nature climate change

PERSPECTIVE

PUBLISHED ONLINE: 25 JUNE 2013 | DOI: 10.1038/NCLIMATE1786

Characteristics of low-carbon data centres

Eric Masanet[1]*, Arman Shehabi[2] and Jonathan Koomey[3]

Data centre services hold promise for reducing societal carbon emissions, but an imperfect and evolving portfolio of performance metrics obscures which data centre characteristics correspond to low-carbon operations. Meanwhile, policymakers face a pressing question: can we identify and promote tangible characteristics that reliably represent low-carbon data centres today while the world awaits better metrics? Fortunately, data centre energy models can provide actionable guidance. Here, we present results that identify such characteristics and illuminate the factors that govern a data centre's actual carbon performance. These results can help public and private sector policymakers accelerate the transition to a low-carbon Internet by aligning data centre incentives with factors that truly matter.

The world is becoming increasingly reliant on the Internet for commerce, social interaction, communications, entertainment and the generation and dissemination of knowledge. As global development continues, our reliance promises to grow even further[1]. Although the Internet consists of many interlinked parts — including modems, routers and communication lines — data centres are its most energy- and carbon-intensive component[2]. A combination of recent explosive growth in data centre demand and numerous, well-documented disincentives for energy efficiency has resulted in a global population of data centres that together use 1–2% of the world's electricity and are often layered with inefficiencies[3–8].

However, an increasing number of studies suggest that information services delivered over the Internet might lead to reductions in societal carbon emissions. Such reductions can occur through decreased demand for physical goods, more efficient resource use and replacement of (comparatively inefficient) local computations[9–14]. Furthermore, the information technology (IT) industry is continuously improving the efficiency of IT devices and data centre operations[15–21], while in some cases shifting towards greater use of more efficient and/or renewable electricity[22–26] — transitions that should gradually increase the potency of Internet-based services for climate change mitigation.

Public and private climate policies that provide financial incentives can help accelerate these transitions (for example, preferential purchasing arrangements for low-carbon data centre services)[5]. However, the success of such policies depends critically on having a consistent set of credible, quantitative metrics for rewarding data centres that minimize carbon emissions[27–28]. Unfortunately, decision makers now face a plethora of evolving and interrelated data centre metrics[29–32] — and various environmental claims based on them[20,33] — that creates confusion over what constitutes a truly low-carbon data centre.

Mathematically, the ideal metrics would be expressed as intensity quotients of a data centre's carbon emissions per unit of information service output[28,31]. In spite of much research effort, such metrics have proved elusive owing to the conceptually and analytically challenging nature of quantifying data centre information services. Moreover, research is beginning to address the indirect (that is, life-cycle) environmental emissions and impacts of data centres, which expands the scope of what policymakers must consider[2,34–36].

Amid this landscape of evolving and imperfect information, policymakers face a pressing question: can we identify and promote tangible characteristics that reliably represent low-carbon data centres today while the world awaits credible metrics? Fortunately, data centre energy models can provide actionable guidance. Here, we present results that identify such characteristics and illuminate the factors that govern a data centre's actual carbon performance. We employ US data, but our methods and findings can be generalized to data centres globally.

The data centre carbon footprint

A useful starting point is the data centre carbon footprint, which we define on an annual basis as the sum of carbon emissions from three components: (1) data centre operations; (2) manufacture of new or replacement IT devices (that is, servers, external storage and network switches) and (3) the disposal and/or recycling of IT devices. The last two components are hereafter referred to as 'embodied' emissions. The high carbon intensity of IT-device manufacturing has been thoroughly documented[37,38], as has the importance of recycling for environmental protection[39]. Footprint analysis therefore provides a complete picture of a data centre's direct and indirect emissions, while also indicating each component's importance as both a source of emissions and a potential opportunity for emissions reductions.

Figure 1 summarizes our results for a single prototypical US data centre along with three additional cases that reflect improved energy and carbon management. Our prototype is a representative US enterprise-class facility, with 20,000 volume servers, 40,000 external hard disk drives, 2,060 network switches and a US average power utilization effectiveness (PUE) of 1.8 (refs 6,36). In all cases, embodied emissions were estimated using the stock turnover model in ref. 40 with average values derived from IT-device life-cycle inventory data. Despite acknowledged imperfections in the use of PUE as a singular metric for data centre energy efficiency, the IT community has converged on PUE as a near-term proxy metric for data centre energy efficiency[28,41]. The PUE is calculated by dividing a data centre's total electricity use (kWh yr^{-1}) by the electricity used by the data centre's IT devices (kWh yr^{-1}). PUE quantifies how efficiently a data centre's infrastructure provides power to its IT devices and forms the basis of the current US ENERGY STAR rating scheme for data centres[42]. Operational emissions were estimated using the data centre energy model in ref. 6, assuming the

[1]McCormick School of Engineering, Northwestern University, 2145 Sheridan Road, Room L494, Evanston, Illinois 60208, USA, [2]Environmental Energy Technologies Division, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Building 90R2000, Berkeley, California 94720, USA, [3]Steyer-Taylor Center for Energy Policy and Finance, Stanford University, PO Box 1545, Burlingame, California 94010, USA.
*e-mail: eric.masanet@northwestern.edu

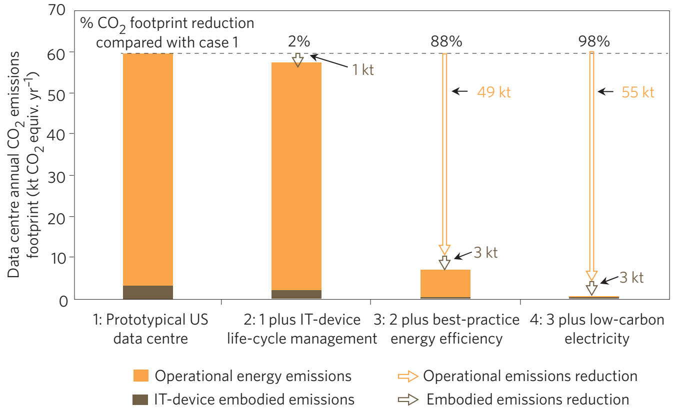

Figure 1 | Prototypical US data centre carbon footprint and plausible reductions. The prototypical data centre footprint is 59 kt CO_2 equiv. yr^{-1}, using average US carbon intensity for electricity (0.6 kt CO_2 equiv. GWh^{-1}). Three additional, additive cases illustrate plausible reductions in operational and embodied emissions (expressed by coloured arrows). Life-cycle management adds IT-device lifetime extension and 100% recycling. Best-practice energy efficiency adds state-of-the-art IT-device and facilities energy management, with a PUE of 1.1. Low-carbon electricity adds average renewable power. See section S1 of the Supplementary Information for details.

US average electricity mix. See section S1 of the Supplementary Information for further details.

The results suggest that, although not negligible, embodied emissions represent only a small contribution to the total footprint, in terms of both absolute and relative emissions. Adding IT-device life-cycle management led to a small footprint reduction, whereas adding improved operational energy efficiency and low-carbon electricity produced comparatively large savings. Interestingly, improved energy efficiency resulted in greater reductions in embodied emissions than did IT-device life-cycle management itself. The explanation for this effect is that optimal IT-device efficiency necessitates reducing the number of IT devices to their practical minimum while maximizing the utilization of the remaining devices. Such reductions, which can be achieved through, for example, server virtualization, application consolidation and capacity maximization, lead in turn to a significant decrease in embodied emissions through avoided purchases and disposal of IT devices.

Note that the operational emissions results in Fig. 1 are based on the US average CO_2 intensity for electricity. As such, the relative and absolute emissions associated with operations in Fig. 1 would vary by electricity mix. However, embodied emissions do not vary with a data centre's electricity mix. Therefore, the primary conclusions of Fig. 1 — that embodied emissions are likely to be small on an absolute basis and the fastest path to reducing such emissions is reducing device counts through operational efficiency — hold for any data centre.

Thus, it is clear that operational energy and carbon performance represent the most fruitful paths to low-carbon data centres, even from an embodied emissions perspective. It follows that metrics based on IT-device life-cycle management — although recommended from environmental and social responsibility perspectives — are not accurate proxies for low-carbon data centres. Case in point, Fig. 1 shows that good life-cycle management does not always imply a low carbon footprint, nor, conversely, does poor life-cycle management always imply a high carbon footprint.

Operational energy and carbon performance

Given the importance of operational energy and carbon performance, two logical policy-relevant questions follow: (1) which tangible characteristics consistently correspond to low-carbon operations? and (2) which metrics quantify such characteristics? To help answer these questions, in Fig. 2 we introduce a conceptually simple — yet illustratively powerful — two-dimensional graphic, which we refer to as the data centre energy–carbon performance map. The map illustrates how operational energy use (*y* axis) and carbon emissions (*x* axis) can vary based on a data centre's technology, location and electricity source characteristics. For simplicity, we index performance to a 'worst case' baseline point at the upper right vertex of the map; points leftwards and downwards from the baseline indicate reduced carbon intensity and reduced energy use, respectively. The relative positions of points indicate how the energy use and carbon emissions of different data centre configurations compare, while the map's area bounds the possible range of performance. The map can be tailored to a specific data centre or region based on the performance bounds and/or incremental electrical supply improvements specific to that plant or region.

Figure 2 depicts the feasible performance range for our prototypical US data centre. The baseline point (6D) assumes typical energy efficiency (see Fig. 1) coupled with carbon-intensive coal-fired electricity. Using the model in ref. 6, we populated the map with five additional energy-use levels, which are based on plausible variations in US data centre location, IT-device efficiency and PUE. See section S2 of the Supplementary Information for details. For ease of interpretation, our configurations include either minimal or maximal IT-device efficiency (but not the many possible levels in between). The electricity sources in Fig. 2 capture a 50-fold span in carbon emissions from feasible US electricity options and include natural gas solid oxide fuel cells (SOFCs), which have been employed in several west coast data centres, as well as centralized renewable electricity sources being used by one or more data centres as reported in the literature[22–26]. Point 1A marks the technical minimums for energy use and carbon emissions, which represent a 10-fold reduction in energy use and a 500-fold reduction in CO_2 emissions compared with the baseline.

Figure 2 illustrates the critical role of maximal IT-device efficiency in achieving both low-energy and low-carbon data centres. Although the importance of IT efficiency is well documented[5], our results further show that this axiom holds regardless of PUE, location, or electrical power source. Indeed, Fig. 2 shows that improved PUE and reduced electrical carbon intensity account for only small improvements when applied to the maximal IT-efficiency case. These results indicate that characteristics implying IT-device efficiency — including the presence of efficient-rated devices, high-capacity utilization, IT power management and virtualization — are probably the best proxy metrics for low-carbon data centres in the near term[5,6,32].

The map also highlights an important distinction: although the use of renewable electricity can lead to significantly better carbon performance, it does nothing to improve energy performance. Indeed, low-carbon electricity without low-energy operations is suboptimal because, given their high energy intensity, data centres now using renewable power almost exclusively draw from centralized and resource-constrained supplies (notably hydropower, wind, geothermal and biogas). In other words, high-energy data centres unnecessarily tie up 'low-carbon' electrons that might otherwise be used to reduce emissions from other customers in a region. For example, a low-energy data centre operating at point 1A would consume one-tenth as much 'low-carbon' electricity as a high-energy data centre operating at point 6A.

Figure 2 also suggests that PUE — despite its status as the *de facto* data centre energy-efficiency metric — is by itself a suboptimal proxy for both absolute energy and carbon performance. Although a low PUE can significantly reduce energy use when coupled with minimal IT-device efficiency, this effect is drastically diminished when IT-device efficiency is maximized. Thus, data centres with the same PUE value (for example, points 3C and 1C) can have

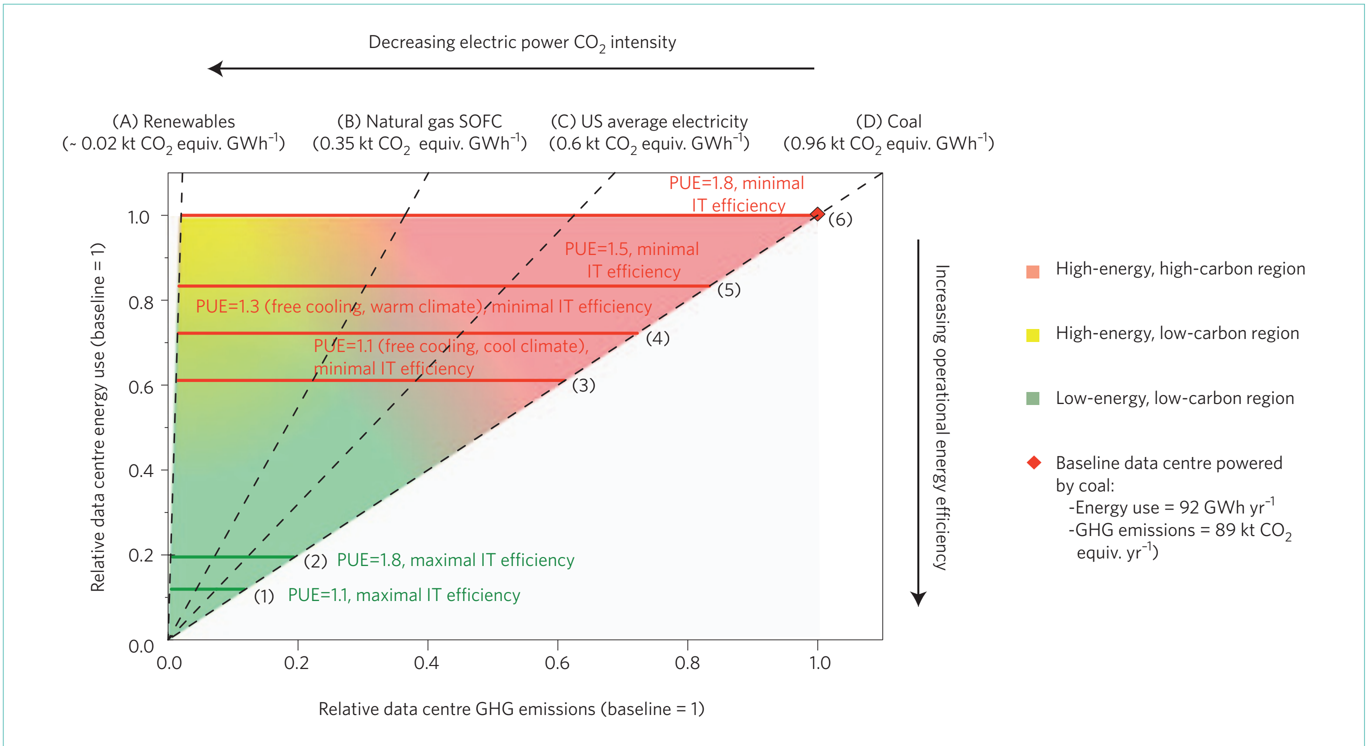

Figure 2 | Energy–carbon performance map. The shaded area bounds the potential operational energy and carbon performance range of a prototype US data centre and illustrates the relative performance of different data centre characteristics. Coloured areas indicate general regions of energy–carbon performance. For some data centres, only subareas of this map will apply depending on equipment and electrical power constraints. Numbered points are discussed in the text. See section S2 of the Supplementary Information for details. GHG, greenhouse gas.

drastically different energy and carbon performance. Moreover, a data centre with a 'poor' PUE of 1.8 (for example, point 2D) can exhibit much lower energy use and carbon emissions than one with a best-in-class PUE of 1.1 (for example, point 3D).

Data centre location is another characteristic that is often seen as critical[22,33], as it can govern the carbon intensity of purchased electricity as well as the availability and quality of 'free cooling'. Free cooling has been shown to reduce PUE by cooling IT equipment with outside air that is directly routed into a data centre during favourable weather conditions[43]. Here again, Fig. 2 shows that these location-specific characteristics are most important for data centres with minimal IT-device efficiency. For data centres with maximal IT efficiency, optimal location delivers much smaller (albeit non-trivial) absolute energy and carbon performance benefits. Note also that although our cases do not explicitly include the possibility of using waste heat from a SOFC to provide facility cooling through an absorption or adsorption chiller — a strategy that would lower a data centre's PUE — one could easily assess that possibility by assuming the appropriate PUE improvement that might coincide with a shift to SOFCs.

Towards low-carbon data centres

Our results and previous analyses suggest that IT-device efficiency is the most important characteristic to be reinforced through low-carbon data centre incentives[5,6,44–45]. Although renewable energy can significantly reduce a data centre's carbon emissions, an inefficient (that is, high-energy) data centre will use far more low-carbon electricity than is technically required. Data centres that operate in the low-energy, low-carbon region of Fig. 2 — which requires maximal IT-device efficiency — will minimize demand on constrained centralized renewable sources of energy. Maximizing the reach of renewable energy through demand reduction is a critical policy goal. Numerous measures can serve as potential proxies for maximal IT-device efficiency in the near term — such as high-capacity utilization and the presence of high-efficiency IT devices — but efforts to develop quantitative metrics for consistent assessment of IT-device efficiency across data centres should be redoubled[27,28,46]. While we await more precise intensity quotients, other extant metrics such as PUE and electricity source can play a role in assessing carbon performance when used concurrently.

Furthermore, data centre energy–carbon performance maps help expose various degrees of freedom under local design and retrofit constraints, which can allow for flexible policy design. Here we offer the following recommendations to policymakers who seek to design effective incentives for low-carbon data centres: all existing data centres should maximize IT-device efficiency, especially as these devices can turn over quickly and thereby deliver rapid improvements. Decisions regarding when to upgrade remaining devices to more efficient models can be informed in part by a break-even analysis of the embodied emissions required to manufacture new devices versus the operational energy savings that would be realized[47]. New data centres should locate in areas with ample free cooling and/or low-carbon electricity grids to further push operations towards better energy and carbon performance. In new or existing facilities where optimal IT-device efficiency is not feasible, significant reductions in PUE critically rise in importance as a policy aim (but still result in higher energy-use levels than efficient IT devices would deliver). Where such PUE reductions are constrained by location (for example, a lack of free cooling), procuring low-carbon electricity — either from local electricity providers or through the installation of reduced-carbon self-generation such as SOFCs — becomes the next chief lever after energy efficiency has reached its practical limit. With these insights in mind, public- and private-sector policymakers can accelerate the transition to a low-carbon Internet by aligning their incentives with data centre characteristics that matter.

Received 26 July 2012; accepted 21 November 2012; published online 25 June 2013

References

1. Hilbert, M. & López, P. The world's technological capacity to store, communicate, and compute information. *Science* **332,** 60–65 (2011).
2. The Climate Group *SMART 2020: Enabling the Low Carbon Economy in the Information Age* (Creative Commons, 2008); available at http://www.smart2020.org/_assets/files/02_Smart2020Report.pdf
3. Koomey, J. Worldwide electricity used in data centers. *Environ. Res. Lett.* **3**, 034008 (2008).
4. Koomey, J. G. *Growth in Data Center Electricity Use 2005 to 2010* (Analytics Press, 2011); available at http://www.analyticspress.com/datacenters.html
5. Brown, R. *et al. Report to Congress on Server and Data Center Energy Efficiency: Public Law 109–431.* Report No. LBNL-363E (Lawrence Berkeley National Laboratory, 2007).
6. Masanet, E., Brown, R. E., Shehabi, A., Koomey, J. G. & Nordman, B. Estimating the energy use and efficiency potential of U.S. data centers. *Proc. IEEE* **99,** 1440–1453 (2011).
7. Koomey, J. G., Brill, K. G. & Turner, W. P. *A Simple Model for Determining True Total Cost of Ownership for Data Centers* (Uptime Institute, 2007).
8. Glanz, J. Power, pollution, and the Internet. *The New York Times* A1 (23 September 2012).
9. Erdmann, L. & Hilty, L. Scenario analysis: Exploring the macroeconomic impacts of information and communication technologies on greenhouse gas emissions. *J. Ind. Ecol.* **14,** 826–843 (2010).
10. Verdantix *Cloud Computing –The IT Solution for the 21st Century* (Carbon Disclosure Project, 2011).
11. Weber, C. L., Koomey, J. G. & Matthews, H. S. The energy and climate change implications of different music delivery methods. *J. Ind. Ecol.* **14,** 754–769 (2010).
12. Salesforce.com in collaboration with WSP Environment & Energy *Salesforce.com and the Environment: Reducing Carbon Emissions in the Cloud* (WSP Environment & Energy, 2011); available at http://www.salesforce.com/assets/pdf/misc/WP_WSP_Salesforce_Environment.pdf
13. Accenture in collaboration with WSP Environment & Energy *Cloud Computing and Sustainability: The Environmental Benefits of Moving to the Cloud* (Accenture and WSP Environment & Energy, 2010); available at http://www.accenture.com/us-en/Pages/insight-environmental-benefits-moving-cloud.aspx
14. Google *Google's Green Computing: Efficiency at Scale* (Google, 2011); available at www.google.com/green/pdfs/google-green-computing.pdf
15. Koomey, J. G., Berard, S., Sanchez, M. & Wong, H. Implications of historical trends in the electrical efficiency of computing. *Ann. Hist. Comput.* **33,** 46–54 (2011).
16. US Department of Energy *Best Practices Guide for Energy-Efficient Data Center Design* (Federal Energy Management Program, 2011); available at http://www1.eere.energy.gov/femp/pdfs/eedatacenterbestpractices.pdf
17. Samson, T. The green IT stars of 2010. *InfoWorld* (22 April 2010); available at http://www.infoworld.com/d/green-it/the-green-it-stars-2010-454
18. Lammers, H. At NREL, even the ones and zeros are green. *NREL Newsroom* (6 July 2011); available at http://www.nrel.gov/news/features/feature_detail.cfm/feature_id=1505
19. Smalley, G. 2011: The year data centers turned green. *Wired* (30 December 2011); available at http://www.wired.com/wiredenterprise/2011/12/green-data-centers-of-2011
20. Gelber, R. Facebook showcases green datacenter. *HPCwire* (26 April 2012); available at http://www.hpcwire.com/hpcwire/2012-04-26/facebook_showcases_green_datacenter.html
21. Google *Our Energy-Saving Data Centers* (Google, accessed 11 July 2012); available at http://www.google.com/about/datacenters/index.html
22. Laursen, L. Iceland exports energy as data. *MIT Technology Review* (11 April 2012); available at http://www.technologyreview.com/news/427468/iceland-exports-energy-as-data
23. Kessler, R. A. NextEra wind farm supplying power for Google data center. *Recharge News* (29 November 2011); available at http://www.rechargenews.com/energy/wind/article291445.ece
24. SustainableBusiness.com Two large California concentrating solar projects get $2B loan guarantee. *SustainableBusiness.com* (16 June 2011); available at http://www.sustainablebusiness.com/index.cfm/go/news.display/id/22553
25. Sayer, P. New eBay data center will run on fuel cells. *PC World* (21 June 2012); available at http://www.pcworld.com/businesscenter/article/258098/new_ebay_data_center_will_run_on_fuel_cells.html
26. LaMonica, M. Apple data center helps fuel Bloom Energy move to east coast. *CNET* (30 April 2012); available at http://news.cnet.com/8301-11386_3-57424037-76/apple-data-center-helps-fuel-bloom-energy-move-to-east-coast
27. BCS *Future of Data Centre Efficiency Metrics: A Review of Progress so Far and a Suggested Direction for Further Development* (BCS, 2010).
28. The Green Grid *Harmonizing Global Metrics for Data Center Energy Efficiency* (The Green Grid, 2011).; available at http://www.thegreengrid.org/en/Global/Content/Reports/HarmonizingGlobalMetricsForDataCenterEnergyEfficiency
29. Wang, L. & Khan, S. U. Review of performance metrics for green data centers: a taxonomy study. *J. Supercomput*. **63,** 639–656 (2011).
30. Belady, C., Azevedo, D., Patterson, M., Pouchet, J. & Tipley, R. *Carbon Usage Effectiveness (CUE): A Green Grid Data Center Sustainability Metric.* White Paper No. 32 (Green Grid, 2010).
31. Stanley, J. R., Brill, K. G. & Koomey, J. G. *Four Metrics Define Data Center "Greenness"* (Uptime Institute, 2007).
32. Haas, J. *et al. Proxy Proposals for Measuring Data Center Productivity*. White Paper No. 17 (Green Grid, 2009).
33. Cook, G. & Van Horn, J. *How Dirty is your Data? A Look at the Energy Choices that Power Cloud Computing* (Greenpeace International, 2011); available at http://www.greenpeace.org/international/Global/international/publications/climate/2011/Cool%20IT/dirty-data-report-greenpeace.pdf
34. Emerson Network Power *Recycling Ratios: The Next Step for Data Center Sustainability* (Emerson Electric, 2011); available at http://www.emersonnetworkpower.com/en-US/Documents/EDC/Market%20Position%20Paper%20-%20Recycling%20Ratios_6.pdf
35. Bash, C. *et al.Cloud Sustainability Dashboard, Dynamically Assessing Sustainability of Data Centers and Clouds* Report No. HPL-2011-148 (HP Labs, 2011); available at http://www.hpl.hp.com/techreports/2011/HPL-2011-148.pdf
36. Mahadeven, P., Shah, A. & Bash, C. Reducing lifecycle energy use of network switches. *Proc. 2010 IEEE Int. Symp. Sust. Syst. Tech.* (Washington DC, May 2010).
37. Fraunhofer, I. Z. M. & Europe, P. E. *Ecodesign of EuP Products: Preparatory Studies LOT 5: Consumer Electronics: TV* (European Commission, 2007); available at http://ec.europa.eu/energy/efficiency/studies/ecodesign_en.htm
38. Williams E., Ayres, R. U. & Heller, M. The 1.7 kilogram microchip: Energy and material use in the production of semiconductor devices. *Environ. Sci. Technol.* **36,** 504–5510 (2002).
39. Williams, E. *et al.* Environmental, social, and economic implications of global reuse and recycling of personal computers. *Environ. Sci. Technol.* **42,** 6446–6454 (2009).
40. Masanet, E. & Horvath, A. Enterprise strategies for reducing the environmental impacts of personal computers. *Proc. 2006 IEEE Int. Symp. Electron. Environ.* (San Francisco, California, IEEE 2006).
41. Belady, C., Rawson, A., Pfleuger, J. & Cader, T. *Green Grid Data Center Power Efficiency Metrics: PUE and DCIE*. White Paper No. 6 (Green Grid, 2008).
42. United States Environmental Protection Agency *ENERGY STAR Data Center Energy Efficiency Initiatives* (United States Environmental Protection Agency, 2012); available at http://www.energystar.gov/index.cfm?c=prod_development.server_efficiency
43. Shehabi, A. *et al.* Data center design and location: Consequences for electricity use and greenhouse-gas emissions. *Build. Environ.* **46,** 990–998 (2011).
44. Eubank, H., Swisher, J., Burns, C., Seal, J. & Emerson, B. *Design Recommendations for High-Performance Data Centers: Report of the Integrated Design Charrette (conducted 2–4 February 2003).* (Rocky Mountain Institute, 2004).
45. Lovins, A. B. *Energy End-Use Efficiency* (Rocky Mountain Institute, 2005); available at http://www.rmi.org/images/other/Energy/E05-16_EnergyEndUseEff.pdf
46. Stanley, J. & Koomey, J. *The Science of Measurement: Improving Data Center Performance with Continuous Monitoring and Measurement of Site Infrastructure* (Analytics Press, 2009); available at http://www.analyticspress.com/scienceofmeasurement.html
47. Prakash, S., Liu, R., Schischke, K. & Stobbe, L. *Timely Replacement of a Notebook under Consideration of Environmental Aspects* (German Federal Environment Agency, 2012); available at http://www.oeko.de/oekodoc/1584/2012-440-en.pdf

Acknowledgements

Lawrence Berkeley National Laboratory is supported by the US Department of Energy under contract number DE-AC02-05CH11231.

Additional information

Supplementary Information is available in the online version of this paper.

Competing financial interests

The authors declare no competing financial interests.