)** | **Electricity use/subscription (kWh)** | **Total electricity use (TWh)** |
|---|---|---|---|
| *Wireless Mobile Connection* | 5,9 | 17 | 100 |
| *Wired Home Access Network* | 1,65 | 45 | 74 |
| *Wired Enterprise Network* | 0,73 | 35 | 26 |
| | | | 200 |

Malmodin et al.'s life cycle assessment of ICT from 2014 presented among other things a study of the operational electricity use of ICT. The operational electricity demand of transferring 1 GB data are presented in Table 6 below, the calculations are made for Swedish conditions. The estimates include access network, data transmission, IP core network and an Atlantic submarine cable. The transfer for fixed broadband also include routers at home. It is the modems and Wi-Fi at home that constitute the majority of the fixed broadband's electricity use. For mobile transfer the energy guzzler is the base stations.

Table 6. Electricity consumption per transferred gigabyte depending on transfer type. (Malmodin, et al., 2014)

| *Transfer type* | kWh/GB |
|---|---|
| *Fixed broadband* | 0.48 |
| *Average 3G* | 3 |
| *Average 2G* | 37.1 |

Baliga et al. made a study 2009 also describing fixed broadband's electricity demand and arrived at 0,17 kWh/GB, the numbers are calculated using theoretical values and is solely based on data transfer by optic cables (Malmodin, et al., 2014). Malmodin et al. describes Baliga et al.'s study as a future state-of-the art because of the theoretical calculations.

### 5.2.1 MOBILE TRANSFER

The 2G mobile network covered 85 % of the world population in 2012, the 3G and 4G net are not as widespread (Ericsson, 2013). All forms of mobile networks are increasing its coverage though, as will be described more closely in chapter 0. In Table 7 below the coverage of the mobile networks in 2012 is displayed.

Table 7. The mobile networks coverage of the world's population in 2012 (Ericsson, 2013, p. 14)

| *Mobile Network* | Coverage of world population (%) |
|---|---|
| *2G (GSM/EDGE)* | 85 |
| *3G (WCDMA/HSPA)* | 55 |
| *4G (LTE)* | 10 |

In rural areas, where the electricity grid not is prevalent, diesel is often used as energy source when producing electricity to local grids and thus local cell towers (Greenpeace, 2012). One country where these kinds of installations are abundant is India - over half of the cell towers in India are powered by diesel (GeSI, 2012).

There were 1.7 billion mobile broadband subscriptions in the world in the beginning of 2013 and this number is expected to increase to 7 billion by 2018 (Ericsson, 2013). Mobile broadband is mainly used by smartphones but there are for example also tablets with mobile broadband subscriptions. If all mobile SIM enabled connections where to be included, i.e. not only mobile broadband, these will reach almost 10 billion in 2017 and about 13 % of these connections will be machines to machines (GSMA, 2013). If mapping the world population and how many who have SIM enabled connection this number will reach 4 billion in 2018 according to GSMA[9] (2013). The reason as to why there are less subscribers than subscriptions is that many people have multiple subscriptions and that M2M subscriptions are included. To place the mobile data transfer in perspective, GSMA describes that during 2012 more data was transferred by the mobile network than all the years before – together.

### 5.2.2 WIRED TRANSFER

Transferring data 27 000 km by cables consumes 0.2 kWh/GB according to a study by Coroama et al. in 2013. This distance is more than halfway across the globe and Coroama et al. states that their number can be used when calculating energy usage of data transfer anywhere as it is an overestimation and the data transfer will then not be underestimated. (Coroama, et al., 2013)

## 5.3 END USER DEVICES

The *SMARTer 2020* report calculated carbon emissions worldwide from *inter alia* the end user devices PCs, tablets, smart phones and mobile phones (GeSI, 2012). From the report the electricity use per device and number of devices can be seen, here represented and multiplied in Table 8, for sources and assumptions see Appendix section 13.3.2.

Table 8. Showing number of devices and electricity use from end user devices worldwide during 2011 (GeSI, 2012).

| *Device* | Quantity (billion) | Electricity use (kWh/device) | Total electricity use (TWh) |
|---|---|---|---|
| *PCs (desktops and laptops)* | 1,53 | 219,0 | 335,0 |
| *Tablets* | 0,07 | 15,6 | 1,1 |
| *Smart Phones* | 0,66 | 5,5 | 3,6 |
| *Mobile Phones* | 3,65 | 2,5 | 9,1 |
| | | | 348,9 |

A life cycle assessment which covered the environmental impact from the smartphone Sony Xperia™ T arrived at 117 kg carbon dioxide equivalents during its lifetime of three years. The number included network and accessories and assuming global electricity emission factors. The network alone accounted for 30 % of the total greenhouse gas emissions during the phones lifetime if used in Sweden. During the user phase the network emissions constituted a majority of emissions. (Ercan, 2013)

IDC have made a forecast of how many computers, servers, tablets and cell phones that are to be shipped during 2014. The complete report have not been studied, but in the abstract the numbers presented in Table 9 is shown. The forecast covered 49 countries. (IDC, 2014)

Table 9. Forecasted quantities of equipment that will be shipped during 2014 (IDC, 2014)

| *Equipment (billions)* | Quantity (billion) |
|---|---|
| *PCs* | 0.30 |
| *Tablets* | 0.27 |
| *Cell Phones* | 1.70 |

---

[9] Global System Mobile Association

Laptops manufactured during 2008 and 2009 could perform more than $10^{15}$ computations per kWh (Koomey, et al., 2011 b). The computations per energy usage increases as a constant rate, more about this in section 6.5.

## 5.4 ICT

There are several studies investigating the ICT sectors emission of greenhouse gases and electricity usage. In this section the greenhouse gas emission will be presented first and then the electricity demand will be investigated.

### 5.4.1 GREENHOUSE GAS EMISSIONS

A paper in the Journal of Industrial Ecology in 2010 presented estimates of the ICT sector's GHG emissions. The estimations were based on life cycle assessments scaled up to global magnitude, which is a bottom-up method. The results showed that ICT emitted 1.3 % of the world total GHG emissions in 2007 (Malmodin, et al., 2010). The disposal of devices was not included in the assessment, but raw material output, manufacturing, shipment and usage were. Malmodin et al. included computers, modems, phones, telecom and data networks and data centers in their study.

In GeSI's *SMARTer* 2020 report (2012) it is stated that the ICT sector emitted 0.91 Gt $CO_2$- equivalents in 2011 and this number is estimated to increase to 1.27 Gt $CO_2$ equivalents in 2020. This was 1.9 % of the world's total GHG emissions in 2011. Compared to Malmodin et al. (2010) the boundaries of ICT was the same apart from printed media, which GeSI (2012) included but Malmodin et al. (2010) did not. GeSI used second hand-sources for their calculations. Both studies used the world mean carbon emission factor of 0.6 kg $CO_2$-eq per kWh electricity. In Figure 6 below the GHG emissions in 2011 calculated by GeSI (2012) are presented, divided into end user devices, voice & data networks and data centers.

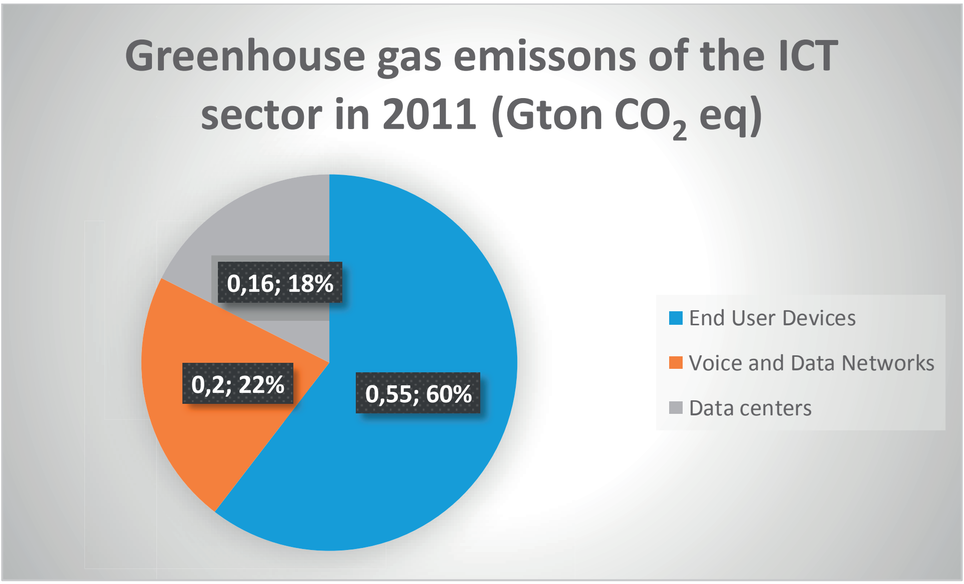

Figure 6. The greenhouse gas emission of the ICT sector according to the *SMARTer 2020* report by GeSI (2012).

The total emissions of the ICT sector are increasing each year, but the emission rate is expected to decrease in the future according to GeSI (2012); from 6.1 % increase between 2002 and 2011 to 3.8 % increase between 2011 and 2020. This depends mainly on efficiency improvement in all categories and increased usage of laptops instead of desktop computers in the end user category. The largest contributor to greenhouse gas emissions of the ICT sector is end user devices, as can be seen in Figure 6.

In an earlier report, called the *SMART 2020,* by the GeSI (2010) the emissions from the ICT sector during manufacturing, disposal and transport was estimated and compared to the user phase, see Figure 7 below. The authors mentions that only public data were used in the analysis when available, so some of the information might not be included. The results are similar to what Malmodin et al. (2010) arrived at. However, as Malmodin et al. describes in a newer report from 2014, the electricity sources play an important role when determining the hot spots of the life cycle. If a global GHG emission factor is used it is the usage phase that emits the most but for certain countries with much renewables it is on the contrary the manufacturing, disposal and transport that emit the absolute majority of emissions.

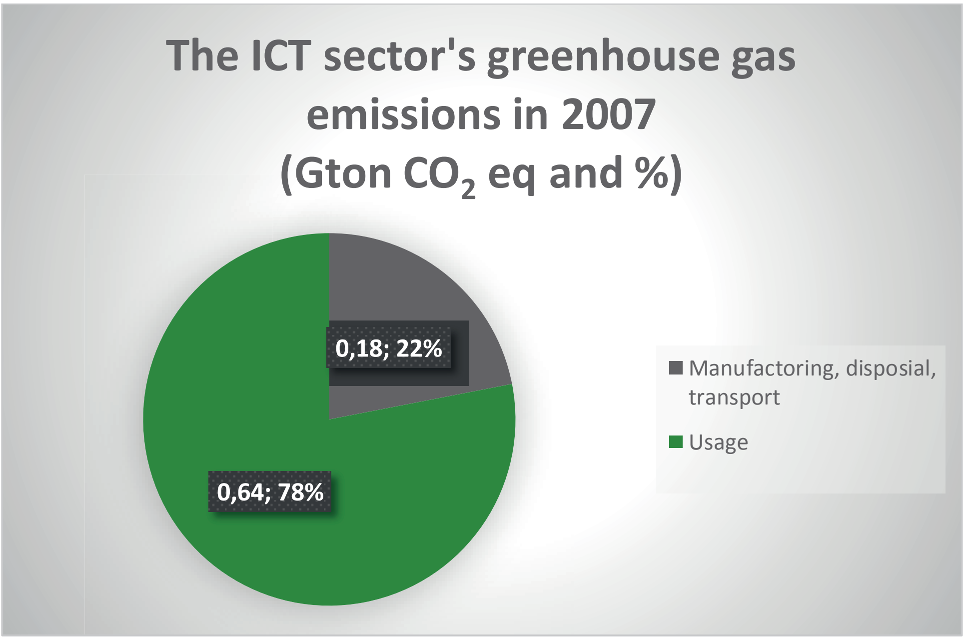

Figure 7. The GHG emissions of the ICT sector in 2007 according to GeSI *SMART 2020* (GeSI, 2010)

### 5.4.2 ELECTRICITY USE

The total global electricity consumption in 2010 was 18112 TWh (IEA, 2012, p. II.11). Three sources which have made estimations of the electricity use of ICT will be presented in this section. The different estimates are collected in Figure 8 below.

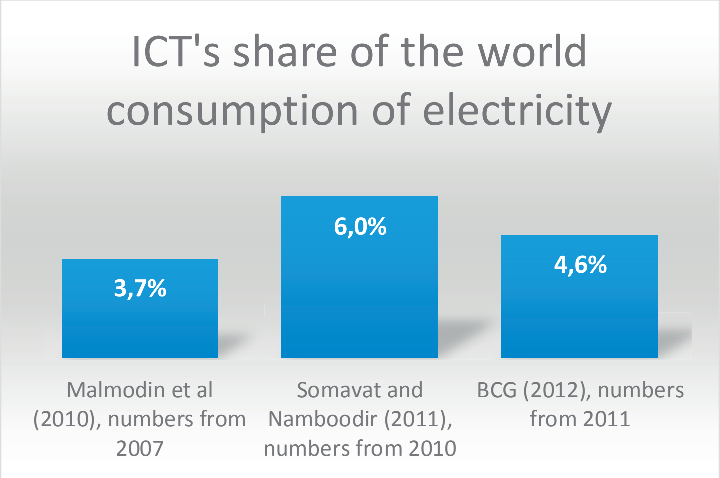

Figure 8. ICT's share of the world consumption of electricity, according to three different sources.

Malmodin et al. (2010) also calculated the operational electricity use of the ICT sector to 3.7 % of the world's electricity usage in 2007. In absolute numbers the electricity demand correspond to 710 TWh for the ICT sector.

Somavat and Namboodiri (2011) estimated the ICT sector's electricity demand to 1078 TWh during 2010. In the study the ICT sector consisted of data centers, internet and computing networks, mobile phones, laptops, desktop computers and mobile infrastructure. Only the operational consumptions were considered. This corresponds to about 6 % of the global electricity consumption according to the report. (Somavat & Namboodiri, 2011)

In the *GeSI SMARTer* 2020 report the assumptions used when calculating greenhouse gas emissions are presented. The electricity consumption of ICT could then be derived using the numbers stated, see Table 10. The exact calculations for these estimates can be seen in appendix, section 13.3. The electricity use of ICT was calculated to 4.6 % of the world's electricity consumption using the electricity demand of the world in 2010.

Table 10. The electricity use of ICT calculated using the assumptions in the *SMARTer 2020* report by GeSI (2012).

| | Electricity use 2011 (TWh) | Electricity use 2020 (TWh) |
|---|---|---|
| ***End user devices*** | 362 | 353 |
| ***Data transfer*** | 244 | 292 |
| ***Data centers*** | 237 | 441 |
| ***Total ICT*** | 843 | 1085 |

Malmodin et al. (2014) present an electricity consumption in data centers of 1 kWh/GB user traffic. If the access and transmission networks are to be included this number increases to 1.5 kWh/GB user traffic. The numbers are calculated for Sweden but the authors state that the Swedish network is a good representative of the world mean.

## 5.5 THE CLOUD

There are few reports specifying the cloud energy use. Greenpeace have in two reports calculated the energy use of the cloud based on assumptions in *SMART 2020* and *SMARTer 2020* reports. The numbers differ some compared to Table 10, which also is based on the *GeSI SMARTer 2020* report, the cause of this difference will be discussed later in the analyze of this chapter.

The report *Click Green* by Greenpeace (2014) used the estimates in GeSI's report *SMARTer 2020* to calculate the electricity demand of cloud services worldwide. Greenpeace arrived at 684 TWh to power the cloud in 2011, the cloud in this case consisting of networks and data centers. Compared to countries electricity use, the cloud would be the sixth largest according to Greenpeace (2014), see Figure 9 below.

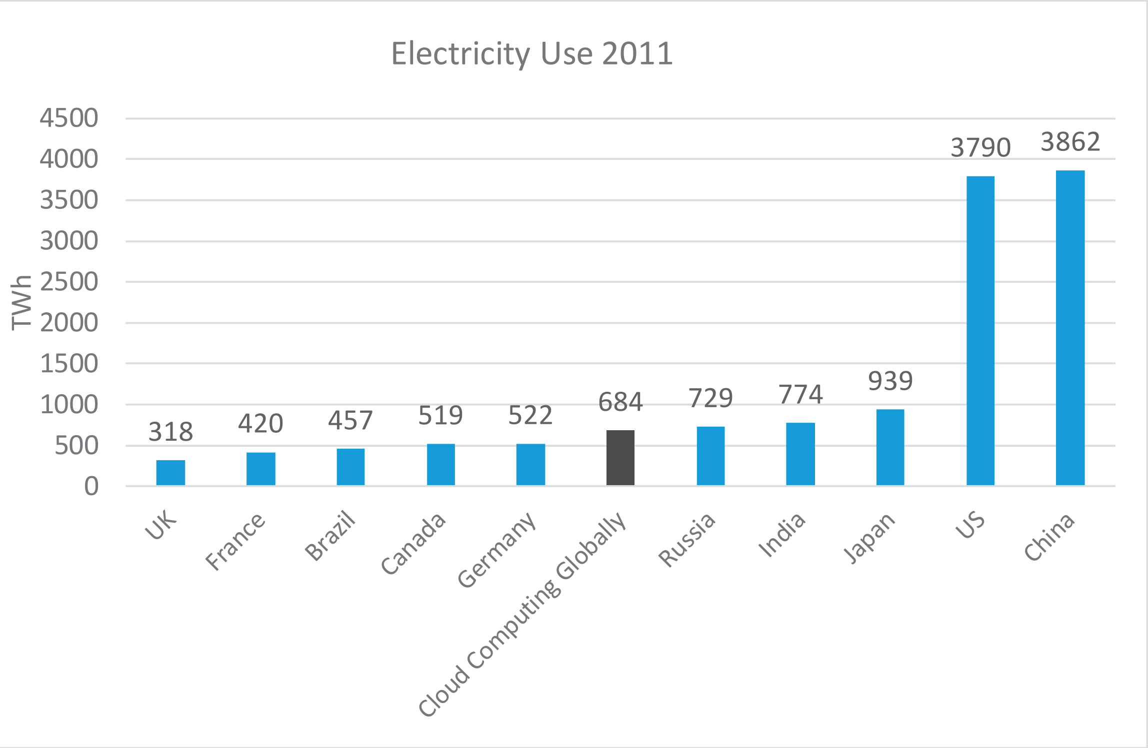

Figure 9. The countries with the largest electricity consumption in 2011 and the clouds electricity consumption globally in that same year. Statistics from (Greenpeace, 2014, p. 11)

Another Greenpeace report, *Make IT Green*, from 2010 estimated the cloud energy demand to 662 TWh in 2007. In this report the clouds electricity demand in 2020 was predicted to 1964 TWh. The prediction is based on a growth in data center of 9 % CAGR[10] from the *SMART 2020* report by GeSI (2010) and an increase in telecommunications of 9.5 % CAGR from Gartner.

Gantz and Reinsel (2011) have summarized a study by IDC and further describes that about 10 % of all servers were virtualized in 2011 and that is a step toward cloud computing, though more attributes have to be applied to fully be considered cloud computing, as pay-per-user, scalability and self-service. Today the investments spent on the cloud is about 2 % of total IT investments (Gantz & Reinsel, 2011).

A study by Koshy et al. in 2012 tried to answer the question if offloading a computation from a mobile device to the cloud, using Wi-Fi connected to broadband, could save energy. The computation was to sort numbers. Their results showed that it was more energy efficient to offload the computing to the cloud instead of letting the mobile device sort the numbers. In the basic case offloading saved just over 50 % energy. (Koshy, et al., 2012)

## 5.6 ANALYSIS

### Data centers

The number of shipped servers during 2014 are forecasted to 9.2 million. This can be compared with the total number of servers in 2011; 32.7 million. The shipment during 2014 is thus 28 % of the total number of servers in 2011. One can draw the conclusion from this that the number of servers have not reached a steady state, but is increasing rapidly each year.

Only Koomey (2011a) presents how much of the global electricity demand data centers stand for. The other reports studied in this paper referred to Koomey's reports when mentioning the global electricity demand. The estimation cannot be compared to any other which increases the uncertainty,

---

[10] Compound Annual Growth Rate, see section 8.2

more sources had been an advantage. Short et al. (2011) pointed out that the statistics from IDC normally do not include server shipments from big companies such as Google and Microsoft, which build their own special ordered servers. As Koomey uses IDC as a source the electricity use estimation might be an underestimation. Though, in a chapter of Koomey's latest report, he calculated Google's part of the world electricity use and thus actually have number on this, which indicate that his estimation might include these companies as well.

The different reports describing the electricity usage should also be set in relation to who has written the report. GeSI describes the possibilities of using ICT to reduce emissions of greenhouse gases in many sectors. Greenpeace points out the growing electricity use from data centers. These two actors have two interests diverging from one other and have therefore made different assumptions when calculating energy usage. Greenpeace is sceptic to the increased energy usage and call attention to which energy sources big data centers are using while GeSI uses a more optimistic tone.

### Data transfer

The energy usage of fixed transmission are less than mobile, in the Malmodin study 6 times more energy is used when using the 3G mobile network compared to fixed broadband. 2G uses a lot more electricity per transferred GB data than 3G. However the amount transferred data by the 3G net is much larger and thus contribute to a greater deal of the traffic. No study of how much energy is used per transferred GB via the 4G mobile network has been found. This is unfortunate as the 4G network will increase its coverage in the future and probably compose a lot of data traffic. But as 3G uses less energy than 2G per byte, this might also be true for 4G compared to 3G.

Coroama et al. (2013) arrived at 0,2 kWh /GB transferred bytes via cables and this can be compared with Malmodin et al.'s results of 0,46 kWh/GB transmitted data by broadband. The reason to the differences can be that Malmodin et al.'s value is a mean of all transfer in Sweden while Coroama et al studied a specific transfer. Malmodin et al. also included modems and routers in households, which Coroama did not, and it was this equipment which consumed the majority of the electricity according to Malmodin et al.

### End user devices

It is interesting to compare the numbers of Table 8 and Table 9 regarding shipped devices and devices in use. There are for example more tablets shipped during 2014 than was in use 2011. 1.7 billion Cellphones will be shipped during 2014. This is almost 40 % of the number of cellphones on the market in 2011 if adding up smart phones and mobile phones into one category; cell phones. But the lifetime of a cell phone is not long and this can partly explain the large numbers of new cell phones entering the market each year. There is also an increase in number of users and thus the total number of cellphones in use increases as well.

Koshy's statement that energy is saved by offloading from a mobile phone to the cloud might not be true in all cases. It depends on the connection and speed of the connection and the server used. Simpler computations might be better calculated on the phone instead of offloading and if wireless connections are used instead of wired the energy consumption of the transfer increases.

### ICT

The electricity demand of ICT differ a lot between the studies. One interesting value to compare is Somavat and Namboodiri (2011) who estimated the ICT sector's electricity demand to 1078 TWh during 2010 and GeSI (2010) whose assumptions led to 1085 TWh in 2020. The boundaries seem to be almost the same when comparing the studies, though GeSI also included printers in their estimations. Despite this the electricity demand reaches the same amount but differs ten years. This shows the uncertainties of the different estimations and that it is hard to make estimations of an ever-changing market where energy efficiency is improved each year and at the same time usage increases.

## The cloud

There are few studies specifying the cloud energy usage, and this is not too surprising as there are different views of what the cloud consists of. It is easier and more convenient to specify mobile networks', data centers' or PCs' energy usage than the cloud.

Greenpeace (2011 and 2014) used estimations from GeSI to calculate the cloud's electricity demand for the years 2007, 2011 and 2020. The cloud was assumed to consist of data centers and networks in Greenpeace's reports. Greenpeace, however, does not explain how they calculated the numbers in both reports and even though the author of this report has sent an email to them for explanation no answer has been received. Greenpeace's estimations of the cloud energy demand in 2011 was 684 TWh and 1964 TWh in 2020. If data centers and networks electricity usage in Table 10 are added up, based on GeSI's assumptions, the number 481 TWh is reached for the year 2011. This is much smaller than what Greenpeace arrived at for the same year. It is hard to tell how they made their calculations and it is unfortunate that they chose not to disclose them. Greenpeace might have made other assumptions that enhance the electricity demand. One can also have in mind that Greenpeace reports point out the environmental input the cloud have while the reports by GeSI point out the benefits the ICT sector offers. Diverging interest might lead to diverging input assumptions.

# 6 TRENDS AND FORECASTS

The cloud is growing as the number of tenants and users increases while the storage and processing each tenant uses increases as well. At the same time energy efficiency improvements and indirect saving lowers the specific energy usage. Consumers and companies are also becoming more conscious and start to demand more sustainable services (GreenTouch, 2013). The question is if the total energy usage will increase or decrease in the future, and how much. In this chapter the driving forces for and against moving into the cloud will be addressed and future predictions of energy efficiency, data transfer, processing and storage will be investigated.

## 6.1 DRIVING FORCES...

### *6.1.1 ... FOR MOVING INTO THE CLOUD*

The drivers for moving into the cloud are here divided into economic and environmental benefits. There could also be other drivers as laws requesting low GHG emissions, as the cloud often use less energy than traditional IT services this can lead into increased usage. Regulations of the cloud will be described later in chapter 7. Some of the advantages of using the cloud is embedded in the definition of the cloud, as scalability, reduce time to market and virtualization.

#### Environmental benefits

Using the cloud causes direct emissions of *inter alia* GHGs and demands energy, but at the same time even more emissions and energy might be saved by the indirect effects. The indirect effects arise when, for example, arranging cloud-based video and audio links instead of travelling physically.

#### *Indirect savings*

ICT is, as discussed before, a broader term than the cloud, but some of the indirect savings applied to ICT can belong to the cloud as well. In the report *SMARTer 2020* the total global emission savings by using ICT was estimated to reach 9.1 Gton carbon dioxide equivalents by 2020 (GeSI, 2012). The abatement potential can be compared with the annual direct emissions from ICT which GeSI estimated to 1.27 Gt $CO_2$ equivalents in 2020. This means that ICT have the potential to reduce GHG emissions 7 times its own emissions.

With real time traffic data, drivers can use their mobile phones to avoid queues caused by accidents, roadwork or rush hours. This reduces the time on the road and reduces the emissions. ICT can help to introduce renewables to the electricity grid by forecasting production from wind and solar power. ICT is an essential part of the smart grid as well. Farmers can through satellite pictures see how their crops are doing and optimize the amount pesticides and fertilizers in order to create a larger yield. (GeSI, 2012)

Instead of traveling, video conferencing can be an alternative. Arranging a video conference instead of traveling by air or by road saves a lot of carbon dioxide emissions (GeSI, 2012; Coroama, et al., 2013). Also to sell products via internet reduces travel as the byer does not have to go to the store and the product does not have to be brought to the store, instead the product can be sent to the costumer directly from the manufacturing place or the central warehouse. Some products such as films and music do not have to be physically transported at all as they can be transferred directly via internet to the customer. (GeSI, 2012).

#### *Direct savings*

Big data centers are more energy efficient compared to small on premise data servers. Data centers on industrial scale operate their servers at optimal temperature and usage in order to save energy as the energy cost is an essential expense of their business. Small enterprises can reduce their GHG emissions by 90 % when moving into the cloud instead of having their own servers (Accenture and WSP, 2010). Koshy et al.'s (2012) study, earlier mentioned, found out that offloading a task to the cloud could save energy compared to letting the mobile phone itself perform the task, in this case sorting a large amount of numbers.

### Economic benefits

Cost reduction is achieved when moving into the cloud as the IT department can be reduced and no big investments in infrastructure are needed. You pay for what you use, an operational cost, which reduces the overall cost. (Verdantix, 2011) The cloud provider manages installation of new hardware and upgrades programs instead of each company having their own IT management doing this (Verdantix, 2011) . This reduces the costs for the tenant. The level of how much the cloud provider manages is decided by which services model the tenant has signed up to.

A company´s local data server must be dimensioned for its maximum load (Verdantix, 2011; Kihl, 2014; Accenture and WSP, 2010; Rhoton, 2011). It is hard to predict the maximum load and it is therefore often overestimated in order to be on the safe side. Cloud computing's scalability solves this problem when the company moves into the cloud (Accenture and WSP, 2010). In the same way that water pipes are not dimensioned for everyone in a building complex to use their tap water at the same time, big data center do not need to be dimensioned for each tenants maximum load happening at the same time. Bigger data centers can also turn off unused servers and look at their customers work pattern in order to find companies that has peaks at different hours (Kihl, 2014). Hence, scalability result in less amount of servers needed and thus saves energy.

Some companies uses private cloud to save energy and costs. This means that the different departments of a company use the same data center, protected by a firewall, with the cloud's functionalities. The scalability is then utilized within the company. (Verdantix, 2011)

The time to market is decreased using the cloud. To buy and install new resources can take over a month for regular companies; using the cloud dramatically reduces this time to seconds (Verdantix, 2011). The lack of human contact when ordering new services also speeds up the process.

### *6.1.2 … AGAINST MOVING INTO THE CLOUD*

There are some concerns for a company when choosing whether or not to move into the cloud. One of them is security; fearing security breach and leakage of information. This particularly concerns financial and other institutions that must have very high security. Companies are also worried that the knowledge of how their systems work will not stay inside their company, and instead they start to be very dependent on their cloud provider. If the cloud provider decides to raise the cost for the service, the tenant does not have so much choose but to agree if dependent on the cloud provider. Another reason not to move into the cloud is failure to provide the service from the cloud provider, as happened for Amazon Web Services in 2011 when services went very slow, affecting a lot of clients. (Verdantix, 2011)

## 6.2 THE CLOUD

Even though the cloud energy use is not specified very often and there are many definitions of the cloud, Cisco has made some calculations of the clouds future development. Cisco uses the NIST's definition of the cloud when estimating the future of the cloud (Cisco, 2014). 66 % of all workloads and 66 % of global data center traffic will belong to the cloud in 2017 (Cisco, 2013 b). The data center traffic of the cloud will grow fast, with a CAGR of 35 % between the years 2012 to 2017. The data center traffic included traffic within data centers, between datacenters and between end user and data centers. The majority of the traffic will occur within the data centers.

## 6.3 DATA CENTERS

The number of data centers worldwide are increasing meanwhile they get more efficient. This section will describe the increased need for storage and processing as well as efficiency improvements incentives.

### Storage

As has been mentioned before data managed by enterprise data centers are expected to increase 50 times from 2011 to 2020. In 2011 the available storage capacity was 1.8 zetabytes (Gantz & Reinsel, 2011). Taylor (2014) thinks that we only have scratched the surface of how much storage is needed in the future. The IDC report by Gantz & Reinsel (2011) states that 20 % of the information in the world is expected to be stored or processed by a cloud sometime during its lifetime in 2015 and about 10 % of the information will be stored in the cloud.

The amount of files data centers handle are expected to increase 75 times during the next ten years, meanwhile the IT professionals maintaining the data centers are expected to increase 1.5 times (Gantz & Reinsel, 2011). Gantz and Reinsel states that the solution to this problem is a flexible environment; namely cloud computing.

Cloud services will increase according to Maria Kihl, associate professor at Lund University (2014). More of our lives are on the internet and with wearable, as bracelets that record your health and mobile phones that collect information of your achievements on your weekly running round, the information increases even more. Kihl asks herself if there should be a limit of how much we can be store. For example there should be time limits on how long a YouTube film is stored according to Kihl.

### Energy Efficiency

The driving force for data centers to be more energy efficient is to save money (Kihl, 2014; Taylor, 2014; Bergquist, 2014). Jay Taylor from the IT Business Unit at Schneider Electric describes that the costs is divided into operational costs and capital costs. The operational cost is expected to rise as the electricity price will increase two or three times in the next two decades according to Taylor. The capital costs consist of among other things buildings, servers, infrastructure and the renewable energy system. In the future the demand for digital storage will increase as government start to store medical records digitally, Smart Cities are developed, Smart Traffic initiatives are applied, Internet of Things is implemented and the use of computing Big Data, for example to analyze weather and public health, increases. As both the electricity cost and the demand for storage and processing power are anticipated to increase, this creates incentives for making data centers more energy efficient. (Taylor, 2014)

In year 2006 the US Environmental Protection Agency started to raise the question of more energy efficient servers and Koomey (2011) states that it is from this point server manufactures started to improve the energy performance of their products. The increased electricity usage from data centers doubled between the years 2000 and 2005, while “only” increasing with 56 % between the next five years, 2005 to 2010. Koomey explains this change in growth by the economic crises in 2009, more energy effective servers, the adoption of effective cloud computing and the virtualization of serves. Virtualization enables multiple users to use the same physical server at the same time through a software application, creating virtual servers for each user. This means that fewer numbers of servers are needed. (Koomey, 2011 a)

The cost of storing and processing data is one sixth in 2011 compared 2005, this is due to efficiency improvements and better compression of files (Mearian, 2011).

### Open Compute Project

Open Compute Project is a platform where engineers can share their best practice of how to build the most efficient data centers using scalable computing. The site is publically available and anyone how wants can join and share ideas. Facebook announced the Open Compute Project in 2011 and is so far the only company that describes how they built their data center, in this case in Oregon, to be efficient. (Open Compute Project, 2014)

## 6.4 DATA TRANSFER

The network system company Cisco has made a forecast of future global annual internet traffic, arriving at 1 zetabytes by 2015 and 1.4 Zetabytes by 2017 (Cisco, 2013 a). Cisco's method was a combination of using statistics, estimates and analysis. The statistics were for example number of subscriptions and broadband connections which were provided by several companies. Cisco themselves estimated for example the time a user watches a video and the average amount of data transferred during watching.

Internet data traffic has grown 100 times during the decade 2000 to 2010 but are expected to grow only 16 times during the next decade, between 2010 and 2020 (Korotky, 2013). The growth rate is thus expected to decrease. Korotky's method was to make a hyperbolic curve to fit the global data traffic from 1990 to 2010 and then determine the global data traffic for 2020 following the curve. In the year 2020 his estimations of annual data traffic was 3 zetabytes. Further following the hyperbolic curve, traffic are expected to increase even more before reaching a steady state (Korotky, 2013). Korotky used among others sources a forecast by Cisco between the years 2012 and 2016 as a base in his calculations.

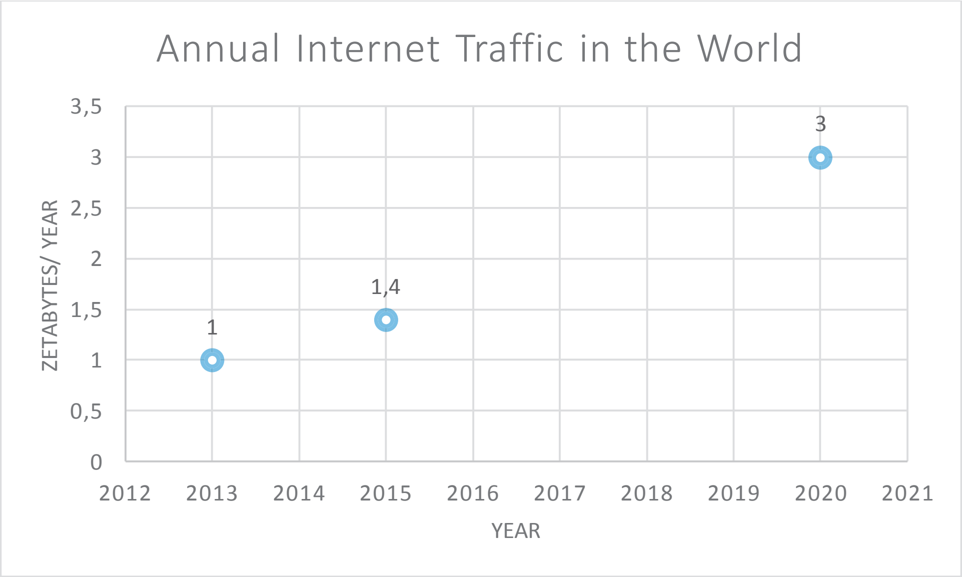

Figure 10. Shows estimates of the annual internet traffic in different years. For the year 2013 and 2015 the source is Cisco (2013) and for the year 2020 the source is Korotky (2013).

In Figure 10 above the internet traffic from user to data center and between users are included but not the traffic within data centers and between data centers. For 2017 the traffic between data centers was estimated to 0.5 ZB and the traffic within data centers to 5.9 ZB (Cisco, 2013 b). There are some trends of internet traffic usage observed by various reports, here compiled in Table 11 below.

Table 11. Internet traffic trends according to Cisco (2013) and Ericsson (2013).

| *Trend* | More specified | Sources |
|---|---|---|
| *The majority of internet traffic will consist of video, and it is the fastest growing sector of internet traffic.* | Cisco - 73 % by 2017<br>Ericsson – 50 % by 2018, annual growth of 60 % | (Cisco, 2013 a, p. 2)<br>(Ericsson, 2013, p. 12) |
| *Other devices than PCs will constitute the majority of internet traffic.* | Cisco – nearly 50 % will origin from non-PCs (TVs, tablets, mobiles, machines to machines and other devices) | (Cisco, 2013 a, p. 2) |
| *Wireless and mobile connections will be used more than wired.* | Cisco – by 2017<br>Ericsson – In many markets already today | (Cisco, 2013 a, p. 2)<br>(Ericsson, 2013, p. 6) |
| *Fixed global broadband speed will almost quadruple.* | Cisco – by 2017 | (Cisco, 2013 a, p. 9) |

One reason to the increased traffic is that consumers tend to use online applications instead of offline, which substantially increases traffic. Especially streaming games where graphics are transferred over the internet instead of being processed on the players own computer. Cisco expects that this category will make a large part of internet traffic if streaming games gain in popularity. (Cisco, 2013 a)

### IP addresses

IP addresses, which were developed in the 70s, are not enough to provide todays users with unique addresses (NE, 2014 c). More people, more machines and more things get connected, all needing a unique IP address. Therefore a new version of IP addresses, IPv6, which can have far more unique users have been developed. In Cisco's forecast mentioned earlier 12 % of fixed and mobile connection and devices were IPv6 capable 2012 and this number will increase to 42 %, or 8 billion, in 2017. (Cisco, 2013 a). There were 2.4 billion people connected to internet year 2012 (NE, 2014 c).

### Mobile broadband coverage

More than 90 % of the world's population will be covered by the 2G mobile network in 2018. 3G and 4G will cover 85 % respectively 60 % of the world's population in 2018 (Ericsson, 2013, p. 14). The 4G network exist on all continents except from Antarctica. Mobile broadband subscription will in some countries be more abundant than wired broadband connections (Ericsson, 2013). In 2013 the 4G network constituted 30 % of the mobile traffic, despite the fact that only 3 % of mobile networks were 4G networks (Cisco, 2014). Cisco states that 4G traffic will exceed 3G and 2G traffic by 2018. Seen to number of subscriptions 4G will not increase in the same amounts, only about 10 % of the subscriptions are forecasted to be 4G subscriptions by 2017 but 3G and 4G together will account for more than half of all subscriptions that same year. (GSMA, 2013)

### Increased network efficiency in the future

GreenTouch is a consortium of companies with the mission to provide a roadmap of how to accomplish energy efficient networks. The cooperation was founded in 2010. GreenTouch compiled their first results in the report *Smart Meter Research Study* published in June 2013. Among the results was the statement that network traffic can be 90 % more energy efficient in 2020 compared to 2010 even though total traffic will increase exponentially. (GreenTouch, 2013)

There are several assumptions in GreenTouch model and the consortium explains that their model for 2020 is not fully realistic as to be in operation 2020. They have not included standardization and development times but they state that it is possible to achieve this energy efficient network in time. The GreenTouch estimated that mobile networks could be 1043 times more energy efficient by 2020 compared to 2010. The corresponding number for fixed access networks and core networks are 449 and 64, respectively.

The roadmap to this energy efficient network consist of several improvements and innovations. There is no single solution as how to reach this energy efficient network but instead several solutions are proposed. Among the improvements are sleep modes, new energy efficient components, optic cables and optimization of traffic. GreenTouch also predicts that the mobile infrastructure will be shared in the future, in their scenario for 2020 they calculate for a scenario of cooperation between mobile operators and one single infrastructure. This leads to lower electricity demand as the overlap occurring today no longer exist and thus fewer base stations are needed (GreenTouch, 2013).

An example of efficiency improvements concerning mobile data transfer is the company *Eta devices* which is developing a new amplifier to base stations. Today the amplifier usually consumes 65 % of the energy in base stations but with this new technology the total energy consumption can be reduced by half. The company states that today almost 1 % of the world electricity were consumed by base stations in 2012. (Talbot, 2012)

## 6.5 MOORE'S LAW

In 1965 G.E Moore states that "The complexity [of integrated circuits] for minimum component costs has increased at a rate of roughly a factor of two per year.'' Moore´s statement have been transformed several times and proven to apply to energy efficiency as well. Koomey et al. (2011b) analyzed computer efficiency from 1945 to 2010 and arrived at one main conclusion: the energy efficiency of PCs improves every 1.5 year – for the same amount of energy the number of computations doubles. This trends is likely to continue in the future according to Koomey et al. The trend towards smaller chips and transistors in order to fit more computing into smaller devices also reduce energy usage; a smaller transistor uses less energy than a bigger. (Koomey, et al., 2011 b)

## 6.6 INCREASED CONSUMPTION OF BYTES

If consumed information was to be expressed in compressed bytes the consumption increased with 350 % from 1980 to 2008 (Short, et al., 2011) In Figure 11 below the number of compressed bytes consumed related to task is illustrated for an US citizen in 2008, this corresponds to 34 gigabytes per day and person. The circle diagram have been calculated by Short et al. by estimating number of hours spend on each task each day and multiplied with an average consumption of bytes depending on task.

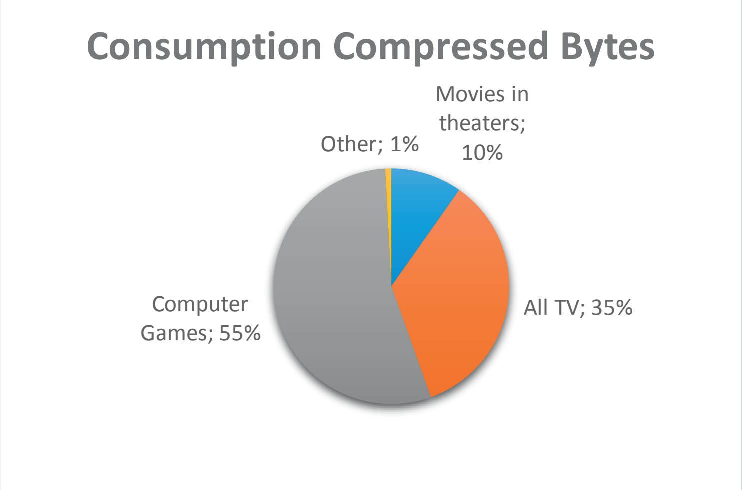

Figure 11. US citizens' average consumption of compressed bytes, total 36 GB per day in 2008 (Short, et al., 2011).

It can be seen that moving pictures, i.e. video as computer games, movies and TV constitute almost all consumption of bytes. Information consumed during work is not included. The section “other” consist of reading magazines and book, using computers to other things than playing games, listen to music, talking on the phone and listen to radio.

If everyone would engage in online gaming and on-demand IPTV, these compressed bytes would have to be transferred via internet, creating a huge demand of fast connections that can handle enormous quantities of data each day and data centers that can process and store the information.

## 6.7 THE WORLD’S FUTURE POPULATION NUMBER

The growth of the human population is also a measurement to take into consideration when estimating the future development of cloud computing. The more users - the higher the energy demand. In Figure 12 below the world population is shown, numbers provided by the United Nation’s Department of Economic and Social Affairs. Three projections are made; one low, one medium and one high. The highest indicate a population over 16 billion people in 2100 and the lower indicate a population under 7 billion the same year. There are thus big differences in the projections.

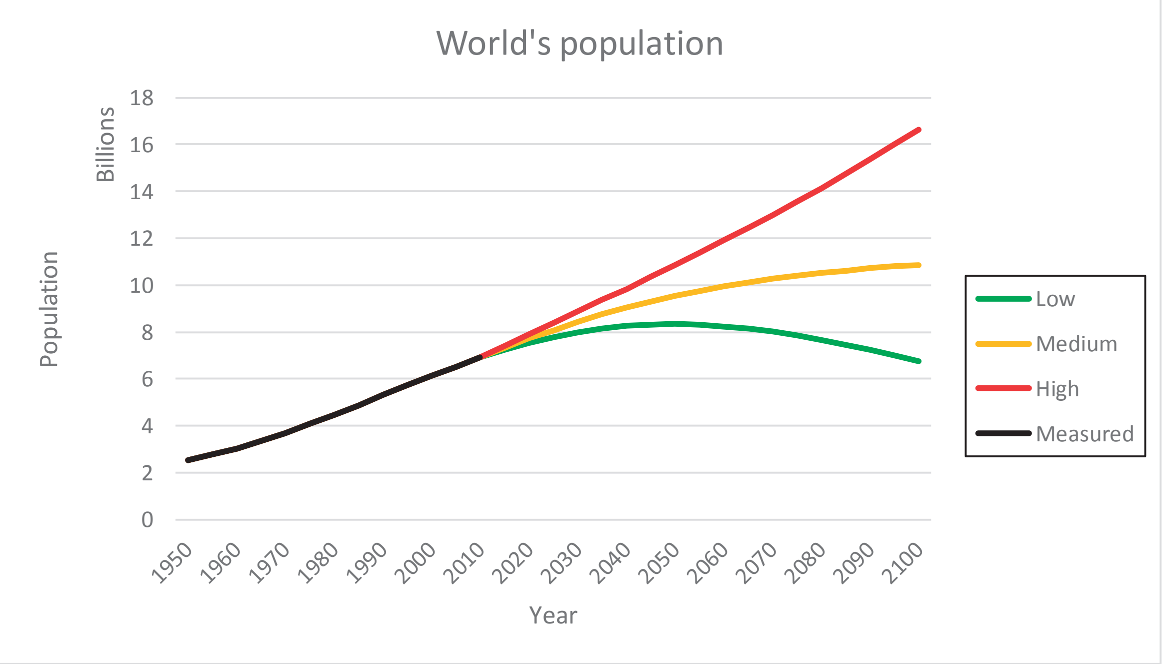

Figure 12. Shows the worlds measured population until 2010 and then three forecasts of how many people might inhabit the world in the future. The numbers are provided by the United Nation’s Department of Economic and Social Affairs (UN DESA, 2014)

Another aspect to consider is how much electronics and bandwidth each person consumes. As has been seen before the usage of mobile broadband increases and it is not uncommon to own two mobile phones, one computer and one tablet per person in some parts of the world. Even if the world population will decrease in the future the total usage of ICT will most probably continue to increase due to the increased usage.

# 7 REGULATING THE CLOUD

What regulations exist for limiting the energy usage of the cloud and ICT? The US Energy Star is a voluntary program to increase energy efficiency that is used all over the world. In Europe there is the Ecodesign directive that among other things also aims at improving energy efficiency. The subject of making ICT more energy efficient has received more attention, most recently by IEA's[11] report *More data, Less energy* from 2014. Meanwhile ICT contributes to reduce emissions globally by the indirect effects stated in chapter 0, the connected devices are very energy inefficient themselves – often consuming the largest amounts of electricity when idle in standby mode (IEA, 2014). IEA states that there can be 100 billion such network-connected devices by 2030 compared to 14 billion in 2013 and the need for regulations is huge. GSMA (2013) is on the same track and believes that it is also important to make regulations predictable and transparent, considering the increase in mobile broadband and connections which will demand investments in infrastructure. The investors will not hesitate to build new infrastructure that can support increased data transfer if the future regulations are transparent and predictable. The market for mobile subscriptions are very competitive and the cost for subscribers have decreased each year. This creates small margins for the network operators to invest in new infrastructure since the usage increases but the income from each subscriber lags behind. (GSMA, 2013) Malmodin et al. (2014) point out the end user devices, including routers and switches, and data centers as the main areas where efficiency improvements can lower the energy demand of ICT the most.

## 7.1 ENERGY INTENSIVE INDUSTRY

Electricity intensive manufacturing industry can get exemptions from electricity taxes in Sweden, in order to make the companies competitive towards companies in other countries which have lower electricity cost. In exchange the company will improve its energy performance and work toward lowering the total energy demand. However, no such benefit exist for data centers, even though they use a lot of electricity and the market is very international (Björklund, 2014; Bergquist, 2014). In 2012, however, the program for tax reduction are not possible to participate in any longer as the law was suspended but the Swedish Energy Agency is investigating new ways to help improve the energy efficiency of the industry (Energimyndigheten, 2014). Landström (2014), employee of Tieto in Sweden, states that data center industry is an immature market and there is no regulation as to how energy efficient data centers have to be. Bergquist (2014), employee of Bahnhof in Sweden, explains that it is in their own interest to have energy efficient data centers as it reduces energy costs. A newly released article in the Swedish magazine *Dagens industri* states that the Swedish government reviews a proposal of lowering the electricity taxes for data centers in Sweden in order to attract more data center companies (Svensson, 2014).

## 7.2 ENERGY STAR PROGRAM

In 1992 the US Environmental Protection Agency launched the Energy Star program, which is a voluntary program for identifying and encouraging energy efficient products and buildings. The Energy Star label sprung from the program in order to make it easier for companies and customers to identify energy efficient products and thus reduce their carbon emissions. A product labelled with the energy star is energy efficient compared to conventional products and have been tested and approved in third party labs. The requirements are tightened regularly in order to continue the development of even

---

[11] International Energy Agency

more energy efficient products. Examples of products that can earn the Energy Star label are computers, servers, UPS[12], table boxes, television, routers and modems. (Energy Star, 2014)

In 2010 data centers where included in the energy star program and can earn a label if rated above 75 on a scale of 1 to 100. The score is calculated by dividing the actual PUE value with a predicted PUE value, the lower the ratio the higher the score. The predicted PUE is based on average PUEs of hypothetical buildings that operate the same way as the data center evaluated. (Energy Star, 2014)

## 7.3 ECODESIGN DIRECTIVE

In Europe the Ecodesign Directive sets mandatory requirements on products in order to lower the energy usage and emissions. The requirements can for example include life time, energy usage both in standby and active state and usage of hazardous substances. The objective is to exclude the most energy consuming products from the market. From the 1st of July 2014 the Ecodesign Directive also covers computers, servers and tablets. The products must fulfill the requirements of the Energy Star label or else they cannot be sold in Europe. The Swedish Energy Agency stresses that especially computer manufacturers did not have any initiatives to lower the energy demand before as the experience of the product is not related to energy consumption. Tablets, phones and laptops, on the other hand, need long battery time for the customers to be satisfied, and this creates an incentive to make the products more energy efficient. For servers it is only the power supply unit that will be demanded to have the Energy Star label. (Energimyndigheten, 2014)

---

[12] Uninterruptable power supply

# 8 FUTURE PROJECTIONS OF THE CLOUD

In this section future electricity demand from the cloud and ICT will be calculated based on growth rates and numbers from previous chapters. Two approaches will be used, one bottom-up and one top-down.

The compound annual growth rate, CAGR, is often used to describe growth rates of ICT and will be used here to calculate future electricity demand. To calculate CAGR the following equation is used:

$$CAGR = \left(\frac{E_{End\ value}}{E_{Start\ value}}\right)^{\left(\frac{1}{n}\right)} - 1 \qquad (6)$$

$E_{End\ value}$ = The electricity demand in the most recent estimation/prediction
$E_{Start\ value}$ = The electricity demand in the oldest estimation
$n$ = Number of years between the two estimations

To predict future electricity demand equation (6) is transformed to equation (7).

$$E_{End\ value} = E_{Start\ value} * (1 + CAGR)^n \qquad (7)$$

The cloud is difficult to separate from ICT when estimating energy usage. IDC said in a study previously mentioned, that around 20 % of all data will be touched by the cloud in 2015 and 10 % will be stored in the cloud. Cisco estimated in another study that 66 % of all workloads and 66 % of all data center traffic will belong to the cloud in 2017. These numbers will however not be used when calculating the future energy demand of the cloud and ICT, but they can be kept in mind to know approximately the clouds part of the energy usage. The reason as to why not including the numbers is that they will change over time as the cloud increases its portion of ICT and this increase is hard to predict.

Note that there is a difference between data center traffic and IP traffic. Data center traffic includes traffic *within data cen*ters as well as *between different data centers* and *between end user and data center* while IP traffic includes *traffic between end users* and *between data centers and end users*. The data center traffic is much larger than IP traffic, especially the traffic within data centers that represent 77 % of all data center traffic (Cisco, 2014).

## 8.1 BOTTOM-UP

In this bottom-up analysis all data traffic and data centers, including data rooms and data closets, will be included. Saying that the calculations represent the cloud is actually an overestimation, but at least the clouds energy demand will not be underestimated when it comes to what is the cloud and what is not. There are many other uncertainties that makes the result unsure and it is not meant to illustrate the absolute numbers but to give the reader an overview of the future development. The end user devices are not taken into account as it is assumed not to be part of the cloud. Equation (1) from chapter 5 describe one way to calculate the cloud's energy demand and it will be used in this section.

$$E_{cloud} = E_s + E_p + E_t = (n_s \cdot B_s) + (n_p \cdot B_p) + (n_t \cdot B_t) \qquad (1)$$

Where
$E$ = Energy use per year
$n$ =Number of bytes per year
$B$ =Energy use per byte per year

The subscriptions describe which type of operations that occur:
$s$ = Storage
$p$ =Processing
$t$ =Transfer

To calculate the total energy usage of the cloud peripherals must be included. This is done using PUE values on the IT equipment belonging to datacenters according to equation (9) below.

$$E_{total} = E_{cloud} * PUE \tag{9}$$

### 8.1.1 DATA CENTERS

Data centers energy usage can be divided into storage, processing, transfer and infrastructure's energy usage. This section will describe the different parts of data centers energy usage, using equation (1).

#### Storage

The first term in equation (1) is energy use storage per year, $E_s$, and it can be represented by multiplying stored bytes each year, $n_s$, and energy use per stored byte, $B_s$. First the number of stored bytes will be investigated. Servers worldwide had a total storage capacity of 1 ZB in 2010 and will reach nearly 8 ZB in 2015 according to IDC. This corresponds to CAGR of 48 %. It is also predicted that the amount created information 2020 will reach 50 times the amount created in 2011, corresponding to a growth rate of 51 %. These two different CAGR's will be used in two scenarios when estimating future energy demand.

A modern enterprise hard drive[13] at 3 TB is used for calculations. The hard drive uses 7.7 W when idle (Seagate, 2012). The power consumption when in read/write mode is higher and this must be taken into consideration. The magnitude of this increase in power consumption can vary between 20 and 60 %, depending on type of hard drive used (Li, et al., 2012). It is assumed that the average power consumption increases with 40 % when the server is in read/write mode and that the hard drive is in idle state half of the time and in active read/write half of the time. The servers are assumed to be always on. The mean power consumption of the studied enterprise hard drive is then 9.24 W. This corresponds to an energy consumption, $B_s$, of 0.027 kWh/GB and year.

#### Processing

Processing of information is often measured in computations per second. An average laptop in 2009 performed $10^{15}$ computations per kWh. Unfortunately, no information of how big numbers this corresponds to in global scale has been found. Therefore the energy use of processing information is simply achieved by applying Koomey's relation between server power and storage power: storage electricity use is 25 % of server power; that is a ratio of 1:4. The server power, or also called processing power, $E_p$, is then (0.027*4 =) 0.108 kWh/stored GB and year.

#### Network

There are also large amounts of data transfers within data centers to take into consideration. Koomey's assumptions are used here as well; communications energy usage is 15 % of processing power. The network equipment inside a data center consumes (0.108*0.15=) 0.016 kWh/stored GB and year, which is part of $E_t$ in equation (1). The other part of energy usage from data transfer is the transfer that occurs outside data centers.

#### Total IT equipment

In total; storage, processing and communication within a data center uses approximately (0.027 + 0.108 + 0.016 =) 0.151 kWh per stored GB and year.

---

[13] Constellation

### Infrastructure

The infrastructure power consumption varies a lot between different data centers. The trend is that PUE is decreasing. Two scenarios are used; one state-of-the-art using a PUE value of 1.05 and one all-trends-continue of 2. This lead to an infrastructure power consumption varying between 0.0076 and 0.151 kWh/stored GB and year. Total data center power consumption per stored GB one year range between 0.16 kWh and 0.30 kWh.

### Efficiency improvements

If Moore's law is applied on servers the energy needed for the same performance is halved every 1.5 year. This means the energy demand is reduced by 37 % each year. This will be taken into consideration in one scenario. Mearian mentions that the costs for managing data have been reduced to one sixth of what it was in 2005 compared to 2011. This corresponds to an annual reduction of cost, or in other words efficiency improvement, of 26 % each year. This CAGR will also be used in one scenario.

### Scenarios

For the data center electricity demand there are two things that are expected to change over time. The first is the amount of data that is managed and the second is more efficient hardware and software. In order to predict future data center energy usage four scenarios have been used. The first two do not take energy efficiency into consideration.

- **A:** Using an annual growth rate of stored bytes of 52 % and a PUE value of 2.
- **B:** Using an annual growth of stored bytes of 48 % and a PUE value of 1.05.
- **C:** Using an annual growth of stored bytes of 52 %, a PUE value of 2 and an efficiency improvement of 26 % each year (Mearian).
- **D:** Using an annual growth of storage needed of 52 %, a PUE value of 2 and an efficiency improvement of 37 % each year (Moore's law).

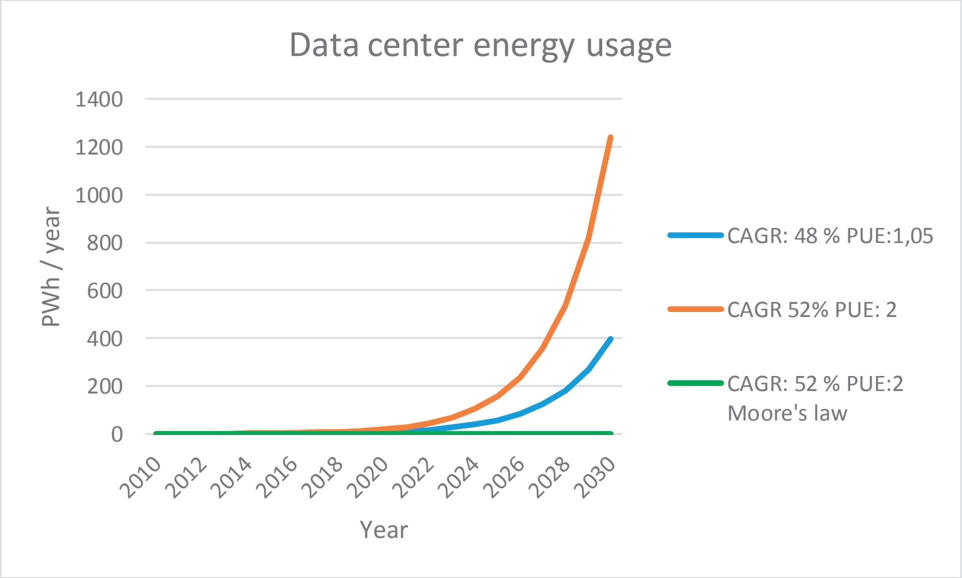

Figure 13. The predicted data center energy use according to scenario A, B and D.

Scenario C is not included in Figure 13 as it would overlap scenario D, both these scenarios result in considerable lower energy demand, and the magnitude cannot be seen in the figure as the scale is too large. It is however highly unlikely that no efficiency improvements would occur in the future and the scenarios C and D probably more realistic. These two scenarios will be illustrated more detailed in the next section.

## Clarification

To further clarify the calculations above three more figures have been produced. In Figure 14 below the number of stored bytes are projected according to the two different CAGRs previous mentioned. Compared to 1 ZB in 2010 the storage in the world will increase fast according to both projections.

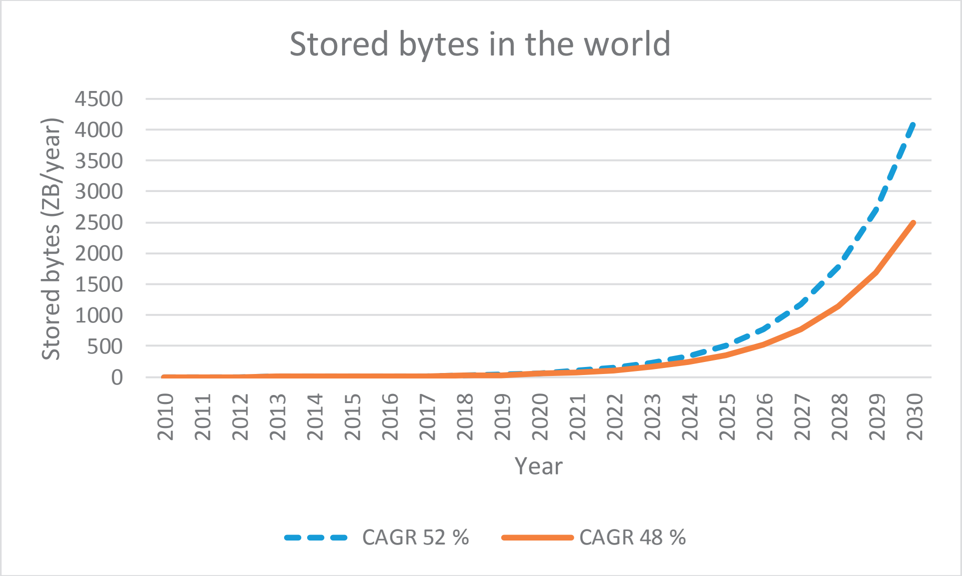

Figure 14. Shows the projected number of stored bytes worldwide according to two growth scenarios.

Figure 15 below shows the electricity usage in data centers per stored byte and year if the efficiency improvements of Moore's law (37 %) and Mearian (26 %) is applied. The start value is 302 TWh per ZB and year and includes the electricity usage for storing, processing and transferring bytes within a data center plus the peripherals (PUE = 2). Note that the efficiency improvements are assumed to be valid for the entire data center.

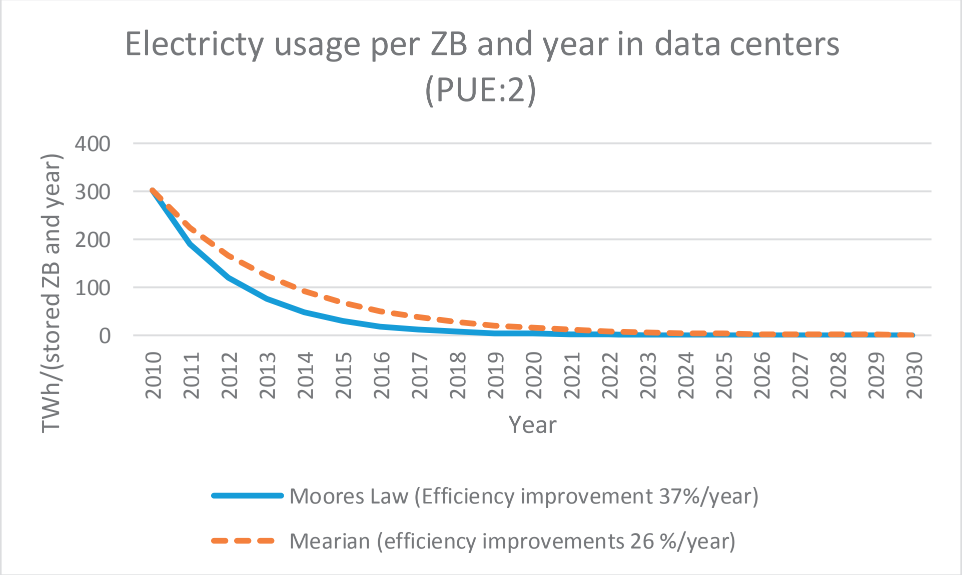

Figure 15. Shows the electricity usage per ZB and year in data centers if efficiency improvements are included.

Figure 16 below is the same as Figure 13 but now only scenario C and D are shown. As can be noted in scenario D the total energy usage actually decreases if the efficiency improvements are as great as 37 % per year. The figure are calculated by multiplying the number of stored bytes (case CAGR 52%) from Figure 14 and the two scenarios of electricity usage per byte in Figure 15.

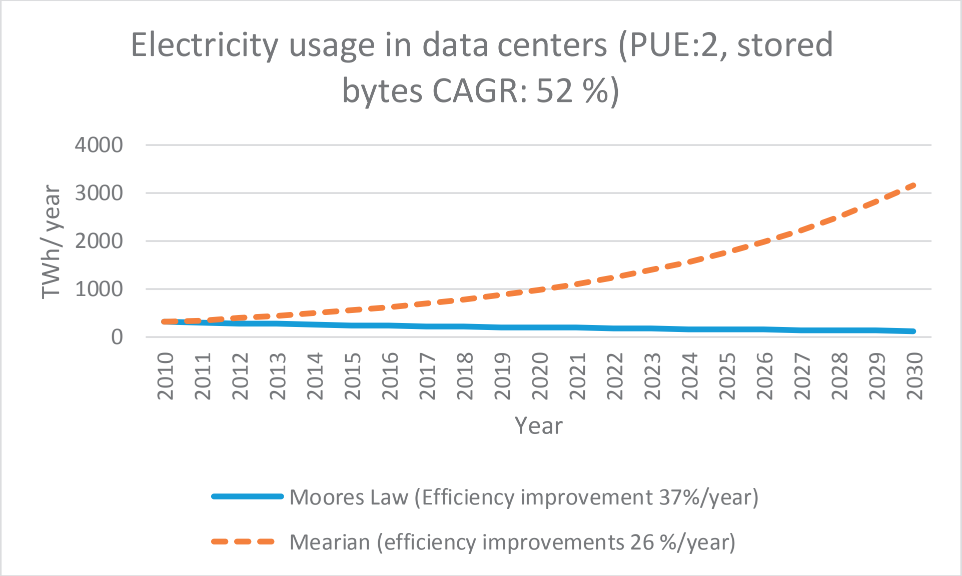

Figure 16. Shows the electricity usage per year in data centers according to two different efficiency improvements

### 8.1.2 DATA TRANSFER

The energy use from data transfer outside data centers, or more specific IP traffic will be calculated in this section. First the number of bytes transferred per year, $n_t$, will be estimated and then the energy usage per transferred byte, $B_t$, will be studied.

Projections of IP traffic have been made by Cisco and Korotky for the years 2015, 2017 and 2020. These projections will be used to identify $n_t$. The estimations made by Cisco and Korotky are as follow:

| | |
|---|---|
| 2015 | 1 ZB |
| 2017 | 1.4 ZB |
| 2020 | 3 ZB |

CAGR is 18 % between the years 2015 and 2017. This can be compared to 25 % between the years 2017 and 2020. Both these CAGRs will be used in two different scenarios when calculating future network energy usage.

The network traffic is assumed to consist of 50 % mobile 3G traffic and 50 % fixed broadband traffic. There are two scenarios used when calculating energy demand of data transfer, $B_t$, one high estimate and one low. The high is the numbers presented by Malmodin et al. (2014). The low is the fixed broadband presented by Baliga. It is assumed that the relation between Malmodin's fixed broadband and Baliga's fixed broadband can be applied on average 3G as well, see Table 12 below.

Table 12. Energy usage of network traffic, $B_t$, used in calculations.

| Scenario | Transfer type | kWh/GB |
|---|---|---|
| High estimate | Fixed broadband | 0,48 |
| High estimate | Average 3G | 3 |
| Low estimate | Fixed broadband | 0,17 |
| Low estimate | Average 3G (calculated) | 1,65 |

### Scenarios

Four scenarios are used when calculating future IP network energy usage:

- **A:** CAGR of 25 % and high energy use per transferred byte
- **B:** CAGR of 18 % and high energy use per transferred byte
- **C:** CAGR of 25 % and low energy use per transferred byte
- **D**: CAGR of 18 % and low energy use per transferred byte

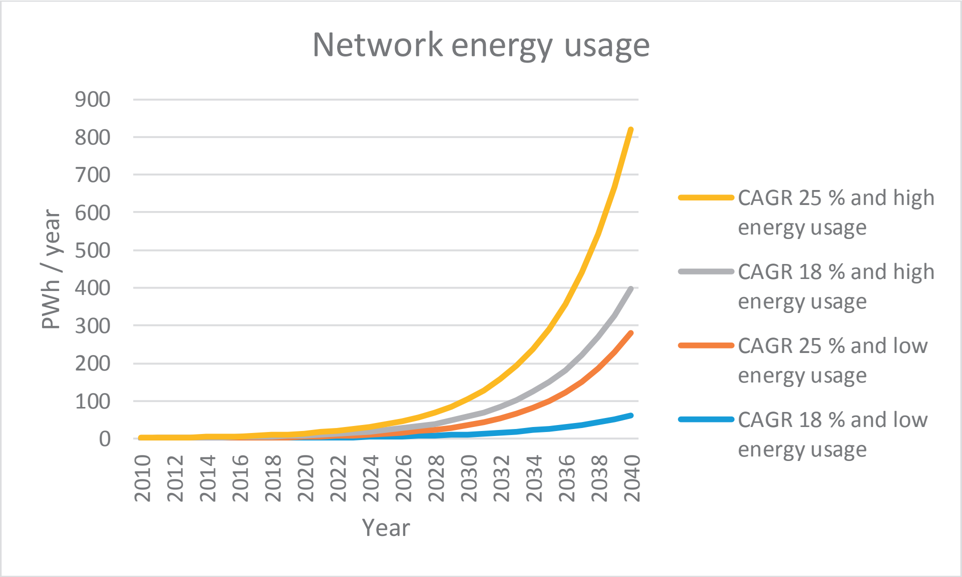

Figure 17. Projections of IP traffic energy usage, assumed half 3G mobile connections and half fixed broadband, according to four different scenarious.

In Figure 17 above the four scenarios are illustrated. It should be noted that efficiency improvements are not accounted for, which mean that the energy use per transferred megabyte in each scenario has not been calculated to improve, but held constant.

### 8.1.3 COMMENTS ON THE RESULT

The energy consumption of the cloud are calculated by adding up the energy usage from data centers, closets and rooms with the energy usage from the global IP traffic. Then all the parameters of equation (1) are considered. This is however a very broad definition of the cloud and to achieve more exact calculations the cloud should be narrowed down.

A modern enterprise server is used to calculate storage energy demand. The energy use of this server play a main role in the calculations. If the server would have used three times as much electricity, the electricity usage illustrated in Figure 13 would have increased with a magnitude of three as well. To make the results more reliable more hard drives' energy usage ought to be studied and compared. In the future this server might be very out of date and scenario A and B, where no efficiency improvements were included, are not likely to reflect reality. At the same time these scenarios shows the importance to make data centers more efficient – or else the energy demand will skyrocket.

The energy efficiency described by Moore's law applied for PC's computations. In this section Moore's law has been applied to server's storage and processing, transfer within datacenters and peripherals. There is an uncertainty if this assumption is correct and if this wide implementation can actually should be applied. The energy efficiency for servers might be even better or worse in reality.

Traffic between data centers is not included in the calculations above, Cisco stated that 0.5 ZB will be transferred between data centers in 2017. The reason to not include this traffic is that it is not well represented by Malmodin's and Baliga's estimations of energy usage. It is the customer premise equipment, CPE, that contributes to the absolute majority of the energy demand of broadband and for mobile transfer it is the base stations. As the traffic between data centers do not use CPEs or base stations these numbers would be misleading. However the total energy demand of data traffic should be approximately 4 % larger if including the traffic between data centers. This percentage was calculated using an electricity demand of 0.1 kWh/GB transferred data, numbers found in from Malmodin's (2011a) study.

The number of stored bytes in the future are the only variable in equations (1), as the processing and transfer of data are depending on stored bytes as well by using Koomey's relationships between storage, networks and processing. These relations might change in the future, as more processing than storage might be needed or the other way around. This might alter the calculations considerably.

The assumptions that data traffic constitute 50 % of wireless 3G traffic and 50 % of broadband traffic can also be discussed. Today the wired data transfer exceed the wireless but this will change in the future to the other way around. How much mobile traffic will exceed wired traffic is however not clear. Moreover wired traffic is not only carried out by broadband and wireless traffic do not only constitute of 3G traffic. The energy usage of the wired traffic is larger than illustrated in the beginning as less efficient cables and connections are used, however in the future the energy efficiency improvements will most certainly reduce the energy usage per transferred byte. 4G traffic might be more efficient per transferred byte than 3G traffic and the energy usage should be lower due to this. In nearest future the scenarios using high energy usage, A and B, might reflect reality the most while further into the future the scenarios with lower energy use is more correct as energy efficiency improves. GreenTouch estimations of 90 % less energy usage in networks 2020 would drastically change the results and even scenario D might be an overestimation if GreenTouch predictions appear to be true.

## 8.2 TOP-DOWN

A top-down method are also calculated, using reports' estimations of ICT's energy consumption and applying different CAGR's to calculate future energy demand. In Table 13 below, GeSI's (2012) estimations of ICT energy usage in 2011 and 2020 are illustrated and the corresponding compound annual growth rate are calculated.

Table 13. The electricity demand and CAGR of ICT, estimations based on GeSI (2012).

**Electricity use of ICT (TWh)**

| | 2011 | 2020 | CAGR (2011-2020) |
|---|---|---|---|
| ***End user devices*** | 362 | 353 | – 0.3 % |
| ***Data transfer*** | 244 | 292 | 2.0 % |
| ***Data centers*** | 237 | 441 | 7.1 % |
| ***Total ICT*** | 843 | 1085 | 2.9 % |

Greenpeace used a growth in networks of CAGR 9.5 % and a growth in data centers of CAGR 9 % when calculating the future energy use of the cloud. This corresponds to a mean CAGR for Greenpeace (2010) estimates of 9.2 %. This is higher than the mean for GeSI which is 2.9 %. The end user devices are however not accounted for in Greenpeace's reports and that is the only category that decreases its total emissions in GeSI's estimations.

To summarize the electricity demand of ICT two scenarios are constructed:

- **A:** One lower bound where GreenTouch estimations of 90 % less electricity used 2020 compared to 2010 for networks are used. For the years beyond 2020 the electricity demand of networks is assumed flat. Data centers and end user devices uses the CAGR of Table 13 as it is the lowest estimations of CAGR found. The lower bound also starts at GeSI's low values of ICT electricity use.

- **B:** One upper bound assemble all higher estimates found; using the mean CAGR of 9.2 % for all ICT and starting with using Somavat and Namboodiri (2011) high estimations of 1078 TWh in 2010. The result is shown in Figure 18 below.

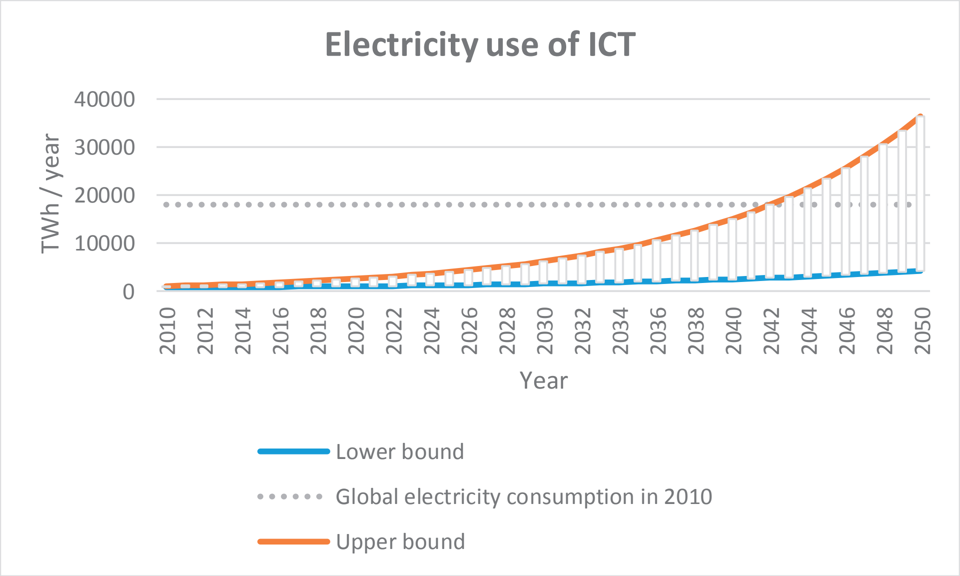

Figure 18. ICT energy demand in the future according to one upper bound and one lower bound.

### 8.2.1 COMMENTS ON THE RESULT

Figure 18 is more of a sensitivity analysis than actual results, as the gap between the lower and the upper bound is huge. Moore's law states that the energy efficiency will improve every 1.5 year, which has not been taken into account in the lower and upper bounds. The total electricity use in the world was in 2010 around 18 000 TWh, a number which the upper bound exceed soon after 2040. Even though the ICT's part of world's electricity use will increase this rate seem unrealistic.

The lower bound can also be discussed. Moore's law is not used for this scenario either, but GreenTouch estimations of 90 % less energy use in network traffic was. This is however a theoretical improvement and might not be realistic to happen in 2020. However, the assumption of flat development of electricity usage in networks can be seen as realistic if GreenTouch's estimations can; the data traffic increases rapidly and if the energy usage can decrease despite this fact until 2020 it is not a big leap to assume it flat in after 2020. The only result that can be extracted from Figure 18 is that electricity use of ICT will probably be between the lower bound and the upper bound.

The upper bound CAGR is a mean for networks and data centers, not end user devices, but it has been applied to end user devices as well. GeSI's number shows that end user devices total electricity usage are going to decrease, but as more people get connected and more devices are used it is likely that the electricity use instead increases. If this increase is as high as Greenpeace's mean CAGR of 9.2 % can be discussed though.

## 8.3 SUMMARIZING COMPARISON

As can be seen from previous diagrams the unit on the y-axis is different on the top-down cases compared to the bottom-up cases. The magnitude of electricity use is much greater for the bottom-up cases, measured in PWh instead of TWh. If the different approaches are to be combined only the lower bounds of the bottom-up cases can be used. This seems reasonable as the lower bounds includes some energy efficiency improvements which the upper bounds do not. In Figure 19 below the bottom-up cases are named “Cloud” to simplify the figure, but is actually all data center and networks. In the top-down case the end user equipment is included as well.

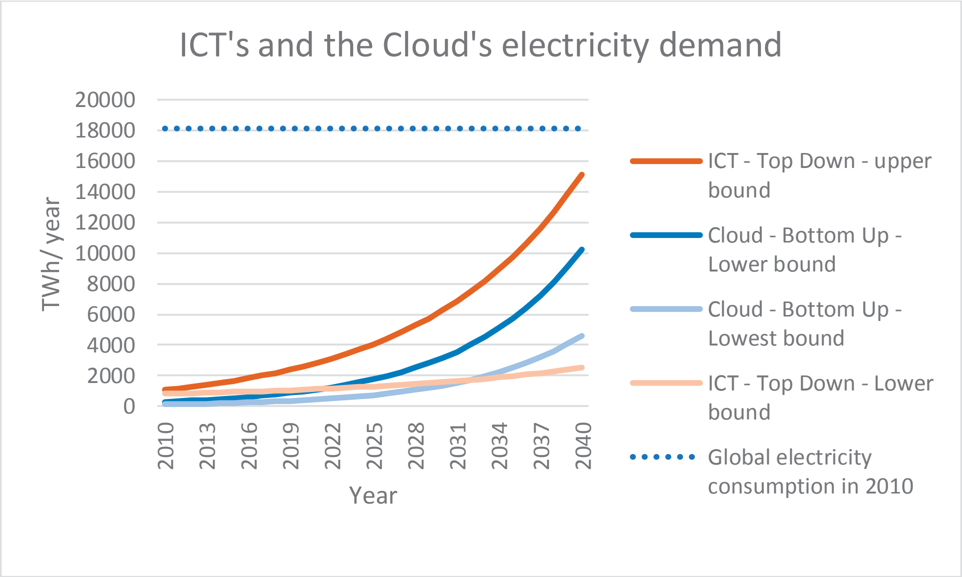

Figure 19. Comparison of the bottom-up and top-down cases.

The “Cloud – Bottom-Up – Lower bound” uses the data center scenario C and the network scenario D while the “Cloud – Bottom-Up – Lowest bound” uses the data center scenario D and the network scenario D.

Malmodin et al. (2014) estimated the energy usage of data centers and networks based on transferred bytes to 1.5 kWh/GB transferred data. It is not comparable to the numbers calculated in the bottom up case as they were based in stored bytes as well as transferred bytes. But the energy usage is in the same range as the same unit is used and the network traffic is partly based on numbers from Malmodin et al.

The precise numbers should be used with care; the main result one can get from the results is that energy efficiency is very important to improve in order to keep energy consumption low while usage increases. In the bottom-up cases when no energy efficiency is included, the world electricity usage of 18 000 TWh for 2011 is reached before 2020 – these scenarios are thus not likely to reflect reality.

None of the estimates result in a decrease in energy usage, apart from the network part in the lower bound of the top-down case and the data center part, scenario D, in the bottom-up case. The decrease of networks energy usage in the top-down case in Figure 18 is hidden by the increase of data center electricity usage. If these two scenarios are combined the energy usage actually decreases while the usage increases.

It is also interesting to notice that the results using the bottom-up and top-down cases are in the same range. This indicates that the numbers are not totally misleading.

It can also be discussed if the CAGR model, which describes an exponential growth is correct. Korotky for example used a hyperbolic curve to describe internet network traffic – resulting in a maximum value of internet traffic instead of a never-ending growth. But at the same time the data transfer, data storage and data processing increases according to all sources. The question is if the energy efficiency improvements can keep up the pace, making the total energy consumption manageable. The scenarios with no efficiency improvements in the bottom-up case states that energy consumption can fast get out of reach if allowed.

# 9 DISCUSSION

The modern person's way of life has consequences for the environment on many levels and we are facing a big threat of upcoming and ongoing climate change. There are certainly not many people who reflect on the fact that watching an episode of their favorite series over the internet further contribute to the greenhouse gas emissions. How many of the videos posted on YouTube every day are actually necessary? For how long time should a clip be available? For how many generations should we keep birthday photos and films? All studies show increased amounts of transferred, stored and processed data. Maybe it is about time we, as consumers, start thinking green in the digital world as well. Data centers, mobile and broadband operators can contribute to more efficient energy use, higher grade of compressed bytes and ensure that electricity is green and renewable - but the actual use is more difficult to streamline. Today the users have no incentives to lower her usage of the cloud – and why should they? Maybe there should be regulations of how much we can use the cloud as our usage in the end emits greenhouse gases and therefore contribute to global warming.

At the same time this question is very controversial. Internet has made it possible to spread the word from a single person to the rest of the world. It has resulted in fast communication, contributed to equality and the right to speak one's mind. If you place restrictions on what can be uploaded and shared, and for how long – it might infringe on human rights and be a step in the wrong direction. To make restrictions on the internet can be comparable to not allowing people to say what they want or to travel where they want. But to make people aware of the environmental impact and thus decrease their usage themselves is another thing.

Online gaming is predicted to contribute to a large part of future data transfer. Is it wise to let something that we only use for having fun contribute to global warming? But by following this track one can say that charter trips should be forbidden as they are only for fun as well. The best way to approach this problem is to make the activity as energy efficient as possible without destroying the experience of the service. There are also many advantages with online gaming such as people learning languages, cooperating with people and meeting new people.

The increased usage is based on, among other things, the development of the Smart Grid and M2Ms and more mobile and broadband subscriptions with increased usage per subscription. The usage of the cloud can come to explode in the future and it is important for the industry to be prepared for this and start working towards efficient data centers and data transfer already today, if not yesterday. Future projections have shown that there is a large potential to make the cloud more energy efficient.

The energy cost is one incentive to make energy efficient data centers. But is this incentive enough? Is there a need for regulations to reduce the energy usage of data centers and data traffic as well? Bahnhof states that regulations on data centers efficiency would not impact them as they already are very energy efficient, but at the same time they can be even more energy efficient than they are. Their PUE of 1.2 can be improved to 1.05. There are also other incentives as competition between companies – if one company improves its environmental footprint it can be used for promoting the company. Other companies might follow to get the same advantage and a positive feedback loop is created. But for this to happen one company must start. Customers have also started to demand greener services which is yet another incentive to reduce the energy usage. Data centers have been pointed out as the new base industry and it is surprising that there are no regulations as to how energy efficient they should be today. The European Union has decided to include servers' power supply unit to the Ecodesign directive and that is a step in the right direction.

The cloud has many advantages compared to traditional IT, among others it has been proven to be more energy efficient. It is better to use the cloud than to use traditional IT and if everyone would move into the cloud today the overall energy usage would decrease. Maybe it is not that interesting to talk about the cloud energy demand when it is traditional IT that is less efficient and need improvements the most. The main question to be answered was: “What is the energy demand for cloud services now and in the future?” As the cloud has proven to be more energy efficient than general ICT a more relevant question could be: ”What is the energy demand of ICT now and in the future?”. ICT has been studied in more papers and covers a broader area but are also including all components of the cloud.

But on the other hand an environmental friendly car, for instance, tend to be used more as it emits so little greenhouse gas emissions. All in all the environmental performance can be worse than before because of increased usage. The rebound effect can be applied to the cloud as well. If each picture you store in the cloud uses very little energy you tend to store more and more files. Especially since the cloud seem unlimited and it is very cheap to buy more storage. If the cloud continues to grow as fast as it does today, its energy demand will exceed traditional IT. One solution is to not make everything available 24-7. It is possible to store data without the constant use of electricity – for example there are USB sticks for personal usage. The same technology might be scaled up and used for data centers.

The incentives to owners of data centers to improve energy efficiency have been discussed – but what incentives exist for developers of apps, network owners, manufactures of mobile phones or M2M? It has not been closely studied which regulations of energy efficiency that exist for network owners or end user devices. The energy cost can of course be one incentive to reduce energy usage of networks but the questions is if the usage is so large today that this is a sufficient incentive. The main focus might instead be to provide a fast and reliable network. When manufacturing mobile phones the battery time is a selling argument and thus it is in the interest of the manufacturer to provide a phone that is energy efficient and can last long between the charges. For developers of apps it can also be in their interest to program energy efficient apps as they might be used by more people if it does not influence the battery time so much. But the main focus here is most probably not energy efficiency but to create a useful and/or joyful app.

ICT can lead to decreased travelling when physical travelling is replaced by video conferences but at the same time ICT also make people from different part of the world get in contact and be friends – which creates more travelling. Instead of knowing people in your neighborhood acquaintances are created all over the world and the travel distance is increased.

This report have only analyzed the operational electricity use of the cloud – other part of the lifecycle as manufacturing and treatment of waste is not included. The future for cloud computing is also dependent on resources, such as earth minerals that is necessary for the manufacturing. Shortage of these mineral can cause the development to slow down.

# 10 CONCLUSIONS

The questions stated in the beginning of this paper will now be answered. The main question is merged into the second question in this section.

- What is the definition of the cloud?

The most passable and used definitions of the cloud is defined from a user perspective. This definition is best illustrated in Figure 20 below, using attributes, deployment types and service models to describe the functionalities of the cloud. The cloud constitutes of IT resources, i.e. storage and processing, that are always accessible and seem unlimited to the user. The resources are supplied according to demand and the cost are proportional to the usage of the cloud.

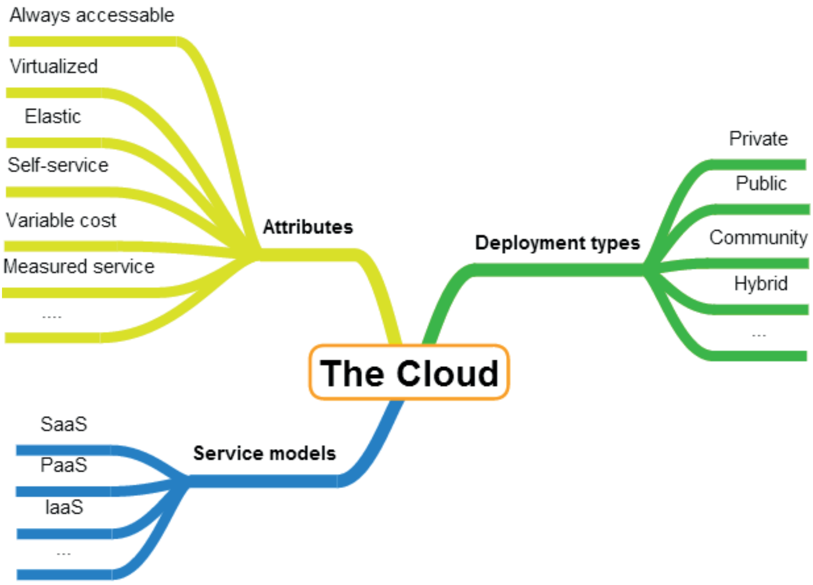

Figure 20. The definition of the cloud from a user perspective.

- What is the energy demand for cloud services now and in the future? How will cloud services evolve in the future?

The cloud is anticipated to grow exponentially in the future. However, it is better the cloud grows than traditional IT as the cloud consumes less energy per task. The energy demand of the cloud, using a very wide definition of the cloud as all network traffic and data centers are included, can range between 5000 and 10 000 TWh per year in 2040, if taking energy efficiency improvements into consideration. For comparison the current global electricity use is approximately 18 000 TWh. Without energy efficiency improvements the energy consumption will skyrocket. There are huge potentials to make ICT and the cloud more efficient in the future.

- Which incentives exist for lowering the energy demand concerning the cloud?

Today there are few regulations concerning the cloud's efficiency. There are mostly economic incentives to make data centers more efficient. One common label is the voluntary Energy Star that promotes energy efficient buildings and equipment. The European union have started to include computers, tablets and server's power supply units to their Ecodesign directive so it seems like the authorities have started to see the upcoming problem. But all in all companies have to take matters into their own hands and act proactive as the usage of the cloud increases in an explosive pace.

# 11 REFERENCES

Accenture and WSP, 2010. *Cloud Computing and Sustainability: The Environmental Benefits of Moving to the Cloud,* s.l.: WSP Environment & Energy, Accenture.

Amazon Web Service, 2014. *AWS Elastic Beanstalk.* [Online] Available at: http://aws.amazon.com/elasticbeanstalk/ [Accessed 24 March 2014].

Bahnhof, 2014. *Om Bahnhof.* [Online] Available at: https://www.bahnhof.se/om/ [Accessed 7 April 2014].

Barroso, L. A. & Hölzle, U., 2009. *The Datacenter as a Computer: An Introduction to the Design of Warehouse-Scale Machines ISBN: 9781598295573,* Madison: Morgan & Claypool Publishers.

Bergquist, G., 2014. *Technical Manager at Bahnhof AB* [Interview] (12 May 2014).

Björklund, D., 2014. *Responsible for Data Center Facilities and Network Department in Sweden for Tieto* [Interview] (17 April 2014).

Business Sweden, 2013. *Data Centre Risk Index - Informing global investment decisions; written by source8, hurleypalmerflatt, cushman & wakefield.* [Online] Available at: http://www.business-sweden.se/PageFiles/9118/DCRI%202013.pdf [Accessed 25 June 2014].

Business Sweden, 2014. *The Swedish Data Center Initiative.* [Online] Available at: http://www.business-sweden.se/en/Invest/Industry-Opportunities/ICT/The-Data-Center-Initiative/ [Accessed 10 March 2014].

Cisco, 2013 a. *The Zettabyte Era—Trends and Analysis (White paper).* [Online] Available at: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/VNI_Hyperconnectivity_WP.html [Accessed 21 March 2014].

Cisco, 2013 b. *Cisco Global Cloud Index: Forecast and Methodology, 2012–2017.* [Online] Available at: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/Cloud_Index_White_Paper.html [Accessed 16 June 2014].

Cisco, 2014. *Cisco Global Cloud Index Q&A.* [Online] Available at: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/Cloud_Index_QA.html [Accessed 10 July 2014].

Cisco, 2014. *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2013–2018 (White Paper).* [Online] Available at: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.html?CAMPAIGN=MobileVNI2014&COUNTRY_SITE=us&CREATIVE=MobileVNI+2014+Cisco.com+&POSITION=sl&REFERRING_SITE=cisco.com+homepage [Accessed 2 June 2014].

Coroama, V. C., Hilty, L. M., Heiri, E. & Horn, F. M., 2013. *The Direkt Energy Demand of Internt Data Flows,* Lissabon: Journal of Industrial Ecology.

Data Center Map, 2014. *Data Center Map.* [Online] Available at: http://www.datacentermap.com/ [Accessed 25 June 2014].

Energimyndigheten, 2014. *Datorer, surfplattor och servrar blir mer energisnåla.* [Online] Available at: http://www.energimyndigheten.se/Press/Nyheter/Datorer-surfplattor-och-servrar-blir-mer-energisnala/ [Accessed 8 July 2014].

Energimyndigheten, 2014. *Program för energieffektivisering i energiintensiv industri.* [Online] Available at: http://www.energimyndigheten.se/pfe [Accessed 9 July 2014].

Energy Star, 2014. *About ENERGY STAR.* [Online] Available at: http://www.energystar.gov/about [Accessed 8 July 2014].

EPA, 2007. *Report to Congress on Server and Data Center Energy Efficiency Public Law 109-431,* s.l.: US Environmental Protection Agency and ENERGY STAR Program.

Ercan, M., 2013. *Life cycle assessment of global warming potential for a smart phone - Master thesis,* Stockholm: Ericsson and KTH.

Ericsson, 2013. *Ericsson Mobility Report. On the Pulse of the Network Society.,* Stockholm: Douglas Gilstrap, Senior Vice President and Head of Strategy, Ericsson.

European Commission, 2010. *The Future of Cloud Computing,* Brussels: European Commission - Information Society and Media.

Facebook, 2014. *Starting work on Luleå building 2.* [Online]
Available at: https://www.facebook.com/notes/lule%C3%A5-data-center/starting-work-on-lule%C3%A5-building-2/587772454624780
[Accessed 14 March 2014].

Fredriksson, J., 2014. *Consultant at The Northern Link, working with Facebook in Luleå, Sweden* [Interview] (22 April 2014).

Gantz, J. & Reinsel, D., 2011. *Extracting Value from Chaos, IDC iView,* Framingham: IDC Go-to-Market Services.

GeSI, 2010. *SMART 2020: Enabling the low carbon economy in the information age,* s.l.: The Climate Group on behalf of the Global eSustainability Initiative (GeSI).

GeSI, 2012. *SMARTer 2020 - The role of ICT in Drivning a Sustainable Future,* Boston: Global e-sustainability Initiative (GeSI) and The Boston Consulting Group.

Google Official Blog, 2009. *Powering a Google search.* [Online]
Available at: http://googleblog.blogspot.se/2009/01/powering-google-search.html
[Accessed 24 March 2014].

Greenpeace, 2010. *Make IT Green,* Amsterdam: Greenpeace International.

Greenpeace, 2011. *How Dirty is Your Data?,* Amsterdam: Greenpeace International.

Greenpeace, 2012. *How Green Is Your Cloud,* Amsterdam: Greenpeace International.

Greenpeace, 2014. *Click Green: How Companies are Creating the Green Internet,* Washington: Greenpeace Inc..

GreenTouch, 2013. *GreenTouch Green Meter Research Study: Reducing the Net Energy Consumption in Communications Networks by up to 90% by 2020,* Amsterdam: GreenTouch White Paper, Stichting GreenTouch.

GSMA, 2013. *The Mobile Economy (ATKearney),* London: GSMA.

Holmqvist, J. K., 2014. *dr, Senior Environmental Specialist, Sony Mobile Communication* [Interview] (21 May 2014).

IDC, 2014. *IDC's Worldwide Information Technology Market Map: January 1, 2015.* [Online]
Available at: http://www.idc.com/getdoc.jsp?containerId=247540
[Accessed 22 May 2014].

IEA, 2012. *Electricity consumption, ISBN 978-92-64-17468-9,* Paris: International Electricity Agency Statistics.

IEA, 2013. *CO2 emissions from fuel combustion - highlights,* Paris: International Energy Agency (IEA).

IEA, 2014. *More Data, Less Energy - Making Network Standby More Efficient in Billions of Connected Devices,* Paris: OECD/IEA.

Imaizumi, H. & Morikawa, H., 2009. *Directions Towards Future Greeen Internet,* Tokyo: The University of Tokyo.

IPCC, 2011. *Annex II: Methodology. In IPCC Special Report on Renewable Energy Sources and Climate Change Mitigation,* Cambridge: Cambridge University Press.

IVA, 2013. *Nationell agenda - Internet of Things (ISSN:1102-8254),* Stockholm: Kungl. Ingenjörsvetenskapsakademien.

Kihl, M., 2014. *Associate Professor, PhD, Department of Electrical and Information Technology at Lund University* [Interview] (17 March 2014).

Koomey, J., Berard, S., Sanchez, M. & Wong, H., 2011 b. *Implications of Historical Trends in the Electrical Efficiency of Computing,* s.l.: IEEE Computer Society.

Koomey, J. G., 2011 a. *Growth in Data Center Electricity Use 2005 to 2010,* Oakland: Analytic Press.

Korotky, S. K., 2013. Semi-Empirical Description and Projections of Internet Traffic Trends Using a Hyperbolic Compound Annual Growth Rate. *Bell Labs Technical Journal,* 18(3), pp. 5-21.

Koshy, I. K., Juby, M. A., Namboodiri, V. & Overcash, M., 2012. *Can Cloud Computing Lead to Increased Sustainability for Mobile Devices?,* Kansas: Electrical Engineering and Computer Science & Industrial and Manufacturing Engineering Department.

Landström, J., 2014. *Strategist for data center facilities at Tieto, Sweden* [Interview] (22 April 2014).

Li, Z., Greenan, K. M., Leung, A. W. & Zadok, E., 2012. *Power Consumption in Enterprise-Scale Backup Storage Systems,* s.l.: Stony Brook University.

Luleå Näringsliv AB, 2014. *Facebook bygger sitt andra datacenter i Luleå.* [Online]
Available at: http://www.luleanaringsliv.se/nyhetsarkiv/facebook-bygger-sitt-andra-datacenter-i-luleaa/
[Accessed 11 March 2014].

Länsstyrelsen Norrbotten, 2014. *Strategi för att skapa en världsledande teknikregion i Norrbotten för klimatsmarta effektiva datacenter,* Luleå: Länsstyrelsen Norrbotten.

Länsstyrelsen Norrbotten, 2014. *Strategi för att skapa en världsledande teknikregion i Norrbotten för klimatsmarta effektiva datacenter,* Luleå: Länsstyrelsen Norrbotten.

Malmodin, Jens, Lundén, D. & Lövehagen, N., 2013. *Methodology for Life Cycle Based Assessments of the CO2 Reduction Potential of ICT Services,* Stockholm: Ericsson and Telia Sonera.

Malmodin, J., Lundén, D. & Andersson, G., 2011. *Life-cycle assessment of ICT networks and primary subscription services in Sweden.* s.l.:Telia SOnera and Ericsson.

Malmodin, J. et al., 2014. Life Cycle Assessment of ICT; Carbon Footprint and Operational Electricity Use from the Operator, National, and Subscriber Perspective in Sweden. *Journal of Industrial Ecology*, Only available online yet, pp. 1-17.

Malmodin, J. et al., 2010. Greenhouse Gas Emissions and Operational Electricity Use in the ICT and Entertainment & Media Sector. *Journal of Industrial Ecology,* 14(5), pp. 770-790.

Mearian, L., 2011. *Computerworld.* [Online]
Available at: http://www.computerworld.com/s/article/print/9217988/World_s_data_will_grow_by_50X_in_next_decade_IDC_study_predicts
[Accessed 30 May 2014].

Mell, P. & Grance, T., 2011. *The NIST Definition of Cloud Computing 800-145,* Gaithersburg: NIST-National Institute of of Standards and Technology .

Miljöaktuellt, 2014. *Google storsatsar på svensk vindkraft.* [Online]
Available at: http://miljoaktuellt.idg.se/2.1845/1.543770/google-storsatsar-pa-svensk-vindkraft#disqus_thread
[Accessed 11 March 2013].

Mills, M. P., 2013. *The Cloud Begins With Coal,* s.l.: The Digital Power Group (sponsored by National mining Assosiation and American Coalition for Clean Coal Energy).

NE, 2014 a. *Molnet (Nationalencyklopedin).* [Online]
Available at: http://www.ne.se/molnet
[Accessed 28 January 2014].

NE, 2014 b. *IT (Nationalencyklopedin).* [Online]
Available at: http://www.ne.se/lang/it/214244
[Accessed 17 March 2014].

NE, 2014 c. *Internet (Nationalencyklopedin).* [Online]
Available at: http://www.ne.se/internet
[Accessed 24 March 2014].

NE, 2014 d. *Bredbandskommunikation (Nationalencyklopedin).* [Online]
Available at: http://www.ne.se/lang/bredbandskommunikation?i_h_word=DSL
[Accessed 2 April 2014].

NE, 2014 e. *Smartmobil (Nationalencyklopedin).* [Online]
Available at: http://www.ne.se/lang/smartmobil
[Accessed 26 March 2014].

Nielsen, K., 2014. *Confused About the Cloud? Deciphering Cloud Services for Consumers.* [Online]
Available at: http://cloud-services-review.toptenreviews.com/confused-about-the-cloud-deciphering-cloud-services-for-consumers.html
[Accessed 24 March 2014].

Open Compute Project, 2014. *Open Compute Project.* [Online]
Available at: http://www.opencompute.org/
[Accessed 13 May 2014].

Pinnacle Sweden AB, 2011. *Tillstånd enligt miljöbalken för uppförande och drift av reservkraftsgeneratorer på fastigheterna Porsön 1:3 och Porsön 1:416, Luleå kommun,* Stockholm: Miljöprövningsdelegationen.

Rhoton, J., 2011. *Cloud Computing Explained ISBN-10: 0-9563556-0-9.* Second edition ed. s.l.:Recursive Press.

Rountree, D. & Castrillo, I., 2013. *The Basics of Cloud Computing: Understanding the Fundamentals of Cloud Computing in Therory and Practice.,* Burlington: Elservier Science.

Seagate, 2012. *Seagate.* [Online] Available at: http://www.seagate.com/files/www-content/product-content/_cross-product/en-gb/docs/storage-solutions-guide-nov2012-sg1351-11-1211emea.pdf [Accessed 4 June 2014].

Short, J. E., Bohn, R. E. & Baru, C., 2011. *How much information? 2010 Report on Enterprise Server Information,* San Diego: UC San Diego.

Sjöström, L., 2013. *Energianvändning i Serverhallar, Energiflöden och kylning, Examensarbete 30 hp,* Uppsala: Uppsala Universitet.

Somavat, P. & Namboodiri, V., 2011. *Energy Consumption of Personal Computers Including Portable Communication Devices.* [Online] Available at: http://www.cs.wichita.edu/~vnambood/mypubs/JGEPavel2011.pdf [Accessed 24 March 2014].

Svenska Dagbladet, 2013. *Så mycket drar Facebook i Luleå.* [Online] Available at: http://www.svd.se/naringsliv/digitalt/sa-mycket-el-drar-facebook-i-lulea_8259950.svd [Accessed 11 March 2014].

Svensson, K., 2014. Borg flörtar med Facebook. *Dagens industri*, 20 May, p. 11.

Sydsvenskan, 2014. *Google köper skånsk vindkraft.* [Online] Available at: http://www.sydsvenskan.se/ekonomi/google-koper-skansk-vindkraft/ [Accessed 11 March 2014].

Talbot, D., 2012. *MIT Technology Review - Efficiency Breakthrough Promises Smartphones that Use Half the Power.* [Online] Available at: http://www.technologyreview.com/news/506491/efficiency-breakthrough-promises-smartphones-that-use-half-the-power/ [Accessed 1 July 2014].

Taylor, J., 2014. *Subject matter expert for the Data Center topic within the IT Business unit for Schneider Electric* [Interview] (9 April 2014).

TGG, 2010. *Carbon Usage Effectiveness (CUE): A Green Grid Data Center Sustainability Metric,* s.l.: The Green Grid (TGG).

TGG, 2014. *Harmonizing Global Metrics for Data Center Energy Efficiency (The Green Grid - TGG).* [Online] Available at: http://www.thegreengrid.org/Global/Content/Regulatory-Activities/HarmonizingGlobalMetricsForDataCenterEnergyEfficiency_DCeP [Accessed 31 March 2014].

Thomond, D. P., Mackenzie, D. I., Williams, D. & Velkov, A., 2013. *The Enabling Technologies of a Low Carbon-Economy: A Focus on Cloud Computing,* s.l.: Think, Play Do Group, GeSI, Microsoft.

Tieto, 2014. *See what cloud could do to you - Uncomplicating cloud business,* s.l.: Tieto Corporation.

UN DESA, 2014. *World Population Prospects: The 2012 Revision.* [Online] Available at: http://esa.un.org/wpp/unpp/panel_population.htm [Accessed 3 April 2014].

Vaquero, L. M., Rodero-Merino, L., Caceres, J. & Lindner, M., 2009. A Break in the Clouds: Towards a Cloud Definition. *ACM SIGCOMM Computer Communication Review*, Januari, 39(1), pp. 50-55.

Vattenfall, 2011. *Certified Environmental Product Declaration EPD® of Electricity from Vattenfall’s Nordic Hydropower,* s.l.: Vattenfall AB.

Vattenfall, 2013 a. *Certified Environmental Product Declaration EPD® Electricity from Forsmark Nuclear Power Plant, Version 1.2,* Forsmark: Vattenfall AB.

Vattenfall, 2013 b. *Certified Environmental Product Declaration EPD® of Electricity from Vattenfall’s Nordic Wind Farms,* s.l.: Vattenfall AB.

Verdantix, 2011. *Cloud Computing - The IT Solution for the 21th Century,* s.l.: Carbon Disclosure Project.

Vermesan, O., Friess, P. & Patrick, G., 2011. *Internet of Things Strategic Research Roadmap.* [Online]

Available at: http://www.internet-of-things-research.eu/pdf/IoT_Cluster_Strategic_Research_Agenda_2011.pdf [Accessed 2 April 2014].

Vesset, D. et al., 2012. *Worldwide Big Data Technology and Services 2012-2015 Forecast #233485,* Framingham: IDC.

Williams, D. R., Thomond, P. & Mackenzie, I., 2013. *The Greenhouse Gas Abatement Potential of Enterprise Cloud Computing,* Zürich: ICT4S 2013.

Yuventi, J. & Mehdizadeh, R., 2013. *A critical analysis of Power Usage Effectiveness and its use in communicating data center energy consumption,* Stanford: Elsevier B.V.

## 11.1 PICTURES

Jastremski, M., 2004 (legacy.openphoto.net) [CC-BY-SA-2.0 (http://creativecommons.org/licenses/by-sa/2.0)], via Wikimedia Commons. [Online] Available at: http://upload.wikimedia.org/wikipedia/commons/b/b5/Cumulus_clouds_in_fair_weather.jpeg [Accessed 18 March 2014].

U.S Army photo, 1947 (Public domain) [Online] Available at: http://en.wikipedia.org/wiki/Eniac#mediaviewer/File:Eniac.jpg [Accessed 13 June 2014]

# 12 GLOSSARY

**API** – Stands for Application Programming Interface and it is a set of rules that allows different programs to communicate with each other.

**Big data** – Big data is the concept of managing and analyzing huge amount of data relatively fast (Kihl, 2014). Vesset et al. (2012) describes big data with four V's, or attributes; velocity, volume, variety and value in a report published by IDC. Another description of big data mentioned in the IDC report is "data sets that grow so large or so fast that they are difficult to handle using traditional technology" (Vesset, et al., 2012, p. 1). IDC state that this is a common definition used by data vendors and analytics.

**Bit** – The smallest storing capacity for a computer, it can either have the value 0 or 1.

**Bit/s** – Measurement of data transfer, also called bps (bits per second). 1 bit/s is 1 bit per second transferred data. Not to be confused with byte/s.

**Byte** – 1 byte represents 8 bits, and can have (2^8 =) 256 different values. In order to describe greater amounts of data prefixes are used and these can be quite confusing. There exists, for example, three different meanings of megabyte, depending on which definition used.

- 1 megabyte can mean 1,000,000 bytes if using the decimal notation, which often is used when referring to physical disk storage.
- 1 megabyte can also mean 1024^2 bytes when using the binary notation, which is utilized when referring to virtual storage. The IEC, International Electrotechnical Commission, has made a standard naming this mebibyte in order to part it from the decimal notation. This standard is winning ground, but still there are reports using the word megabyte instead.
- Furthermore 1 megabyte can mean 1024*1000 bytes, which is a mixture of the two other definitions, used to describe a floppy disks storage capacity.

Table 14. Shows different notations of byte prefixes.

| Decimal notation | Number of bytes | Binary notation | Number of bytes |
|---|---|---|---|
| Kilobyte (kB) | 1000 | Kilobyte (KB), Kibibyte (KiB) | 1024 |
| Megabyte (MB) | $1000^2$ | Megabyte (MB), Mebibyte (MiB) | $1024^2$ |
| Gigabyte (GB) | $1000^3$ | Gigabyte (GB), Gibibyte (GiB) | $1024^3$ |
| Terabyte (TB) | $1000^4$ | Tebibyte (TiB), Terabyte (TB) | $1024^4$ |
| Petabyte (PB) | $1000^5$ | Pebibyte (PiB), Petabyte (PB) | $1024^5$ |
| Exabyte (EB) | $1000^6$ | Exbibyte (EiB), Exabyte (EB) | $1024^6$ |
| Zettabyte (ZB) | $1000^7$ | Zebibyte (ZiB), Zettabyte (ZB) | $1024^7$ |

Source: http://wikipedia.com/ and http://whatsabyte.com/

**Byte/s –** Measurement of data transfer. 1 byte/s is 1 byte per second transferred data or 8 bits per second transferred data.

**Carbon dioxide equivalents –** There are several molecules that contribute to global warming, for example carbon dioxide, methane, nitrous oxide and CFC's. To measure total emission of greenhouse gases from an industry for example, each molecules impact is normalized to carbon dioxide. Methane for example is a 25 times stronger greenhouse gas than carbon dioxide.

**CAGR** - Compound Annual Growth Rate, more about this in section 8.2

**GSM** – 2G

**Hops** – When a data transfer sequence has done one hop it has passed between two routers.

**IaaS** – Infrastructure as a Service

**ICT** – Information and Communication Technology

**IoT** – Internet of Things

**IoE** – Internet of Energy

**LCA** – Life Cycle Assessment

**LTE** – Long Term Evolution – 4G

**PaaS** – Platform as a Service

**PUE** – Stands for Power Usage Effectiveness. It is a measurement of data center efficiency. It describes how much of the utilized power that goes to the actual servers and not to cooling, lights and peripheral. The equation is as follows:

$$PUE = \frac{P_{total}}{P_{IT\ equipment}}$$

If all power goes to the IT equipment, the ideal value 1 is achieved. Very efficient data centers can have values slightly over 1.1.

**SaaS** – Software as a Service

**Virtualization** – One server can be used by multiple tenants with the assistance of virtualization; the creation of multiple virtual operating systems.

# 13 APPENDIX

## 13.1 VARIOUS CLOUD DEFINITIONS

Table 15 below presents various definitions and explanations of the terms *cloud* and *cloud computing* that appear in reports discussing the subject.

Table 15. Definitions and explanations of the terms *cloud* and *cloud computing* from various reports.

| Author and Title | Definition: The cloud | Source |
|---|---|---|
| Greenpeace<br>*How dirty is your data?* | “… this report will use the term ‘cloud’ to describe energy and resources used broadly with online services, and will refer as needed to ‘cloud computing’ as a type of IT computing services for hire within the online ecosystem. “ | (Greenpeace, 2011, p. 5) |
| Greenpeace<br>*Make IT Green* | “The term cloud, or cloud computing, used as a metaphor for the internet, is based on an infrastructure and business model whereby - rather than being stored on your own device - data, entertainment, news and other products and services are delivered to your device, in real time, from the internet.” | (Greenpeace, 2010, p. 2) |
| European Commission<br>*The future of cloud computing* | “… a 'cloud' is an elastic execution environment of resources involving multiple stakeholders and providing a metered service at multiple granularities for a specified level of quality (of service). “ | (European Commission, 2010, p. 8) |
| Nationalencyklopedin (webpage)<br>*Molnet* | Translated from Swedish:<br>“The cloud … external server space for computer services that is available via the Internet from individual computers.” | (NE, 2014 a) |
| | **Definition: Cloud computing** | |
| The Global e-Sustainability Initiative<br>*SMARTer 2020: The Role of ICT in Driving a Sustainable Future* | “System of computing in which the computing resources being accessed are typically owned and operated by third-party providers on a consolidated basis in data centre locations.” | (GeSI, 2012, p. 239) |
| Thomond P., Williams D. R, and Mackenzie I and Velkov A.<br>*The enabling technologies of a low-carbon economy:* | “Cloud computing enables computing services (software, platforms, infrastructures) that are traditionally provisioned on- | (Thomond, et al., 2013, p. 5) |

| | | |
|---|---|---|
| *A focus on cloud computing* | site within enterprises[1] to be delivered across the internet, on-demand from purpose built data centres."<br>"1. For the purpose of this report, we use the term *enterprise* to refer to all private, public and non-governmental organisations." | |
| Williams D. R., Thomond P and Mackenzie I.<br>*The Greenhouse Gas Abatement Potential of*<br>*Enterprise Cloud Computing* | "Cloud computing enables computing services (software, platforms, infrastructures etc.) that are traditionally provisioned on-premise within organisations to be delivered on-demand from purpose built data centres across the internet. | (Williams, et al., 2013, p. 21) |
| Accenture and WSP<br>*Cloud Computing and Sustainability:*<br>*The Environmental Benefits of Moving to the Cloud* | "Cloud computing—large-scale, shared IT infrastructure available over the internet-…." | (Accenture and WSP, 2010, p. 2) |
| Koshy K., Juby A., Namboodiri V. and Overcash M.<br>*Can Cloud Computing Lead to Increased Sustainability of Mobile Devices?* | "Cloud computing is a model which is omni-present allowing on-demand network access to a pool of computing resources that are shared." | (Koshy, et al., 2012, p. 1) |
| Mell P. and Grance T.<br>*The NIST Definition of Cloud Computing* | "Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g. networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction." | (Mell & Grance, 2011, p. 2) |

## 13.2 KOOMEY'S CALCULATIONS

When Koomey presented data center electricity use for 2010 he first divided servers into volume, mid-range and high-end. Then he made four scenarios; all trends continue, best guess 2007 and a lower bound and an upper bound. Only the two latter scenarios will be presented here as the other scenarios are not likely to reflect reality. The all trends continue scenario uses the trends from an earlier report by Koomey, investigating the years 2000 to 2005, and extrapolate number of servers and energy usage to 2010. The best guess 2007 scenario also uses the energy trends from the earlier report and an IDC forecast from 2007 to calculate number of servers. The IT market has evolved a lot since 2005 which these scenarios do not reflect. The electricity use of servers in the lower and upper bound scenarios are presented in Table 16 and Table 17.

Table 16. The number of servers, power consumption and total server electricity consumption during 2010 is presented, using the scenario lower bound.

| Lower bound | Type of server | | | |
|---|---|---|---|---|
| | **Volume** | **Mid- range** | **High-end** | **Total** |
| **Number of servers** | 31 620 000 | 991000 | 117200 | - |
| **Average power consumption per server (W)** | 242,6 | 587,5 | 7597,4 | - |
| **Total server energy consumption 2010 (TWh)** | 67,2 | 5,1 | 7,8 | **80,1** |

Table 17. The number of servers, power consumption and total server electricity consumption during 2010 is presented, using the scenario upper bound.

| Upper bound | Type of server | | | |
|---|---|---|---|---|
| | **Volume** | **Mid- range** | **High-end** | **Total** |
| **Number of servers** | 31 620 000 | 991000 | 117200 | - |
| **Average power consumption per server (W)** | 294,2 | 840,9 | 12662,3 | - |
| **Total server energy consumption 2010 (TWh)** | 81,5 | 7,3 | 13 | **101,8** |

Koomey then calculated the electricity use of storage and communication by adding 25 % versus 15 % to the server usage. The total IT equipment electricity usage could then be compiled by summarizing server, storage and communication. To account for the infrastructure electricity use a PUE value of 1.83 (lower bound) versus 1.92 (upper bound) was used. The total IT equipment electricity usage was multiplied with PUE -1 to calculate the infrastructure electricity use. These calculations are summarized in Table 18. The mean value of 237.7 TWh is the value that has been used by GeSI, Greenpeace and Malmodin.

Table 18. The server, storage, communication, infrastructure and total electricity use of data centers worldwide according to Koomey (2011).

| TWh during 2010 | | | |
|---|---|---|---|
| | **Upper bound** | **Lower bound** | **Comment** |
| **Servers** | 101,8 | 80,1 | - |
| **Storage** | 24,4 | 19,2 | 25% of server power |
| **Communication** | 15,3 | 12,0 | 14 % of server power |
| **Infrastructure** | 130,2 | 92,4 | Based on PUE |
| **Total** | **271,7** | **203,8** | Mean 237,7 |

## 13.3 SMARTER 2020 CALCULATIONS

In GeSI’s (2012) report *SMARTer 2020* the assumptions behind the presented carbon dioxide emission are described. The carbon dioxide emission are calculated inter alia by using the operational electricity demand and the world carbon emission factor of 0.6 kg $CO_2$ e/kWh. This section present the operational electricity demand of ICT, using the number of equipment, electricity use per equipment and overhead used in GeSI’s report. As it is only operational electricity that is regarded in this section, shipment and embodied emission are not included. The sources to the different values are presented as well.

### 13.3.1 DATA CENTERS

As has been mentioned before GeSI used Koomey's (2011) estimation of electricity use from data centers to calculate carbon dioxide emissions from data centers. The estimation Koomey presented was for the year 2010. As for the estimations of data center emissions in 2020 the increase of carbon dioxide emission was assumed to 86 percentages, see Figure 21. This estimation is solely based on electricity use and can thus be used to calculate the future electricity demand from data centers.

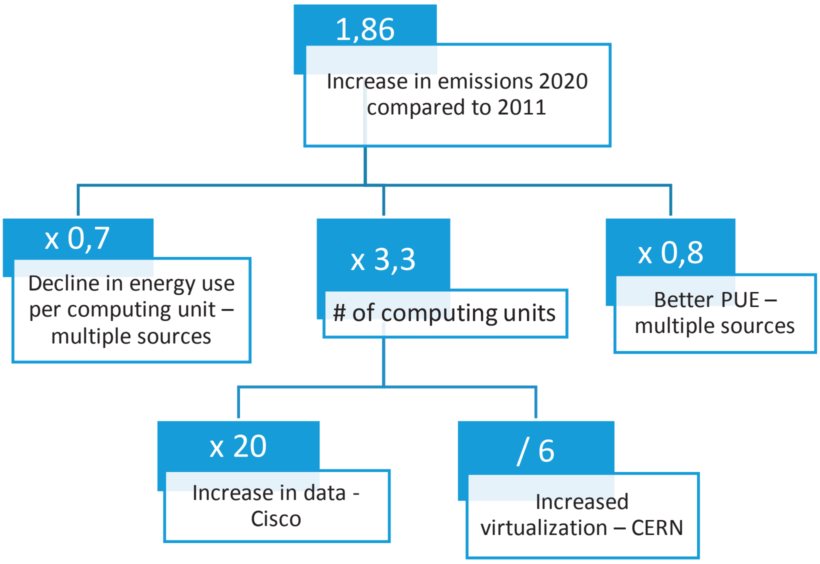

Figure 21. Assumptions behind the estimations of data center electricity use 2020 in the *SMARTer 2020* report by GeSI (2012)

### 13.3.1 DATA TRANSFER

Table 19. The estimated energy usage of data transfer in 2011, using the assumptions from the *SMARTer 2020* report.

| | 2011 | | | | | | |
|---|---|---|---|---|---|---|---|
| **Network** | **Subscriptions (billion sub)** | **Source** | **Electricity use (kWh/sub)** | **Source** | Overhead factor | **Source** | **Total electricity use (TWh)** |
| Wireless Mobile Connection | 5,9 | GreenTouch | 17 | GSMA: Mobile's Green Agenda 2012 | 1,2 | (Malmodin et al., 2010) | 120 |
| Wired Home Access Network | 1,63 | GreenTouch and Ovum | 45 | (Malmodin et al., 2010) | 1,25 | (Malmodin et al., 2010) | 92 |
| Wired Enterprise Network | 0,73 | Gartner | 35 | (Malmodin et al., 2010) | 1,25 | Assumption | 32 |
| | | | | | | | **244** |

Table 20. The predicted energy usage of data transfer in 2020, using the assumptions from the *SMARTer 2020* report.

| | 2020 | | | | |
|---|---|---|---|---|---|
| **Network** | **Subscriptions (billion sub)** | **Source** | **Electricity use/ sub (kWh/sub)** | **Source** | **Total electricity use (TWh)** |
| Wireless Mobile Connection | + 62 % of total emissions** | | | | 195 |
| Wired Home Access Network | 1,74 | + 0,5 %, GreenTouch and BCG research | 34 | Expert interview | 59,2 |
| Wired Enterprise Network | 1,4 | BCG research | 27 | Expert interview | 37,8 |
| | | | | | **292** |

** The number of subscriptions and electricity use was not presented in *SMARTer 2020*, the calculation was only given in carbon dioxide emission. Thus the increase of 63 % emissions are not directly the same as 63 % increased electricity use as GeSI also weighed in increased use of green energy sources of 4 %. But for a roughly estimation the same percentage are used.

### 13.3.2 END USER DEVICES

Table 21. The estimated energy usage of end user devices in 2011, using the assumptions from *the SMARTer 2020* report.

| | 2011 | | | | | |
|---|---|---|---|---|---|---|
| **Device** | **Number (billion)** | **Source** | **Electricity use (kWh)/device** | **Source** | **Total electricity use (TWh)** | Source |
| **PCs (desktops and laptops)** | 1,53 | 45 % laptops, Data from Gartner and Forrester | 218,98 | Boston Consulting Group Research | 335 | - |
| **Tablets** | 0,07 | Gartner and Forrester | 15,6 | Boston Consulting Group Research | 1,1 | - |
| **Smart Phones** | 0,66 | Gartner and Forrester | 5,45 | Boston Consulting Group Research | 3,6 | - |
| **Mobile Phones** | 3,65 | Gartner and Forrester | 2,5 | Boston Consulting Group Research | 9,1 | - |
| **Printers** | - | - | - | - | 13 | Gartner |
| **Total** | | | | | **362** | |

Table 22. The predicted energy usage of end user devices in 2020, using the assumptions from the *SMARTer 2020* report.

| | 2020 | | | | | |
|---|---|---|---|---|---|---|
| **Device** | **Number (billion)** | **Assumption** | **Electricity use (kWh)** | **Assumption** | **Total electricity use (TWh)** | **Assumption** |
| **PCs (desktops and laptops)** | 2,81 | + 7 %, 70 % laptops | 101,66 | Extrapolations from industry trends, assuming more laptops, more efficient PCs and less usage | 286 | - |
| **Tablets** | 1,81 | + 37 % | 15,6 | Extrapolations from industry trends (Same as 2011) | 28 | - |
| **Smart Phones** | 2,8 | + 17 % | 5,5 | Extrapolations from industry trends | 15,4 | - |
| **Mobile Phones** | 4,17 | + 2 % | 2,5 | Extrapolations from industry trends (Same as 2011) | 10 | - |
| **Printers** | - | - | - | - | 13 | Same as 2011 |
| **Total** | | | | | **353** | |

*The increase of end user devices is derived using global penetration rates and a world population of 7.6 billion in 2020.

## 13.4 CLOUD PROVIDERS IN SWEDEN

Some cloud providers operating in Sweden have been studied and interviewed in order to get an insight to their approach to energy efficiency, usage of renewable resources and how they think the cloud will evolve in the future.

### *13.4.1 GOOGLE*

The search engine Google is constantly looking for renewable energy sources for its data centers. In Sweden the company has signed a contract with two wind power companies, Eolus Vind and O2, to serve its data centers in Finland with renewable energy. Eolus Vind's four wind farms with 29 wind turbines in southern of Sweden will provide electricity to Googles data centers for 10 years according to the contract.[14] O2 energy will build 24 wind turbines in northern Sweden on behalf of Google. Francois Sterin, who is responsible for the global infrastructure at Google, says that the excess electricity will be sold to the electricity market. (Sydsvenskan, 2014; Miljöaktuellt, 2014)

### *13.4.2 FACEBOOK*

One company that has decided to build data centers in Sweden is Facebook; in March 2014 the first building of three was finished. This was the company's first data center outside of the USA. (Luleå Näringsliv AB, 2014). The first hall has a PUE of 1.05 - almost all electricity used in the facility goes to powering the server and network equipment (Facebook, 2014). Facebook has decided to start building the second data center during 2014 as well. The buildings are/will be powered by renewable hydropower from Lule River according to Facebook. When all three buildings are finished and operating, the electricity demand will reach 0.47 TWh per year, assuming normal operation according to the environmental permit (Pinnacle Sweden AB, 2011). Other sources report that the electricity

---

[14] Note that the actual electricity mix that are used by the data centers most probably have another mix than was purchased. There is a different between actual consumption and what you buy. By for example purchasing 100 % renewable electricity the total portion of renewables increases in the country/region even though the electricity you consume at your facility might be from the local coal fired plant.

demand can reach 1 TWh per year (Svenska Dagbladet, 2013). The magazine also report that Facebook has over 800 million users and the data centers in Luleå are planned to serve a large part of them. The excess warm air from the servers will partially be used to warm up office spaces during winter but most of it will be released to the outside air (Pinnacle Sweden AB, 2011).

Facebook is sharing how they are building their data centers through the Open Compute Project. It is a platform where ideas and processes of how to build an energy efficient data center are shared and input can be given and received. Facebook are also striving to store files more efficient and thus lowering the electricity demand that way. To the question if Facebook ever thought about limiting the amount of or time pictures and videos are shared they answer that it is up to each user to decide what and how long they want to share their media. (Fredriksson, 2014)

### *13.4.3 TIETO*

Tieto is operating in twenty countries and offers cloud services of all three service types; IaaS, SaaS and PaaS, along with traditional services and legacy. The headquarters are in Helsinki in Finland and the company advertises itself as the largest IT service company in Northern Europe. (Tieto, 2014) Two coworkers at Tieto in Sweden have been interviewed; Björklund, responsible for data center facilities, and Landström, strategist for data center locations in the future.

#### Data centers

Tieto has two data centers in Sweden; one older in Älvsjö and one newer build 2010 in Tullinge (Björklund, 2014). Both data centers are using 100 % renewable electricity according to Björklund. In Älvsjö free cooling is not used, but in Tullinge it is (Landström, 2014). According to Björklund, Tieto are prepared for an increased demand of storage space from their clients. Traditional IT is expected to decrease while cloud services are expected to increase at a rapid pace. Tieto started offering cloud services in 2012 (Björklund, 2014). Today around 20 % of Tieto’s server capacity is used by cloud services according to Landström. Even more, around 65 %, of the servers have cloud features as virtualization (Landström, 2014).

#### Efficiency measurements

Björklund says that Tieto has chosen not to calculate PUE of their data centers as it is a misleading term, not saying anything about their energy reuse. In Älvsjö excess heat from the data center is reused in their own offices and in Tullinge excess heat will be transmitted to the nearby residential area. Tieto has not calculated any of the other measurements presented in section 0 either, but Landström says that Tieto has started working on it and will present the different efficiency measurements in the future. Landström also states that clients during recent years have started to request environmental friendly business and efficient data centers.

There are some recommendations of how a data centers should be built according to Landström, but no actual laws of how energy efficient data centers have to be. Landström describes data centers as an immature market compared to the manufacturing industry when it comes to laws and regulations.

#### Economy

Electricity demanding manufacturing industry can get tax reduction of the electricity bill in Sweden, in order to make the companies competitive towards companies in other countries which have lower electricity cost. Björklund says that no such tax reduction exist for data centers, even though they as well use a lot of electricity and compete with companies in other countries. Björklund choose not to confide the total electricity use of Tieto’s data centers. (Björklund, 2014)

### *13.4.4 BAHNHOF*

Bahnhof delivers services within data and telecommunication in mainly Sweden. To protect integrity and build secure data centers are important parts of Bahnhof's business according to the company's homepage. (Bahnhof, 2014) Gustav Bergquist, an employee and technical manager of Bahnhof, has answered some questions of Bahnhof's operations, the answers are compiled below.

#### Data centers

Bahnhof has four big data centers in Sweden, sited in Stockholm, Uppsala and Malmoe. The newest data center, finished in 2012 is named Thüle and is sited in Stockholm and one even newer is under construction in Malmö, named Sparven. About 10-20 % of Bahnhof's own servers are visualized, according to Bergquist. There are however many servers leased by Bahnhof to banks and insurance companies that Bahnhof themselves do not control the configuration of and thus do not know the visualization degree of. (Bergquist, 2014)

Only a small fraction of Bahnhof's business is to provide cloud services according to Bergquist, but he points out that if the customers demand more cloud services then Bahnhof will also provide it. Today three server racks are used for Bahnhof's public cloud. Bergquist continues to describe that it is mostly small or medium companies that benefit from a public cloud. Bigger companies have the strength and incentives to make their data halls effective themselves, so called private clouds. Bergquist believes in the growth of hybrid clouds, where a company owns and controls the hardware and firewalls but can rent out excess resources as a public cloud and thus benefit from the cloud idea (Bergquist, 2014)

#### Energy sources

Bahnhof power their data centers with water power only. They use therefore 100 % renewable electricity sources. The have also plans to buy wind farms and thus benefit from tax reduction. If the generated electricity from those wind farms correlates to the electricity consumption of Bahnhof's data centers, they do not have to pay any electricity taxes. The Swedish government are discussing to set a boundary of the tax reduction though and therefore Bahnhof sits tight and awaits the results. (Bergquist, 2014)

#### Efficiency measurements

The data center that is under construction in Malmoe will probably have an PUE under 1.2 and the newest data center in Stockholm, built 2012, have an PUE of 1.2 according to Bergquist. He points out though that he believes PUE to be a falsely representative of energy efficiency. He states that it is easy to achieve good values by for example buying remote cooling from another company and not include the energy it takes to produce that cool water in the efficiency calculations. Also, if Bahnhof chose to deliver heat to the local district heating system they have to install heat pumps that in the end will enhance their PUE value. Bahnhof have not calculated other efficiency measurements as ERF or CUE. (Bergquist, 2014)

There are two companies in Sweden which construct data centers, but they do this in the exact same way as they did ten years ago and thus the results are not so effective according to Bergquist. Therefore Bahnhof projects, designs and maintain their own data centers. (Bergquist, 2014)

There are mainly two incentives to make a data center more effective according to Bergquist:

- There are market advantages to be seen as ”green”. It is good for business nowadays to be energy effective and use renewable energy sources.
- There are money to be saved when improving energy performance.

### Law and regulations

There are no tax reductions for data centers today and no law as to how effective a data center have to be. A future law of how effective a data center have to be would probably not concern Bahnhof as the company already are very energy efficient and have worked with that kind of questions for a long time according to Bergquist. If such law would come to pass it would mainly affect companies with no profit making business, other companies would already have made the acquired efficiency improvements. (Bergquist, 2014)

### Core and access networks

Bahnhof owns and maintains some networks as well but they are not working at making them more efficient as it is a very small total electricity cost. A router or a switch serving 300 households uses about 70 W and it is too small profit, if any, to make them more energy efficient. (Bergquist, 2014)