Data Centres and Power: Fact or Fiction?

Introduction

Data centres are relatively new features on the UK's urban landscape. In fact, many people are unaware that they even exist. Often occupying boring-looking industrial buildings or nondescript offices, data centres take the form of rooms or halls packed full of computing equipment, patrolled by small numbers of geeky looking individuals who talk in a language only vaguely related to English. Nobody seems to understand why they are there or what they actually do, which in most cases appears to be shrouded in secrecy. As with any poorly-understood phenomenon, a whole mythology has built up around them. This mythology in particular focuses on the way data centres use energy.   These notes identify and address some of the common assumptions about the way that data centres use power. These are expressed as a series of assertions that have actually been made about data centres.   Against each assertion we have explained whether it is true or false and why.

The assertions are loosely grouped into themes. These include power use (the amount and type of energy that data centres use), efficiency (how effectively energy is deployed in data centres), consumer perceptions (our limited understanding of how networked devices really work), future trends (how the industry is – or should be – developing) and calculations (worked examples to demonstrate that we need to get our sums right if we are going to make claims one way or the other).

Each heading in the list below takes you to a detailed explanation. The points aren't intended to be read in order but sampled like a smorgasbord.

I	Power use: the amount and type of energy that data centres use	
1	**"Data centres use lots of power"**	**True, and here's why...**
2	**"Data centre power demand is growing exponentially"**	**False, but.....**
3	**"Data centre operators don't care about power because they simply pass through power costs to their customers".**	**False**
4	**"Data centres use noisy polluting generators"**	**0.1% True, 99.9% False**
5	**"Data centres use dirty power"**	**It depends....**
6	**"Data centres are going to gobble up all the electricity until there is none left for anyone else"**	**False**
II	Efficiency: how effectively is energy deployed in data centres	
7	**"Data centres are inefficient"**	**It depends....**
8	**"Data centres don't use power productively"**	**False**
9	**"Data centres are full of servers doing nothing except using energy"**	**Generally false but...**
10	**"Data centre operators don't care about energy efficiency"**	**False**
11	**"Servers are inefficient because they can only operate within a very small temperature range"**	**False – but....**
III	Consumer perceptions: understanding how our networked devices really work	

Section I   Power use:
The amount and type of energy that data centres use

1   "Data centres use lots of power"

True: Data centres use lots of power. And this is why: Firstly, data centres support and enable service economies in the way that heavy industries supported manufacturing economies so you should expect some energy intensive activity to be taking place. Secondly, data centres represent the consolidation of computing demand. This means they are super-concentrating lots of activities that individually are not very energy intensive but when consolidated become extremely energy intensive simply because of the degree to which they are concentrated. Thirdly, data centres use power in two ways: they need power to run the IT equipment that they house (ie servers which execute the digital transactions that we rely on) and, because servers emit heat when they are working[1], they need power to keep those servers cool enough to function reliably (see point 2 below). A very large data centre may consume 30GWh of power in a year, costing its operator around £3,000,000 for electricity alone. A handful of sites in the UK consume even more than this although the majority of sites consume far less.   The total power demand of the UK data centre sector is between 2-3TWh per year[2]. Energy is usually the largest single element of operating costs for data centres, varying from 25-60%.

---

[1] Servers actually emit heat even when they are idle, which means that they are on but not doing anything much.

[2] This is the power consumed by what we call "recognisable facilities" by which we mean data centres, and does not include power consumed by server rooms and cupboards and other forms of distributed computing.

2   "Data centre power demand is growing exponentially"

False: But there's no room for complacency. Data centre power demand is **not** growing exponentially, but the demand for digital data **is** growing exponentially (see diagram). To keep up with this demand, one might expect that data centre power would increase in line but fortunately it doesn't, because of Moore's Law[3]. However, despite Moore's law, data centre power demand is growing year on year. Figures from DCD global census suggest an increase of around 14% a year. There are a number of reasons for this: Firstly the rate of increase in demand for digital data is faster than the savings achieved through Moore's Law. Secondly Moore's Law applies to leading edge technology and state of the art computing equipment will not be installed in every data centre[4] so there will always be a lag. Thirdly the ICT sector is gradually moving away from distributed computing (qv) towards the consolidation of computing within purpose built facilities (i.e. data centres!) so power use that was previously hidden is now appearing in these figures, even though there is a net energy saving[5].

Nerd Notes: Moore's Law applied to Power

Chart showing (simplistically speaking) the power required by a chip to process a given amount of data. This has decreased by 6 orders of magnitude in 30 years. (Note logarithmic scale on Y-axis).   Source: Intel
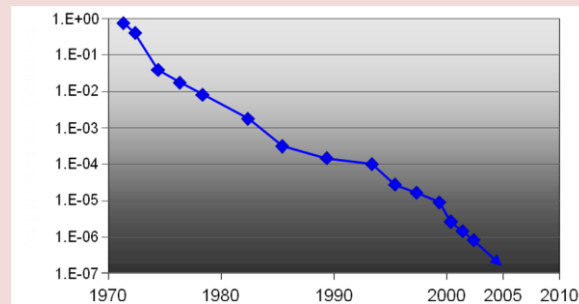
Diagram comparing growth rates in data centre energy capacity with growth rate in data demand.
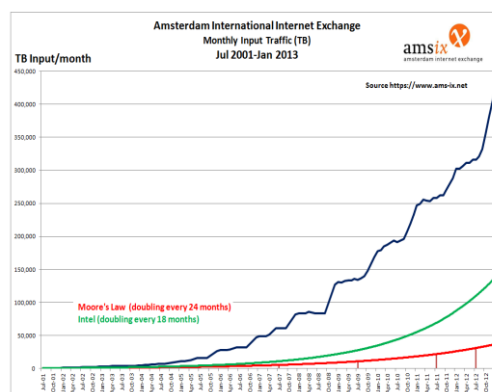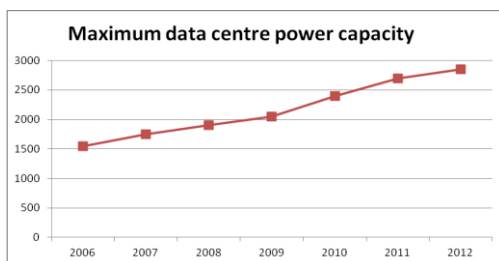
Increase in total power capacity of data centres 2006-13 (Source: DCD) vs increase in demand for digital data, 2001–13 (Source: Dr I Bitterlin CEng/Amsix)

---

[3] Moore's Law says that the processing power of a chip will double every 18-24 months,. Moore's law also applies to energy. Simplistically the energy required to process a given amount of data is halving every 18-24 months.

[4] – IT refresh cycles have to take into account embodied energy as well as energy efficiency in use

[5] It's worth mentioning that new semiconductor devices are being developed all the time with lower power consumption. A new type of semiconductor could completely transform the power requirements of microprocessors (chips). Traditionally, processing speed was the ultimate objective – now energy efficiency is the driver and this may result in a new generation of processors.

**3	"Data centre operators don't care about power because they simply pass through power costs to their customers"**

False: Data centres may pass power costs to their customers but that doesn't mean they don't care about power. In fact, this is one of the reasons that they **do** care so much about power. Data centres provide highly commoditised services for location agnostic[6] organisations. Data, after all, is the most mobile commodity on earth and the majority of companies needing data services can choose where they buy them from depending on cost. Energy is a very high proportion of data centre cost and the higher that energy cost is, the less competitive that operator is in a global market.

**4	"Data centres use noisy polluting generators"**

0.1% true: 99.9% false. Data centres rely on electricity from the grid. Very occasionally, there is a power cut and grid electricity is not available. Data centres have to continue running irrespective of interruptions in power supply so they have batteries to supply immediate power in the event of short term grid failures and they have generators to provide power for longer outages.   These generators have to be test fired on a regular basis to ensure they will work if they are ever needed. The number of times data centres actually run "for real" on diesel generators will be dictated by the number of significant grid power outages. How many of those can you remember in the last year? The answer, in case you wondered, is very close to zero in the UK. Evidence to support this comes from industry data collated by Intellect in 2012. The power used by generators in our sample was well below 0.1% of the total power consumption.

Very occasionally a data centre operator may have an agreement with their electricity supplier to use their generators for short times to reduce load on the grid when peak demand becomes critical.   Since this allows the National Grid to manage fluctuations in demand without expanding our generating infrastructure, this practice has a beneficial impact by smoothing demand.
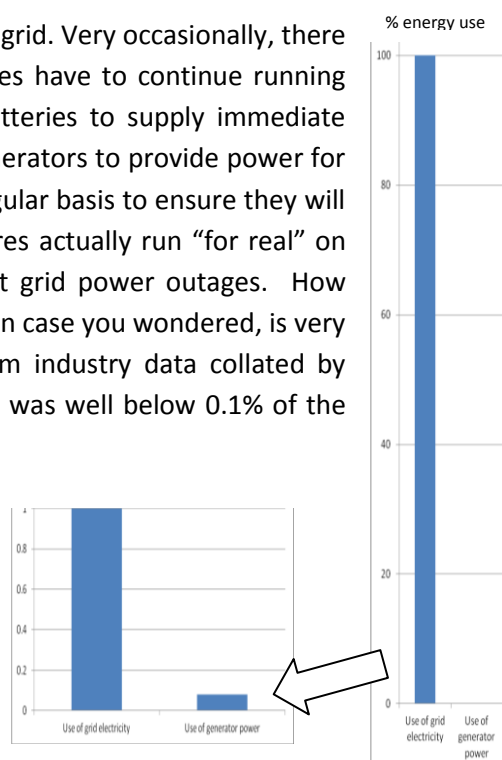
**5	"Data centres use dirty power"**

It depends: Data centres use electricity so a data centre is as clean as the electricity supplied to it.   Powering data centres from renewables or low carbon generation on site is problematic because data centres need constant and reliable power supply and this is not yet feasible using current technologies. That doesn't mean that it will not be possible in future and some companies are siting new data centres near to renewable power sources. In the meantime many take the view that the next best thing is to drive demand for renewables by buying green power indirectly. You can pay a premium for certified green power from your energy supplier. Interestingly, 72% of the energy used by the ten largest data centre operators in the UK is certified green power, for which they have chosen to pay this premium. What kind of power do you use?

---

[6] Organisations that can locate their operations anywhere in the world. This would contrast, say with a coal mine, which has to locate its operation close to a source of coal. Coal mining is therefore a location dependent activity.

6   "Data centres are going to gobble up all the electricity until there is none left for anyone else"

False (drivel). Our insatiable demand for digital data will not lead to system meltdown. Why not? Because the movement of digital data is price elastic[7]. If power is either not available or the market price for power increases, then people will think twice about what they upload or download. If we ever got to the stage when a Google search really did consume the same amount of power as boiling a kettle (contrary to the Guardian's estimate, see below) then there would be a strong price signal against frivolous activities involving the movement, storage or processing of digital data.


Section II   Efficiency
How effectively is energy deployed in data centres?

7   "Data centres are inefficient"

There are two answers to this question.

Firstly: False, because data centres are FAR more efficient than the alternative. Data centres are invariably more efficient than the alternative – distributed computing. In distributed computing, the computing functions that support all ICT-enabled processes are done by servers[8] that are housed in the same buildings as the organisations that rely on them. Distributed computing is very inefficient but it is also very common, and we estimate that over half the servers in the UK are accommodated in this way. This means that they are spread around office buildings in cupboards, box rooms and basements that are not purpose built and usually need powerful fans to keep the computing equipment cool. Very significant energy and financial savings can be achieved by collecting all this activity together in specialised, purpose-built facilities. The problem is that as soon as you do this, then the energy use associated with these computing functions, which was hidden across organisations, becomes apparent and quantifiable. That is of course a good thing because it is much easier to monitor and manage energy use that you know about than energy you don't know about. However, data centres attract criticism for being power hungry when in fact they are far less power hungry than the same amount of distributed computing. The difference is that you can see the power.

Inappropriate uses for a server cupboard

#67 - drying laundry

---
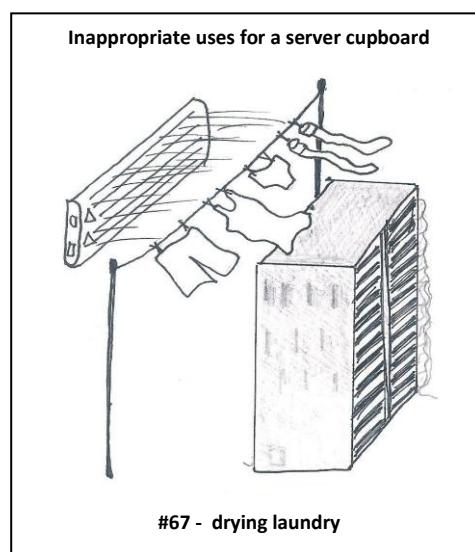
[7] If a product or service is price elastic this means that the cheaper it is, the more people buy it. Communications services are very price elastic. If a product is price inelastic then reducing its price will not make much difference to how much people buy. An example of a price inelastic commodity is a coffin. However cheap coffins are, total demand for coffins is unlikely to rise.

[8] Simplistically speaking, a server is generally a computer that is remote from its operator and usually performs specialised functions. This distinguishes it from a desktop which is a personal computer that performs lots of different functions for an individual user or users.

Secondly: It depends: Data centres vary in their efficiency. It is true that a significant proportion of power delivered to a data centre may be used for cooling and other ancillary functions rather than being delivered to the IT equipment within it. Strictly speaking this energy is not being wasted because it is used to maintain conditions of temperature and humidity within which the IT hardware can be guaranteed to operate reliably.

Nerd Notes: What is PUE?
PUE, or Power Usage Effectiveness is the relationship between the overall energy demand of a data centre and the energy used purely to power the IT activity. PUE was developed by the Green Grid (qv) in 2007 and although it is not perfect it is the best metric we have to measure data centre efficiency. The Green Grid has also developed additional metrics which include WUE (Water Use Effectiveness), CUE (Carbon Use Effectiveness – measuring the carbon impact of energy generated for data centre operations), GEC/REC (Green Energy Coefficient/ Renewable Energy Coefficient which is the proportion of energy consumed by data centre that is renewable/ green) , ERE (Energy Reuse Effectiveness – which takes into consideration alternative uses for waste energy from data centre operations), and EDE (Electronics Disposal Efficiency – which assesses the disposal of the IT hardware through known responsible entities). Other approaches under discussion include DCcE (Data Centre Compute Efficiency, which enables operators to assess the efficiency of their computing resources) and ScE (Server Compute Efficiency, a sub-set of DCcE)
The ISO is also developing important standards in this area which will similarly cover PUE, WUE, CUE. They are also developing ITEE (IT energy efficiency) and ITEU (IT Energy use) and two further standards to cover use of renewable power and re-use of waste heat.

However, it is not productive energy and data centre operators strive to minimise it. The ratio between total energy delivered to a site and energy used by the IT is known as PUE: Power Usage Effectiveness. A PUE of 2 means that 50% of the power delivered to the site is used to run the IT within it. A PUE of 1.25 means that 80% of the energy delivered to the site is used to power the IT (or in other words that for each 1Wh of energy used by the IT, another 0.25Wh is consumed by ancillary equipment). The closer the PUE is to 1, the more efficient the site. A leading-edge, highly efficient data centre might have a PUE of 1.2 and average good practice is around 2. A high PUE of 3 and above, means that the data centre is either new and is only partially full of IT equipment[9], or that it is inefficient. There is some trade-off between resilience and efficiency.

Nerd Notes: Efficiency vs. resilience
There is a degree of trade-off between efficiency and resilience. Highly resilient facilities tend to have slightly higher PUEs. This is why PUE is best used to compare the efficiency of a single data centre over time. If you use PUE to compare data centres then you need to take into account the level of resilience each one offers, as well as a range of other factors.

Moreover, a PUE of 2 is actually not that bad when you look around. How much of the power of an incandescent light bulb emerges as light?
About 5% - now you can see why they are banned! How much of the heat from an open fire actually makes it into the room? About 8% (the other 92% goes up the chimney).   How much of the power used to heat homes actually heats rooms that people are in at the time? A pitifully small proportion.   And a glance inside the average teenager's bedroom will tell you some very sad things about how productively our domestic electricity supply is being used. So data centres are in fact much better at using power than lots of other things. The difference is that data centres bother to measure and report on it.

---

[9] This is true of older data centres but not necessarily true in a modern data centre where power used for cooling should be more closely proportional to power used for the IT.

**8	"Data centres don't use energy productively: they consume lots of power without doing anything useful"**

False: Data centres are not just power hungry sheds full of computers that do nothing more than store pictures of people teaching their dogs to tap-dance. Data centres are the engines of the digital economy and no modern economy can function without them. As well as supporting critical national infrastructures, telecommunications, government operations and security, they deliver corporate IT services, e-government services, financial services, digital media, mobile computing and knowledge management. Yes, they enable social networking, internet access, online shopping and gaming but they also underpin all the technologies that enable a low carbon economy, such as smart grid, intelligent transport systems, smart buildings and smart cities. They even facilitate a whole range of virtual substitutes for carbon intensive activities, such as virtual presence and transport substitution technologies, e-commerce and access to creative content through digital downloads. Data centres pervade and enable our everyday activities, from commuting to shopping. If you want to understand what data centres do then read our paper "So What have Data Centres Ever Done For Us?" If you want to understand how much of your own day to day existence is enabled by a data centre, then see our related publication "Data Centres: A Day in YOUR Life" or our illustrated example "Buying a Hat Online".

And don't forget that (as covered in point 2) data centres actually help to improve the efficiency of ICT by concentrating it into purpose built facilities. So on the one hand, data centres help reduce the energy impact of ICT and on the other hand they underpin a range of critical functions.   They are not there just to satisfy the whims of budding Dog Astaires.

**9	"Data centres are full of servers doing nothing except using energy"**

Generally (but not always) false – i.e. it depends: Utilisation[10] rates in servers vary. In data centres where work can be scheduled, utilisation rates should be high. Maintaining high utilisation in other areas is more problematic unless demand is very predictable. In many cases, however, utilisation rates are low and could be improved. This may be due to reasons like institutional memory loss (people simply don't know what those servers are doing so daren't turn them off – they could be storing records of shoe sizes of people from Wolverhampton in 1972 or they could be managing all the applications that support an active communications or IT infrastructure) or risk-averse planning (people use a new server for every new application because they are worried about compromising an existing server's performance). However, in a country like the UK where energy costs are relatively high it would be surprising to find a commercial operator with a data centre full of comatose servers. It is increasingly common practice to audit data centre activity regularly and to identify and locate comatose servers, and to make better use of servers by consolidating computing activity.

---

[10] Utilisation really means how busy the server is. If utilisation rates are high then the server is busy performing the tasks for which it is designed. If utilisation rates are low then the server is under-employed.
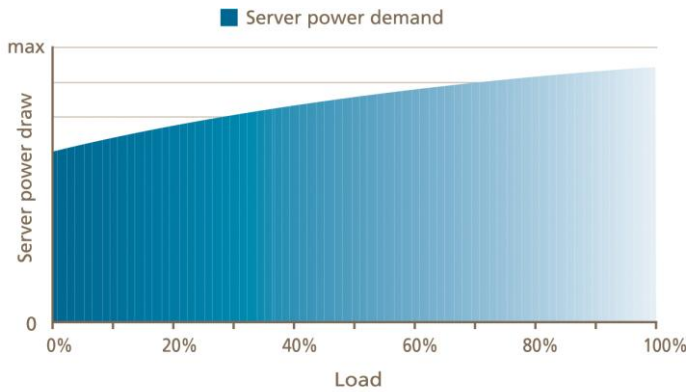
Server utilisation is a very serious issue because a typical server does not use much more power when it is working at capacity than when it is idling (see diagram). It is true that technological advances are improving server performance and that approaches like virtualisation[11] and cloud computing (qv) can dramatically improve server utilisation, but these solutions have not been applied widely enough yet.

10   "Data centre operators aren't interested in energy efficiency"

False: Rubbish in fact: Nothing interests a data centre operator more than how he uses energy. If you ever find yourself sitting next to one at a dinner party, we dare you to mention the word energy. His or her eyes will light up and you will be trapped, helpless while the food congeals on your plate and your host or hostess curses you for ruining a perfectly nice evening. The only way to escape is to feign death. If this seems too risky a social experiment for you, then just cast an eye down the topic list at any data centre conference. It's generally dominated by

concerns about power and how to use less of it. This is because power accounts for 25-60% of operating costs for data centres and 15-30% of total turnover.
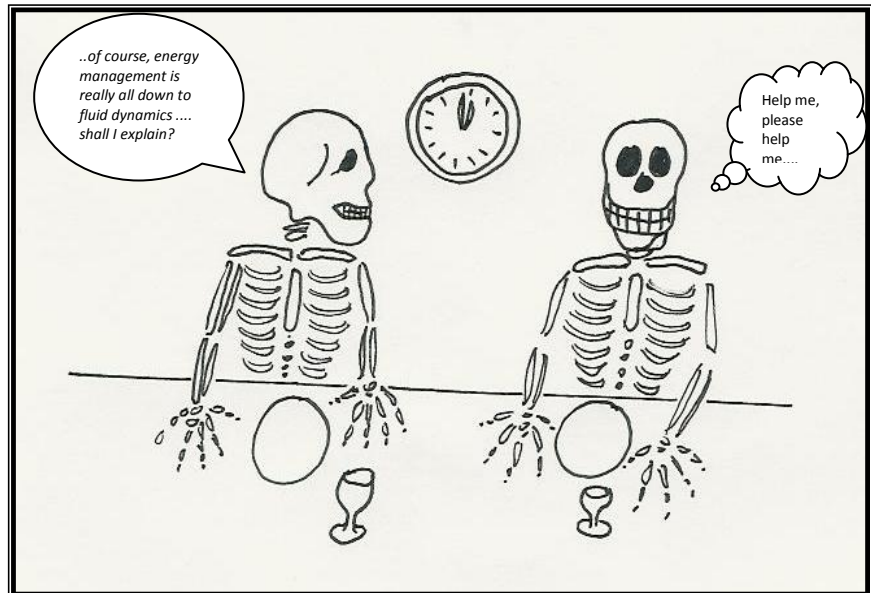
Moreover, if data centres didn't care about energy efficiency there would not be such a strong industry focus on improving energy use in the data centre. Highly conscious of the cost and availability of power, data centres have some of the most advanced tools of any sector for improving energy efficiency. These include The Green Grid(qv) and the Green Grid Maturity Model (qv), the European Code of Conduct(qv) for data centres, the work of the Uptime Institute, and the development of a range of industry standards and certifications for energy efficient data centre operations. The sector also makes widespread use of generic standards like ISO 9001, 14001, 27001 and 50001.

---

[11] At a server level, virtualisation turns a physical server into a virtual one. Virtual servers can share hardware resources without interfering with each other. This enables the consolidation of a number of virtual servers onto one physical server, whilst maintaining the individual properties of each of those systems.

11   "Servers are inefficient because they can only operate within a very small temperature range"

False – but they are often treated as though this is true....

One traditional concern of data centre operators was that certain operating conditions might lead to the servers breaking down or malfunctioning. So in 2004 ASHRAE (the American Society of Heating, Refrigeration and Air Conditioning Engineers) defined a common set of guidelines for operating conditions in data centres that would not invalidate the warranties provided by server manufacturers. In 2004 the
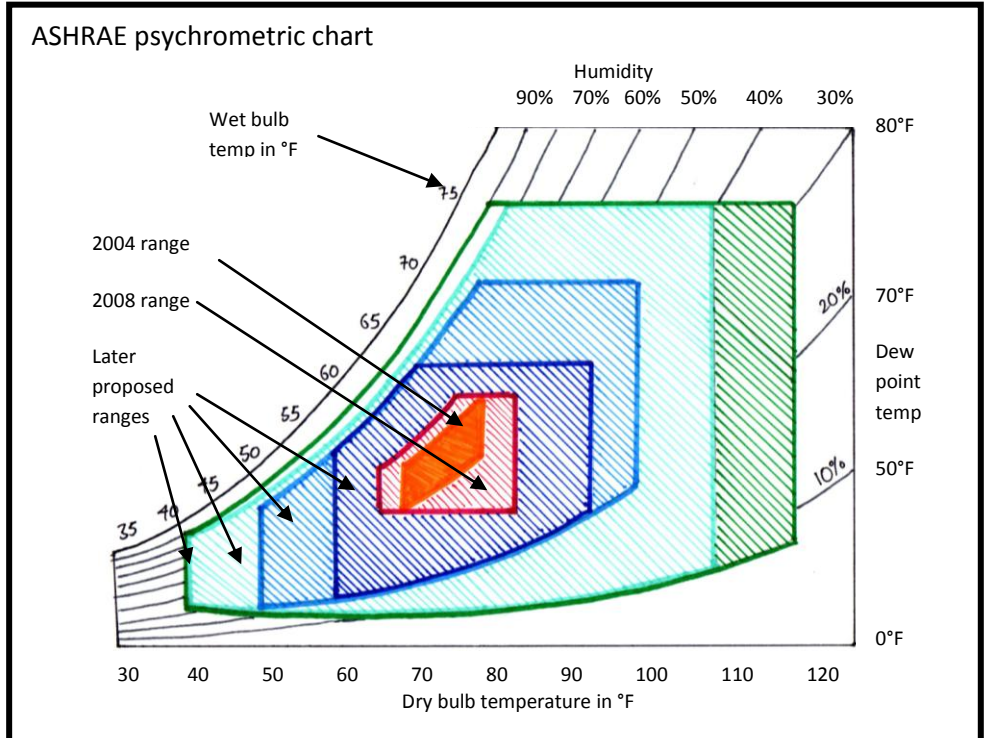
original envelope proposed was 20°C to 25°C (68-77F). In 2008 the range expanded from 18°C

ASHRAE psychrometric chart

to 27°C (64.4-80.6F). In 2011 the envelope was challenged again and allowable operating ranges as wide as 5° to 45°C (41 to 113 F) have been considered (see the ASHRAE Psychrometric[12] Chart, adapted from Don

Beaty's version, and the accompanying nerd notes). Many manufacturers guarantee their equipment to work reliably beyond of the 2008 ASHRAE ranges, and experiments at the University of Hertfordshire have

> **Nerd Notes: ASHRAE Envelopes**
> The chart shows both temperature and relative humidity. Dry bulb is what a thermometer reads, usually in the shade. Wet bulb is the effective surface temperature when moisture is evaporated (if you get out of the swimming pool when it is sunny but a breeze is blowing- your skin feels the cold as the water is evaporated. It used to be measured by wrapping the bulb of the thermometer in absorbent cloth, soaking it in water and then spinning it in a thing like a football rattle).
> The small rhombus in the middle represents the 2004 recommended operating envelope. The larger red polygon is 2008 and still remains the recommended range for any class of IT hardware in data centres (essentially the lowest common denominator). However, some systems and equipment are designed to tolerate wide ranges of temperature and humidity. If these are deployed in data centres then the operating envelope can be expanded. The larger blue and green polygons represent allowable ranges for different classes of hardware. For instance, C is the range for "ruggedised" systems which should operate reliably from 5-40°C.

demonstrated that server equipments works perfectly reliably well in quite extreme ranges of both temperature and humidity.

So there is no need for temperatures within data centres to be maintained so strictly. Yet we often find that they **are** operated within this range. Why? Because "safe operating" ranges of temperature and humidity are set by service level agreements (SLAs) between operator and customer which require the data centre to operate within unnecessarily constrained environmental conditions. In other words, operators are contractually obliged to waste energy on cooling. This is an important issue because cooling is very energy intensive – more energy intensive than heating. Relaxing temperature constraints by even 1°C could reduce the power used by cooling by up to 4%.

---

[12] Psychrometric means that it shows temperature, humidity and water vapour. Not to be confused with Psychometric which relates to personality profiling.

Section III   Consumer perceptions:

Understanding how our networked devices really work

12   "I've bought an energy efficient laptop so all my computing is green"

Er, not necessarily unless you never use it to access the internet, email, an online service like ebanking, or any other external digital network. Even then, there are other factors to take into account (See nerd notes). Most laptops are used for a combination of functions including teleworking, email, internet access and social networking. Teleworking, email, internet access and social networking are all activities that rely on external networks which in turn rely on server rooms and data centres. Your facebook photo of your tap-dancing dog does not get uploaded by magic – it needs a sophisticated communications infrastructure and at least one data centre, both of which use energy. Your computer therefore relies on any number of digital infrastructures (see Data centres: a Day in YOUR Life) to perform even the most ordinary operations. These infrastructures may be very efficient – or they may not. The problem is that for many of the things you do online, you simply can't tell. The same applies to the kind of energy used by that infrastructure – the data centre could be running on low grade coal in Oregon or on hydroelectricity in Norway. Many data centres do run on certified green power but because you can't easily see which parts of the infrastructure you are using at a particular time then you can't really be sure what kind of power you are using. So you have to be aware of the indirect energy demand of your device as well as its direct energy demand. One is visible to you (via your portable appliance tester or your electricity bill) and the other is not.   To cut a long story short you just need to understand that the energy impacts of your laptop – or any other networked device - extend far beyond the device itself. Controlling its direct power use only deals with part of the problem.

> **Nerd Notes: Energy performance of laptops**
> It is a very good thing to exercise your power as a consumer and choose an energy efficient laptop because this will drive competition around energy efficiency at the design stage. In fact the relative efficiency of laptops (compared, say to desktops) is evidence of this already happening, because there has been fierce competition around extending the battery life of laptops (length of time they can run before they need recharging). But the energy used by the laptop when in use is only part of its energy impact – the embodied energy (the energy taken to manufacture and distribute it to point of sale), is also significant. There is also an environmental impact at the disposal stage. So it would also be good to look for a computer designed for disassembly and with relatively low embodied energy. Eco-labels like ENERGY STAR or Blue Angel can help you identify better performing devices.

13   "Some people use the internet as though power is free and unlimited"

True: Currently the price signal on a single internet search in terms of its energy cost is too small to see, firstly because that incremental energy cost itself is miniscule and secondly because only part of it is met directly by the user. In fact only part of the energy used by a search is even visible to the user who only sees the power used by his computer. What he doesn't see is the power used by the communications network or the data centres that store, process or manage his information. That is because the only power he pays for directly is that used by his computer. Some is paid by the network operator (who passes it back to the user in the form of his telephone bill) and some is covered by the business model of the search engine or service provider. In this "freemium" model the energy costs are covered by advertising. The problem arises from the cumulative total of many millions of searches and online activities that everyone regards as free. And as we all know from Jevons[13], the cheaper a commodity is, the more we use it. So there is an issue here and we need to do more to help people understand the indirect energy impact of their online activities.

---

[13] Jevons Paradox is called a paradox because at first it seems to run counter to common sense. It states that when technological progress leads to an increase in efficiency in resource use the end result is increased consumption of the resource rather than decreased consumption. This sounds

Section IV   The future:
How the industry is (or should be) developing

14   "Cloud computing won't save energy"

False – provided that it is implemented intelligently and appropriately: Cloud computing was originally envisaged as a way of making use of existing distributed computing resource (your home PC for instance) when it wasn't active, and using this as part of a global network to share out computing activities. For many reasons it has not developed in this way and cloud computing will rely on large purpose built data centres, so however virtualised an application is, it will still run on servers. Data centres designed for cloud operations are built for high server utilisation, high virtualisation and low PUE so the efficiency should be much better than most other options. Cloud computing is not a panacea – it is just a different business model with the capability to deliver very significant improvements in energy efficiency as well as other economies of scale.
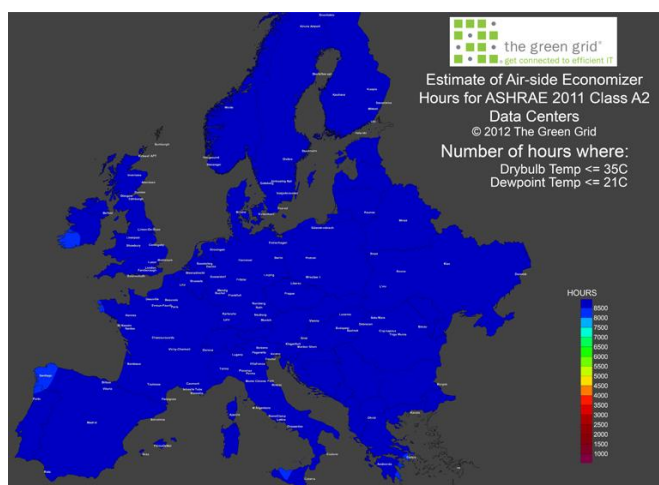
15   "Data centres should all be moved to Iceland"

 False – but... Since data centres use a lot of power and many of the services they provide are not dependent on a particular geographic location, it might seem sensible to look for places where there is a high
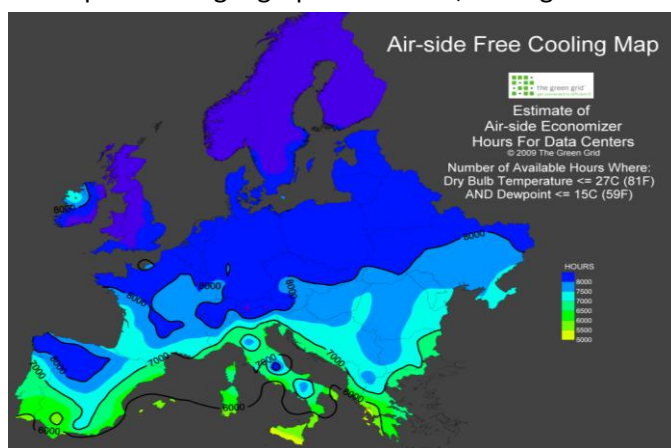
proportion of renewable power and where ambient temperatures are low. Low ambient temperatures are important because this enables a data centre to be cooled by fresh air through ventilation (this is known as free cooling) instead of using energy intensive refrigeration units (chillers) to physically cool the air before circulating it around the data centre. Iceland is certainly an attractive location from the point of view of renewable power since they have a significant and reliable source of geothermal

power. However, Iceland is no different from most of Northern Europe in terms of the scope to use free cooling. Look at the two maps. In both maps the areas that enjoy average temperatures that enable optimum free cooling are shaded dark blue. The 2011 map defines the area according to IT hardware tolerances agreed by ASHRAE (see point 11). Those tolerances have since expanded, hence the 2012 map is nearly all dark blue, which indicates that free cooling can be used almost anywhere in Europe. So in terms of ambient temperature, Iceland has no advantage over the UK.   The UK, however, does have several advantages over Iceland. Firstly, data centres need two things to function: power and connectivity (access to high bandwidth

---

strange but it makes sense because increased efficiency makes a resource cheaper, so people are likely to use more of it, not less. Technological innovation makes ICT cheaper and more efficient and the result is that we use more of. See also point 17.

communications infrastructure), and the UK has far better connectivity than Iceland, effectively acting as the node connecting the US to the rest of Europe. Some communications such as stock market transactions have to be almost instantaneous and do need to be physically close to trading platforms. This proximity is termed "low latency". Other data is required by law to remain within national boundaries. Secondly, the UK is not vulnerable to permafrost in winter which can affect data centre operation. Thirdly some operators rule out Iceland because its geological volatility increases the risk of particulates in the air.

16   "Big Data has no impact on energy use"

False, though it depends on what you mean by Big Data. Big data is one of those terms that mean different things to different people. Industry tends to define big data as increasingly large and complex datasets where the exponential growth in the quantity of data means that we need new tools to make sense of it – traditional database management tools are inadequate. Big data presents enormous opportunities to aid decision making and enable innovation. However, there are looser terms applied more widely. Many perceive big data as the increasing availability and accessibility of digital data – in other words, more and more and more data, all of which is at our fingertips, immediately accessible.

Storing more and more data does not necessarily present an energy challenge because vast amounts of data can be stored passively with a negligible energy impact. The issue arises from our need, real or perceived, for instant access. If we want to access data immediately, then it has to be stored "live", in data centres, on the spinning disks of servers that all use power. So for instance if you feel a compelling need to find out the shoe sizes of people from Wolverhampton, the energy impact of your quest depends on how quickly you want that data to be available.

So big data does have an energy impact and that impact is a function of the amount of data we are storing and the speed at which we want to access it.

17   "We will reduce the total power demand of data centres if we make them more efficient"

Sadly the answer is "not necessarily". There are a number of reasons for this including the drivers for growth in data centres and Jevons' Paradox.

Let's start with growth. As we mentioned in point 3, the UK data centre sector is growing actively. One reason for this growth is that distributed ICT (servers that are spread around offices in server cupboards) is being consolidated into specialised, purpose built facilities (data centres) which are much more efficient. So by doing this, the number of data centres increases BUT there is a net decrease in energy demand. Secondly, many carbon intensive physical activities are being replaced by digital alternatives, a process called dematerialisation. Although this too drives growth in data centres, there is again a net reduction in energy demand.

However, much of the increased growth in data centres is not the result of one inefficient activity being replaced by a more efficient one, but by an overall increase in the demand for digital data. Ironically, due to massive improvements in the energy efficiency of computer processors, the storage, transfer and manipulation of digital data has become cheaper and cheaper. The result is that we use more and more of it. This is Jevons' Paradox- which states that improvements in efficiency can actually drive an increase in overall demand. So by making data centres more efficient we actually drive growth. The important thing is to be able to differentiate the growth in data centres that is actually enabling savings elsewhere and the growth that is additional. But that is another story....

Section V   Getting the sums right:
Using worked examples to examine assertions

18   "You can calculate the total power a data centre uses in a year if you know its power rating" (the power available to it)

False: It's a common mistake to look at the power supply to a data centre (measured in MW) and multiply it by the number of hours in a year (8760) to estimate the annual power demand of that facility (measured in MWh). Yes, it is true that a data centre with 20MW of provisioned power[14] will almost certainly use more power than a data centre with 1MW of provisioned power, but the relationship is not exact.   In reality, the actual power used may bear little relation to the power rating of the facility, particularly if that facility is immature. A graph of many facilities would show a loose association between the two.

There are several reasons for this.

- Firstly, data centres have to provision for the maximum power demand that could theoretically be imposed on the facility from a combination of the IT function within it and the demands of the ancillary equipment that maintain the environment to enable that IT to function reliably. So a data centre must be able to provide enough power for every server in every rack to be working at full load on the hottest day of the year (i.e. when both the IT and the cooling equipment are running flat out).
- Secondly the equipment is always allocated more power than it will use to provide a safety margin. This means that for most of the year a data centre is "over-provisioned".
- Thirdly, data centres may not be full so will not be running at capacity.
- Fourthly, data centre operators also provision power to allow for future expansion rather than what they need at the time. This is arguably the most common reason for over-provisioning – data centres are typically multi-million pound investments built for a 10-20 year life, so businesses would invest, build and provision power on the basis of 20 years of growth. Telecoms equipment itself has a longer life expectancy than IT equipment so planning horizons are longer.   It is also much harder and more expensive to upgrade the power supply retrospectively.

19   "A Google Search takes the same amount of power as boiling a kettle" (The Guardian)

False. A Google search takes less than 0.2% of the power used by boiling a kettle. Let's do the sums: A 2KW kettle takes about 4.5 minutes to boil when full which equates to about 150Wh (you can test with your own kettle, a watch and a portable appliance tester). Average Google query, including building the search index, uses 0.3Wh. 150Wh/0.3Wh = 500, or 0.2% (See

http://googleblog.blogspot.co.uk/2009/01/powering-google-search.html#!/2009/01/powering-google-search.html

---

[14] Provisioned power means the power available to that site. Power ratings for data centres refer to the power available to that site, irrespective of whether it is used or not.

20   "A medium sized server has the same annual carbon footprint as an SUV doing 15 mpg covering 15,000 km". (Global Action Plan/Green ICT)

False. Around an order of magnitude out. A server's carbon footprint is only a fraction of that of an SUV. Neither the maths nor the assumptions bear scrutiny.   Let's look at each in turn.

Firstly let's do the SUV carbon footprint.   The mixed units suggest a rather sluttish approach to the problem but we can deal with them. Converting gallons to litres, miles to km you get the SUV doing 5.3km per litre of fuel and therefore using 2830 litres of fuel in a year. Using the Carbon Trust's conversion factors you end up with the SUV emitting 6.542 tonnes CO_2 if it is petrol and 7.550 tonnes CO_2 if it is diesel (fuel was unspecified). Obviously we haven't factored in the embodied carbon.

> **Nerd Notes: Conversion Factors**
> We have used the following conversion factors which you can download from the Carbon Trust. These tell you how much CO2 is emitted when you consume a unit of fuel and are very important in enabling us to compare emissions irrespective of the fuel type used.
> Conversion factor for diesel is 2.668 Kg CO_2 per litre
> Conversion factor for petrol is 2.312 kg CO2 per litre.
> Conversion factor for electricity is 0.526kgCo2 per KWh..

Now let's look at the server. A single unit (1U) server[15] uses 80-120W constantly and a double unit (2U) server uses 160-240W continuously (Source: Memset, actual measurements). So the smaller server has emissions around 450kg CO_2 per server and the large one around 900kg.

These calculations can only provide ballpark figures. Comparing carbon impacts in a meaningful way is far, far more complex. If you want to understand just how difficult it is to evaluate the carbon impact of a product or service over its life then have a look at our paper: *"Evaluating the carbon impact of ICT: The Answer to Life, the Universe and Everything".*

> **Nerd Notes: Assumptions:**
>
> - That an SUV does 15,000 km in a year when in fact the average for an SUV is 13,700 *miles* per year.   So after fiddling with the units you should really multiply the SUV footprint by 1.47 which takes it up to 9615kg CO_2 or 11095kg CO_2 depending on whether it is petrol or diesel.
>
> - That the energy used is equivalent - but it isn't. Servers use electricity so their carbon footprint depends on the way that electricity was generated. For the purpose of this exercise we have assumed UK grid average but in reality a significant proportion of data centre operators buy green power.
>
> - Don't forget that Moore's Law applies to servers but not to SUVs. So in two years time a medium sized server will be twice as efficient as today's model but the SUV won't. And don't forget that the server is working 24 hours a day every day of the year. The SUV is only active for a small proportion of the time.   And if 50% of the energy reaching the data centre is being used by the server, at least 90% of the energy reaching the SUV is used to move the SUV around; less than 10% is used to move the people or goods inside it.
>
> - Consider the economic contribution that the server is making compared to that of the SUV. An SUV can only support one type of activity – moving people from one place to another, whereas the server could be enabling delivery logistics for a fleet management company or facilitating teleworking or e-Commerce. The server can have an enabling effect, the SUV cannot.
>
> - There is also the embodied energy to take into account, for both the server and the SUV. And the carbon footprint of the infrastructures they depend on. In both cases the embodied energy in the SUV dwarfs that of the server.

---

[15] A 1 U server or single unit server looks like a pizza box. A 2U server is simply twice the depth with double the processing power.

References

The Green Grid: The Green Grid is a global consortium of IT companies and professionals aiming to improve energy efficiency in data centres and in business computing systems around the world. The objective is to develop a standardised set of metrics, processes, methods and technologies which will drive improved efficiency in data centre operations. www.thegreengrid.org

The Green Grid Data Centre Maturity Model: The Data Center Maturity Model (DCMM) enables users to benchmark their current performance, determine their levels of maturity, and identify the ongoing steps and innovations necessary to achieve greater energy efficiency and sustainability, both today and into the future. The maturity model touches upon every aspect of the data centre including power, cooling, compute, storage and network. The levels of the model outline current best practices and a 5-year roadmap for the industry.

The EU Code of Conduct for data centres: The Code was created to help address the growing energy consumption of data centres and by so doing address related issues of cost, security of energy supply and carbon emissions.   Its objective is to encourage data centre operators to reduce energy consumption cost effectively without compromising the mission critical function of data centres. The Code aims to improve the level of understanding of energy use in data centres and provides recommendations for operators in the form of best practice and reduction targets.

http://re.jrc.ec.europa.eu/energyefficiency/pdf/CoC%20data%20centres%20nov2008/CoC%20DC%20v%201.0%20FINAL.pdf

Other initiatives include work by the Uptime Institute (http://uptimeinstitute.com/) and the development of ISO standards to measure efficiency in data centres, which is ongoing.

For more information please contact:

Emma Fryer, Associate Director, Climate Change Programmes: emma.fryer@intellectuk.org

James Harbidge, Programme Manager, Energy and Environment: james.harbidge@intellectuk.org

Contributors and Reviewers

Text by: Emma Fryer

Reviewers	Data / images kindly provided by
Darren Mansfield, CBRE	Ian Bitterlin, Emerson Network Power
Sylvie Feindt, DigitalEurope	Billy McHallum, Equinix
Steve Strutt, IBM	Henry Wong, Intel
Alice Harrison, Intellect	Kate Craig-Wood, Memset
Stephen Hearnden, Intellect	Chris Williams, Next Generation Data
Mark Bailey, Speechly Bircham	Melissa Gray, Rackspace
Harkeeret Singh, Thomson Reuters	Felicia Asiedu, TelecityGroup

Lay Readers: Robert Burdall, Carole Fryer, Paul Vincent

Further Information

Other publications in this series or discussed above include:

- Er, What IS a Data Centre?

- Data Centres: A Day in YOUR Life

- So What Have Data Centres Ever Done For Us?

Generic publications relevant to this topic include:

- **Data Centres: The Backbone of the UK Economy**

- We Need to Talk About Jevons (final draft, publication imminent)

- **Evaluating the carbon impact of ICT: The Answer to Life, the Universe and Everything**

- **High Tech: Low Carbon**