2013-05-01

Optimizing Data Centre Energy and Environmental Costs

Aikema, David

Aikema, D. (2013). Optimizing Data Centre Energy and Environmental Costs (Doctoral thesis, University of Calgary, Calgary, Canada). Retrieved from https://prism.ucalgary.ca. doi:10.11575/PRISM/27215
http://hdl.handle.net/11023/664

UNIVERSITY OF CALGARY

Optimizing Data Centre Energy and Environmental Costs

by

David Hendrik Aikema

A DISSERTATION
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE
DEGREE OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF COMPUTER SCIENCE

CALGARY, ALBERTA
April, 2013

Abstract

Data centres use an estimated 2% of US electrical power which accounts for much of their total cost of ownership. This consumption continues to grow, further straining power grids attempting to integrate more renewable energy. This dissertation focuses on assessing and reducing data centre environmental and financial costs.

Emissions of projects undertaken to lower the data centre environmental footprints can be assessed and the emission reduction projects compared using an ISO-14064-2-compliant greenhouse gas reduction protocol outlined herein. I was closely involved with the development of the protocol. Full lifecycle analysis and verifying that projects exceed business-as-usual expectations are addressed, and a test project is described.

Consuming power when it is low cost or when renewable energy is available can be used to reduce the financial and environmental costs of computing. Adaptation based on the power price showed 10–50% potential savings in typical cases, and local renewable energy use could be increased by 10–80%. Allowing a fraction of high-priority tasks to proceed unimpeded still allows significant savings.

Power grid operators use mechanisms called ancillary services to address variation and system failures, paying organizations to alter power consumption on request. By bidding to offer these services, data centres may be able to lower their energy costs while reducing their environmental impact. If providing contingency reserves which require only infrequent action, savings of up to 12% were seen in simulations. Greater power cost savings are possible for those ceding more control to the power grid operator.

Coordinating multiple data centres adds overhead, and altering at which data centre requests are processed based on changes in the financial or environmental costs of power is likely to increase this overhead. Tests of virtual machine migrations showed that in some cases there was no visible increase in power use while in others power use rose by 20–30W. Estimates of how migration was likely to impact other services used in current cloud environments were derived.

Acknowledgements

It’s been a great few years and I have so many people to thank for their support throughout them.

Thanks to my family for the support and for putting up with my occasional eccentricities.

To my friends both in and outside Calgary: I owe you my gratitude. Thanks for the tea (particularly on Tuesdays), and for causing to me to question whether or not the laws of physics also apply in squash courts. Thanks for all the conversations over the years.

Thanks are also due to everyone who I had the opportunity to work with at the Grid Research Centre during my time at the U of C, and particularly to Rob for the support over the years.

Table of Contents

List of Tables

List of Figures

List of Terms and Abbreviations

AB Alberta, Canada.

Additionality The need to exceed business-as-usual standards.

AESO Alberta Electrical System Operator; operator of the Alberta power grid.

Affected (SSR) An SSR which is indirectly related to a project.

Ancillary Services Used to maintain the short term stability of the electrical grid and provide resilience against power plant failures.

ASHRAE The American Society of Heating, Refrigerating and Air-Conditioning Engineers.

Balancing Area A region over which power grid stability is maintained.

Baseload The minimum amount of power used in an electrical grid.

Baseline Scenario A scenario representing a business-as-usual course of action against which to evaluate emission reduction projects.

BC British Columbia, Canada.

Blade A high density server housed in a blade enclosure.

Blade Enclosure Houses blade servers and contains shared infrastructure components enabling blades to operate more efficiently.

BPA Bonneville Power Administration, an entity operating hydro dams and managing the power grid in parts of the Northwestern United States.

CANARIE Operator of a Canada-wide network for academic and research organizations and funder of the GreenStar Network project [112].

Capacity Factor The average fraction of generator capacity available when needed.

Carbon Intensity A ratio of the CO_2 emissions produced per unit of energy.

Carbon Usage Effectiveness A rating of the quantity of CO_2 emissions produced per unit of energy consumed in a data centre.

CDN Content distribution network; A storage cache distributed across multiple data centres.

Cloud Computing Services which offer resources to users on demand, often in the form of virtual machines.

CO_2 Carbon Dioxide.

CO_2e The volume of CO_2 emissions having the same climate change impact as the sum of all greenhouse gases produced by an activity.

Contingency Reserve An ancillary service used to respond to power plant failure or other conditions causing a sudden mismatch between the supply of and demand for energy.

Controlled (SSR) An SSR whose emissions are under direct control of the project.

CRAC Computer Room Air Conditioner.

CUE See Carbon Usage Effectiveness.

Cycling (power plants) Rapid adjustment of output up and down.

DA Day-ahead.

Data Center Infrastructure Efficiency The ratio of the power supplied to computing equipment to the total power used by a data centre.

DCIE See Data Center Infrastructure Efficiency.

Demand-Response The lowering of demand by power consumers when requested to improve the stability of the power grid.

Downstream (SSR) An SSR which occurs outside project boundaries following project activity.

DVFS Dynamic voltage and frequency scaling.

Elasticity The rate at which a cloud computing service is able to provision new virtual machines.

EDE Electronics Disposal Effectiveness.

Energy Proportionality The degree to which a computer's energy consumption varies based upon system load.

Energy Reuse Effectiveness A revision to PUE which subtracts reused energy in the calculation.

Energy Reuse Factor The fraction of input energy which is reused.

Emission Intensity The level of emissions per unit of activity.

EPA The US Environmental Protection Agency.

ERE See Energy Reuse Effectiveness.

ERF See Energy Reuse Factor.

Ex-Post Pricing A system in which prices are not finally fixed until after consumption.

Feed-In Tariff The payment of guaranteed premium rates for renewable energy in order to boost renewable energy production.

FERC The US Federal Energy Regulatory Commission.

Gamut An artificial workload generator for computer systems [176].

GEC Green Energy Coefficient.

GeoChronos A portal for the earth observation science community incorporating access to resources with traditional social network functionality.

GHG GreenHouse Gas.

GigaCenter A data centre operated by RackForce in Kelowna, BC.

GPU Graphics Processing Unit.

Green500 A reordering of the TOP500 list by energy efficiency.

GreenStar Network A testbed which was created to evaluate follow-the-renewables computing architectures. It was composed of servers, distributed across Canada as well as other countries and powered by renewable energy sources and connected by high-speed networks. The development of the GreenHouse Gas Reduction protocol for the ICT sector described in this dissertation was also funded by CANARIE as part of the GreenStar Network project.

GW Gigawatt.

HA Hour-ahead.

HPC High performance computing.

HVAC Heating, ventilation, and air conditioning.

Hypervisor Used to create and manage virtual machines.

I/O Input/Output.

IOzone A benchmark suite for evaluating server I/O performance.

ICT Information and Communications Technology.

ISO-14064-2 An ISO standard outlining how quantification, monitoring, and reporting of greenhouse gas emission reduction projects should be done.

LAN Local Area Network.

KVM Kernel Virtual Machine. A commonly-used hypervisor.

kWh Kilowatt-hour.

MW Megawatt.

MWh Megawatt-hour.

NERC The North American Electrical Reliability Corporation; an entity with legal authority in the US and parts of Canada over power transmission reliability standards.

Non-Spinning Reserve See Non-Synchronous Reserve.

Non-Synchronous Reserve An ancillary service used to respond to power plant or other system failures but which permits a longer reaction time than spinning reserves.

NREL The US National Renewable Energy Laboratory.

NYISO The New York Independent System Operator, responsible for operating the power grid in the state of New York.

PDU A power distribution unit.

Peaking (power plant) Used to meet peak load and typically operated at changing output levels.

Post-copy live migration A virtual machine migration method where execution is transferred to the destination machine prior to the memory contents of the virtual machine being transferred.

Power Usage Effectiveness A ratio of total energy consumption to the total energy consumption of IT equipment in a data centre.

Pre-copy live migration A virtual machine migration method where execution of the virtual machine is not transferred to the destination host until attempts have been made to copy memory contents over in hope of reducing downtime.

Project In the context generally used in this dissertation, an emission reductions project and the activities associated with managing and monitoring it.

Project Scenario A plan to be evaluated against a baseline scenario when estimating emission reductions.

Protocol For emission reduction projects this is a specification of the procedures and requirements used to assess the type of project in question.

PUE See Power Usage Effectiveness.

QoS Quality of Service.

RackForce A data centre operator based in Kelowna, British Columbia.

Ramp Rate The rate at which a power plant is able to increase its output.

Regulating Reserve An ancillary service used to manage normal variation in electrical demand, with entities providing this service receiving commands every few seconds from power grid operators.

Related (SSR) An SSR associated with the project but not directly involved in operation. It may also refer to design, construction, or decommissioning activity.

RT Real-time.

Sharding (Databases) The separation of the rows of a database table across multiple machines.

Spinning Reserve An ancillary service used to respond to power plant or other grid failures by rapidly supplying additional power or reducing demand on the system, requiring a relatively quick reaction time.

SSR Sources, Sinks, and Reservoirs (of CO_2e emissions).

Stream A benchmark for evaluating a computer's main memory.

TOP500 A list of the world's 500 highest-performance supercomputers.

UPS Uninterruptible power supply used to protect against power grid failures in data centres by providing power from energy storage for short durations.

Upstream (SSR) An SSR whose output occurs prior to project activity.

US United States of America.

VM Virtual Machine.

VMM Virtual Machine Manager. See hypervisor.

Water Usage Effectiveness The rate of water consumption per unit of energy used by a data centre.

WAN Wide-area network.

WUE See Water Usage Effectiveness.

Xen A commonly used hypervisor.

Chapter 1

Introduction

> When written in Chinese, the word crisis is composed of two characters. One represents danger and the other represents opportunity.
>
> John F. Kennedy, Speech to United Negro College Fund

In spite of continued improvement in the energy efficiency of computing equipment over the past decade, data centres continue to consume more and more energy. This energy accounts for a large fraction of the total cost of ownership of computer systems and, together with the embedded energy and emissions from the manufacture of servers and supporting equipment, results in computing posing a substantial environmental burden.

Climate risk and energy security questions have prompted increased efforts in recent years by power grid operators to integrate more volatile renewable energy resources like solar and wind power. Though power grids have long had to address variation in power supply and demand, and respond to system failures, the drive to integrate more renewable energy further strains the system.

As power grids are becoming more complicated due to energy integration issues, so too are computer systems. Virtualization and cloud computing have increased the flexibility of computing infrastructure. The power consumption of servers now varies significantly from when they are at idle to under peak load whereas just a few years ago servers consumed nearly as much power at idle as they did when under full load.

This dissertation outlines and evaluates strategies to reduce both the financial and environmental costs associated with data centres. It details how the increased flexibility of computing infrastructure and the need for power grids to integrate more volatile renewable energy sources can produce synergy, reducing costs for data centres while enabling volatile renewables to be more effectively integrated. This chapter provides motivation for the chapters ahead, outlining each and reviewing contributions made. It begins in Section 1.1 by reviewing the costs associated with computing from a number of perspectives and outlining the growth trajectory. Power grids and the challenges of integrating renewable energy are then briefly examined in Section 1.2. Section 1.3 then describes how computing infrastructure has become increasingly flexible, enabling it to be used in ways which can lower its environmental and financial costs. Section 1.4 then provides an overview of the chapters ahead and Section 1.5 outlines contributions made. Summarizing thoughts in Section 1.6 then end the chapter.

1.1 The Costs of Computing

The costs associated with computing can be examined from a variety of angles. Data centres have operational costs and other, non-operational costs associated with the manufacture, acquisition and installation of computing equipment. Costs can be measured in dollars spent, in joules of energy consumed, or in environmental impacts, e.g., a volume of greenhouse gas emissions produced. Computing costs can be examined in isolation, or the impact of the use of computing on other activities can also be considered. This dissertation takes a broad view of computing costs, addressing each of the above variants of cost to some extent, and this section briefly quantifies data centre costs.

In spite of an increase in server power efficiency observable in publicly available results from benchmarks like SPECpower_ssj2008 [223], data centre power consumption continues to increase. In 2007 the Environmental Protection Agency (EPA) [89] concluded that data centre power consumption had more than doubled between 2000 and 2006 and predicted that power consumption would double again by 2011. A mid-2011 report [146] argued that a doubling had not occurred but still noted a 56% increase in global IT power consumption between 2005 and 2010.

The costs to manufacture and deliver equipment like servers account for a non-trivial fraction of their lifecycle. The US Environmental Protection Agency estimated [89] that, beginning around 2008, the financial cost of power for a 1U server would exceed its purchase price. A 2008 study [116] concluded that approximately 24% of the lifecycle energy expenditure of a server evaluated came from the non-operational phase of its life. Manufacturing has significantly improved in efficiency in recent years, and further improvements to manufacturing efficiency are being explored. However, efforts to reduce embedded energy have also resulted in decreased operational energy consumption [46].

The energy costs of computing are significant but expanded use of computing may lower overall emissions. Information and Communications Technology accounts for 4–10% of total carbon emissions in countries with knowledge-based economies and 2–3% worldwide [44, 212]. Data centres alone consume approximately 2% of US electricity [146]. Despite this impact, studies have sometimes concluded that expanded use of computing can produce environmental benefits far in excess of the sector's own footprint [13, 128]. Various ways in which computing can reduce emissions include the optimization of business processes, the enabling of dynamic pricing systems, dematerialization, and substitution effects such as when computer communication replaces business travel [186].

1.2 Power Grids

Power consumption accounts for a large fraction of the total cost of ownership of servers, whether evaluated from a financial or environmental cost perspective. Power grids vary significantly in their emission intensity, though, with the carbon intensity of the power grids of different Canadian provinces differing more than 200-fold in some instances [87]. Where data centres are built can thus have a great impact on their environmental footprint, but in carbon-intensive locations data centres may be able to act to enable volatile renewables like wind and solar to be more effectively integrated into the local power grids.

At the same time as power grids are integrating more renewable energy, they have also been introducing finer-grained metering often accompanied by more dynamic systems of power pricing. With these systems in place, data centres may be rewarded for acting in such a fashion as to make renewable energy integration more feasible and can also see a significant reduction in their power costs. Utility companies often offer incentive programs to customers to encourage them to improve their energy efficiency [45]. However, energy efficiency increases may often increase energy demand rather than reducing it — an effect known as Jevons Paradox [133] or the Khazzoom-Brookes Postulate [215] — and do not directly address the challenges of integrating renewable energy. Power grids require a fraction of their users to respond to requests to alter power consumption and reward them for doing so. Providing users with incentives to respond to power grid operator requests more directly addresses renewable energy integration challenges. How data centres are particularly well equipped to meet these challenges and how they can provide these services in a more-efficient manner than the powerplants that have traditionally provided them are addressed in the chapters ahead.

1.3 Increasingly Flexible Computing Infrastructure

In recent years the use of virtual machines and cloud computing has brought increased flexibility to computing infrastructure.

Virtualization using virtual machines (VMs) [55] enables multiple operating system instances to execute concurrently on a single physical server while providing resource isolation guarantees preventing a single instance from monopolizing resources or causing other virtual machines to crash if it fails. Without virtualization servers will likely spend most of their lives at 20–30% utilization or below, where they exhibit comparatively low energy efficiency [206, 58]. Workloads with semi-independent usage patterns can be consolidated on a single physical machine through the use of virtualization. Consolidation can reduce both power costs and hardware expenses by enabling fewer servers in more energy efficient states to be used to process workloads.

Migration is widely supported by virtual machine managers and enables virtual machines to be relocated to other physical machines. It can be used to allow hardware maintenance to be performed on the original host as well as for a variety of other purposes. Migration can often be done with minimal impact to quality of service, resulting in downtime on the order of hundreds of milliseconds [75]. Researchers have also addressed other approaches able to achieve similar results such as the migration of groups of processes [189]. Some applications also support checkpointing current progress to a file on one machine and then resuming from this checkpoint on another, thereby achieving similar results to virtual machine and process migration. Migration and checkpoint/restart functionality may be used to enable physical machines to be powered down to reduce energy consumption during times of low system load by consolidating operations on a subset of the available machines. Migration can also shift workload to a different location with more clean energy available. A 1988 paper [79] argued that migration could achieve, at best, minor improvements in performance, but with power costs accounting for a much larger fraction of the total cost of ownership and exhibiting significant fluctuations there may now be a stronger case for migration.

Virtualization has enabled the development of cloud computing services like Amazon's Elastic Compute Cloud [80]. This sector is rapidly growing, with annual spending predicted to exceed $100 billion between 2012 and 2016 [31, 36]. Cloud computing services offer users access to dynamically resizable pools of resources upon which to execute their tasks. Various reports have concluded that using such services can result in significant decreases in energy consumption, carbon emissions and financial costs [38, 65, 21, 167]. Large-scale cloud computing providers are able to exploit economies of scale to deliver services at a low cost to users who no longer have to pay up front or commit to a long-term contract [47].

Virtual machines are not necessary to provide cloud computing services but are commonly used to do so. Even if virtualization is not used, many of the challenges in delivering services remain the same. Users may not directly deal with finding a specific server upon which to execute, but cloud computing services need to be able to take a request and schedule and provision a server to fulfill it.

Large commercial cloud operators typically operate from a number of data centres. Cloud providers may allocate resources in any of a number of locations in response to resource requests depending on varying financial and carbon costs at each. Users retain the flexibility to allocate resources on demand, while cloud providers may introduce incentives such as offering several versions of a service in which higher prices are charged for higher quality of service guarantees. Doing so may balance user quality of service requirements against the demands of the power grid and its need to integrate renewable energy and manage variability as well as respond to system failure.

1.4 Overview

This dissertation explores a variety of ways in which the financial costs and environmental footprint of data centres can be reduced. Chapter 2 outlines related work, beginning with an overview of how power grids are changing and proceeding to examine how to evaluate the efficiency of servers and data centres. The chapter then concludes with an overview of how to operate computing equipment more efficiently. Chapter 3 addresses the challenges of and means by which to assess the impact of projects aimed at reducing carbon emissions in the computing sector. It demonstrates how to produce credible, auditable figures for inclusion in corporate sustainability reports or for sale in carbon markets. Reducing power costs and computing's environmental footprint by adjusting consumption based on the price of energy, estimates of its carbon intensity, and the availability of local renewable energy are all addressed in Chapter 4. This chapter examines the potential benefits of such strategies and their impact on quality of service. Chapter 5 continues to address renewable energy integration and reducing the price of computing services, but examines ceding a degree of control over operations to power grid operators rather than independent action. This chapter outlines the compensation available for providing such services as well as the impact of this on the workload being processed. Workload distribution across multiple data centres is a reality for a variety of technical and policy reasons, and Chapter 6 examines how the added energy overhead of workload distribution across multiple facilities can be minimized. This chapter surveys existing cloud services and then uses data from virtual machine migration experiments to examine how to effectively use multiple data centres in a computing cloud. Avenues for further investigation are then presented in Chapter 7 alongside concluding remarks.

1.5 Contributions

Portions of the research presented in this dissertation have been previously published as described in this section. Both the contributions of each work and my role in it are highlighted.

Chapter 3 is based upon work that has thus far led to the following publications:

- ICT Greenhouse Gas Reduction Project Protocol: Quantification and Reporting, Canadian Standards Associations, Protocol Version 1, 2012.
- Steenhof, Paul, Weber, Chris, Brooks, Martin, Spence, John, Robinson, Randall, Simmonds, Rob, Kiddle, Cameron, **Aikema, David**, Savoie, Michel, Ho, Bobby, Lemay, Mathieu, Cherriet, Mohamed and Fung, Jonathan. (2012). A protocol for quantifying the carbon reductions achieved through the provision of low or zero carbon ICT services. *Sustainable Computing: Informatics and Systems*, 2(1), 23–32.

A formal, standards-compliant Greenhouse Gas Reduction protocol is often required to receive recognition in carbon capping and carbon trading systems and also provides credibility for use in corporate sustainability reporting. The ICT Greenhouse Gas Reduction Project Protocol was the first such protocol produced for the computing sector compliant with the ISO-14064-2 standard [129] for evaluating emission reduction projects.

I was a member of both the Carbon Assessment Working Group responsible for the protocol's development and the smaller team which evaluated the GeoChronos relocation project that served as a test case for the protocol. I was closely involved throughout, from the outlining of the sources, sinks, and reservoirs of carbon for data centres at a Montreal workshop to the completion of the relocation project assessment. Activities with which I was closely involved include, but are not limited to: examining ways in which to compare baseline and project power consumption, performing risk analysis for the project and evaluating how to determine a data centre efficiency estimate given limited information in a multi-purpose facility.

Chapter 4 is based on work which has also led to the following publications:

- **Aikema, David** and Simmonds, Rob. (2011). Electrical cost savings and clean energy usage potential for HPC workloads. In *Proceedings of the 2011 IEEE International Symposium on Sustainable Systems and Technology.*
- **Aikema, David**, Kiddle, Cameron and Simmonds, Rob. (2011). Energy-cost-aware scheduling of HPC workloads. In *Proceedings of the 2011 IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM).*

I was the first author of each of these papers and also responsible for the development of the time-step tool and the simulator used in the tests as well as the running of the tests themselves. Cameron Kiddle offered feedback and advice on the preparation of the latter paper. Though several papers had addressed price-responsive behaviour and the use of local renewable energy for web and enterprise computing workloads, exploiting dynamic pricing and energy availability had not been addressed for deferrable high-performance computing when “Electrical cost savings and clean energy usage potential for HPC workloads” was submitted for publication. Unlike web and enterprise workloads which are typically not deferrable, require immediate access to resources, and have latency requirements restricting data centre location, high-performance computing workloads are typically run unattended and may be deferred to a later point in time rather than execution beginning immediately. Such attributes may enable high-performance computing workloads to take greater advantage of dynamic pricing and varying levels of renewable energy availability as Chapter 4 explores.

Chapter 5 is based upon work that has thus far led to the following publications:

- **Aikema, David**, Simmonds, Rob and Zareipour, Hamidreza. (2012). Data Centres in the Ancillary Services Market. In *Proceedings of the 2012 International Green Computing Conference*. San Jose, CA.
- **Aikema, David**, Simmonds, Rob and Zareipour, Hamidreza. (2013). Delivering Ancillary Services with Data Centres. In *Sustainable Computing: Informatics and Systems* (an extended version of the previous paper; to appear).

I was the first author for these papers and responsible for the development of the simulator used and the simulations run. Hamidreza Zareipour aided in understanding the ancillary services presented in the paper and in interpreting the market rules and compensation available to organizations providing ancillary services. Though price-responsiveness and varying levels of renewable energy availability had been previously investigated, the ancillary services used to ensure power grid stability and reliability as demand changes and systems fail were largely unstudied and the compensation for providing such services unaddressed. This chapter surveys categories of ancillary services and the compensation available to those providing them. It also details simulations conducted to evaluate the potential of these ancillary services to reduce energy costs and the corresponding impact upon workload quality of service.

Chapter 6 describes research which has thus far led to the following publication:

- **Aikema, David**, Mirtchovski, Andrey, Kiddle, Cameron and Simmonds, Rob. (2012). Green Cloud VM Migration: Power Use Analysis. In *Proceedings of the 2012 International Green Computing Conference*.

I was the first author for this paper, creator of the test engine which executed the tests, and responsible for the execution of the tests themselves. Andrey Mirtchovski performed the initial environment setup, assisted in determining the origin of problems affecting some of the workloads, and also aided in collecting monitoring information. Cameron Kiddle liaised with the other GreenStar Network projects and provided suggestions and feedback when conducting the experiments and during analysis of the results. A variety of prior work had explored the idea of reducing system power consumption or energy costs through performing VM migrations in LAN or WAN environments. These earlier works did little to account for added energy consumption from the migration process though this additional overhead will reduce energy savings and may even result in a net increase in energy consumption. Based upon migrations of predictable workloads in a metered environment, the added energy consumption of migrating different types of workloads in different ways was assessed. These results were then used to provide guidance as to what workloads are most suitable for migration as well as when and how to migrate them.

1.6 Summary

Computing has a complex relationship to energy and carbon emissions. Data centre power consumption is rapidly growing despite the also-rapid increase in computing's energy efficiency. However, the increased use of computing services holds the potential to reduce emissions in other areas. Though expanded use of computing increases its energy use and carbon footprint, the net effect may be a decline in each.

The power grid is changing, with market-based approaches becoming more common and power grids being tasked with incorporating additional renewable energy from variable and volatile sources. The challenge of integrating these additional renewables is one that data centres may be well suited to aid in. Data centres typically possess both local energy storage and local generation capacity. In addition they may have the ability to either defer the execution of workloads or migrate them to a different locale. Altering its operational patterns to suit the power grid may open new revenue streams for data centres or reduce their energy costs.

This dissertation investigates how to reduce the environmental and financial costs of computing by focusing on a variety of mechanisms to do so. Computing projects can be made more efficient, not only generating positive press for the organizations involved but also leading to possible sales of carbon credits and reducing vulnerability to changes in fossil-fuel energy costs. Both financial and environmental savings can result from efforts to shift energy consumption to points in time in which power is available at low cost and from clean energy sources. Data centres may also be able to lower power costs while enabling power grid operators to better integrate more renewable energy by responding to their requests to alter power consumption. Doing so is likely to necessitate increased cooperation between data centres and additional synchronization activity, heightening the need to understand how to perform such activities in an efficient manner.

Chapter 2

Background & Related Work

While early research in computer systems typically focused on achieving maximum throughput, reducing power consumption has become of increasing importance. The US Environmental Protection Agency estimated [89] that, beginning around 2008, the total cost to power a 1U server would exceed its purchase price. In recent years the cost focus has expanded to encompass not just financial but also environmental costs and total lifecycle costs rather than focusing only on the operational phase. With costs during a server's operational phase due largely to power plant emissions and power grids incorporating added renewables, an understanding of power grids is also key to reducing computing's financial and environmental costs. This chapter provides an overview of these subjects, beginning with power grids and proceeding to discuss efficiency measures for servers and data centres as well as outlining efficient ways to operate them.

This chapter is organized as follows. It begins in Section 2.1 by surveying relevant aspects of power grids. Section 2.2 then outlines metrics and standards for evaluating servers and data centre efficiency. How to efficiently operate computing equipment and data centres is the focus of Section 2.3. Summary remarks in Section 2.4 then close the chapter.

2.1 Power Grid Overview

A reliable supply of energy has been of particular importance to data centres over the years, leading to the widespread installation of uninterruptible power supply equipment in data centres. Reliability concerns have attracted broader attention in recent years, however, due to the challenges associated with increasing renewable energy use. Power costs are one of the largest bills that data centres pay and with energy costs rising and many renewables currently requiring subsidies and preferential treatment to operate profitably, ways to reduce power costs are thus of great relevance.

This section outlines the changes taking place to the power grid, providing a basis for the work in this dissertation on how data centres can face and even embrace them. Section 2.1.1 describes the dynamic, market-based approach to power pricing which has become more prevalent with finer-grained power metering. Section 2.1.2 then outlines low-emission energy sources being incorporated into power grids as well as the challenges they impose. Integrating increased and often volatile renewable energy into the grid strains reliability, but power grid operators have always had to confront challenges of this nature. Section 2.1.3 describes how power grids address both normal variation and the failure of transmission lines or power plants and how they propose to address reliability challenges imposed by added renewables. How data centres may be able to save energy and money by adjusting how and where they connect to the grid is then addressed in Section 2.1.4. Energy storage is finally discussed in Section 2.1.5, since it is likely to become more common in future power grids.

2.1.1 Growth of Markets

Individual consumers may pay a fixed price for each unit of electricity consumed, or possibly peak and off-peak rates, but large industrial consumers may experience much more variable power prices. Demand for power varies throughout a typical day and substantial variations in power price often exist even on such a short timescale.

Different types of power markets exist, with market structure varying from region to region. For example, a single hourly price is determined for power purchases in Alberta [131]. In New York state on the other hand, consumers can participate in day-ahead, hour-ahead, or real time markets with different prices for different parts of the state [142]. Committing to purchasing power closer to the time of delivery appears to produce lower average prices but higher volatility [202]. The price of power is often not yet fixed at the time of consumption in many markets — a phenomena known as ex-post pricing [90]. Alberta prices are set based upon an average of 60 1-minute prices [131], meaning that a spike in price late during an hour can skew the hourly price.

Negative prices due to an intermittent oversupply of power are occurring with increasing frequency. They are often due to policies limiting power grid operators' ability to curtail certain renewable energy sources as well as production subsidies that enable such sources to operate profitably when paid negative rates. Other generators may not be able to easily adjust their output as supply and demand changes, resulting in them operating at a loss in such instances [125].

2.1.2 Lowering Emission Intensity

Fossil-fuel based power plants account for a large fraction of world pollution. According to the US Environmental Protection Agency [1], US power plants account for 40% of the nation's carbon footprint, produce 67% of its sulphur dioxide emissions and emit 23% of its nitrogen oxide. Displacing fossil-fuel-intensive energy sources with lower-emission generators is key to reducing this pollution.

The most suitable renewable energy sources vary in prominence from region to region but renewable energy sources are often more challenging to incorporate into the power grid than traditional, carbon-intensive power plants. A majority of electrical power in Canada is produced using hydro plants but coal-intensive Alberta has a greater level of wind turbine capacity installed than many other Canadian jurisdictions [61]. Other types of low-carbon energy sources include nuclear, solar, and geothermal power. Burning biomass derived from waste or crops farmed for the purpose is often considered to be carbon neutral though this is not without some controversy [135].

Lowering emissions of fossil fuel power plants is another approach being tried. Capturing and sequestering carbon emissions from power plants is being attempted. Combined heat and power plants, also known as cogeneration plants, reuse waste heat and district heating and cooling projects can also improve efficiency. Though carbon-intensive fuel sources are finite and responsible for much pollution, such fuels are relatively easily transported and made available on demand in the quantities required, unlike many renewable energy sources whose availability varies significantly over time and which may be most plentiful in remote areas lacking transmission capacity.

Even in hydro reservoirs, where power grid operators have a high level of short-term control, water must still be available upstream and operators may also be forced to deal with high-water challenges. Overdependence on hydro reservoirs may be problematic during drought [192], requiring the creation of large energy storage facilities [54]. Hydro dams may also harm aquatic life during times of overabundance by increasing dissolved gases in the water [69].

A deeper-than-surface-level examination of energy sources is needed to understand the environmental costs of each. There are risks with carbon sequestration approaches [248] and wind turbines have been found to increase bird deaths [57]. Hydro development destroys animal habitat, alters waterflow patterns, and kills plant life that might otherwise act as a carbon store, leaving it instead to rot [210]. Geothermal power generation also imposes risks, including, for example, increased earthquake activity [230].

Carbon emissions may also be significantly impacted by short-term variation in power demand or measures intended to preserve the reliability of the power grid against transmission line or power plant failures. Thus the ancillary services which address these challenges are now discussed.

2.1.3 Ancillary Services

Ensuring reliable power has long been a priority for power grids, but the drive to increase use of renewables heightens the challenge. Ancillary services is the name given to the set of services power grids use to ensure reliability. These services are being reevaluated and restructured to facilitate greater renewable energy use. This section surveys existing ancillary services, the compensation available for providing them, and changes to ancillary services being discussed [182, 183]. Ancillary services have traditionally been delivered primarily by power plants but in recent years demand-response has become increasingly important.

Providing ancillary services in an environmentally friendly manner is important to reduce power grid emission intensity. Altering the output of baseload natural gas [139] and coal [64] power plants to accommodate rapidly shifting renewables like wind may in some instances increase rather than decrease emissions. Other consequences include a potential increase in fuel consumption and extra wear to power plants. Though power plants can be built to more efficiently alter output, power plants built to operate at constant output are generally more efficient. If able to accommodate variability in power production and consumption with energy storage devices and by having power users alter consumption on request, a more efficient grid is likely to result, reducing energy prices and the amount of reserves required [92].

There exist inconsistencies among power markets in both the terminology and details of current ancillary services, but some generalizations may be made. Services typically are composed of regulating reserves, contingency reserve, replacement reserves and a few other services such as black start and voltage control.

Regulating reserves address normal minute-to-minute changes in power consumption as devices are powered on or off or the consumption of a device changes. Regulating reserves use automatic generator controls which receive fine-grained commands from grid operators as to what their state should be. Instances exist of batteries [218] and flywheels [153] providing regulation.

Contingency reserves address faults such as power plant or transmission line failures, and are typically separated into spinning and non-spinning (or non-synchronous) categories. The primary difference between spinning and non-spinning reserves are response time requirements with spinning reserves required to begin acting more quickly. Replacement reserves are used to replenish other reserves once they have been activated.

When integrating volatile renewables like wind, better forecasting can reduce but not eliminate the added strain on reserves. A variety of approaches to address this added strain are under discussion. Power grids are increasingly permitting consumers to provide ancillary services. Purchasing larger volumes of existing reserves is another option, with regulating reserve purchases in Germany doubling in recent years due to added wind and solar power [51]. A greater role for energy storage to provide ancillary services is also seen [27]. Other approaches include reducing the length of scheduling intervals, increasing transmission capacity, and merging balancing areas over which stability is maintained [181]. Ramping services and wind “firming” services are also being discussed [40]. Ramping services provide capacity to accommodate shifts up and down in generation and could smooth transitions associated with normal diurnal patterns. Wind-firming involves facilities altering power consumption to even out wind energy fluctuations.

Payment for the provision of ancillary services may account for a large component of a power plant’s profits [242]. Power prices already exhibit high volatility, but, although this may differ depending on market structure and location, the pricing of ancillary services appears to exhibit higher volatility, more frequent jumps, and more extreme price spikes (although lower average prices). Extreme prices in energy and ancillary services only occur simultaneously about 20% of the time. Outside extreme prices, prices paid for ancillary services typically exhibit the same peak hours as in the energy market.

2.1.4 Connecting to the Grid

In addition to the unit cost of purchased energy and ancillary service fees and payments various other charges may be found on power bills. Capacity costs associated with a consumer's maximum power consumption may account for a significant proportion of its power bill [107]. This section examines ways in which to reduce capacity costs, reduce losses in transmission, and improve reliability of data centre power.

Locating data centres at the site of a renewable energy source reduces transmission system load, lessening the need for new transmission lines. Renewable energy resources are often located in remote areas with poor connectivity to the power grid [61, 144]. High-voltage transmission lines are difficult to build due to regulatory obstacles, possess limited capacity, and cost more than a million dollars per kilometre to construct [144]. When local renewable energy sources are inactive, transmission lines normally used to deliver energy from a location may instead be repurposed to deliver energy to a facility there.

If alternate approaches to charging for transmission capacity are adopted, such as basing capacity charges on the incremental transmission capacity required to support it [161], facilities located near wind farms may be able to dramatically cut such costs. In Alberta entities willing to curtail consumption when necessary can use otherwise surplus transmission capacity at reduced cost [39]. As the primary transmission bottleneck is in delivering energy from turbines to distant consumers [185], it seems unlikely that a data centre located near turbines would be unable to obtain power from the grid during lulls in local energy availability.

Connecting to a local renewable energy source may increase efficiency due both to reduced transmission losses and fewer energy transformations needed [214]. Data centres located at renewable energy sources may also be able to share some of the costs of a grid connection through mechanisms such as the Alberta Electrical System Operator's primary service credit [39]. In opposition to earlier research [234], a 2012 study concluded that there was little difference in efficiency between state-of-the-art AC and DC methods of power distribution in data centres [208]. However, using DC to connect data centres to nearby renewables may still increase efficiency. DC power is natively produced by photovoltaic cells [236] and the "wild AC" current produced by wind turbines must be stabilized by first converting it to DC [72]. If DC is used for transmission to the data centre, this eliminates conversions in which energy is lost and may increase reliability and reduce costs. Reducing the number of inverters and rectifiers required to convert between AC and DC simplifies the system and eliminates potential points of failure. Using DC rather than AC transmission enables flow control, preventing an external, malicious, or faulty entity from bringing down the grid [211]. It eliminates the requirement to harmonize multiple AC sources to the same frequency [178], making it easier to connect a variety of power sources to a data centre. Renewables can also serve as a second, independent source of power with eBay using biogas as a primary power source in new construction [29] and forgoing local energy storage due to having the power grid as a backup energy source.

2.1.5 Energy Storage

Sourcing a higher fraction of energy from renewable sources is likely to require both more and larger energy storage facilities in power grids. This applies particularly to high variability resources like wind and solar. Photovoltaic solar panels are known to vary in output +/- 50% in a 30–90 second time frame and as much as 70% in 5–10 minutes [185]. The rising need for energy storage also applies to hydro-based grids due to seasonal variation and drought as Section 2.1.2 noted.

Capacity factors are assigned to projects based on the fraction of installed capacity usable when needed, providing an indication of the level of volatility which must be accounted for. An analysis of US wind capacity factors [173] notes that these vary widely from 2%-60% depending on location and other factors such as availability during specified peak hours. In Alberta, the regulatory authority has assigned wind projects a 20% capacity factor and hydro projects a 50% capacity factor [61].

Storage makes it possible to deliver reliable power from relatively unreliable sources. Various energy storage options exist, outlined in more detail elsewhere [194, 218]. Supercapacitors, flywheels and batteries can be used to keep unexpected energy shortfalls from affecting running equipment. Groups of flywheels totalling up to 20 MW capacity are currently being tested [153], and the use of batteries in plug-in hybrid vehicles by power grids is also being studied [140]. Compressed air storage is another option involving air compressed into underground caverns being later used to drive natural gas generators at extra-high efficiency. Hydro reservoirs can also be used as batteries, with British Columbia using low-cost Alberta coal power at night to enable it to replenish its reservoirs [170], and Alberta purchasing BC hydro power at premium rates during peak hours. Water may also be pumped to higher elevations to achieve energy storage [94].

Local energy storage using uninterruptible power supplies has long been a part of data centres. Energy storage within data centres and potential synergy with power grid interests will be discussed later in this chapter.

2.2 Evaluating Efficiency

While the TOP500 list [231] of the world's most powerful supercomputers has long been established, the rising importance of energy efficiency can be seen in the emergence of the Green 500 list [109] which ranks supercomputers by efficiency rather than peak performance. The emergence of the Green Grid [110], an industry consortium developing efficiency standards and benchmarks, and Energy Star's added focus on data centres [85] are two more examples of this shift in focus.

Though reducing power grid emission intensity is important to reduce data centre environmental costs, the use of efficient equipment is a more direct way to do so. This section addresses metrics for evaluating the comparative performance of servers and data centres. The remainder of this chapter, by comparison, focuses on operational concerns. This section proceeds as follows. Section 2.2.1 examines standards and benchmarks for comparing the efficiency of two groups of servers and Section 2.2.2 then focuses on comparing data centres.

2.2.1 Servers

Early computing efforts generally focused on metrics like throughput over energy efficiency, but improving server energy efficiency has been an important goal for some time now. The first devices to receive Energy Star efficiency ratings were computers and monitors [6], but assessing server efficiency remains a challenge.

Nameplate power labels are known to poorly reflect actual server power consumption [91, 138, 174]. Other measures must thus be used to assess server efficiency. In 2007 Barroso *et al.* [58] argued for a greater focus on energy proportionality — the extent to which a server's power consumption varies based on load — and in recent years significant progress has been made in this regard [237]. The actual workload processed by servers thus is of key importance in evaluating their efficiency. Average server utilization differs based on the type of workload being executed and may vary from system to system. A 2002 study of web server traces [68] found average utilization ranged from 11–50%, and the Parallel Workloads Archive [190] contains high performance computing traces in which utilization varies from 10.7–87.9%. Given the variation in utilization, with servers also varying in energy proportionality the most energy efficient server for one task may be a poor choice for another.

Benchmarks have emerged addressing the energy efficiency with which different workloads are processed. SPECpower_ssj2008 [223] was a pioneer, evaluating energy efficiency at varied levels of utilization for a Java-based workload, and other benchmarks and metrics focused on energy efficiency have also been developed [17, 199, 209, 220, 224]. Machines may also be evaluated against published efficiency data for similar systems, using sources such as the Green500 reports [109] which address energy efficiency for the world's largest high-performance computing systems.

If using cloud environments instead of owning servers, corporations also need to account for factors like elasticity [219] — the rate at which new virtual machines can be launched and made ready for use. This will influence the amount of spare capacity needed to handle variation in workload.

2.2.2 Data Centres

Data centres vary significantly in overhead. The predominant metric used for data centre efficiency in recent years has been power usage effectiveness (PUE) [16]. PUE is calculated by dividing a data centre's total energy consumption by the total energy consumption of the IT equipment it contains, providing an indication of data centre overhead. PUE's reciprocal is known as data centre infrastructure efficiency (DCIE).

A PUE value of 1 has typically been viewed as ideal, an indication that all input energy is used to power computing equipment. Others have sometimes suggested that PUE < 1 can be achieved through waste heat reuse. Reuse has been captured in other metrics: energy reuse effectiveness (ERE) which incorporates reuse into the PUE equation, and energy reuse factor (ERF) which refers to the fraction of input energy reused [233].

Metrics for other aspects of data centre efficiency have also been devised. Carbon usage effectiveness (CUE) [50] tracks the carbon emissions per unit of energy consumed in the data centre and Green Energy Coefficient (GEC) refers to the fraction of energy coming from renewable sources [33]. Water Usage Effectiveness (WUE) [49] addresses water use in cooling. There is often a trade-off between water use and energy use, and a decrease in data centre water use may increase total water usage as power generation is also often water-intensive. WUE_{source} was thus introduced alongside WUE, incorporating water used in power generation in addition to the data centre's direct water use. Another metric, Electronics Disposal Effectiveness (EDE) [30] notes the percentage, by weight, of decommissioned equipment disposed of through organizations which have been certified environmentally friendly.

Computing the earlier metrics allows an individual data centre's efficiency to be measured. To determine if a data centre is "efficient," however, requires that it be compared to others. Comparing the performance of data centres is difficult, with the original PUE specification including the following warning [16]:

> It should be noted that caution must be exercised when an organization wishes to use PUE to compare different data centres, as it is necessary to first conduct appropriate data analyses to ensure that other factors such as levels of reliability and climate are not impacting the PUE.

Quality of service requirements for certain industries such as banking may be higher than others, requiring greater redundancy and resulting in correspondingly lower efficiency. Regulatory requirements and privacy laws may also restrict data centre location, preventing connections to low-emission-intensity power grids. Latency requirements also play a role, perhaps most significantly in high-frequency trading but also affecting web server farms and various other interactive services.

Estimates have been made of typical PUEs. The EPA expected an average PUE of 1.9 in 2011 [85]. A recent survey [77] of three hundred data centres reported an average PUE of 2.8 though, with just 19% reporting a PUE < 2. The average PUE for 2010 was estimated as 1.83-1.92 by Jonathan Koomey [146]. Large data centres can have substantially lower PUEs though, with Google reporting a 12-month average PUE of 1.13 as of June 2012 [9].

Criticism has been raised over focusing on PUE to the exclusion of other metrics. This can penalize data centres for energy-saving actions such as powering down idle servers since this increases the fraction of energy counted as overhead. Data centre PUEs can fluctuate throughout the course of a data centre's life and even well-designed data centres may start out life with a high PUE due to an initially low IT load [67]. Google's average data centre PUE appears to fluctuate by about 0.05-0.07 depending on the time of year. Google has also reported that calculating PUE in different ways can lead to somewhat different results [9]. Although a later version of the PUE specification [26] outlines categories of PUE measurements to reduce the ambiguity, the category of measurement performed is not typically noted in published data.

In addition to the long-standing Energy Star ratings for computer components, Energy Star has in recent years developed a data centre rating system [85], ranking data centres on a percentile basis from least to most efficient, though based primarily on PUE. Some criticism has been raised [14] that Energy Star ranks data centres in part based on their size, expecting lower PUEs for larger facilities. A well-rated small data centre might compare poorly to an average large-scale data centre in its actual efficiency.

Green building rating programs are also adapting to cover data centres. As part of recent updates to the Leadership in Energy and Environmental Design (LEED) program, the U.S. Green Building Council expanded it to address data centres [35]. The Green Building Initiative [101] has also certified a number of data centres [5] as part of its Green Globes program.

2.3 Operating Efficiently

To be most efficient requires not only purchasing efficient equipment but also operating it efficiently. Section 2.3.1 begins with a focus on efficiency in individual servers and is followed by Section 2.3.2 which addresses clusters of servers. Storage system efficiency is then surveyed in Section 2.3.3 and this is followed with an overview of data centre power and cooling infrastructure in Sections 2.3.4 and 2.3.5 respectively. Energy efficient networking is then addressed in Section 2.3.6, and the survey of efficient operations concludes in Section 2.3.7, focused on using multiple data centres. More detailed overviews of server and data centre efficiency are available elsewhere [63, 238].

2.3.1 Individual Servers

Various techniques for operating individual servers efficiently have been developed. In this section the efficient operation of server components such as CPUs and main memory is first discussed and then increasing efficiency through the broader use of graphics processors is outlined. Discussion of efficient data storage systems is deferred to Section 2.3.3. Finally, the exposure of energy consumption information to applications is discussed.

The focus of operational efficiency research for CPUs has been dynamic voltage and frequency scaling (DVFS) [62, 193, 245], the changing of voltage levels supplied to processors and the altering of processor frequency. By manipulating these individually or in combination, processor power consumption can be significantly lowered in many cases. With careful analysis and workload monitoring, a decrease in power consumption can be achieved with little or no reduction in throughput.

RAM power consumption has also been studied [156], with researchers examining how to effectively use the power states of different banks of RAM to save energy. With fast, cheap, non-volatile memory on the horizon, changing the type of memory in servers has implications for operating systems with the potential for significant energy savings [52], enabling servers to be powered down without loss of program state.

In recent years graphics processor units (GPUs) have begun to be used for non-graphics code. Though less flexible than general-purpose processors, graphics processors are able to execute parallel code at high efficiency. Claims of a thousand-fold improvement in energy efficiency appear exaggerated [157] but even the more-modest efficiency improvements achievable appear likely to increase future use of these devices.

The exposure of energy consumption information to applications running on systems has been another focus of energy efficiency research. The Nemesis operating system [160], for example, was extended to implement energy accounting and charging for applications [184]. With such information available, applications may choose to operate in a reduced mode to save power [96].

2.3.2 Multiple Servers

Looking at not just a single machine but groups of machines provides added opportunity to increase efficiency. This section focuses primarily on operational concerns, noting a shift to server blades sharing enclosures [206] as well as the adoption of containers as the basic building blocks of some data centres [240] without dwelling on such issues. This section outlines areas of research including load unbalancing, processor scaling at a cluster level, and the impact of heterogeneous servers. How to effectively utilize local renewable energy is also discussed. Virtual machines were introduced in Section 1.3, but this section expands upon virtual machine migration and the possible energy savings in greater detail due to its relevance to Chapter 6.

In contrast to traditional load balancers which attempt to even the load on available servers, load unbalancing aims to concentrate workload on a small subset of servers, enabling others to be placed in low-power states. A 2001 paper [73] introduced a system named Muse which routed incoming requests to servers based on load information, operating a fraction of servers at near peak load and placing others in a low-power state. A simple QoS scheme outlined in the paper permitted the dropping of requests which could not be satisfied in an energy efficient manner. A second paper that year [196] described the creation of a power-aware cluster web server. A management algorithm was used to migrate work from the least-loaded machines during times of decreasing load so that these machines could be shut down. It then powered machines on and redirected load to them when demand rose. The paper also examined a cluster configured for more general-purpose applications and the applications influencing load unbalancing decisions using an API. Researchers have continued to explore load unbalancing, addressing added wear-and-tear from repeatedly powering servers on and off [74] and also addressing areas like high-performance computing [151]. Even simple heuristics for powering down servers have been found to significantly improve the energy proportionality of clusters [71].

The use of DVFS to reduce the power consumption of individual servers was discussed in the previous section, but greater savings may be achieved by coordinating power scaling across clusters and combining this with load unbalancing. A 2003 evaluation [84] compared voltage scaling on individual servers against coordinated voltage scaling across a cluster, also addressing their effectiveness when combined with load unbalancing. Energy savings for independent voltage scaling on servers ranged between 20–29%. Coordinated voltage scaling offered slightly increased savings, though not enough to justify the added complexity the authors argued. Whether load unbalancing was more effective than voltage scaling varied by workload, but combining the strategies offered the best results. The most effective manner in which to combine load unbalancing and processor scaling varied from application to application with sometimes surprising results. One study [99] found that increasing the number of servers used and aggressively using DVFS sometimes decreased both execution time and energy consumption.

Real-world clusters are often composed of heterogeneous servers, and heterogeneity can be exploited to increase energy efficiency. Heath *et al.* [117] used a model of a cluster's heterogeneity to determine an efficient workload distribution policy. Their tests showed a 45% increase in savings relative to an energy-savings strategy that assumed homogeneous servers.

Making effective use of local renewable energy is another subject that has drawn attention. Maximizing the use of local wind [147] and solar [104, 105] energy when scheduling high-performance computing workloads has been investigated. Due to the volatile output from many renewable energy sources and the need to immediately process requests in many workloads, the use of local renewable energy has most often been considered in the context of multiple data centres, a subject that will be addressed in Section 2.3.7.

Virtual machines have become widely used to enable workload consolidation, as Section 1.3 outlined, and form the basis of many cloud computing platforms. Consolidating multiple workloads on a single physical machine generally improves energy efficiency. Rather than each executing on a separate server requiring some power to operate when idle, multiple VMs can share a single physical machine. As servers typically operate at low average utilization as outlined in Chapter 1, physical resources can often be oversubscribed at low risk of violating quality of service agreements. Means to track the power consumption of individual VMs have also been developed [227].

Migration capabilities not only enable virtual machines to be relocated to enable physical maintenance to be performed, but migrations within and between data centres may be used to establish a more energy efficient configuration. There are three general categories of virtual machine migration: non-live migration, pre-copy live migration, and post-copy live migration.

When a non-live migration starts, the virtual machine is first suspended on the source machine where it was running. Its entire memory contents are then transferred to the destination machine where its execution is resumed. This leads to predictable downtime but results in the VM being suspended for long periods.

Pre-copy live migration is implemented in production virtual machine managers (VMMs), having attracted researchers looking to reduce migration-related downtime [75, 180, 232]. It proceeds as follows. First, in a pre-copy stage the virtual machine continues to execute as before while memory pages are copied to the destination physical machine. The copying process involves iterating through pages of memory used by the virtual machine, copying each to the destination machine. Memory usage of the still-active VM is monitored and overwritten pages of memory are marked “dirty” and flagged for retransmission. Pages may thus be transmitted repeatedly during the migration. Memory later allocated by the virtual machine is also flagged for transmission. Eventually, based on VMM-specific conditions, the VM will be suspended and the second stage of the migration begun. Xen initiates this phase when [41] few pages of memory were modified during the last pre-copy iteration, many pre-copy iterations have taken place, or a large amount of network transfer has occurred. KVM instead begins the second stage of migration when estimated downtime is below a set threshold, and this approach may result in migrations lasting indefinitely while still adversely impacting VM performance due to the overhead of the pre-copy process [126]. Once all allocated pages of memory have been transferred to the destination, the third stage of the migration takes place where the VM begins execution on the destination machine and its remnants are cleared from the originating machine. The second and third stages of the migration process typically result in downtime of a few milliseconds to a few seconds [75, 232].

Researchers have studied post-copy migration [120, 121] but it is not yet available in production systems. Post-copy live migration begins by suspending the execution of the VM on the source machine and then immediately attempting to launch it on the destination machine. The memory contents of the virtual machine are then transferred to the destination, with the VM blocked waiting whenever it attempts to read a memory page not yet transferred. Memory need be copied at most once resulting in a fast migration. However the performance of the migrated virtual machine may be more adversely affected during the transition phase than in a pre-copy live migration due to waiting for the arrival of needed memory pages. Post-copy migration also increases VM failure risk during migration, as if either node fails before the memory transfer is complete then the VM fails, but if the destination machine fails in either non-live or pre-copy live migrations, the VM can be resumed on the original machine.

Key factors affecting VM migration duration include [41] the migration link bandwidth and, for pre-copy live migrations, the rate at which pages of memory are overwritten during the pre-copy migration stage and the conditions determining its end.

Research into reducing energy consumption using migration has been conducted but has focused on energy consumed before and after migration rather than including energy consumption of the migration process itself. Research using migration to place systems in a more efficient configuration has generally focused on metrics such as total migration time, impact on application quality of service [225, 239, 252] or ignored the migration period [132] rather than considering the added overhead of the migration process. A paper introducing the Virt-LM migration benchmark [124] did note that migration overhead ranged from 0.36–11.31% of system resources and varied from hypervisor to hypervisor but did not explore this issue further.

2.3.3 Data Storage

This section outlines various techniques for reducing the energy costs of data centre storage systems. Strategies for reducing storage-related energy consumption have focused on spinning down idle disks and taking advantage of cached, centralized, and replicated data.

Through analysis of user activity [78, 118], disk platters may be spun down when idle, thereby saving significant energy. Since it takes much longer to reactivate disks than to change CPU power states, such algorithms must operate at much coarser granularity than algorithms targeting CPU power states. Disk wake time can be reduced by performing I/O at a variety of disk speeds rather than alternating between stopped and full-speed disks. DRPM [114], for example, adjusted the disk spin rate based on load. Coop-I/O [246] is another approach which can result in disks spun down a larger fraction of the time by enabling applications to specify operations as deferrable or cancellable so that operating systems can group operations more effectively.

Careful use of cached and replicated data provides further opportunity for power savings. Solid-state disk caches can enable significant power savings, with a 1994 analysis [166] concluding that a cache could reduce power consumption by at least 20 percent. As the years have passed larger caches have become more common. Data is generally replicated to two or more locations given the possibility of disk failure. As Pinheiro *et al.* demonstrated [197], storage system power consumption can be lowered by reducing the number of replicas powered at any time, maintaining reliability by caching writes intended for powered-off disks. For archival data it may also be possible to have no replicas actively powered as is the case with removable storage media and tape libraries and may be what enables Amazon's Glacier storage service [2] to achieve a much lower cost than Amazon S3.

2.3.4 Power

Data centres may be able to repurpose local energy storage to reduce peak demand or in other ways lower energy costs. Some locations may have lower average energy costs than others, but the strategies in this section may be applied regardless of data centre location.

Although data centres consume significant power, they may overstate their peak demand despite often paying separate capacity charges for this on their electrical bills. Mitchell-Jackson [174] noted significant differences between the capacity of power connections that data centres requested and their actual power consumption. She noted many reasons for this: nameplate power ratings representing maximum possible power drawn; the use of oversized circuits since circuit capacities are not typically a multiple of system power consumption; the use of redundant power supplies; overestimated cooling needs; and overdesign factors built in by engineers. Similar findings continue including a 2007 paper [91] reporting results from large clusters over a six month period. Even executing the most power-intensive benchmarks, an example server's power consumption reached no higher than 60% of its nameplate power reading. Clusters of thousands of nodes executing different workloads were evaluated with a 7–16% difference found between the possible and actual aggregate power consumption. Smaller-scale differences were observed at rack-level and larger differences seen at the data centre scale.

Power capping to reduce peak power consumption of servers, enclosures, clusters and data centres by dynamically adjusting the component power states is a strategy which has attracted and continues to attract attention [93, 158, 206, 243], with the latter paper in the previous paragraph further confirming its effectiveness. Efficiency could be further improved through systems like RAILS [168] which uses power supplies of various sizes that can be turned off when not needed to meet current load, enabling active power supplies to operate near peak efficiency.

Data centre energy storage has traditionally been used to ensure uninterrupted power long enough for generators to come online in a power grid failure. In recent years researchers have examined how to use energy storage to reduce data centre capacity costs [103, 107, 137], to avoid consuming power when most expensive [235, 113], and to enable under-provisioning of data centre power infrastructure [108, 145, 241].

With Section 2.1.5 outlining a rising need for energy storage in power grids, opportunities for synergy between power grids and data centres arise. Data centres can be compensated for adjusting power consumption or running backup generators in emergency situations, and the demand-response aggregator EnerNOC [86] reports some data centre clients. eBay has taken a different approach [29], using local biogas-powered fuel cells for primary power and, with the power grid serving as a second, independent power source, deeming reliability adequate to forgo local energy storage.

2.3.5 Cooling

Prechilling coolant or overcooling at off-peak hours is a strategy which has been investigated [244] and is now in use in some data centres [172], but it is not the only way to improve data centre cooling efficiency. This section outlines a number of alternative strategies which range from changing the data centre's set temperature, ensuring that chilled air is directed only when and where needed, placing workloads in cooling-efficient locations, and using environmental air for cooling.

With an estimated 2–5% reduction in energy cost per degree if the set temperature of a data centre is moved up a few degrees [81], such a change is a significant opportunity for power savings. Contrary to fears, a 2007 Google study [198] found little correlation between hard disk failure rate and temperature. The American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE) called in 2008 for increased data centre temperatures, though warning of a dramatic increase in server fan power consumption beyond a certain point [48]. A 2012 large-scale review of system temperature, power consumption, and failure rates [81] similarly noted a lower than expected impact of higher temperatures. Component failure rates are often assumed to rise exponentially with temperatures, but increases in failure rates with increasing temperatures were generally at most linear, excepting extremes over 50°C. For a few components failure rates actually decreased as temperature rose. Temperature variation appeared to play a greater role in hardware failure than absolute temperature. The authors also found similar temperature variation across data centres studied. The hottest 5% of servers were found to be more than 5°C above the average temperature, and the top 1% of servers were 8–10°C above the average.

Temperature variation in data centres motivates improved control over air distribution and chiller operation such that cool air is directed when and where needed. This can in part be done by altering the structure of data centres such that they can be cooled more effectively. Strategic use of air barriers in data centres can reduce fan energy by 75% and chilled water demands by 21% according to one study [97], and APC has argued for replacing raised floors with solid floors and ducting [207]. Container-based data centres have also been reported to cut cooling costs by 80% [201].

Sensor networks can be deployed to determine when and where cooling is needed and focusing it there. Air conditioner supply air temperatures and flow rates are now dynamically adjustable based on temperature sensor network data [59, 70, 205] with spot cooling available to treat remaining problem areas [119]. Hot-spot data can also be used by load balancers and schedulers to target cooling-efficient areas of the data centre, with a 2005 paper by Moore *et al.* [177] detailing such an implementation. They found that hot spots could be due to activity in a seemingly unrelated area as, e.g., server fans might suck in exhaust air from racks some distance away if cool air was supplied at inadequate pressure.

So-called “free cooling”, the use of outside air of sufficiently low temperature in place of operating chillers [12], has become common in recent years. Facebook reported achieving a PUE of 1.08 by combining evaporative cooling with use of outside air [98]. Other related approaches to reducing cooling costs exist including, e.g., a cooling loop installed at the University of Calgary using water from the Bow River to improve cooling efficiency. Despite having higher efficiency than more traditional means of cooling, free cooling still requires significant power. Research has thus been conducted into how to incorporate knowledge of the change in efficiency of these systems under different environmental conditions into data centre scheduling [164].

2.3.6 Networking

Networks connect servers within a data centre, data centres to one other, and end users to all. Smart 2020 [217] argued that global telecommunication infrastructure has a larger carbon footprint than data centres, and the networks involved in delivering cloud computing services can significantly affect their energy consumption [53]. Though last-mile access networks presently account for the bulk of network-related energy consumption, this balance is shifting [150]. Networks should thus be addressed when attempting to reduce financial and environmental costs as doing so may require added network capacity for workload migration.

Energy proportionality is still lacking in networking equipment. A 2007 study by Google researchers [91] found that Ethernet switch power consumption varied as little as 2% between idle and fully-loaded states. Examples of network equipment with higher power consumption when idle than when active have even been reported [123]. Improvement in network energy proportionality is possible though. A 2010 Google study [37] found that LAN power consumption might account for up to 50% of cluster power consumption due to server energy proportionality increases, but concluded that network power consumption could be reduced up to 85% using various techniques. Ways to reduce both local- and wide-area power consumption of networks are now summarized based on two surveys [66, 254].

Link-rate adaptation and network consolidation are two high-level strategies for reducing network power consumption. Placing network links in a sleep state and altering the rate of a network link or reducing its bandwidth based upon load are examples of these approaches. Energy-aware routing techniques have also been proposed. Networks are typically overprovisioned and possess a high level of redundancy, enabling energy consumption to be reduced by consolidating traffic onto a subset of links when network load is low.

Offloading computation to the network card or a proxy server can also reduce network-related power consumption. Simple messages such as a ping or address resolution request might be responded to without waking a host system and enhanced network cards could also handle more complex messages such as those seen in file-sharing applications. Further savings might be possible by having an intermediate node act as a proxy, enabling a large number of servers to remain in a sleep state.

Protocol redesign can also improve energy efficiency. Examples include file-sharing protocols incorporating host power state information and integration of sleep states into TCP.

2.3.7 Multiple Data Centres

Having multiple data centres provides additional opportunities to increase efficiency beyond those outlined earlier. Focuses of research include exploiting variations in power pricing between locations, making the best use of the renewable energy available at each location, and stabilization of the power grid.

Early work [202] explored using dynamic power price information to direct web requests to data centres, reducing costs while respecting latency requirements. Another early paper [154] capped non-renewable power use at each data centre during a time period, using this and the power price to determine where to route web requests. Control system engineering techniques like receding horizon control have been used to derive online algorithms for geographical load balancing with good worst-case performance [163]. Research has also addressed migrating virtual machines executing HPC workloads [155] based on power prices in different regions and, since the research in Chapter 4 was first published, also begun to incorporate delay-tolerance into power-cost-aware scheduling algorithms [251].

A more aggressive approach to incorporating renewable energy was proposed in 2009: a follow-the-renewables system architecture [95] in which workload is executed in data centres only when locally generated renewable energy is available. The GreenStar Network [112] implemented a follow-the-renewables testbed and also investigated dynamically allocating lightpaths between data centres based on renewable energy availability [250]. Though I was involved in the GreenStar Network project, I was not involved in implementation of the follow-the-renewables testbed nor the dynamic lightpath allocation aspects of the project.

Section 2.1 outlined the challenges faced by electrical grids in integrating renewables, and some research has addressed the use of workload distribution among data centres to ease integration. A 2010 paper by Mohsenian-Rad *et al.* [175] examined how routing requests to different data centres based on power grid conditions could reduce transmission system overload.

2.4 Summary

This chapter has provided an overview of the changing power grid and addressed server and data centre efficiency metrics as well as how to operate computing infrastructure efficiently.

Power grids are being restructured. Large scale consumers of electricity like data centres are increasingly able to purchase power at dynamic rates, creating opportunity to reduce power bills. Additional renewable energy is being integrated to reduce power generation's high carbon footprint. This task remains challenging due to variation in output of renewable energy sources like wind and solar. This increases the need for ancillary services, mechanisms which address short-term variation in power grids and respond to the failure of power plants and transmission lines. Data centres may be able to reduce capacity charges by connecting to the power grid in strategic locations. With power grids requiring more energy storage, data centres may also benefit from colocating with energy storage facilities or even further integration.

Computing is energy-intensive, with its power consumption growing rapidly. Both acquiring efficient hardware and operating it efficiently are important. This chapter provided an overview of efficiency standards and benchmarks for servers, enabling efficient equipment to be chosen. Means of characterizing the overhead of data centres were also outlined since inefficient facilities may consume more energy as overhead than is directed to the servers inside. The chapter's focus then shifted to outline efficient ways to operate computing infrastructure. The discussion began by addressing a single server, expanded to consider clusters of homogeneous or heterogeneous servers in a single facility, and then later expanded again to consider coordinating multiple data centres for maximum effect. Along the way the efficiency of related infrastructure was also discussed, highlighting how to efficiently operate data storage equipment, power distribution and UPS equipment, cooling systems, and network hardware.

Much work remains to be done to cut the costs and carbon emissions of computing and there are various avenues for exploiting synergy between power grids and data centres. The following chapters explore these issues further, evaluating opportunities to reduce both the financial and environmental costs of data centres.

Chapter 3

A GreenHouse Gas Reduction Protocol and Its Application

> A baseline scenario . . . represents what would occur in the absence of the project . . . The project proponent shall select or establish, justify and apply criteria and procedures for demonstrating that the project results in GHG emissions or removal enhancements that are additional to what would occur in the baseline scenario.
>
> The ISO-14064-2 standard [129]

Organizations are increasingly interested in their environmental footprint and how to lower it. Organizations may face government mandates to release environmental footprint information and this is also of increasing interest to investors and customers. A survey of 93 major investors controlling $12 trillion in assets [32] found that 57% had conducted investigations into the climate risks of their investments and 26% had altered investment strategies as a result. Potential customers often express environmental concerns, and the increased awareness brought by organizations like Greenpeace [111] to computing's environmental costs has led corporations to take measures to reduce their impact. Being able to show credible improvements from projects may not only enable organizations to avoid charges of "green-washing" but also enable them to acquire funding in carbon markets to reduce the cost of projects pursued.

Organizations may be able to reduce their environmental impact by implementing emission reductions projects targeting their computing infrastructure. It is important to measure the effectiveness of any project or potential project. However it is difficult to do so due to the rapid pace of change in the computing sector combined with how standards like ISO-14064-2 [129] calculate emission reductions.

This chapter is organized as follows. First, Section 3.1 overviews the ISO-14064-2 standard used for assessing emission reductions, describing the difficulties imposed by the rapid rate of change in the computer sector. The next sections outline a greenhouse gas (GHG) emission reduction protocol developed to assess projects in the computing sector [34]. Section 3.2 provides an overview of the protocol and the types of projects it addresses. Ensuring that projects meet the ISO-14064-2 requirement that they exceed business-as-usual standards is then discussed in Section 3.3. The creation of baseline scenarios against which projects are to be evaluated is described in Section 3.4, outlining how to define data centre characteristics in such a scenario as well as how to estimate baseline server performance. Section 3.5 then outlines how initial emission reduction estimates can be computed, how the project should be monitored, and how monitoring data may affect the original estimates. An emission reduction project was performed to evaluate the protocol and this project is described in Section 3.6. The chapter closes in Section 3.7 with summarizing remarks. I was closely involved throughout the development of the GHG reduction protocol described herein as well as in the evaluation of the test scenario which implemented the protocol.

3.1 ISO-14064-2 Overview

The ISO-14064-2 standard [129] is widely used and accepted for providing guidelines as to how greenhouse gas reductions resulting from a project can be quantified and reported as well as how monitoring and measuring of projects should take place. Emission reductions projects typically follow ISO-14064-2 compliant protocols applying the standard's principles to a specific area rather than directly employing the standard.

ISO-14062-2 principles are as follows: relevance, completeness, consistency, accuracy, transparency, and conservativeness. Relevant greenhouse gas sources, sinks, and reservoirs (SSRs) should be identified, with the principle of completeness to ensure that all appropriate entities are considered. SSRs should be treated in a consistent manner, enabling comparison, and bias and uncertainty should be minimized in pursuit of accuracy. Transparency should be ensured through the disclosure of adequate information for project evaluation and later project auditing. Estimates of reduced emissions should be conservative, tending to underestimate emission reductions when ambiguity exists.

Due to the computing sector's rapid pace of change, the ISO-14064-2 principle of conservativeness makes assessing computing projects a particular challenge. To maintain this principle, projects being implemented must surpass business-as-usual standards, a concept generally referred to as additionality. New computing equipment generally provides substantially greater performance per watt than its predecessor but may not necessarily satisfy additionality since this equipment generally has a much shorter expected lifespan than typical targets of emissions reduction projects like power plants whose lifespans are often measured in decades. For additionality to hold it must be demonstrated that the project surpasses business-as-usual expectations.

Whether business-as-usual standards are exceeded is assessed according to the ISO-14064-2 standards by evaluating projects against baseline scenarios representing business-as-usual action along with a barriers analysis demonstrating added cost or other obstacles faced by the project. Rather than only considering operational power consumption, as does the carbon usage effectiveness metric for data centres [50], emissions throughout the full lifecycle of projects and baseline scenarios are to be analyzed and compared. This involves consideration of factors including, but not limited to, data centre construction, manufacture of servers and supporting infrastructure, shipment and installation, operational power consumption, and end of life disposal.

The emission reductions from a project are, according to the ISO-14064-2 standard, the difference between the full lifecycle emissions of the project and baseline scenario being considered. These figures are to be verified at the end of the project, and may be revised based upon this assessment.

3.2 Protocol Overview

As part of the GreenStar Network project, an ISO-14064-2-compliant GHG emission reduction protocol was developed for the ICT sector. Alongside the Canadian Standards Association and the University of Calgary Grid Research Centre, a variety of other industry, academic, and government organizations participated in the Carbon Assessment Working Group responsible for the protocol's development. A technical advisory group of experts from the University of Ottawa, The Delphi Group, the University of California San Diego, Ericsson, and Research in Motion offered feedback on the protocol.

The protocol separated GHG reduction projects into two categories. Type 1 projects involve improvements to data centres, whereas Type 2 projects focus on changes to the equipment and processes involved directly in the actual computing services. Type 1 projects consisted of activities making data centres more efficient, typically trying to improve PUE scores but potentially also involving the integration of lower carbon energy sources into the data centre's power mix. A larger variety of Type 2 projects were identified and the descriptions of these use cases, as specified in the protocol, can be found in Table 3.1.

A full lifecycle analysis of a project's emissions must be conducted. This must then be compared against the full lifecycle emissions of one or more baseline scenarios

Table 3.1: Type 2 use cases.

	Use Case	Description
1.	Colocation	The proponent's physical machines are moved from the original facility to a lower carbon facility.
2.	Leased machines	The proponent's original collection of physical machines at the original facility is replaced by a leased collection of physical machines at a lower carbon facility.
3.	Green cloud	The proponent's collection of already-virtualized machines is moved to ≥ 1 physical machines at a lower carbon facility.
4.	Special purpose hardware	The proponent's original collection of physical machines at the original facility is replaced by a collection of high-efficiency physical machines, either at the original or a new facility.
5.	Virtualization and/or consolidation	The proponent's original collection of physical machines at the original facility is replaced by a collection of virtual machines, either at the original or a new facility.
6.	Network improvements	The project proponent improves and installs new network capacity which is able to transport more information using less power while meeting all of the same customer expectations.

to determine emission reductions under ISO-14064-2 guidelines. The construction procedure for baseline scenarios will be described in Section 3.4.

In order to satisfy ISO-14064-2 requirements the project must be determined to satisfy additionality criteria, as discussed in the next section.

3.3 Demonstrating Additionality

ISO-14064-2 principles require projects to show that they will result in lower emissions than would be expected under a business-as-usual approach — a threshold known as additionality. To evaluate this several additionality tests were defined as outlined in Figure 3.1.

Figure 3.1: Types of additionality tests.

The first of these is the legal and regulatory additionality test. Projects undertaken to comply with legal or regulatory requirements are not considered additional unless exceeding these requirements.

Projects involving Type 1 activities face a common-practice additionality test. Increases to the carbon efficiency of energy supplied to the data centre are to be evaluated against a performance standard. The emission intensity of the power consumed must not exceed the carbon intensity of the local electrical grid nor that of high-efficiency combined-cycle natural gas generation. Emission intensity factors are available from Environment Canada for use in Canadian projects.

A barrier / risk analysis test to ensure that the project bears greater risks than alternate scenarios is a manner in which all proposed projects should be tested. This should address all risks and barriers even if some are only applicable to one of the project and baseline scenarios.

Table 3.2 shows a categorization of risks for all aspects of the project and baseline scenario as specified in a risk analysis toolkit included with the protocol. This toolkit should be used to evaluate the project and baseline scenarios. To determine the overall risk for the project and baseline scenarios, the various risk categories should each be assigned a weighting, with the weights of the categories together summing to 100%. Justification needs to be supplied for each such weight. Once this has been completed, the same process can then be applied to the risk factors within each category. The risk factors should also be assigned a weight, with the weights within each of the risk categories summing to 100% with justification provided. The risk category and risk factor weightings are to be the same in both the project and the baseline scenario.

After creating risk category and risk factor weightings, values should be assigned separately to each risk factor in the project and baseline scenarios. The values assigned should range from 0–3 — no risk to high risk — as outlined in Table 3.3, with

Table 3.2: Categories of risk factors.

Risk Category	**Risk Factor**	**Description**
Operational	Planning	Risks associated with the initial planning of the project.
	Deployment	Risks associated with moving to the new site. This includes software-related challenges such as licensing issues or difficulties configuring software for a particular environment. It may also include other factors such as data loss or equipment failure occurring as a result of transporting equipment to a new data centre.
	Management	Risks pertaining to the availability or lack thereof of equipment in the new site to administrative and technical personnel.
Security	Physical Security	Risks associated with the physical location of the new data centre and physical access to it. This includes environmental risks such as the likelihood of flooding or earthquakes as well as the possibility of unauthorized physical access to equipment there.
	Logical security	Risks associated with electronic means of gaining unauthorized access to servers through compromised software or protocols.
Performance	Quality of service	Risks of quality of service problems arising in the baseline or project scenarios.
	"Project" monitoring	Risks associated with the ability to gain sufficient monitoring data to implement the protocol.
Organizational	Organization policies	Risks related to needed changes in organizational policies, such as efforts to keep computing in-house.
	Personnel Resistance	Risks due to opposition to the project from management, a lack of adequate staff, or staff resistance to the project.

Table 3.3: Risk level descriptions.

Risk level	**Description**	**Value**
High	Unwanted changes in the factor would have significant ramifications for the project proponent, including the potential for significant financial loss or disruption to the operation of the entity.	3
Moderate	Unwanted changes in the factor would have moderate ramifications for the project proponent, including the potential for moderate financial loss or disruption to the operation of the entity.	2
Low	Unwanted changes in the factor would have minor ramifications for the project proponent, including the potential for minor financial loss or disruption to the operation of the entity.	1
None	Unwanted changes in the factor would have no ramifications for the project proponent, including no potential for financial loss or disruption to the operation of the entity.	0

justification supplied for each. Risk scores can then be computed for the project and baseline scenarios. Each value should be multiplied by its risk factor weighting as well as its risk category weighting. These figures should then be added separately for the project and baseline scenarios to produce the final risk score for each. To be additional the project risk should be higher than the baseline risk.

Financial additionality testing is only needed if the project fails both the common practice and barrier / risk additionality tests outlined above. This testing would need to demonstrate that the project is not the most cost-effective approach. With energy accounting for a large portion of computing costs, projects may pay for themselves through energy savings and thus not be financially additional. Additional revenue from government programs such as energy efficiency incentives or feed-in tariffs paying premium rates for power from renewable sources must be considered as part of financial additionality testing. Various mechanisms can be used to assess financial additionality. Simple cost analysis is the most basic and applies where the project would come at a cost but the sale of carbon credits would be the only added revenue. Financial additionality can also be demonstrated through benchmark analysis, the comparison of the rate of return in the project to sector-specific typical investment rates-of-return. Finally, a comparative analysis can be used which compares the project cost to the cost of baseline scenarios.

3.4 Baseline Scenario Construction

To comply with ISO-14064-2, a project must be evaluated against a scenario representing a business-as-usual approach in line with industry norms that would be followed in lieu of the project. Moore's Law and related trends coupled with the short service life of computing equipment relative to other industries means that merely consuming less power than equipment being replaced does not satisfy additionality, but that comparing the project against a credible alternative is necessary.

Baseline scenarios may involve the hypothetical purchase and installation of equipment or the renovation or construction of data centres. Such scenarios must be both conservative in nature and also conservative in emissions estimates. Baseline scenario emissions should be underestimated where ambiguity exists to avoid unduly crediting the project scenario with emission reductions.

This section outlines how to construct baseline scenarios for both data centres and computing equipment. Section 3.4.1 focuses on data centres and is relevant to all of the project types outlined in Section 3.2. Principles for characterizing the baseline power consumption of computing equipment are outlined in Section 3.4.2. This section only applies to some protocol project types, with (Type 1) projects involving improvements to data centres themselves assuming server power consumption to be equal in the project and baseline scenarios. Similar assumptions apply to Type 2 projects involving Use Cases 1–3 which involve improvements in efficiency by relocating to, leasing equipment in, or using cloud services hosted in green data centres.

3.4.1 Data Centres

Data centre efficiency varies based on many factors and can be evaluated using a range of metrics as described in Section 2.2.2. The protocol provides guidance as to how to establish data centre efficiency for the baseline scenario, basing this on both Power Usage Effectiveness (PUE) and energy source emission intensity.

Baseline scenario PUE was generally set as the PUE of the pre-project data centre with some exceptions. In legacy data centres it is often difficult to accurately calculate PUE, particularly when the data centre is part of a multi-use facility where office space and other uses take up much of the building. The original PUE specification [16] addressed only dedicated data centre facilities, and the second version [26] required that data centres have dedicated cooling and electrical systems though this may not be the case in practice. PUE in such instances can be calculated by metering the computing equipment itself and separately metering all input energy at the data centre's boundaries, using these figures to determine a power-usage effectiveness reading. PUE can also be estimated by metering the computing equipment and the building as a whole, and then subtracting from building power consumption power use estimates for non-data-centre areas. A third approach is to conservatively estimate PUE using a performance standard, with the 75^{th} percentile reported in the Energy Star for Data Centres ranking system being used here [85].

A year of measurement data should be available from energy meters used to calculate PUE, with those pursuing a project otherwise required to justify its unavailability. Steps must be taken to ensure that conservativeness is maintained when less data is available, accounting for seasonal variation and the data centre's occupancy rate.

The second edition of the PUE specification divides PUE into various subtypes, depending on where it is measured. Categories 1-3 are acceptable, although as available metering increases the somewhat more rigorous Categories 2 and 3 which measure power consumption nearer to the load should be adopted. Care should be taken to ensure that the principle of conservativeness applies when estimating PUE. In the baseline scenario this means acting with a tendency to underestimate emissions and in the project scenario it means acting with a tendency to overestimate emissions.

To translate from PUE to the language of "emissions" requires information about the carbon intensity of energy used. For the purposes of the baseline scenario, rather than simply calculating emissions based on the emission intensity of the energy sources used as is done in calculating the Carbon Usage Effectiveness (CUE) metric, adjustments to emission intensity may be needed to ensure conservativeness is maintained. The emission intensity assumed in the baseline scenario should be the lower of the local power grid's intensity and the emission intensity reported for high-efficiency natural gas generation.

3.4.2 Computing Equipment

For many projects the power consumed by the computing equipment will be assumed equal in the project and baseline scenarios. Projects implementing Type 2 Use Cases 4–6 require further analysis however. These three use cases — the use of uncommonly high-efficiency computing equipment, the replacement of physical machines by virtual machines, and power-saving network improvements — aim to reduce the power consumption of computing equipment beyond business-as-usual expectations.

Two principles guide the estimation of relative power consumption in project and baseline scenarios. First, ISO-14064-2 requires that the project and baseline scenarios be functionally equivalent, handling the same workload and meeting the same quality of service requirements. Second, as servers have become significantly more energy proportional [237], understanding the properties of the workload they execute and how this affects power consumption is key. As Section 2.2.1 noted, average utilizations vary significantly, ranging between 11–50% in a set of web server traces and 11-90%

in available high performance computing traces.

Initial estimates of the emission reductions of a project inherently involve estimates of future workload and may also involve equipment which the project proponent may neither own nor have access to during the project period. Specific hardware should be identified for the baseline scenario, but this represents only purchase patterns in a business-as-usual scenario. An estimate can be used when evaluating the potential of a project, but conservativeness must still apply. At the end of a project, as part of the auditing phase, the actual workload as recorded in logs should be used to reevaluate baseline power consumption estimates.

Mechanisms are becoming available to assess the comparative power consumption of different systems executing different workloads. Section 2.2.1's outline of a variety of different benchmarks is also useful for determining what to measure and what monitoring should be done during the project period. The SPECpower_ssj2008 benchmark [223], SAP's power benchmarks [17], and TPC-Energy [18] provide guidance as to how to measure a computer system's power consumption under different levels of load. For HPC clusters, how effectively the LINPACK benchmark can be used to estimate the power consumption for other workloads has also been evaluated [138], and a Google study [91] found CPU utilization alone provided a reasonable approximation of total system power consumption for particular workloads.

Even if power consumption executing a benchmark is measurable, using this to estimate power consumption over a project's lifespan remains challenging. The application being executed might not be the same as that executed during the benchmark test. The impact of workload distribution among a cluster of servers and the ability to power down idle machines need be accounted for to conservatively estimate baseline power consumption. Such obstacles can be addressed while remaining conservative as will be outlined later in the section.

Benchmarks like SPECpower_ssj2008 can be used to estimate the power consumption of servers under different levels of workload intensity for particular workloads as outlined above. Executing a series of trial runs and recording system metrics and power consumption for both such a benchmark and the actual workload to be executed enables comparison of their power consumption to be made.

Conservative estimates of baseline power consumption should err toward underestimation where ambiguity exists. The more monitoring information available, the more likely it is that a power estimate can closely approximate the actual power consumption were the baseline system to execute the workload recorded over the course of the project. The specific logs to be collected will vary based upon the details and characteristics of the application to be executed, e.g., including application logs in the case of an application like a web server or scheduler logs in the case of a high-performance or high-throughput computing cluster.

Baseline and project power consumption can be compared by computing estimated baseline power consumption over intervals of time and then comparing this against the power consumption reported for the project machines.

For some potential project system configurations, such as sharing an enclosure with equipment belonging to other projects, portions of a project's power consumption may be measurable only in aggregate. The same applies to environments where servers are shared by virtual machines belonging to multiple customers, as is often the case in cloud environments. It would be conservative when developing the baseline scenario to assume that no shared component power use is attributable to the project but this limits potential emission reductions since shared infrastructure can account for a large fraction of total power consumption. Another approach would be to distribute the power consumption of shared components evenly among their users based on the percent of the shared environment they account for. For example, if 25% of the servers in an enclosure would likely be used by the project, and all enclosures are fully populated, approximately 25% of the shared infrastructure could be estimated to be attributable to the project's servers.

Workload distribution affects power consumption and must be accounted for when estimating baseline power use. Figures 3.2 and 3.3 demonstrate how workload distribution policies can influence power consumption using SPECpower_ssj2008 results for two machines [221, 222]. Figure 3.2 plots how the energy efficiency of these machines (in ssj_ops/W) changes with server load (in % of maximum system capacity). The figure shows both a typical case in which efficiency continues to increase as load increases and a second case showing a system that exhibits greatest energy efficiency under partial load.

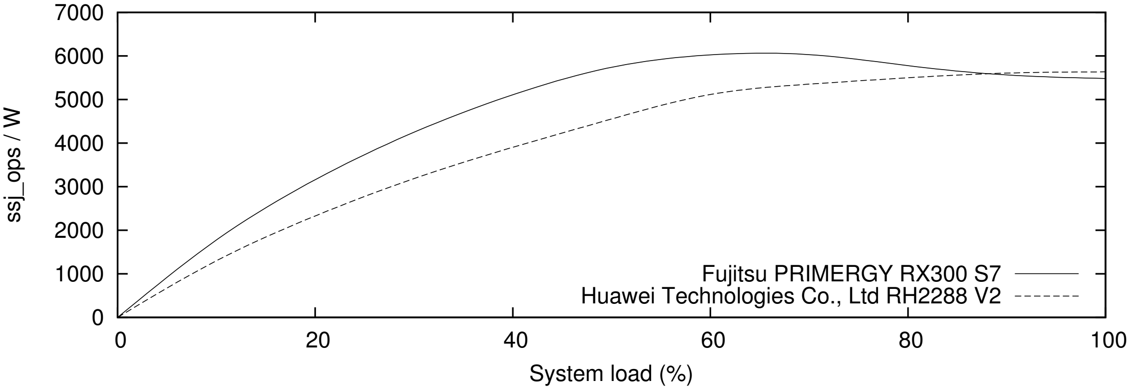


Figure 3.2: Power-performance ratio at different levels of workload intensity.

Depending on the servers in the baseline scenario, greatest efficiency might be found by evenly distributing the workload, but in other cases operating servers at either full load or at near idle might be more efficient. The incremental power consumption required to accommodate an increase in workload may also vary substantially, with Figure 3.3 comparing the incremental power consumption of the same servers as before when load increases.

In the baseline scenario workload should be assumed to be distributed in such a manner that the systems operate relatively energy efficiently. This may involve

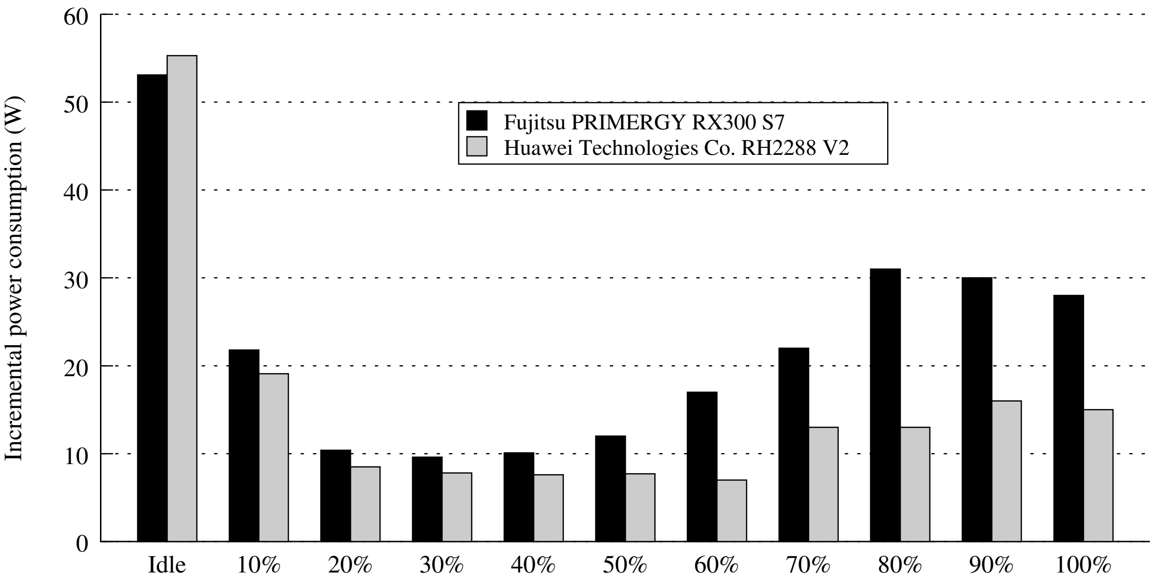


Figure 3.3: Incremental power consumption as load increases in SPECpower_ssj2008.

differences in the amount of workload assigned to each server and the power states of various components. It may involve shutting down servers at times when not needed to meet present demand. Additional servers beyond those needed to process the current workload might also be required to be in an active state as significant variations in demand are experienced with many workloads. Some types of systems such as HPC clusters may have greater flexibility to power down idle equipment since beginning task execution within a few milliseconds is not generally as important a requirement as it would be in applications like typical web servers. Workload characteristics and quality of service requirements should be used to estimate the effectiveness of strategies for powering down idle servers. The amortized cost of shutting down and restarting equipment must also be considered, at least for the project scenario.

Once these factors have been evaluated, a table like Table 3.4 can be prepared. This table shows how estimates can be constructed for both baseline and project scenarios in advance of a project's start and provide an estimate of likely power savings. It requires that power consumption for a variety of different levels of system

Table 3.4: An estimate of average power consumption of baseline and project systems.

Bucket #	**Fraction of Time in Interval**	**Baseline Power Estimate (W)**	**Project Power Estimate (W)**
1	0.1	0	134
2	0.1	260	154
3	0.2	517	179
4	0.3	775	208
5	0.1	1033	311
6	0.05	1290	336
7	0.04	1547	394
8	0.03	1805	413
9	0.03	2063	439
10	0.05	2322	468
Weighted Average Power		824	240
Estimated Power Savings		824 - 240 = **584 W**	

activity be estimated, taking all the above factors into account for both the project and baseline scenarios. It should also be remembered that conservativeness in this situation means a tendency to underestimate baseline emissions and overestimate project emissions. The number of levels of system activity recorded will depend on the specific scenario in question. Estimated power of baseline and project systems at given utilizations can be combined with the estimated frequency that servers spend within a given utilization interval to compute weighted power consumption for each system as well as to estimate power savings.

As future workload is generally difficult to estimate, particularly in cloud computing environments that allow for rapid scaling, these figures will later need to be adjusted in post-project auditing to estimate the actual emission savings based on the recorded power consumption of the project equipment and the workload that was processed by it. It may also be necessary to further revise power estimates based on updated information about the likely performance of equipment used in the baseline scenario. Table 3.5 shows a revised baseline estimate for the project whose pre-project estimate was shown in Table 3.4. In this instance post-project calculations suggest

Table 3.5: Revised estimate of baseline power utilization.

Bucket #	**Fraction of Time in Interval**	**Baseline Power Estimate (W)**
1	0.146	0
2	0.089	260
3	0.205	517
4	0.240	775
5	0.140	1033
6	0.070	1290
7	0.040	1547
8	0.015	1805
9	0.020	2063
10	0.035	2322
Weighted Average Power		762 W

power consumption was likely lower than originally anticipated in the baseline scenario. This figure can then be compared against the recorded power consumption of the project platform to produce a final calculation of power savings.

In some situations direct metering of servers may be extremely difficult or available data may be inadequate, perhaps due to a provider's unwillingness to release aspects of power usage data or lack of equipment to collect it. Fortunately better metering is becoming available. For example, whereas it was challenging to measure power consumption for many devices without separate dedicated monitoring, a Raritan DPXR8-15 PDU used for experiments after the completion of the protocol test project was able to record power consumption of each outlet every three seconds. Means are also being developed to estimate the power consumption of virtual machines based on their activity [148]. Such mechanisms do not account for the power consumed by host machines at idle but this can be addressed by a policy, e.g., accounting VMs responsible for a fraction of idle power corresponding to the fraction of the machine's resources they were allocated.

A similar approach as used to construct the baseline and pre-project power estimates might also be applied to estimate actual power consumption of the project platform during the project period based on workload information. Due to the need to remain conservative, however, potential emission reductions are likely to be significantly underestimated in such a scenario.

3.5 Estimating, Monitoring, and Auditing Reductions

Determining an appropriate PUE for a baseline data centre and estimating the power consumption of servers is not the final stage in producing ISO-10464-2-compliant emission reduction figures. Server power consumption must be converted to total energy consumption by accounting for the length of the project and multiplying by PUE. The carbon intensity of the power used must also be factored in. Emissions from the baseline and project should be calculated separately and then compared to produce a final emission reductions estimate.

Equation 3.1 shows how emissions can be calculated for the project scenario, initially using estimates and then later the final figures:

$$Em(Pr) = P_{ICT}(Pr) \times PUE(Pr) \times EF(Pr) \tag{3.1}$$

In this equation $Em(Pr)$ stands for the emissions associated with the project scenario in kg CO_2e, $P_{ICT}(Pr)$ stands for the power usage of the project in GJ, $PUE(Pr)$ stands for the PUE calculated for the project facility, and $EF(Pr)$ stands for the emission factor of the energy sources used in the project scenario. Source of emission factors for the power grids of different regions as well as for different fuels and other energy sources are supplied in the protocol.

Similarly, Equations 3.2 and 3.3 show how to calculate emissions for the baseline scenario, initially using estimates and then later incorporating information from project monitoring. Equation 3.2 is for Type 1 projects as well as Type 2 projects implementing Use Cases 1–3 where the power consumption of the baseline computing equipment is assumed to be equal to that of the project scenario, whereas Equation 3.3 is for Type 2 Use Case 4–6 projects in which this assumption does not hold.

$$Em(Ba) = P_{ICT}(Pr) \times PUE(Ba) \times EF(Ba) \tag{3.2}$$

$$Em(Ba) = P_{ICT}(Ba) \times PUE(Ba) \times EF(Ba) \tag{3.3}$$

In these equations $Em(Ba)$ stands for the emissions associated with the baseline scenario in kg CO_2e, $P_{ICT}(Pr)$ and $P_{ICT}(Ba)$ stand, respectively, for the power usage of the project and baseline scenarios in GJ, $PUE(Ba)$ stands for the baseline PUE, and $EF(Ba)$ stands for the emission factor of the energy sources used.

Monitoring requirements may vary from project to project dependent on the type of application being run, the computing equipment used, and the power metering available. Power usage of the computing equipment must be measured directly or calculated from other recorded metrics where no energy metering is possible. When only aggregate measurements are available, a conservative fraction of power usage should be estimated. Based on the monitoring data collected, earlier calculations will be revisited at the project's end to reassess its emission reductions.

To use emission reduction figures outside an organization, an external auditor will likely need to audit the project and its documentation to ensure that the figures are reliable and well-supported. This process is typically needed for carbon markets to accept the emission reductions and also enables companies to shield themselves from charges of "greenwashing" corporate sustainability reports, of making the corporation seem more environmentally friendly than it actually is. Guidance to auditing emission reduction projects is provided in the ISO-14064-3 standard [130].

3.6 The GeoChronos Relocation Project

The GeoChronos web portal [100] facilitated communication in the earth observation science community. Using GeoChronos, scientists could interact, contribute to and

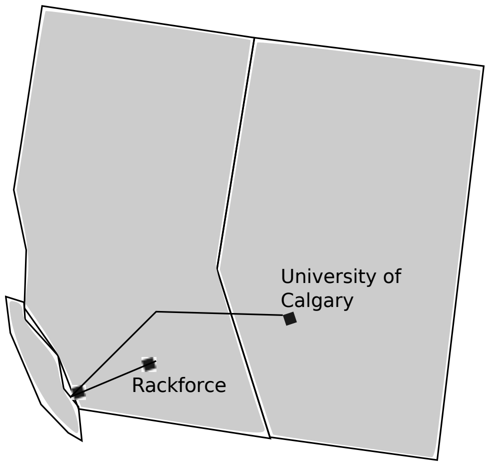


Figure 3.4: A map showing the location of the baseline facility in Calgary, AB, the RackForce project facility in Kelowna, and the network path that carried traffic between Calgary and Kelowna in the project.

use resources such as the site's spectral library, and remotely run applications. The production site ran using a modified version of the Elgg social networking engine [82], with resources to run applications allocated dynamically in a locally-developed cloud environment [165].

The GeoChronos platform shared a cluster of blade servers housed in a legacy data centre on the University of Calgary campus with other applications using the same cloud environment. Relocating GeoChronos from the University of Calgary to a location with a less-carbon-intensive power grid was a test project performed to evaluate the protocol outlined earlier in the chapter. GeoChronos was moved to the RackForce GigaCenter [203] in Kelowna, British Columbia, a location where free cooling was available most of the year and with much lower emission intensity power than Alberta [87].

Chronologically the project proceeded as follows. From mid-June to November 2010, configuration and testing of nodes and cloud infrastructure was performed at the RackForce facility. In December 2010 and January 2011, the GeoChronos development environment was migrated to RackForce and was followed in February 2011 by the production website. From then until September 2011 GeoChronos operated at the RackForce data centre, with monitoring data collected from July to September 2011. At the end of the project, in October 2011, the GeoChronos platform was migrated back to the University of Calgary.

The remainder of this section describes the analysis of the GeoChronos relocation project and how emissions reductions for it were calculated. First Section 3.6.1 outlines how the developed protocol was deemed applicable to the project. Section 3.6.2 then describes the full lifecycle analysis performed for the GeoChronos relocation project, identifying all sources of carbon emissions, all sinks acting to reduce carbon emissions, and all reservoirs acting as temporary storage of emissions. A baseline scenario must be developed against which to compare a project to maintain ISO-14064-2 compliance, and Section 3.6.3 describes how this was done for the GeoChronos project. Whether or not the project satisfied additionality requirements is discussed in Section 3.6.4, and Section 3.6.5 quantifies the project's emissions reductions. Project monitoring is discussed in Section 3.6.6 and the final results are presented in Section 3.6.7.

3.6.1 Project Type

The first step of applying the protocol to the GeoChronos relocation project was determining whether any of the project types outlined in Section 3.2 were applicable. The project did not involve improving the existing data centre which had housed the GeoChronos project, and it was thus not a Type 1 project. Further evaluation was then conducted to determine whether or not it fit any of the Type 2 project scenarios outlined in the protocol.

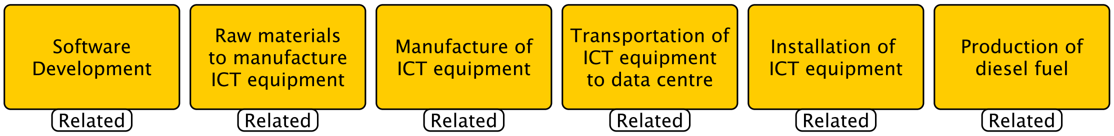
The GeoChronos project was executed in a cloud environment during the pre-project period, using both dynamic and statically allocated resources. The web server and several other services were statically allocated resources, and additional cloud resources were allocated as needed in response to user requests or to additional data being available for analysis. At RackForce the GeoChronos project was allocated space on a storage-area network (SAN), a dedicated file server for managing the storage, a virtual machine on RackForce's cloud, and two blade servers accounting for one quarter of the machines installed in a blade chassis. Since the relocated GeoChronos platform would primarily run in a private cloud environment, rather than a shared cloud as in Calgary, the relocation project was judged to fall in the Leased Machines use case of Type 2 projects (Use Case 2).

3.6.2 Project SSRs

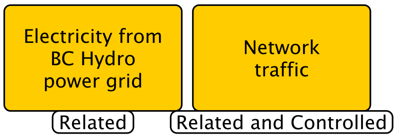
Thorough assessment of project emissions requires consideration of all potential impacts. The protocol requires that all sources, sinks, and reservoirs (SSRs) of carbon emissions be identified for the project scenario. Such an analysis requires tracing the environmental impact of the project upstream to examine the origins of resources as well as downstream. It also requires looking at impacts before, during, and after project activity. Each SSR's relation to the project should also be categorized as controlled by the project, related to the project, or affected by the project.

The SSRs identified for the GeoChronos relocation project can be seen in Figure 3.5. For this analysis the boundary of the project was determined to be the RackForce data centre. The diagram categorizes SSRs as follows: upstream SSRs before the project operation, upstream SSRs during the project operation, onsite SSRs during the project operation, and downstream SSRs following project termination. Below each SSR a label indicates its relation to the project: controlled, related, or affected.

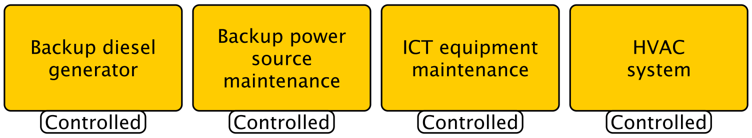
A number of upstream SSRs with impacts prior to the project's start were identified. Software development was a related SSR, representing the development activity required to enable the GeoChronos application to run on the new platform. The resource extraction processes producing the raw materials used in the manufacture of



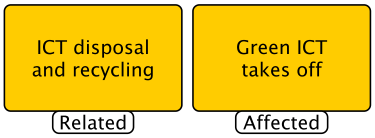
(a) Upstream SSRs before project operation



(b) Upstream SSRs during project operation



(c) Onsite SSRs during project operation.



(d) Downstream SSRs after project termination

Figure 3.5: Carbon sources, sinks, and reservoirs in the project scenario.

computing equipment were also identified, along with the impact of manufacturing itself. Transportation is required to convey equipment to the data centre where the equipment then also required installation. Diesel fuel production for backup generators was the last of the SSRs of this type identified. Each was categorized as a related SSR.

Upstream SSRs during project operation were also identified. These included the electricity drawn from the BC Hydro power grid and the impact of network traffic. Consuming electricity was identified as a related SSR, but network traffic was classified as being in part related and in part controlled by the project. This is due to the project's network configuration as traffic between the GeoChronos platform and the Internet was routed via a dedicated link to the University of Calgary. As the project boundary for this lifecycle analysis was set as the RackForce data centre, emissions from this link and GeoChronos' communication beyond it were classified as upstream usage rather than being divided between upstream and onsite usage as they would have had the RackForce-Calgary link considered inside the project boundary.

There were several onsite SSRs during the project's operation. These included diesel backup generator operation as well as any emissions resulting from generator maintenance. Computing equipment maintenance might also impact emissions. Emissions associated with the HVAC system include not only CO_2 but also potential leakage of any refrigerant used. Each onsite SSR was categorized as being controlled.

Finally, two downstream SSRs following the project's termination were identified. The disposal and demanufacturing of project equipment was identified as a related SSR. The success of the project could spur on Green IT, leading to other similar low carbon computing projects and this was thus identified as an affected SSR.

3.6.3 Baseline Scenario Development and SSRs

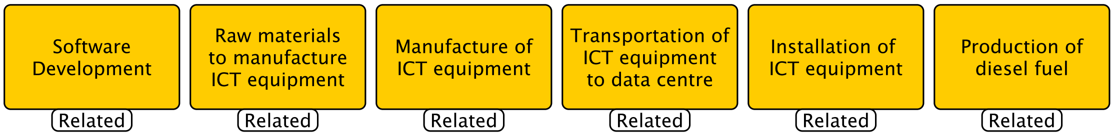
To assess the GeoChronos relocation project's impact a baseline scenario needed to be developed as a basis for comparison. Baseline scenarios represent likely activities should a project not occur and should be functionally equivalent to the project scenario, providing the same service with equivalent quality of service. As in the project scenario, SSRs required identification. This section describes the creation of a baseline scenario and the SSRs identified.

Standard baseline scenarios for some project scenarios were supplied in the protocol. The standard baseline scenario for the use case that the GeoChronos relocation project implemented was as follows:

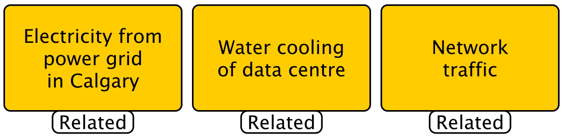
> The standard baseline comprises substitution of the original physical machines with up-to-date commonly available physical machines at the baseline facility, using the project's assignment of software packages if possible.

Without the relocation project it was assumed that the University of Calgary data centre which had been housing GeoChronos would continue to do so. The purchase of new servers to update the aging servers upon which GeoChronos was run was assumed.

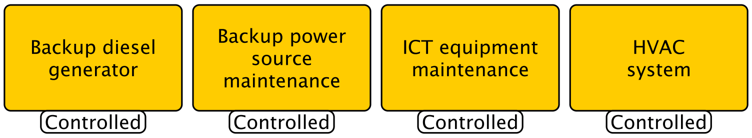
Figure 3.6 outlines SSRs in the baseline scenario, assuming the project boundary to be the University of Calgary data centre. Most identified SSRs also appeared among the SSRs identified for the project scenario. There were some differences in the upstream SSRs during the project's operation though. Electrical energy came from the Alberta power grid rather than BC Hydro. A cooling loop installed at the University of Calgary, which uses Bow River water to aid in cooling was an extra related source not found in the project scenario. As there was no dedicated network capacity carrying traffic to and from RackForce, baseline network traffic was classified as a related SSR rather than a related and controlled one. The SSR representing broader adoption of the protocol outlined in this chapter did not appear among the baseline SSRs.



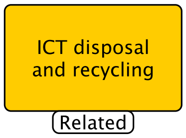
(a) Upstream SSRs before project operation



(b) Upstream SSRs during project operation



(c) Onsite SSRs during project operation



(d) Downstream SSRs after project termination

Figure 3.6: Carbon sources, sinks, and reservoirs in the baseline scenario.

3.6.4 Demonstrating Additionality

As outlined in Section 3.3, projects face several tests to verify that they exceed business-as-usual standards. The government funding received by the GeoChronos relocation project for environmental improvements immediately excluded it from being claimed as additional. However, additionality tests were nonetheless performed

Table 3.6: Risk category and factor weighting.

Category	**Category Weighting**	**Risk Factor**	**Factor Weighting**	**Project Value**	**Baseline Value**
Operational	60%	Planning	20%	2	1
		Deployment	60%	3	2
		Management	20%	1	0
Security	5%	Physical Security	50%	1	1
		Logical Security	50%	1	1
Performance	25%	Quality of Service	50%	2	1
		“Project” monitoring	50%	2	1
Organizational	10%	Organizational policies	0%	0	0
		Personnel resistance	100%	2	1

to evaluate their practicality.

Legal and regulatory requirements for British Columbia and Alberta were reviewed to ensure no government policy, regulations or industry standards required the project be executed. The project also did not receive revenue from feed-in tariffs or energy efficiency grants, amounts that would otherwise need to be accounted for in determining financial additionality.

As a Type 2 project, common practice additionality testing of the emission intensity of energy used at the project facility and the efficiency of the data centre itself was not required. However, a barrier analysis was mandated to evaluate risk in the project and baseline scenarios using the guidelines specified in Section 3.3. Table 3.6 which summarizes the barrier and risk analysis conducted for the GeoChronos relocation project will now be explained more thoroughly. Though in practice risk category and risk factor weightings should be assigned separately from the process of fixing the value of each, they are presented together here for brevity.

Operational risk was assigned a risk category weighting of 60% as the largest type of risk seen by the project. Deploying a service in a new environment may pose many challenges including device driver support, and software may not support the new platform. Licensing issues can also create problems, particularly if moving to a new site. Planning was assigned a risk factor weighting of 20% within the operational risk category. Poor planning can lead to inadequate quality of service, but in this case the system requirements were relatively well understood. The deployment factor was assigned a weighting of 60%, due to concerns about software licensing and the difficulty of moving software and data to new platforms. Operational management risk was assigned a value of 20% as services would continue to be maintained by the Grid Research Centre, with the data centre's support team to handle only issues that could not be addressed remotely. The planning risk was assigned a risk value of 2 in the project and 1 in the baseline due to having easier access to additional nodes at the University of Calgary. A risk factor of 3 was assigned to deployment risk in the project but 2 in the baseline scenario. This is due to moving to new platforms in each scenario but the added challenge in the project scenario of moving to a new location and the corresponding impact on issues such as software licensing. Management risk in the project and baseline were assigned the values 1 and 0, respectively, due to the complications associated with the dedicated network link from Calgary to RackForce over which all GeoChronos traffic was routed.

Security risk for GeoChronos was relatively low and thus assigned a category weighting of 5%. The vast majority of data utilized by GeoChronos was assembled from public domain sources, and an offsite backup copy was maintained. Little confidential information was at risk if security was compromised. Authentication tokens might be obtained by malicious third parties, but even with this information comparatively little damage was likely due again to most data being in the public domain. Security risk was split evenly between the physical and logical security categories. Compromise of the data centre was possible but standard precautions were taken to avoid these including monitored security cameras and requiring access cards for facility access. Malicious actions by data centre staff were also possible but unlikely.

Standard security measures were taken to authenticate users accessing the system and the virtual machine managers and shared storage infrastructure used by the project are widely used in practice and their security thus validated. A risk value of 1 was assigned for both logical and physical security in the project and baseline scenarios.

The performance risk was moderate. Interactive applications were being executed, necessitating quick access to servers and the use of shared equipment might compromise quality of service. Project monitoring risk was also moderate due to the difficulty in obtaining monitoring data from various systems but the ability to replace missing data by applying conservative assumptions to construct estimates from available data. GeoChronos's quality of service and the project monitoring were estimated as facing equal risk. When services are moved to new systems, quality of service may not be easy to predict. It is difficult to collect information to monitor the project, as using shared hardware made it difficult to monitor the precise power consumption. Quality of service was assigned a risk value of 2 in the project scenario and 1 in the baseline scenario due to the added latency and additional point of failure from routing all traffic to Kelowna on a dedicated, non-redundant link.

Organization risk was assigned a risk category weight of 10%. No organizational policy forbade the use of third parties to provide services so organization policies were not a risk and thus assigned a factor weighting of 0%. The organizational risk was concentrated in personnel resistance, as those researchers for whom GeoChronos was operated could oppose movement to a new facility or new computing equipment. The Grid Research Centre is experienced in working with remote equipment. Personnel resistance was assigned a risk value of 2 in the project scenario and 1 in the baseline scenario due to the added complexity of remote equipment in the project scenario.

Using the information summarized in Table 3.6 it was possible to calculate risk levels for the baseline and project and evaluate their relative risk. The risk value for the project was 2.19 whereas the risk value for the baseline was 1.24, showing the project to be of greater risk than the baseline scenario.

Financial additionality tests were also performed to demonstrate that the project was additional to the baseline scenario. As the Grid Research Centre does not earn direct revenues from GeoChronos's users a simple cost analysis could be conducted to show that the project scenario exhibited higher cost than seen otherwise.

Expenses in the baseline and project scenarios were outlined and supporting documentation gathered. Costs were tabulated for both scenarios, separating expenses into acquisition and procurement, operations and maintenance, and end of life management.

Acquisitions and procurement costs covered expenses such as servers, storage equipment, UPSes, PDUs, racks, and network hardware. Costs were also assessed for software packages used by GeoChronos and the modifications necessary in order to use them. Setup of each scenario requires labour expenses, equipment for use in the installation process, shipping costs and setup costs for project monitoring. Administrative expenses for planning the project and travel related to equipment setup were also included.

Operations and maintenance costs included training and support as well as project monitoring. The cost of any needed upgrades and equipment repairs or maintenance while in operation was also assessed, and the dedicated link between RackForce and the University of Calgary included among the expenses. Energy related costs assessed included power costs, fuel for the backup generators, and costs related to coolant. Administrative costs and any project-necessitated travel were also assessed.

End of life management was the third category of costs assessed, covering decommissioning steps such as the sanitization of hard drives. Transportation costs to bring equipment to disposal facilities were assessed, along with recycling-related fees and potential revenue from the sale of old equipment. Finally administrative expenses were added, including any necessary travel.

Whether each expense category was relevant to the project or baseline scenarios required determination and, if so, the related expenses needed to be assessed. Service fees may cover expenses in several categories as, e.g., no separate payment to RackForce was required for power costs or backup generator diesel fuel in the project scenario.

3.6.5 Quantifying Emission Reductions

To accurately assess the emission reductions from a project requires that emissions both in it and the baseline scenario be assessed. This section describes how emission savings were calculated.

An important step in determining emission reductions is the selection of relevant SSRs for quantification or estimation. The SSRs for the project and baseline scenarios were outlined in Sections 3.6.2 and 3.6.3 respectively. Whether or not SSRs were relevant was assessed based on the following principles:

1. If an SSR did not result in emissions directly it was classified as relevant only if it was relevant to determining emissions from another SSR.

2. If an SSR was similar in the project and the baseline, it would be considered irrelevant as changes in its emissions would not impact the project's emission reductions.

3. If an SSR was similar in the project and baseline but resulted in greater emissions in the baseline scenario, it could safely be dismissed though this would limit the emission reductions with which the project could be credited.

SSRs could be spared from quantification if they were deemed minor and if quantification would be too difficult.

Following classification each relevant SSR then needed to be separately assessed in the baseline and project scenarios. The following SSRs were classified as relevant. The power grid in the baseline and project scenarios was classified relevant, with the project data centre powered primarily by hydro and the baseline data centre powered primarily by coal. Water cooling of the baseline data centre was also classified as relevant, but not applicable in the project scenario. Network traffic was classified as a partially relevant SSR in the project scenario due to the dedicated network link between the University of Calgary and RackForce. Backup generator fuel use was also classified as relevant.

Baseline Emission Calculations

Emissions in the baseline scenario were calculated based on the following formula, earlier presented as Equation 3.2:

$$Em(Ba) = P_{ICT}(Pr) \times PUE(Ba) \times EF(Ba) \tag{3.4}$$

$Em(Ba)$ represents the emissions of the baseline scenario in kg CO_2e, $P_{ICT}(Pr)$ represents project power consumption in GJ, $PUE(Ba)$ represents the baseline data centre PUE, and $EF(Ba)$ represents the emission factor of energy consumed. $P_{ICT}(Pr)$ was used since baseline computing power consumption is assumed equal to project computing power consumption for Type 2 Use Case 2 projects. Water cooling of the baseline data centre was factored into PUE and emission factor figures rather than directly represented in Equation 3.4.

The following PUE equation was derived to attempt to account for the different energy sources feeding power into the data centre in the baseline scenario:

$$PUE_{Ba} = 1 + \frac{Energy_{CRAC} + Energy_{CL} + Energy_{other}}{Energy_{ICTequipment}} \tag{3.5}$$

In this equation PUE_{Ba} stands for the PUE of the baseline scenario, $Energy_{CRAC}$ stands for the energy used by the computer room air conditioners (CRACs), $Energy_{CL}$

for the energy supplied by the cooling loop, $Energy_{other}$ for other energy overhead such as from facility lighting, and $Energy_{ICTequipment}$ for the energy consumption of the computing equipment. Determining each of these values could not be done in a cost-effective manner. CRACs from various manufacturers and of varied age were used in the baseline data centre with no monitoring capabilities or published power consumption data available. PDUs and other computing equipment were also often non-meterable. These obstacles could not be circumvented by monitoring the building's electrical system since circuits were not individually metered and the only reading available was of the aggregate power usage of the building. In addition to the baseline data centre, the building's electrical feed powers other equipment rooms, all the lighting and office computers, and even block heater plugs in an adjacent parking lot.

A PUE for the baseline scenario was thus derived using conservative assumptions. This included a review of published data centre efficiency information from sources including Digital Realty Trust [77], Energy Star for Data Centers [85], and the EPA [89] among others and resulted in the PUE in the baseline scenario being set at 1.7.

The emission factor in the baseline scenario also needed to be computed to be able to apply Equation 3.4. Although Alberta electrical grid emission intensity information is available, the data centre also used energy from local backup generators and the campus cooling loop. As most equipment in the baseline scenario was not attached to generator-backed circuits and generators ran only a small fraction of time, their energy use was thus not quantified as it is conservative to underestimate baseline emissions. However, the University of Calgary's cooling loop is more efficient than cooling provided using only grid power and thus required closer examination.

Figure 3.7 shows the structure of the University of Calgary's cooling loop. Water is pumped from the Bow River pumping station to the University of Calgary and a nearby hospital. A heat exchanger in the campus heating and cooling plant uses this water to chill water in a second, closed cooling loop passing through buildings on the university campus. The total power consumption of the Bow River pumping station and the fraction of river water directed to campus was obtained. The power consumption of the heat exchanger and chillers in the campus heating and cooling plant as well as of various pumps along the campus cooling loop was also obtained in addition to the total cooling provided by the cooling loop. As the Alberta electrical grid was used to power this equipment, the available information was adequate to calculate the emission intensity for the campus cooling loop as 180.91 g CO_2e per kWh, a figure far lower than the Alberta power grid's emission intensity. It was estimated by data centre management that about 30% of data centre energy consumption could be attributed to the cooling loop. A weighted emission factor of 680.3 g CO_2e per kWh was thus calculated as the baseline emission intensity factor after factoring in the Alberta electrical grid's emission intensity.

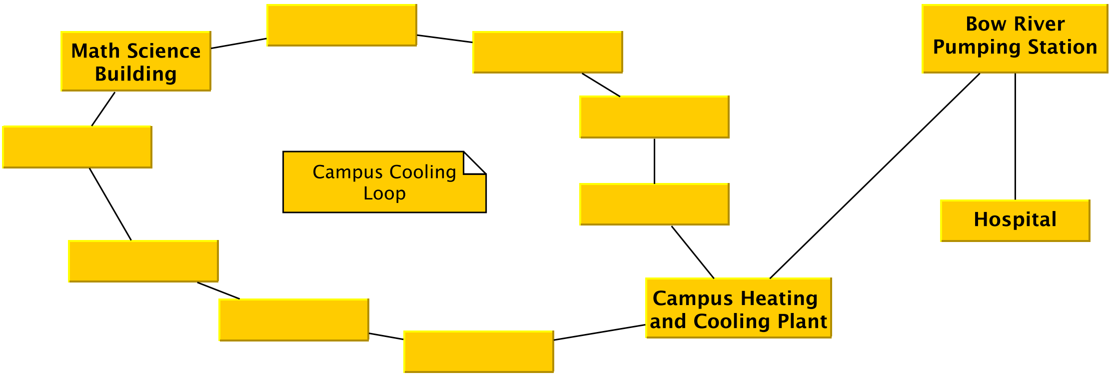


Figure 3.7: Overview of the campus cooling loop.

Project Emission Calculations

The earlier equation for project emissions was revised to incorporate the emissions of the network link between GeoChronos and RackForce:

$$Em(Pr) = P_{ICT}(Pr) \times PUE(Pr) \times EF(Pr) + Energy_{network} \times EF_{network} \quad (3.6)$$

Here $Em(Pr)$ represents the emissions associated with the project scenario. It is calculated using $P_{ICT}(Pr)$ to represent the project scenario's computing power consumption, $PUE(Pr)$ to represent the RackForce data centre's PUE, and $EF(Pr)$ to represent the emission factor in the project scenario. $Energy_{network}$ and $EF_{network}$ represent the power consumption of the network link between RackForce and the University of Calgary and its emission intensity respectively.

Power monitoring was established at the RackForce data centre as will be described in Section 3.6.6. The emission intensity of grid power in the project scenario was determined to be 20 g CO_2e per kWh. PUE was determined by analyzing RackForce's utility bills as well as the load time of the backup generators to determine total facility energy consumption. Due to better metering capabilities in the project data centre than the baseline data centre, the power consumption of the computing equipment could be distinguished from the overhead. This information was used to calculate PUE as being 2.0, a value which is expected to decrease significantly over time as the facility is relatively new and thus still being filled.

The network link between the University of Calgary and RackForce was evaluated as follows. 10% of a 10 Gb/s CANARIE lightpath from Calgary to Vancouver was dedicated to GeoChronos. Once traffic from Calgary had reached Vancouver it was then redirected through a RackForce owned and operated switch to the RackForce data centre in Kelowna over a network link operated by another provider.

Information was retrieved from CANARIE regarding the relay stations between Calgary and Vancouver. In addition to the University of Calgary point-of-presence and RackForce switch in Vancouver, there were a total of 12 other stations through which the signal passed. Each repeater station was assumed to draw power from the local electrical grid and was thus assigned the emission intensity of the corresponding province's power grid. The equipment used was assumed to operate at full power since network equipment generally exhibits very poor energy proportionality. The GeoChronos project was assigned responsibility for power consumption corresponding to the fraction of capacity it was allocated. The link between the University of Calgary and Vancouver was estimated to result in 110.31 kg CO_2e emissions per week of operation. Emissions associated with the network link from the RackForce-owned switch in Vancouver to the RackForce data centre in Kelowna were difficult to assess due to lack of response from the service provider. As a result, emissions from this portion of the network were labelled not relevant when assessing emissions in the project scenario.

Using the monitoring described in the next section the emissions in the project scenario could be determined upon project completion and compared against the baseline scenario to calculate total emission reductions.

3.6.6 Monitoring Project Computing Power Consumption

This section details the monitoring of project computing power consumption, describing the project's monitoring plan and limitations on monitoring current systems.

A variety of computing equipment was used to operate GeoChronos at RackForce, with the monitoring of each type of device described in turn. Devices included SAN storage, a traditional dedicated server, two blade servers accounting for 25% of a blade chassis, and a virtual machine in RackForce's cloud environment.

The SAN power consumption was monitored with Power Xpert software which metered relevant electrical circuits. The project storage's power consumption was calculated based on the fraction of the allocated storage allotted to GeoChronos. Power measurements were recorded weekly during the project.

GeoChronos's dedicated server was connected to a managed power bar and power measurements of each outlet used by the server were read weekly at times when GeoChronos was at an average load in order to be conservative.

The power consumption of the blade servers used by GeoChronos was determined as follows. Sensors accessible through the blade enclosure's management console reported the power consumption of each installed blade and the aggregate power consumption of the chassis. Using these readings the power consumption of the enclosure's shared systems was determined and evenly divided among the blades in the chassis, each of which was also responsible for its own separate power consumption. Measurements were taken manually on a monthly basis.

RackForce operates a cloud environment where GeoChronos was allocated a virtual machine. The total power consumption of the cloud segment on which the VM was executed was monitored, and since RAM was not oversubscribed, the VM was assumed to use energy corresponding to the fraction of total allocated RAM it was assigned.

The low frequency of power measurements possible during the project is a significant concern, but it is hoped that as technology improves more comprehensive monitoring data will be readily collectible. As previously mentioned, power readings were available every three seconds from a PDU used in later experiments, a degree of monitorability which it is hoped will be commonplace in the future.

3.6.7 Final Emission Reductions

GeoChronos ran for over half a year in the project data centre with monitoring data collected for a 13 week project period. This emission reduction project tested the protocol, identifying where it proved successful as well as highlighting certain weaknesses such as the difficulties in measuring power consumption in current data centres.

The total emission reductions were estimated at the start of the project and at the end of the project final emission reduction calculations were performed. The pre-project calculations predicted 1300 kg of CO_2e reductions per month. At the end of the project 2412 kg of CO_2e emission reductions were calculated for the 13 week project period, a figure somewhat lower than anticipated.

3.7 Summary

This chapter detailed the creation of a greenhouse gas reduction protocol for the computing sector and the evaluation of the GeoChronos relocation project which served as a test case to evaluate the protocol. The protocol complies with ISO-14064-2 standards, providing credible emission reduction figures with which to compare projects, use in corporate sustainability reporting, or produce emission reduction credits for sale in carbon markets. The types of emission reduction projects supported by the protocol were outlined, addressing improvements to data centre efficiency as well as other measures to improve the efficiency with which workload is processed. ISO-14064-2 requires that emission reduction projects exceed business-as-usual standards and tests to show that this additionality standard was met were outlined. How to create a business-as-usual baseline scenario against which to compare projects was discussed, providing guidance as to how to estimate the power consumption of servers and efficiency of data centres in such a scenario. Using this information to calculate initial estimates of emission reductions, to outline monitoring requirements for projects, and to guide the auditing process and production of final emission reduction figures were all addressed. How the protocol was applied in the GeoChronos relocation project was then outlined, showing how the protocol could be used in practice.

Chapter 4

Exploiting Variation in Data Centre Power Costs

> . . . He subdues and puts to his service the fierce, devastating spark of Prometheus, the titanic forces of the waterfall, the wind and the tide. . . . He makes the great Sun itself his obedient toiling slave.
>
> Nikola Tesla, “Man’s Greatest Achievement”

Data centres may be able to exploit varying power costs and, in doing so, ease the integration of additional renewable energy into the power grid. This chapter evaluates how data centres might reduce their financial and environmental power costs. Strategies explored in this chapter include adjusting data centre power use based on financial cost, carbon intensity, and the availability of renewable energy from local sources as outlined in Figure 4.1. Another strategy, the provision of ancillary services will be addressed in Chapter 5. I was responsible for the development of the software tools presented in this chapter and for the test cases executed. Existing work at the time of the initial publication of this research had generally focused on routing requests to alternate data centres rather than deferring the execution of tasks to later times.

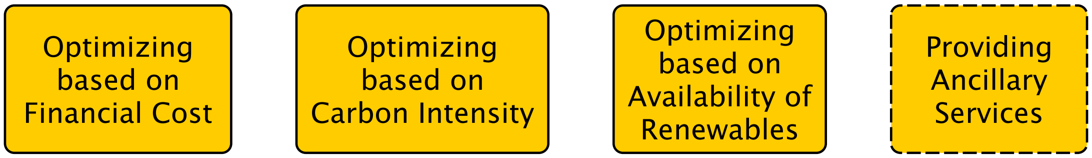


Figure 4.1: Strategies to reduce the environmental and financial costs of data centres.

This chapter is organized as follows. It begins in Section 4.1 with an overview of patterns in power market pricing, renewable energy production, and estimated carbon intensity. Section 4.2 then uses an tool to estimate the potential effectiveness of these strategies. Quality of service requirements vary by workload and this difference is the focus of Section 4.3 which uses a separate simulator to evaluate the impact of treating tasks with different priorities, using heuristics to estimate the relative cost of the upcoming time interval. Section 4.4 concludes with summary remarks.

4.1 Energy Production and Pricing Patterns

Power grids experience diurnal consumption patterns and significant challenges in reducing their carbon footprint by incorporating more energy from volatile renewable power sources. Price is a proxy for energy availability relative to demand and thus shifting power use to low-cost periods may enable an organization to reduce its carbon footprint along with its financial costs as will be described in more detail later in this section. Data centres may also alter their power consumption based on indicators such as local renewable energy availability or the estimated carbon intensity of the power grid.

This section elaborates upon patterns in power market pricing, renewable energy availability and carbon intensity. First, Section 4.1.1 outlines power grid pricing and the impact of adding renewable energy capacity. Patterns in the availability of renewable energy are then explored in Section 4.1.2, and Section 4.1.3 finally estimates the variation in carbon intensity over the course of a day.

4.1.1 Financial Cost of Energy

Power markets with dynamic pricing are increasingly accessible to data centres. This section outlines the structure of and pricing trends in power markets, also noting the impact of added renewables on pricing.

Power grids vary in how they price energy. Some charge a fixed rate per unit of energy consumed whereas others set prices for peak and off-peak times of day or vary prices on a seasonal basis. Still other power grids operate markets of varying degrees of complexity. Alberta possesses a common, province-wide market in which the price is set hourly on an ex-post basis, meaning that the cost of power is not fixed until after its consumption. The New York Independent System Operator runs a more complex market system, separating the state into regions and operating day-ahead, hour-ahead and real-time markets in each.

There is a high degree of price variation in typical power markets. Figure 4.2 plots the average and median power price over the course of a week in Alberta, using 2000–2009 data from the Alberta Electrical System Operator. The figure shows a readily visible diurnal pattern where the median daily peak price is three to four times the median daily trough price. Figure 4.3 examines the probability that a given hour of day is the minimum price or within a threshold of it. The early morning hours frequently hold the minimum price for the day, and midday prices are often four or even eight times the day's minimum.

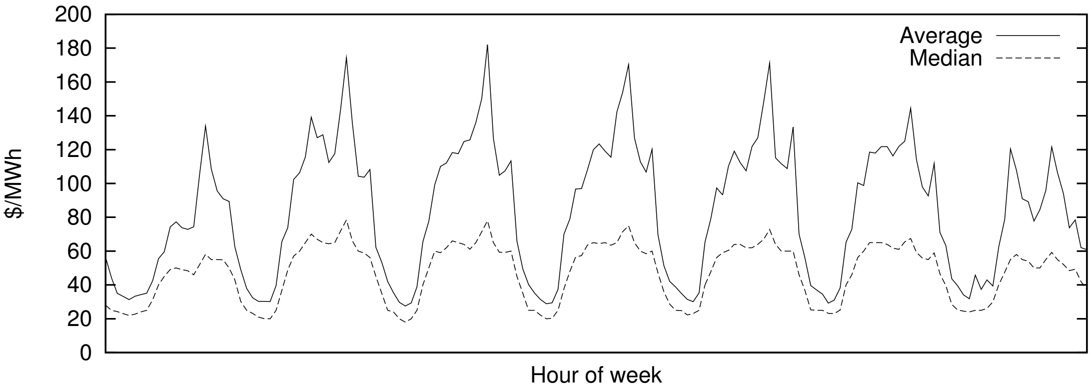


Figure 4.2: Alberta power price versus hour of the week.

Infrequent spikes in the power price, typically due to supply constraints or unex-

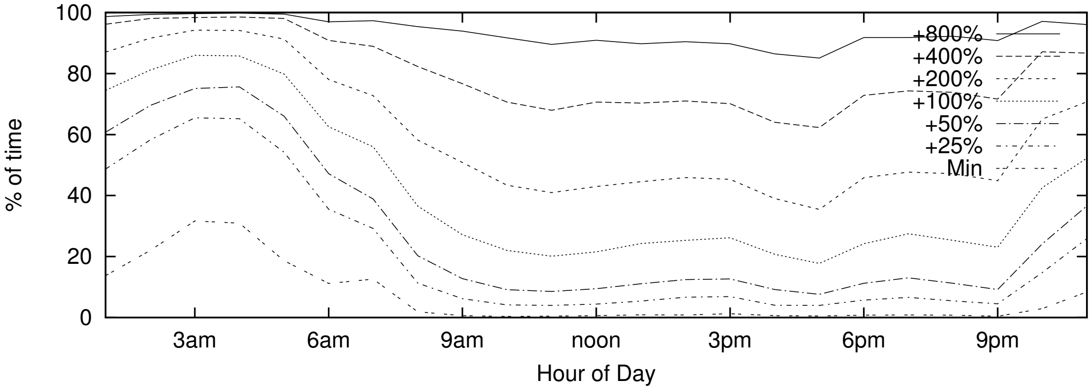


Figure 4.3: The percentage of time that an hourly price is the daily minimum or within a certain threshold of that day's minimum in Alberta in 2000-2009

pected surges in demand, lead to significant differences in average and median power prices as can be seen in Figure 4.2. The price of power from one period to the next can vary by a factor of ten or more. For a facility consuming constant power for the time period graphed in Figure 4.2, 20% of its energy cost would be from just 2.5% of runtime. How this compares to other regions is illustrated in Figure 4.4 which plots Ontario and Texas power prices alongside Alberta's.

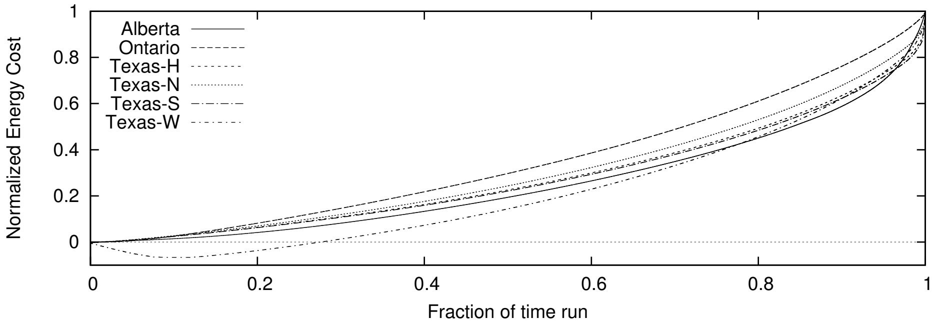


Figure 4.4: How the normalized energy costs vary from region to region based on the fraction of time they consume power, when consuming it in order from low–high cost hours

Attempts to integrate renewable energy and corresponding government incentives have resulted in unusual trends in power pricing. This can be seen in Figure 4.4 where negative power prices have become common in recent years in West Texas as wind turbines can often operate profitably while paying others to take their output due to the subsidies received per unit of energy produced [125].

The ability of power grid operators to curtail unwanted renewables is often limited. Despite environmental damage if water is sent over spillways instead of through generators during high-water periods [69], the Bonneville Power Authority was prohibited from curtailing wind turbines and instead required to implement negative pricing [25]. Damage to aquatic habitat is not the only other potential negative result of restrictions imposed on wind curtailment. In 2010 Ontario's Independent Electrical System Operator stated that by 2013 it might need to curtail other low-carbon power plants 14.5% of the time to accommodate wind. If this drives a nuclear power plant offline, such a facility requires an estimated 72 hours before it can resume supplying power, requiring that other more expensive and more carbon-intensive generators account for any decline in wind production [127].

German feed-in tariffs paying premium, guaranteed rates for low-carbon energy are also distorting power markets, driving daytime prices occasionally to negative levels and forcing other generators to seek profits at night when not competing with solar power [28, 162].

Future shifts in policy may cause power prices to more closely approximate renewable energy availability. In Alberta large-scale emitters are currently charged \$15 per tonne of CO_2-equivalent emitted beyond emission reduction targets [42]. The International Energy Agency's head argued in a 2008 speech to G8 leaders that cutting carbon emissions 50% by 2050 would require a \$200–500 per tonne carbon tax [229]. A more recent study has suggested that the actual social cost of carbon is 2.6–12 times the \$21 per tonne figure used in previous US government reports [136]. If these higher estimates of carbon cost are adopted by government, it would cause market prices of energy to more closely align with its carbon intensity.

4.1.2 Availability of Renewables

This section explores patterns in renewable energy availability, beginning with low-variability renewables and then considering higher-variability ones.

Data centres looking to lower their carbon footprint have often focused on renewable energy sources with low variability. The use of biogas and methane capture have both attracted attention [29, 60, 171, 216]. These offer reliable, low emission intensity power, but the same properties may make them better suited for deployment in the power grid rather than paired specifically with data centres. Geothermal and nuclear power are two other stable sources of low-carbon power, and Iceland is attempting to use geothermal power availability to attract data centre investment [7], but few geothermal plants exist at this point and few new nuclear plants are being constructed.

Hydro, wind and solar power are three other renewable energy sources that can be used to power data centres. When large reservoirs are involved, hydro can generally supply reliable power year round though, as Section 2.1.2 noted, droughts may lead to brownouts and blackouts. Oversupply may also cause environmental damage as the previous section noted. Significant seasonal patterns may be observed where hydro accounts for a moderate amount of grid power, with Figure 4.5 showing how the percent of power from hydro in Alberta changes over the course of a year. Wind power is also shown in the same figure, illustrating a less pronounced degree of seasonal variation. Solar energy also varies significantly by season in Alberta. Figure 4.6 plots the amount of solar energy per unit area by day of year based on fifty years of recorded data [88], illustrating greater variation in received energy than in the hours of sunlight per day over the course of a year.

Though Figure 4.5 shows lower seasonal variation for wind than hydro, high short term variability can still be seen for wind despite the plot averaging all Alberta wind

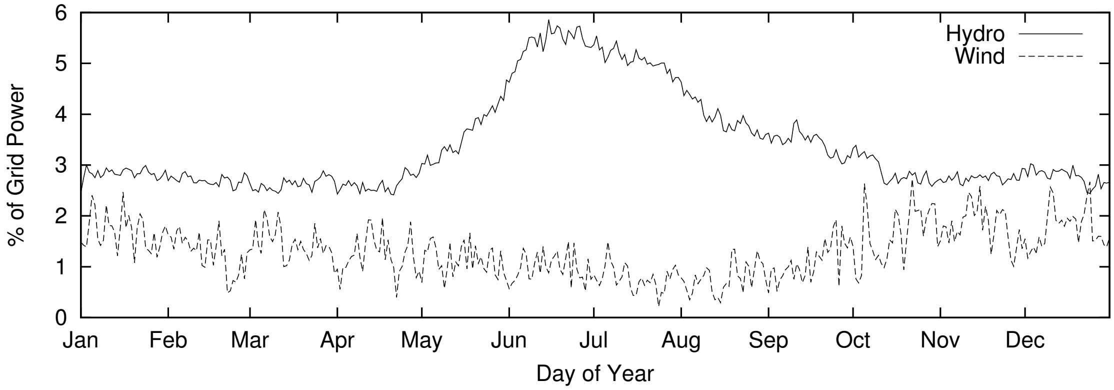


Figure 4.5: Average supply of Alberta grid energy from wind and hydro sources by day of year for 2000–2009

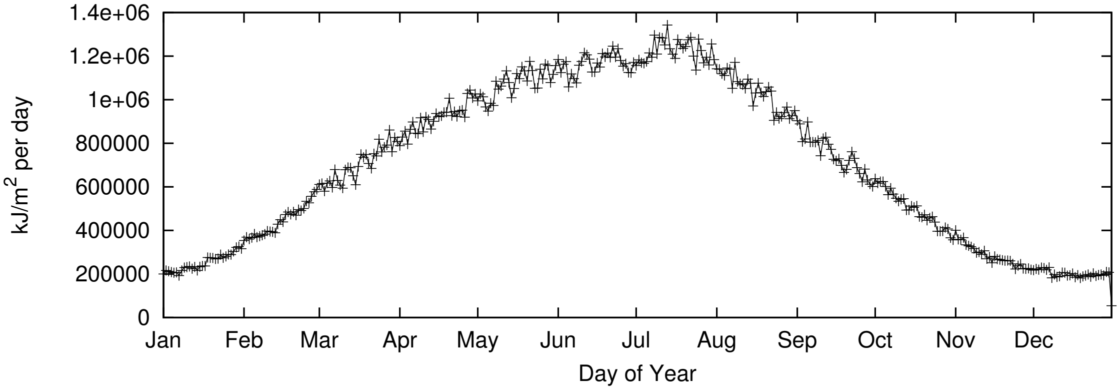


Figure 4.6: Solar energy supplied per square meter per day in Calgary, Alberta

generation per day from 2000 to the end of 2009. As Chapter 2 noted, solar power can vary dramatically on a short-term basis, with output changing as much as 50% in a 30–90 second time frame. How effectively might data centres use such power?

Variability declines when aggregating the output of particular classes of renewables over larger geographical areas. Wind turbines also generate at least a small amount of power most of the time. Section 2.1.5 noted that the capacity factors for wind, indicating the fraction of installed capacity likely to be able to supply energy when requested, varies from 2–60%. In Alberta the regulatory authority has assigned wind a 20% capacity factor and hydro a 50% capacity factor [61]. Figure 4.7 illustrates the probability that turbines generate at least a small amount of energy most of the time, showing that even turbines with a capacity factor below 25% could supply a majority of the power consumed by a customer with capacity one tenth of the wind turbines'. A significant degree of volatility nonetheless remains with wind and solar energy sources however.

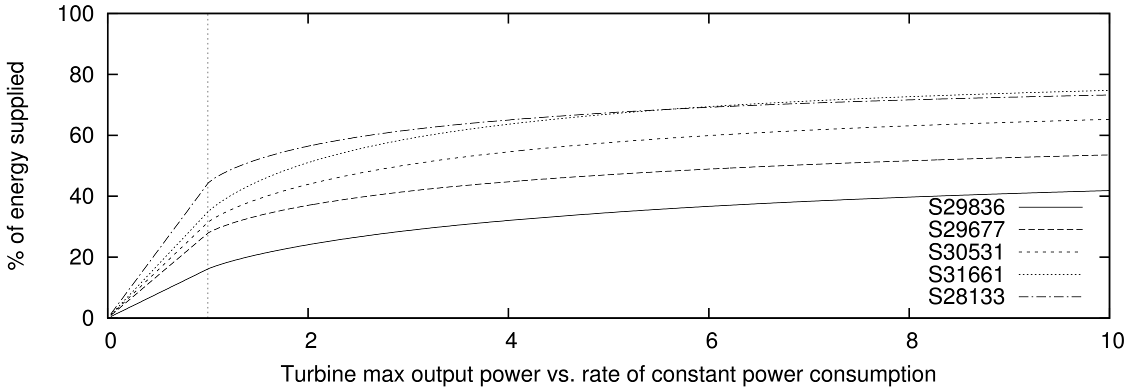


Figure 4.7: Percent of a constant power demand that can be satisfied relative to the capacity of local wind turbine(s). The vertical line at $x = 1$ intersects the curves where the y value is equal to the turbines' capacity factor.

Figure 4.8 compares diurnal patterns in the availability of wind and solar power, plotting these alongside the hourly Alberta power price. In addition to wind energy data from the Alberta Electrical System Operator and the previously mentioned Calgary solar data, wind generation data for Montana was added [179] due to its availability and this state bordering Alberta. Solar power availability occurs primarily at high-demand times of day, suggesting that wind may be more difficult to accommodate. Thus deferring delay tolerant workloads to times in which wind energy is available may be useful. Later work by Lin *et al.* [163] reached a similar conclusion regarding the priority of wind and solar in a paper on geographical load balancing, suggesting the value of wind power over solar power when implementing "follow-the-renewables" routing.

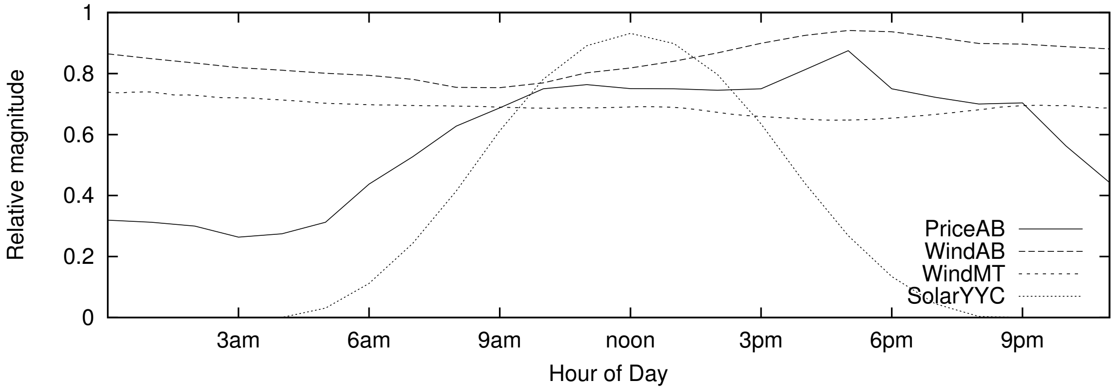


Figure 4.8: Solar availability, wind availability, and power price versus hour of day.

4.1.3 Carbon Intensity

As renewable energy production is more stable over larger geographical areas, using a global rather than a local measure of low-carbon energy availability may be more effective in reducing costs. Attempting to maximize low-carbon energy use regardless of where it is produced may also reduce subsidies needed to keep renewable projects viable by boosting the prices they receive. Power prices are often set based on the marginal cost of the next unit of power required, and thus a small increase in consumption can lead to a larger increase in price.

This section attempts to estimate changes in the power grid's carbon intensity over time by analyzing changes in the power mix, noting significant limitations of this approach which point forward to the discussion of ancillary services in the next chapter.

Hour-to-hour changes in the composition of power were estimated for Alberta based on manual classification of power sources in production data from the Alberta Electrical System Operator. Figure 4.9 plots the average power mix over the course of a day, showing coal dominating and natural gas in second place. Due to the low level of detail in descriptions of these power sources potentially resulting in power source misclassification and validation errors affecting a subset of the data, this dataset is regarded as only an approximation of the actual power mix but believed adequate for the purposes of this chapter. Also, notably, differences in the efficiency of natural gas power plants could not be readily identified nor were power plants operated as cogeneration facilities identified.

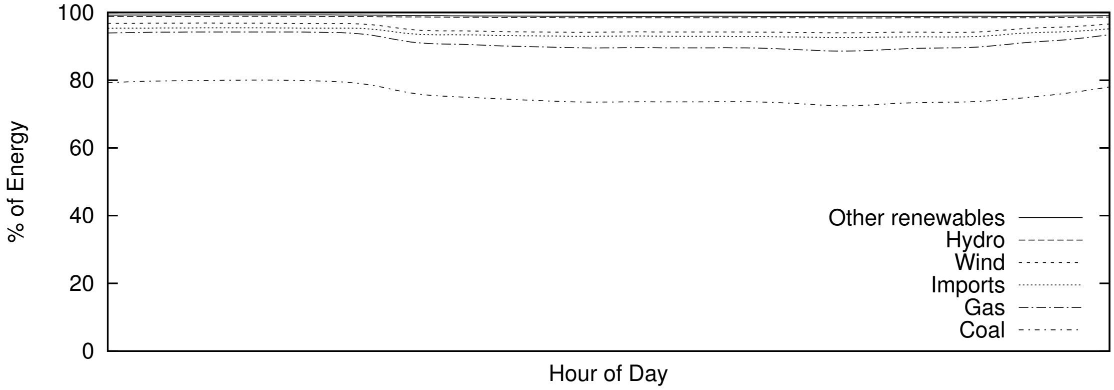


Figure 4.9: The composition of Alberta electrical power by hour of day.

Coal power dominates the Alberta power mix, and this dominance increases when examining estimated carbon emissions. Estimated emission intensity was calculated by multiplying the percentage of power produced by a type of source in an hour with the corresponding emission intensity factor in Table 4.1. Due to much imported power coming from BC during peak hours [170], imported power was assigned BC's average emission intensity [61]. Estimated emission intensities were obtained for other power sources [159], averaging the emission intensities of the different types of plants using each fuel. A run-of-river emission intensity estimate was used for hydro due to its availability though reservoirs may moderately increase emissions [210]. Unclassified power sources were factored out of the data. Biomass was also factored out due to concerns over its carbon neutrality [135] and both the low percentage of energy derived from this source and the low variation in biomass output.

It should be noted that multiplying output by average emission intensity ignores the impact of how each generator is operated. Added emissions from cycling generators up and down are not taken into account in these calculations. The specific

Table 4.1: Estimated emission intensities of generation sources and the corresponding fraction of Alberta electricity derived from each.

Source Type	**Emission Intensity (kg CO_2e/MWh)**	**% of Reported Energy Generated**
Coal	993	75.7
Gas	664	15.5
Hydro	15	3.5
Wind	21	1.3
Import	20	2.7
Other & Biomass	Unknown	1.4

generators called upon varies somewhat based upon the power grid operator's scheduling, accounting not only for bids in the power markets but also reliability constraints. Emissions may be difficult to associate with specific actions, as they may sometimes produce emissions hours in the future as emission control systems adjust [64]. Chapter 5 examines ancillary services, mechanisms for addressing short term variation and failures in power grids, providing an avenue by which to reduce cycling activity.

The carbon intensity of primarily hydro-based power grids which use some fossil fuel generators for peaking power is likely to exhibit greater variation from one hour to the next. However, such a grid would already have relatively low emissions.

4.2 Exploiting Variation Over Time Periods

This section estimates the ability to reduce energy costs by adjusting data centre operations based on the patterns outlined in Section 4.1. How each pattern was addressed is described in Section 4.2.1, which also outlines an experimental plan. The time-step analysis tool used in testing is then introduced in Section 4.2.2. Section 4.2.3 concludes with experiment results.

4.2.1 Strategies and Tests Performed

As outlined in the previous chapters, data centres are able to ensure a reliable flow of power to equipment in the event of interruption or change in the availability of grid power by using energy storage and can take advantage of the significant increases in server energy proportionality in recent years. This section describes strategies a data centre could use to exploit the ability of computers to rapidly alter their power consumption. These strategies were evaluated using Parallel Workloads Archive [190] traces shown in Table 4.2. A more detailed description of each of these traces is available at the Parallel Workloads Archive. The focus of this section is on delay-tolerant workloads. However routing less-delay-tolerant workloads to whichever of a set of data centres has lowest cost power or highest renewable energy availability could similarly reduce the financial or environmental costs of computing.

Table 4.2: Workloads used for testing the potential of different financial and environmental costs of power.

Workload Name	**Avg(Utilization)**	**Avg(Tasks/Day)**	**StdDev(Tasks/Day)**
DAS fs3	10.7%	183.8	553.6
LLNL Atlas	64.1%	225.7	161.7
LLNL Thunder	87.6%	796.3	585.2
LPC EGEE	20.8%	978.7	1015.4
OSC Cluster	43.1%	53.2	42.0
SDSC BLUE	76.7%	248.3	135.4

Financial Cost of Energy

The financial cost of power can be reduced by using the system when power is least expensive. Costs can be reduced by attempting to consume maximum power when the power price falls below zero even if the system does no more than busy-waiting. Experiments were run using the power price data described in Table 4.3.

Table 4.3: Power price data used.

Region	**Timeframe**	**Length of Price Intervals**
Alberta	January 2000 – October 2009	60 minutes
Ontario	May 2002 – November 2010	60 minutes
Texas (4 regions)	January 2008 – September 2010	15 minutes

Availability of Local Renewables

Executing tasks when power is available from local wind turbines is another strategy for reducing energy costs. Montana wind traces in NREL's Western Wind Integration dataset [179] were sorted by capacity factor, and the 1^{st}, 25^{th}, 50^{th}, 75^{th}, and 99^{th} percentile were selected for use in simulations. These traces were linearly scaled so that maximum power production in any time period was 0.25x, 0.5x, 0.75x, 1x, 2x, 4x, 8x, and 10x the peak power consumption of the data centre it was paired with.

Carbon Intensity

Information from the Alberta Electrical System Operator was used to estimate the carbon intensity of grid power during each hour as previously described in Section 4.1.3. To evaluate situations in which wind accounted for a larger percentage of generated power than it presently does in Alberta, a number of additional traces for use in simulation were constructed using Alberta power production data and NREL wind traces. In the modified traces, the demand for electricity during each interval was assumed to remain the same as in the original trace. It was assumed that additional wind power would replace power from gas-fired generators, since these can typically ramp production up and down more readily than coal generators. Additional wind power, beyond what replaced gas power, was assumed to be curtailed. This is only an approximation of the real world, used due to the limited data available on power plant specifics. Power grids with a high fraction of coal power require coal plants to adjust their output, and cogeneration plants fuelled by natural gas may have little flexibility. Historically, peaking plants have often been built using gas for fuel and coal has been commonly used for baseload generation but due to changes in the pricing of natural gas this pattern has weakened. The NREL data contained estimates of the average generation over 10 minute intervals at 2012 candidate wind turbine sites in Montana, immediately south of Alberta. Varying-sized subsets of wind turbines were used to construct traces. Scenarios were also evaluated where a fraction of the coal assumed to be baseload power generation was replaced by peaking gas generation and the same assumptions for wind replacing fossil fuels held.

The same adaptive strategy was then applied as for financial cost, using an environmental cost metric instead of a financial one.

4.2.2 Time-step Analysis Tool

The potential of the strategies outlined in the previous section to reduce the financial and environmental costs of computing were evaluated using a time-step analysis tool developed in C++. Inputs to the tool consisted of a workload trace and financial, renewable energy availability or carbon intensity information.

Data was analyzed period-by-period. Tasks submitted in each period were amalgamated into a demand for a particular number of CPU-hours according to the CPU-time used by each task when run. The simplifying assumption was made that power consumption varied linearly from zero at idle to maximum power under full load. As discussed previously, current servers consume much less power at idle than their predecessors due to mechanisms like dynamic voltage and frequency scaling. Each time period was divided into subintervals whose length corresponded to the length of periods in the cost data used.

For availability-based adaptation the number of CPU hours that could be executed in a period using local renewable energy was calculated based on the power generation in each subinterval and the maximum power draw of the data centre. The execution of queued CPU hours was then simulated by subtracting them from the CPU hours of energy available. Additional CPU hours would require grid energy and the fraction of each period that grid power use was permitted could be capped. A corresponding non-adaptive mode computed the percent of local renewable energy available in each period, bounded in each subinterval by the data centre's maximum power consumption, and assumed that an equivalent fraction of locally-generated energy would be consumed when the system was in operation.

For cost-based adaptation, price data was fixed to each subinterval, and the subintervals were sorted by increasing cost. To provide a non-adaptive estimate, power costs were randomly reordered. This may underestimate prices in some systems since job submissions most often arrive when the price of power is relatively high, though compensated for somewhat by the relatively high utilization of high performance computing systems. This issue will be addressed more closely in Section 4.3. Iteration was performed over the subintervals, deducting from the queued CPU-hour demand the amount that could be executed in the subinterval and updating costs accordingly. In adaptive mode a threshold might be set above which the system would queue tasks until lower cost power was available.

More details of this program are available in Appendix A, Section A.2.

4.2.3 Results

This section presents, in sequence, the results of experiments conducted into the potential benefits of altering data centre scheduling based on the financial cost of energy, local renewable energy availability, and changes in the power grid's carbon intensity.

Financial Cost of Energy

The potential of adjusting operations based on changing financial costs was tested through analysis of day or week-long periods. Figure 4.10 shows a sample of the results where costs are normalized by comparing each situation to an otherwise identical non-adaptive scenario. When analyzing power consumption in one-day increments for the Alberta trace starting in January 2000, costs 25% lower than a non-adaptive approach were seen. Tests were also run in which Alberta trace data was instead started from the summer of 2003. This sometimes led to different results due to variations in power price from year to year in Alberta and how spikes in power prices impact results. When investigating how much costs could be reduced, examining longer time periods led to greater possible savings as the potential for longer deferral compensated for variation in both workload arrival rates and power pricing. The frequency of negative prices regularly occurring overnight in West Texas also produced results in which power costs fell below $0 as longer time periods were considered.

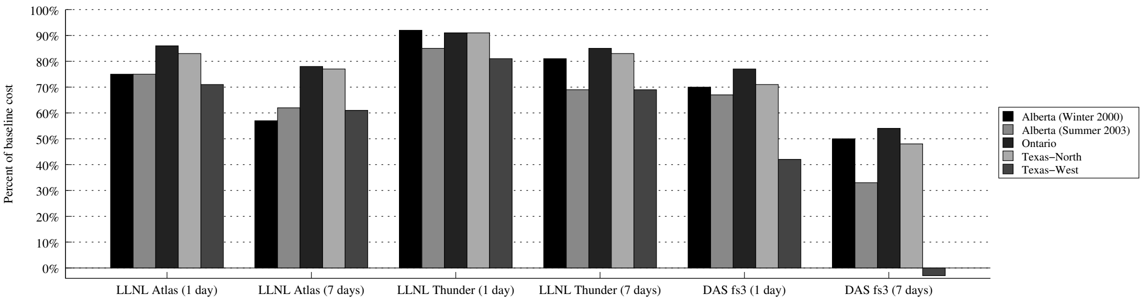


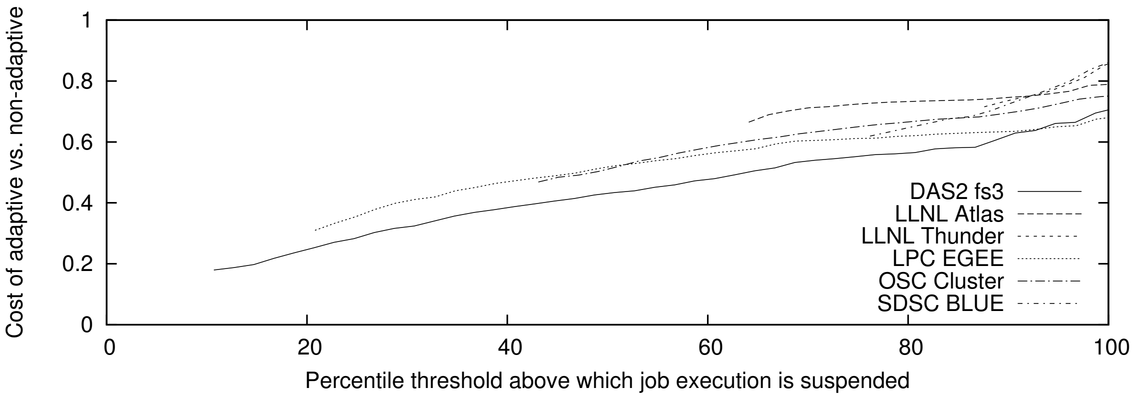
Figure 4.10: Savings when adjusting operations based on changing financial cost.

The impact on costs and workload wait times if the system was suspended when the power cost exceeded a threshold was also evaluated. The remaining workload would then be deferred to later periods in which power costs fell below the threshold. Figures 4.11a and 4.11b illustrate the impact of suspending the system in this manner. Each curve in this and subsequent graphs in this section represents a workload. Curves begin at the x-coordinate representing the corresponding system's long-term average utilization. In Figures 4.11a and 4.11b the curve shows the relationship between the threshold above which tasks are suspended and the cost of processing the workload. The costs presented in the figure were normalized by dividing the results of a simulation conducted in cost-reduction mode against an otherwise identical simulation run in non-adaptive mode. When the x value is 90, this represents the system being suspended when the power price exceeds the 90^{th} percentile of power prices in the region. The y value at this point indicates the corresponding cost of execution. Fig. 4.11a shows results for Alberta where a price floor of $0/MWh exists. West Texas with its frequent negative prices is shown in Figure 4.11b. Where negative prices are a regular, recurring phenomenon total energy cost may fall below zero for low-utilization systems able to defer workload processing.

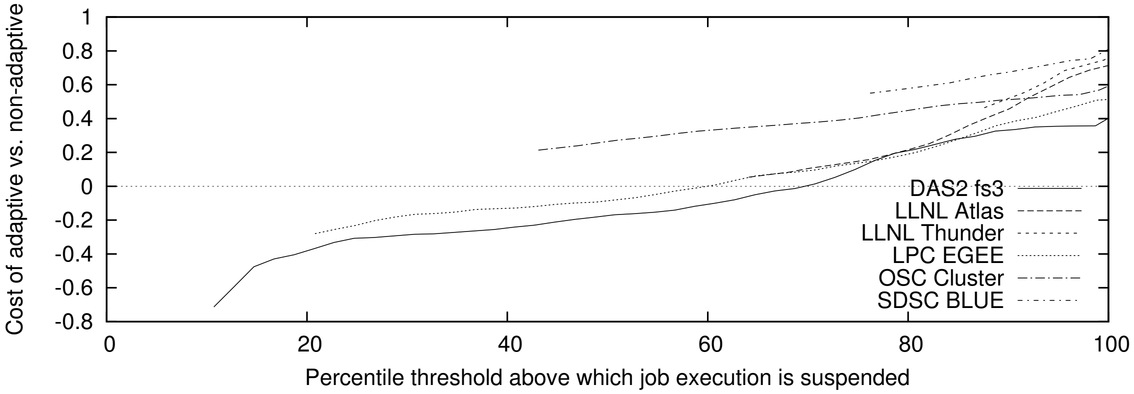
Suspending a system when prices cross a threshold results in extended wait times and a relative wait metric was used to account for this. A wait metric was computed by summing, for each hour of workload processed, the number of time periods it was queued prior to execution. A value of 1 was the queue time given to workload processed in the same period as submitted. Relative wait was calculated as the ratio of wait time of the cluster when the simulator was run in adaptive mode to the same workload executed as early as possible. Waiting time analysis results are shown in Figure 4.12 and correspond to the results presented for cost in Figure 4.11a. For the wait metric evaluated, setting a threshold at the 75^{th} percentile of power prices had little impact on relative wait but showed potential to significantly reduce power costs.

Availability of Local Renewables

Figure 4.13 shows the change in the fraction of energy sourced from local renewables relative to a non-adaptive approach when a data centre was paired with wind turbine capacity equal to its maximum power usage. The wind trace used in Figure 4.13 was the 50^{th} percentile of NREL's Montana wind traces. The x axis plots a cap on the



(a) Alberta



(b) West Texas

Figure 4.11: Relative cost for adaptation across a 1 day period when suspending job execution when the cost of electricity exceeds the xth percentile cost of power prices. Note that the Y axis differs between the two subgraphs due to the negative prices occurring frequently in West Texas.

percentage of each time period that grid power could be used to supplement local renewable energy. If able to adapt in this way even high utilization workloads could increase their utilization of local renewable energy, with the opportunity greater for low-utilization systems.

As described in Section 4.1.2, if the maximum output of local wind turbines is large relative to the data centre's power consumption, the potential to increase the use of locally-generated renewable energy is reduced as wind turbines tend to generate at least a small amount of power most of the time. This would not apply if solar was

Figure 4.12: Relative wait time for Alberta 2003 when pausing execution when the cost of electricity exceeds the Nth percentile.

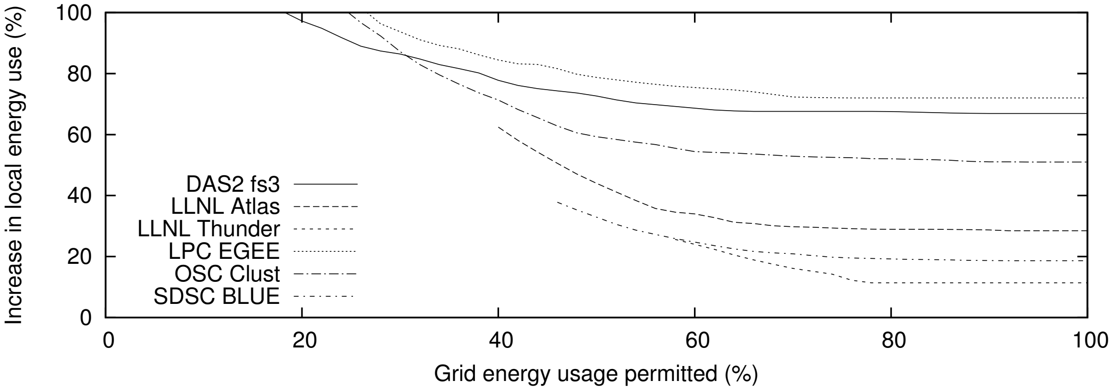


Figure 4.13: Percent increase in the fraction of locally generated energy used relative to the percentage of each time period that grid power was permitted to be used.

the local renewable energy source as solar energy is generally more plentiful when demand for power is greater.

The same wait metric described earlier was also applied to examine the impact on wait times when altering operations based on renewable energy availability. Figure 4.14 presents the wait time results for the scenario previously plotted in Figure 4.13. It is advisable to permit some grid energy to be used to avoid long delays since wind outages have been noted to follow a heavy-tailed distribution [226]. Any cap on grid power consumption should account for both the quality of service requirements of the workload being executed as well as the characteristics of the local renewable energy sources.

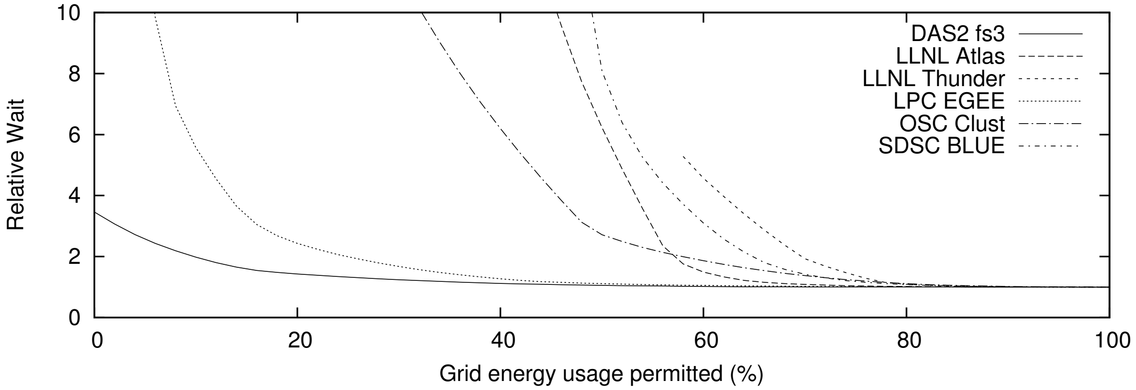
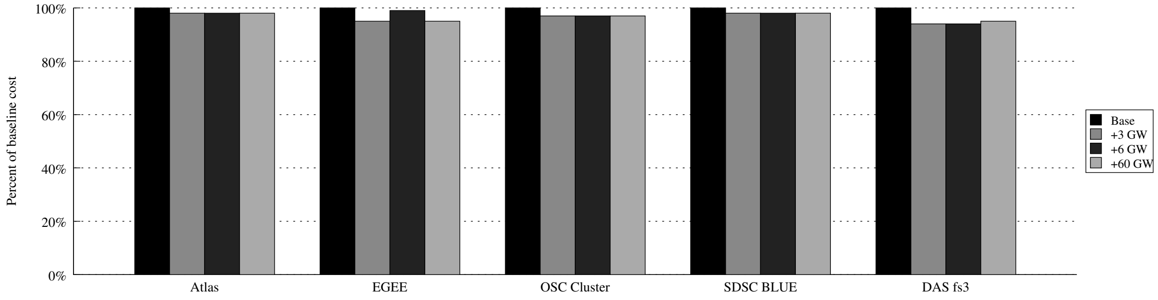


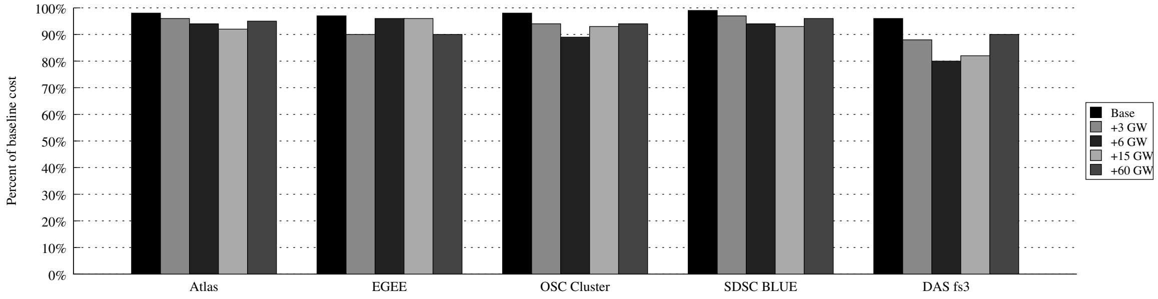
Figure 4.14: Relative wait time as a result of a cap limiting the percentage of each time period that grid power can be used.

Carbon Intensity

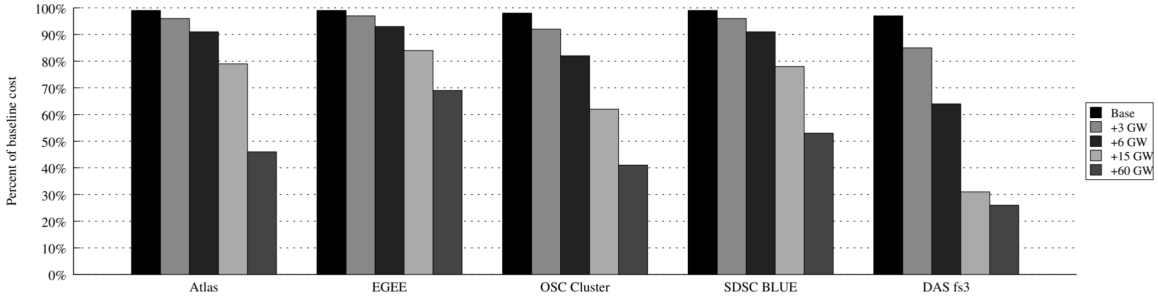
Altering data centre scheduling based on changes in estimated carbon intensity showed lower potential than the other strategies evaluated. Significant changes to the power grid would be needed for this to be effective, requiring large-scale addition of renewable capacity. Figure 4.15 shows the relative emissions estimated for various scenarios. As previously outlined, additional scenarios were evaluated in addition to the original estimates. These scenarios involved combinations of added wind capacity and varying fractions of power plants assumed to be able to rapidly alter their output to enable renewables to be accommodated more effectively. In the figure +3 GW indicates the addition of 3 GW of wind generation capacity, estimated to supply approximately 10% of Alberta power consumption if the grid could accommodate it. Under the original estimate of carbon intensity, with most power plants assumed to operate at constant output, even with significant added wind capacity low savings in carbon emissions were seen. Savings increased somewhat when half of the power plants assumed to operate as baseload power plants were then treated as able to more efficiently adjust their output, but even then savings were small without the addition of a significant amount of wind generation. It was only when all powerplants were assumed to be able to efficiently alter their output and significant wind generation capacity was added that significant savings in carbon emissions were possible.



Assuming original fossil fuel composition



50% fewer powerplants unable to efficiently adjust their output



All powerplants able to efficiently alter their output

Figure 4.15: Savings when altering operations based on carbon intensity.

Changing data centre schedules based on variation in carbon intensity might be a more effective strategy when connected to a power grid based primarily on renewable energy sources. Such grids may occasionally augment the power supply with carbon-intensive generators. These grids already have relatively low emissions however. Thus, despite a larger percentage decrease in carbon emissions, the impact of such a strategy would still be relatively low.

4.3 Handling Tasks of Different Priorities

High-performance computing workloads typically have a higher-degree of delay tolerance than most other workloads. Tasks in such workloads may have varying priorities with some tasks necessitating immediate action. This section examines how high-priority tasks can be accommodated while allowing the remainder of tasks to be deferred when power costs are high.

This section is organized as follows. First, Section 4.3.1 outlines how task arrivals align with power costs, providing motivation for the simulations described in this section. The data centre model used and adaptive strategy employed are then discussed in Section 4.3.2. This is followed by Section 4.3.3 with simulator construction details. Section 4.3.4 outlines experiments performed and their results are presented in Section 4.3.5.

4.3.1 Workload Arrival Patterns

Task requests to high-performance computing clusters are spread over roughly the same portion of the day that power prices reach their peak. The distribution of task request arrivals over the course of the day in various workloads in the Parallel Workloads Archive [190] is shown in Figure 4.16 alongside a scaled version of median Alberta power prices.

The extent to which tasks can be deferred without violating quality of service agreements is limited, but for a large fraction of these tasks a significant degree of time flexibility may exist, enabling power usage to be shifted from peak hours. To ensure quality of service requirements can be met while still enabling cost savings through deferring as many tasks as possible, separating the tasks into different classes with different priorities is important. It is this strategy that was pursued in the simulations discussed here.

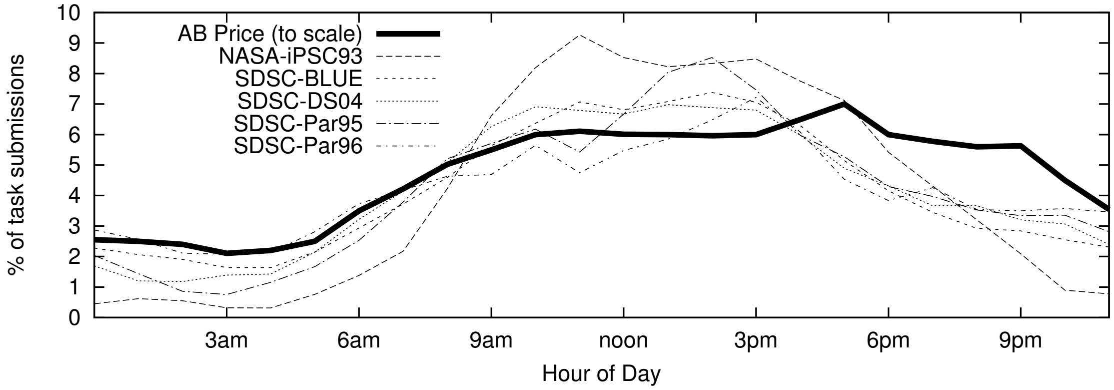


Figure 4.16: Distribution of the submission times of tasks by hour of day.

4.3.2 Model and Adaptive Strategy

This section outlines the model created of computational workload and the adaptive strategy used to reduce financial costs. As in Section 4.2, the simplifying assumption was made that no power was consumed while the system was idle, and that power consumption varied linearly between idle and full load.

The cluster workloads were separated into two categories: interactive tasks and non-interactive tasks. Interactive tasks were modelled individually and launched immediately, preempting any lower-priority, non-interactive tasks that might have been executing. A fluid approximation was used for non-interactive tasks. The volume of fluid in the system at any time was equal to the number of CPU hours awaiting execution. Non-interactive tasks were converted into a fluid volume of required CPU hours upon arrival based on their actual execution time. When non-interactive tasks were executed, fluid was drained at a rate corresponding to the number of CPUs available.

Non-interactive tasks were executed when low-cost power was available and CPUs were not occupied by interactive tasks. If operated in non-adaptive mode, non-interactive tasks were executed subject only to CPU availability. In adaptive mode, the previous week's power prices were used to assess whether or not the current price was too expensive and thus whether or not non-interactive tasks should be suspended. Based on each workload's long-term average utilization a threshold was set, adjusted upward slightly to account for some variability in system utilization. For example, for a system with 75% average utilization over the long-term, non-interactive tasks would be executed when cost fell below the 80^{th} percentile of the previous week's power prices. The assumption was made that the price for the next interval was known when assessing whether or not time periods were expensive, which may result in a moderate decrease in cost savings for power grids using ex-post pricing.

4.3.3 Simulator

A generic, next-event discrete event simulator was created in C++ using the BOOST libraries [3] and Standard Template Library. In each iteration the next event would be selected from a pool of event sources, with each source handling events of a specific type.

An event-handling routine processed arriving events, and was responsible for calling the scheduler routine. Task requests, the arrival of pricing data for a particular time period, and requests to run the scheduler function to handle events such as the completion of tasks were event types handled by the scheduler.

When interactive tasks arrived, they were accepted and allocated processors if enough were available to meet their requirements. An event was then queued to arrive at the time at which the job would be completed. Non-interactive tasks were processed using the fluid model outlined in Section 4.3.2.

Whenever an event arrived the scheduler function was called. The function managed the state of the cluster, marking processors free when interactive tasks completed. It was also responsible for simulating the execution of non-interactive tasks using the fluid model, with the previous week's power prices used to determine whether the power price was sufficiently low for tasks to be executed.

Power price information was processed on arrival. Costs for interactive tasks were tabulated and statistics reported for each completed job. At the end of each simulator run the average power price paid for interactive and non-interactive tasks were reported in addition to other summary information.

More details of this simulator are available in Appendix A, Section A.3.

4.3.4 Experiments Performed

Experiments used traces from the Parallel Workloads Archive outlined in Table 4.4. These traces were selected since they contained and separately identified interactive and non-interactive tasks. Further details of the traces are available from the Parallel Workloads Archive [190]. The systems simulated were assumed to be homogeneous, containing an equal number of processors as in the original systems. Price data came from the sources outlined in Table 4.5.

Table 4.4: Workloads used for experiments separating high and low priority tasks.

Workload Name	**Avg(Utilization)**	**Interactive Tasks/Day**		**Non-Interactive Tasks/Day**	
		Avg	**StdDev**	**Avg**	**StdDev**
NASA iPSC	46.6%	11.2	11.5	186.9	125.5
SDSC BLUE	76.7%	93.2	77.1	171.9	93.8
SDSC DS	63.1%	18.2	57.6	228.5	140.7
SDSC Par95	71.6%	95.1	96.2	53.0	42.2
SDSC Par96	75.6%	37.4	51.4	49.7	44.8

Table 4.5: Power price data used in experiments separating interactive and non-interactive tasks.

Region	**Timeframe**
Alberta	January 2000 – October 2009
Ontario	May 2002 – November 2010
Texas (4 regions)	January 2008 – September 2010

Workload start times were adjusted, such that the local time and day of week in the region from which the power data was obtained matched the workload trace. Workload start dates were adjusted to fit within the periods for which pricing data was available. Several additional tests were conducted, varying the time of day in the power data relative to the workload trace to estimate the impact of executing workload in a data centre in a different time zone.

A variety of tests were conducted as outlined above. The tests examined the cost of interactive tasks as well as the impact on power costs of deferring non-interactive tasks when the price of power was high. The average power costs of interactive tasks were compared to the overall pattern of power costs, and the impact of varying the threshold at which the execution of non-interactive tasks would be suspended was also evaluated. How delays faced by non-interactive tasks were affected by potentially deferring their execution to a later time period due to high power prices were not explored in the simulations due to the loss of detail from using a fluid approximation to represent them. The effect of using a data centre in a time zone different from that of the cluster was evaluated, though, to examine the impact on cost of using remote data centres.

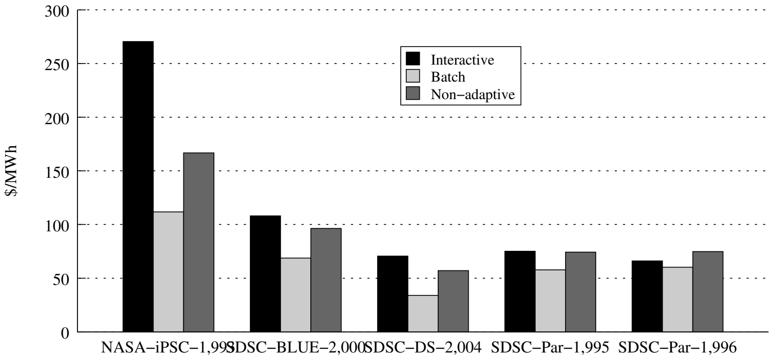
4.3.5 Results

Figure 4.17 shows power costs for interactive tasks as well as the power costs of processing non-interactive tasks in both adaptive and cost-agnostic modes. For example, the average power price paid for interactive tasks in the NASA-iPSC-1993 workload using Alberta pricing data was \$270.45/MWh. By comparison, \$111.71/MWh was paid for non-interactive tasks in adaptive mode and \$166.73/MWh paid for cost-agnostic execution of the non-interactive tasks. The average power cost for interactive tasks typically exceeded that of the cost-agnostic processing of non-interactive tasks in simulator runs, possibly due in part to non-interactive tasks arriving in bulk. If executing non-interactive workloads while aiming to minimize power costs, savings of 25–50% were often seen.

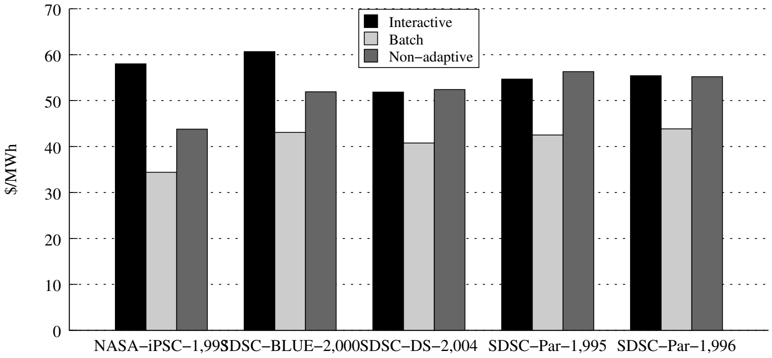
Testing was also conducted to evaluate the impact of changing the threshold above which non-interactive tasks would be suspended and the results are shown in Figure 4.18. The x coordinate represents the power price percentile threshold above which the execution of non-interactive tasks would be suspended to a future time period. A line was plotted for each workload beginning from the x coordinate representing the workload's long-term utilization and proceeding until non-interactive tasks would never be suspended. In the results the average power price varied in an approximately linear manner with the threshold selected, with the slope of the line varying based both on the workload and the power grid on which it was executed. Changing the date on which the workload was started altered the specific costs but not the trend. As in the earlier results, variations in the cost of power in Alberta resulted in one trace showing much higher average costs.

Figure 4.16 in Section 4.3.1 showed that task requests tended to arrive when power prices were high. This can also be seen in Figure 4.19, illustrated using the SDSC-Par-1995 and SDSC-Par-1996 workloads and Alberta pricing data. It plots the percentage of time that power prices fell below a particular threshold in the Alberta power pricing data. Alongside this is plotted the percentage of interactive tasks whose average power price fell below the threshold.

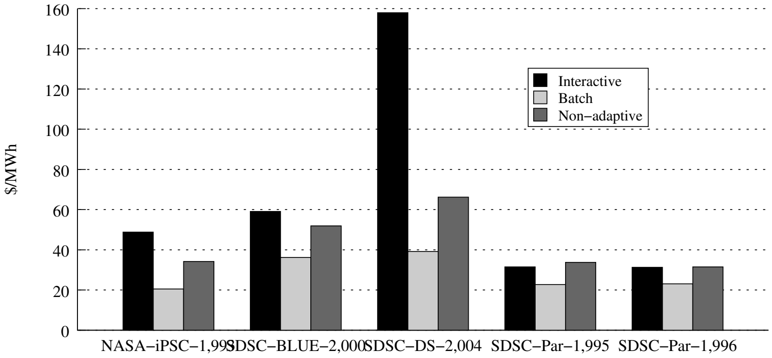
The impact of locating data centres in different time zones was investigated by adjusting the start time of the workloads. This was done by changing the start time in 1 hour increments to cover a 24 hour period. Figure 4.20 shows the potential cost savings from such an approach, illustrating that savings vary from workload to workload with some showing approximately 10% savings and others showing little effect. To effectively use such a data centre, however, the impact of added latency



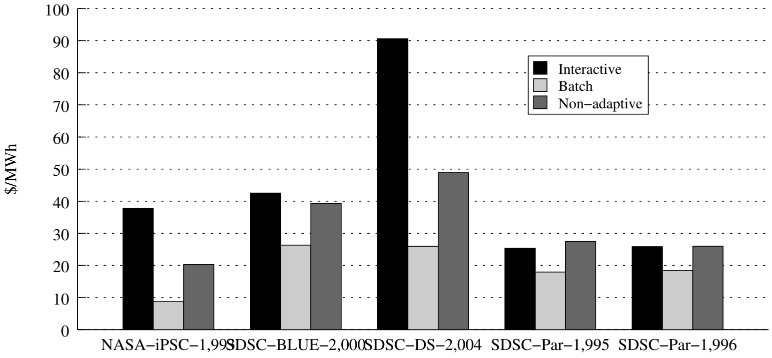
Alberta



Ontario

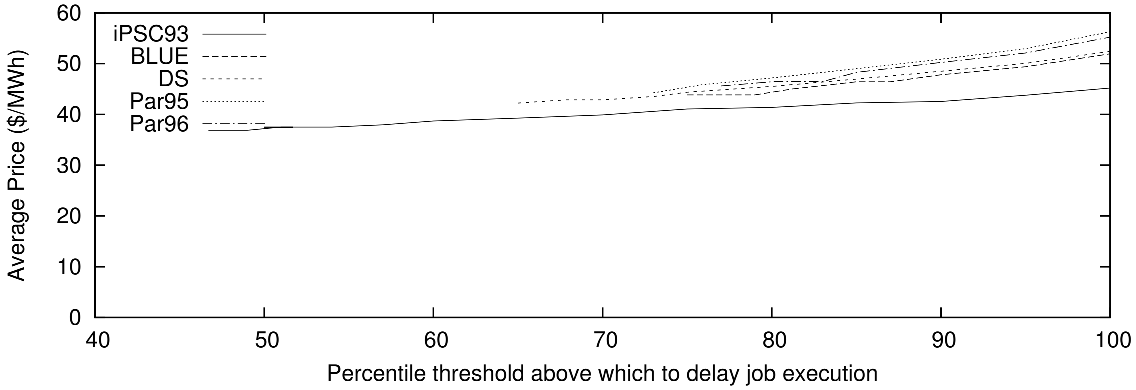


Texas−Houston

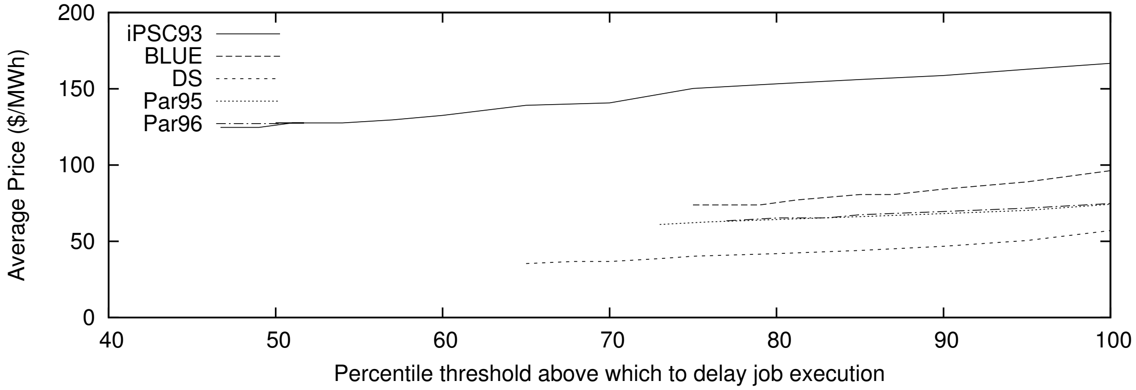


Texas−West

Figure 4.17: Simulation results — Avg. Power Cost ($/MWh).



(a) Ontario



(b) Alberta

Figure 4.18: Effect of changing the threshold above which non-interactive tasks were suspended. The differences in average power cost and volatility in Alberta and Ontario account for the differences in scale between graphs.

and network transfer costs would also need to be accounted for.

4.4 Summary

This chapter examined the potential for data centres to reduce their financial and environmental costs by accounting for change in these costs over time.

Section 4.1 first explored patterns in the financial cost of energy, renewable energy availability, and the power grid's carbon intensity. Regions using dynamic pricing show prices following a diurnal pattern. Prices are generally much higher during daytime hours. Power prices may vary rapidly, changing more than ten-fold between

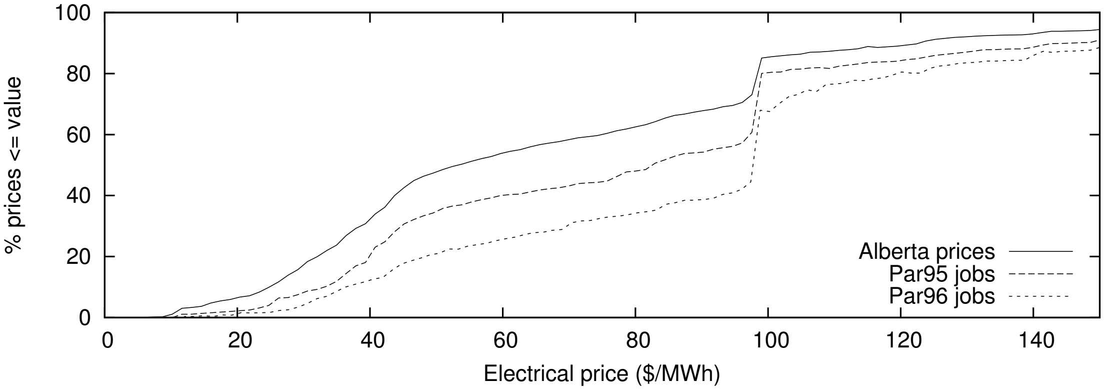


Figure 4.19: Percent of the time that prices were at most the specified value for interactive tasks in SDSC-Par95/96 traces and corresponding Alberta data.

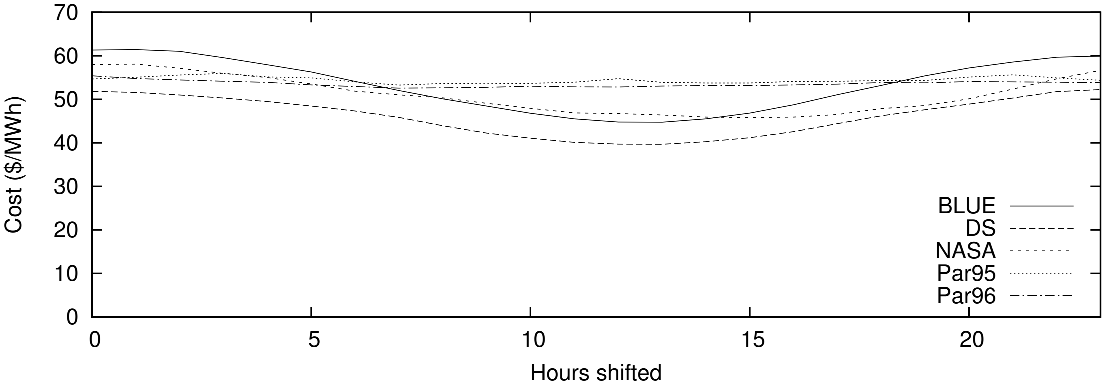


Figure 4.20: Average Ontario interactive job energy prices when the start time is shifted by the specified number of hours.

adjacent time intervals in some instances. The difficulties of renewable energy integration are straining power grids. Supply adequacy problems may occur with renewables like hydro but are a far greater challenge with wind and solar power. Renewable power generation is somewhat more stable when examining the aggregate output of renewable energy in a region. However, if the capacity of a wind farm is much larger than a local data centre's maximum power demand, the data centre may still gain an efficiency boost as energy from local renewables is likely to be available most of the time without having to alter data centre scheduling. The carbon intensity of the power grid was also evaluated, using Alberta as an example. With carbon-intensive coal power acting as a baseload power source, the estimated intensity did not vary from one hour to the next hour as much as power prices did. There might be proportionally larger shifts in emission intensity over time in grids powered primarily by renewables, using fossil fuel generators primarily for extra capacity during peak demand, but such grids would already have low emissions.

Section 4.2 estimated potential savings from altering data centre operations based on the criteria explored in Section 4.1. Changing data centre operations based on varying financial cost or changes in local renewable energy availability showed the most promising results. Bounds varied depending on the power grid involved, the characteristics of the workload to be executed, and the extent of delay tolerance. Results of analysis into adapting operations based on varying financial cost led to savings of typically 10–50%, with costs dropping further where oversupply challenges pushed prices occasionally to negative levels. Suspending the execution of tasks during periods of high energy prices could significantly reduce costs, but the threshold above which to suspend execution should be set based on both workload quality of service requirements and the characteristics of the power grid. If a data centre had local wind capacity equal to the data centre's maximum power demand, a 10–80% increase in the fraction of energy obtained from local renewables could be seen. However, lulls in wind activity can lead to excessive delays unless grid power is used to fill gaps. If the wind capacity installed nearby the data centre was much larger than the data centre's power demand, the potential to increase the use of local renewables fell significantly.

Section 4.3 further explored adjusting data centre operations based on the financial cost of power. A simulator was constructed which executed high-priority, interactive tasks immediately while deferring execution of the remaining tasks when power prices were high. Cost savings during execution of low-priority tasks typically ranged between 25–50% while continuing to provide high-priority tasks rapid access to resources.

Adjusting execution patterns based on the changing financial cost of power appears to be a promising way not only to reduce energy costs for data centres, but also to ease the integration of additional renewable energy into the grid. It seems likely that in the future power pricing will become a stronger proxy for renewable energy availability. This can already be seen in various locations as oversupply pushes energy prices down below zero at times. Altering operations based on changes in local energy availability may also enable more effective use of local renewable energy, but large wind farms generate at least a small amount of power most of the time and thus a data centre located there could obtain most of its power from the wind farm without any need to adjust its operation.

Chapter 5

Ancillary services

> What is needed most in the production of power is steadiness. Tide-mills work intermittently and, no matter how built, the intermittent feature of their work cannot be done away with. . . . The expensive use of wind power, save for pumping purposes is hedged about by somewhat similar obstacles and certainly will not come until all other more economical and convenient ways of making Nature service man have been exhausted.
>
> Thomas Edison, "Mr. Edison's Ideas on Fuel From Trees"

Data centres may be able to work with power grid operators to address the natural variation in power use over time and power plant and transmission line failures. The ancillary services used by power grids to ensure reliability were introduced in Chapter 1 and are examined in greater depth here. This chapter also examines the compensation paid for providing these services and the potential role of data centres in providing them. I collected and analyzed the power price data contained in this chapter, built the simulator used and executed the test cases presented. Hamidreza Zareipour aided in understanding the market rules applicable in the scenarios investigated. I was also responsible for the analysis of wind farms in New York state presented near the end of this chapter.

This chapter proceeds as follows. First, Section 5.1 reviews different types of ancillary services and the compensation available for providing them. Data centres may be able to reduce their power costs by providing ancillary services and this may also have other positive impacts. The particular suitability of data centres for providing ancillary services is also addressed. Section 5.2 introduces a simulator, distinct from the tools discussed in Chapter 4, for evaluating the cost savings potential and the impact on quality of service of providing ancillary services. How the various ancillary services were implemented in the simulator is discussed in Section 5.3 along with how backup generators could be used to increase profitability. An experimental plan is then outlined in Section 5.4 and followed in Section 5.5 with simulation results. Section 5.6 contrasts a follow-the-renewables strategy with providing ancillary services, outlining why the latter approach may be more effective at integrating renewables. Summarizing remarks in Section 5.7 then close the chapter.

5.1 Overview

Rather than acting independently to attempt to ease renewable energy integration, facilities like data centres may instead be fed signals by power grid operators describing how to act in order to better achieve this goal and also help preserve the power grid's stability. Section 2.1.3 outlined various services used in power grids to ensure that the goals of stable and reliable power are met, and this section provides more details of these services, including the compensation available for providing them. The suitability of data centres for providing ancillary services is also discussed.

The details of and compensation for ancillary services vary somewhat between power grids but many have ancillary service markets providing compensation for actions taken to balance the supply and demand of electricity. Power prices already exhibit high volatility and price spikes, with a relatively small fraction of time accounting for a large portion of power bills. Ancillary service pricing exhibits even higher volatility, more frequent jumps, and more extreme price spikes, though with lower average prices [242]. Prices paid for ancillary services exhibit approximately the same peak hours as in the energy markets, providing an opportunity for those providing these services to operate at lower cost during peak hours. Spikes in energy and ancillary service pricing only overlap about 20% of the time.

A general overview of ancillary services is provided [182, 183] in this section, with details of compensation available based on market data from the New York Independent System Operator (NYISO) [19, 20]. Figure 5.1 outlines the different types of programs and services discussed in this chapter, illustrating that power grids incorporate more than just energy markets.

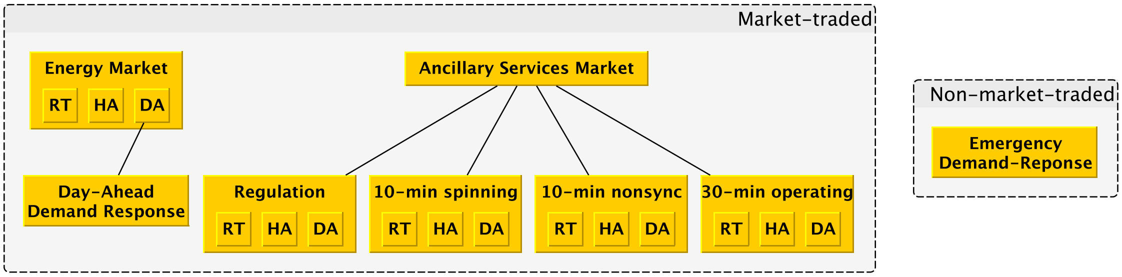


Figure 5.1: Overview of NYISO services mentioned in this chapter.

Figure 5.2 shows the average NYISO day-ahead market prices for energy and various types of ancillary services. Already differences in price are apparent with energy prices highest, regulation somewhat below, and the remaining ancillary services far beneath. These differences are due to the nature of regulation service and the higher level of commitment involved compared to the other services which form a class called contingency reserves. The regulating reserves category into which the NYISO's regulation service falls is described in Section 5.1.1, outlining how these are used to balance short-term supply and demand for power. Section 5.1.2 then elaborates on what purpose contingency reserves serve and why their compensation differs from that of regulating reserves.

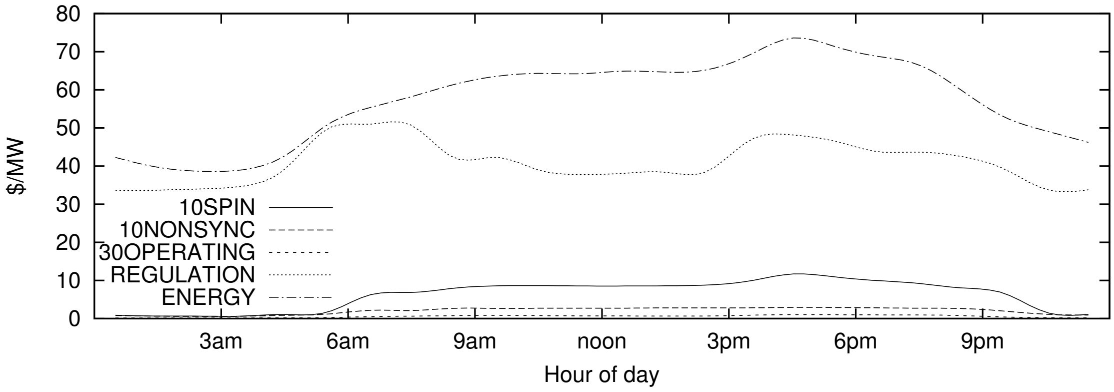


Figure 5.2: Average day-ahead energy and ancillary service availability payments.

5.1.1 Regulating Reserves

Normal minute-to-minute variations in power grid load are addressed using regulating reserves, also known as regulation. Facilities providing regulation service receive fine-grained commands from the power grid operator, typically every few seconds, instructing them how much energy to produce or consume as of that moment. The variation in power consumption expected over the short time scales addressed by regulation service is small relative to total power consumption. The NYISO requires during each time period a volume of regulating reserves which currently amount to less than 1% of the system's historical peak power consumption [15, 200]. By comparison, data centres have been estimated to consume approximately 2% of US electricity [146].

The high level of commitment required to provide regulation service results in relatively high compensation. Ancillary services are traded by the NYISO in day-ahead (DA), hour-ahead (HA), and real-time (RT) markets, and the average payments for regulation service in each can be seen in Figure 5.3. Payment is similar in each of these markets, with slightly higher average prices paid for commitments made in the hour-ahead and real-time markets.

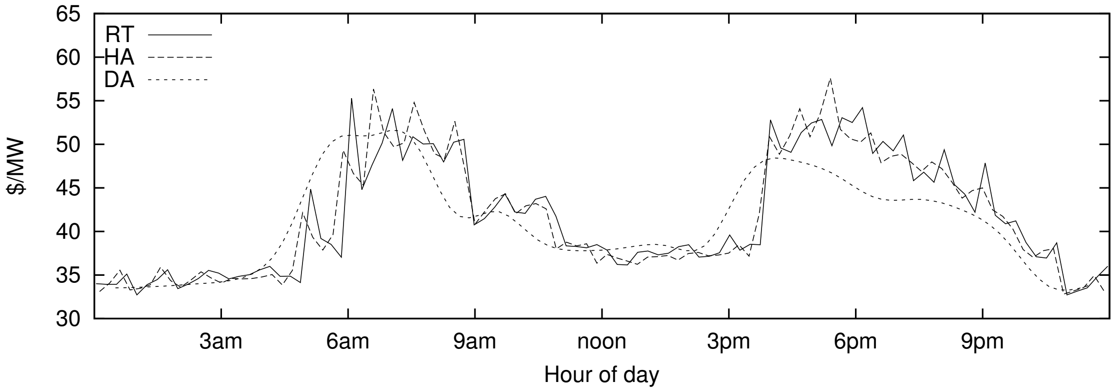


Figure 5.3: Average regulation payments by hour of day for real-time, hour-ahead, and day-ahead markets.

At least in the NYISO, providers of regulation service are paid based on the regulation capacity provided. The regulation capacity is the size of the range within which the power grid operator is able to issue fine-grained commands specifying particular levels of power production or consumption by the entity providing the service. When it provides regulation service, a facility can expect to be requested to operate power levels that average over the long-term the mid-point of the capacity range offered. Though entities providing regulation service are paid for the regulation capacity offered, they are also paid or charged for the energy in the regulation service range that they produce or consume. If a power consumer were to be paid for 1 MW of regulation service, it could also expect to be charged for approximately 500 kW of related power consumption. Regulation prices have fallen somewhat in the NYISO in recent years, but in recent history a facility able to efficiently and rapidly adjust its power consumption could expect payment for providing regulation service to, on average, equal or exceed the cost of regulation-related energy consumption during this time.

Regulation service is typically provided by fossil-fuel power plants. Conventional power plants are most efficient when operating at a constant level. Baseload power plants are designed to operate in such a manner and perform poorly when required to rapidly scale their output up or down. Peaking power plants are designed to scale their output up and down more effectively than baseload power plants but are overall less fuel-efficient.

Increasing the use of renewable power from wind and solar sources significantly increases the short-term variation that must be compensated for. However, forcing baseload natural gas [139] and coal [64] plants to accommodate this by rapidly altering output can have significant negative consequences. Potential consequences include significant increases in NO_x and SO_x emissions, increased fuel consumption, additional wear and tear, and an increase in CO_2 emissions which may occasionally exceed the CO_2 savings from incorporating a volume of renewable energy.

Finding alternative, more environmentally friendly ways to supply the needed regulation capacity is desirable. Demand-side resources are permitted to provide regulation service in an increasing number of power grids including the NYISO, although they currently only supply a small fraction of it [195]. Both battery-based [218] and flywheel-based [152] energy storage have also been used to supply regulation capacity. With data centres typically possessing limited local energy storage in UPS form, using UPS capacity as a buffer while adjusting data centre power states at a coarser granularity may enable data centres to provide regulation capacity more effectively than others, while potentially reducing battery lifetime challenges.

The Federal Energy Regulatory Commission (FERC) noted that efficiency increases from using faster-responding resources like energy storage and demand-response for regulation service can reduce energy prices and lower levels of regulation reserves required to maintain grid stability [92]. The FERC also ruled that paying providers of regulation service solely on the basis of capacity offered is discriminatory against such faster-responding resources and proposed to revise rules so that providers would also be paid for the speed of power consumption or generation changes made in response to power grid operator requests. Although the NYISO already calculates a “regulation performance index” to partially compensate for fast-responding resources, this FERC ruling may result in higher revenue available to facilities like data centres which are able to rapidly adjust their power use.

5.1.2 Contingency Reserves

Whereas regulating reserves address normal variation in the power grid, contingency reserves are intended to respond to power plant or transmission line failures. Adequate volumes of contingency reserves must generally be purchased to handle the failure of several of the largest power plants on the grid. Compared to regulating reserves, contingency reserves need only wait until called upon by the grid operator on rare occasions.

Contingency reserves are differentiated primarily by the time it takes them to begin responding to requests as well as the time period over which they must be able to deliver their full commitment to supply power to the grid or reduce demand upon it. Reserves able to begin responding immediately are known as spinning reserves, whereas reserves that require more time to synchronize before being usable are known as non-spinning or non-synchronous reserves. The NYISO has both “10-minute spinning” and “10-minute non-synchronous” reserves. Replacement reserves are used to replenish other reserve categories in the event they are required to act. The NYISO also possesses a volume of 30-minute operating reserves for longer term reactions.

The ancillary service prices shown in Figure 5.2 include the three types of contingency reserves described in the previous paragraph, but their prices are lower than energy or regulation service since they are seldom activated. Figure 5.4 shows a zoomed-in view of the contingency reserves.

The higher the requirements on the resource in question, the higher the compen-

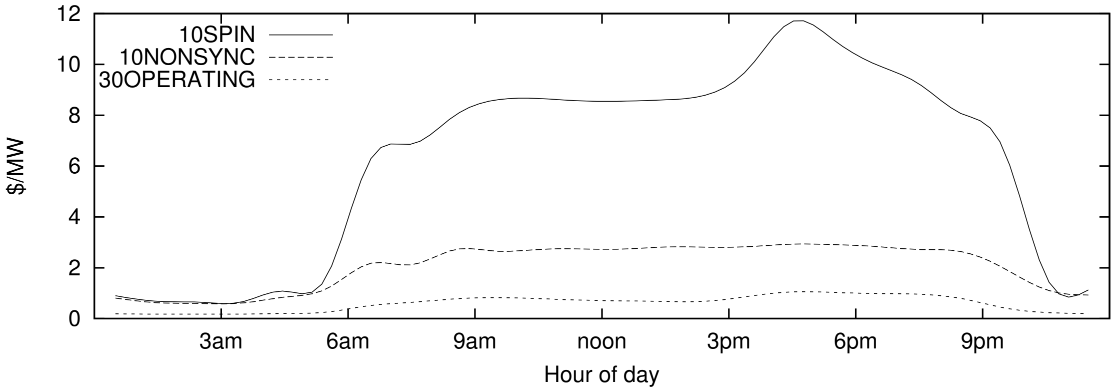
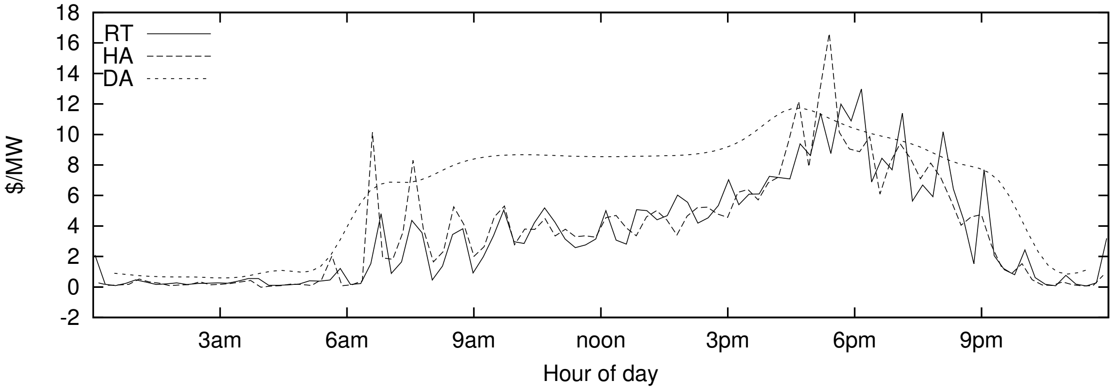


Figure 5.4: Average day-ahead ancillary availability payments for non-regulation ancillary services by time of day.

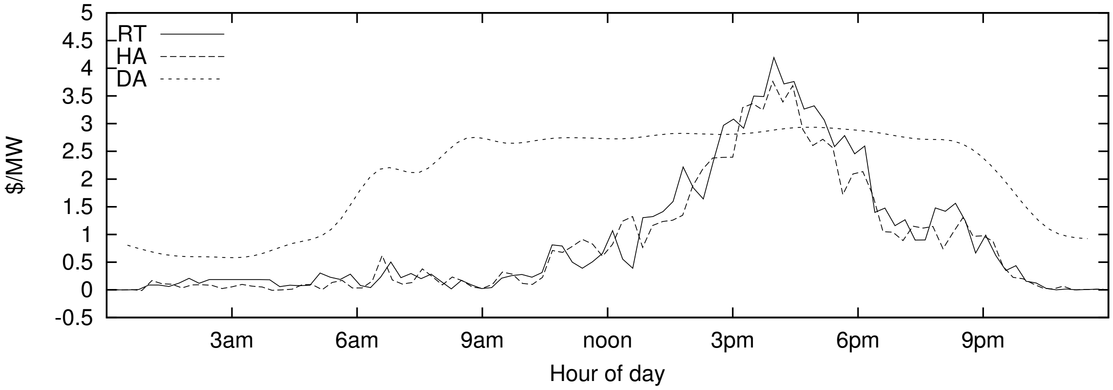
sation received. This applies as well to the different time horizons over which the NYISO purchases contingency reserves. Figure 5.5 shows the difference in compensation received by the 10-minute spinning, 10-minute non-synchronous, and 30-minute operating reserves in the day-ahead, hour-ahead, and real-time markets. Those committing further in advance to reduce their power consumption upon request during a particular time interval are compensated more highly. For non-synchronous contingency reserves, what little compensation was available in markets other than the day-ahead was concentrated primarily from mid-afternoon to the early evening.

The premise that the higher the requirements upon the resource the greater the compensation explains not only the difference between the different classes of contingency reserves used by the NYISO but also the difference between pricing of contingency reserves and regulating reserves. Regulating reserves receive commands every few seconds instead of being activated infrequently like contingency reserves.

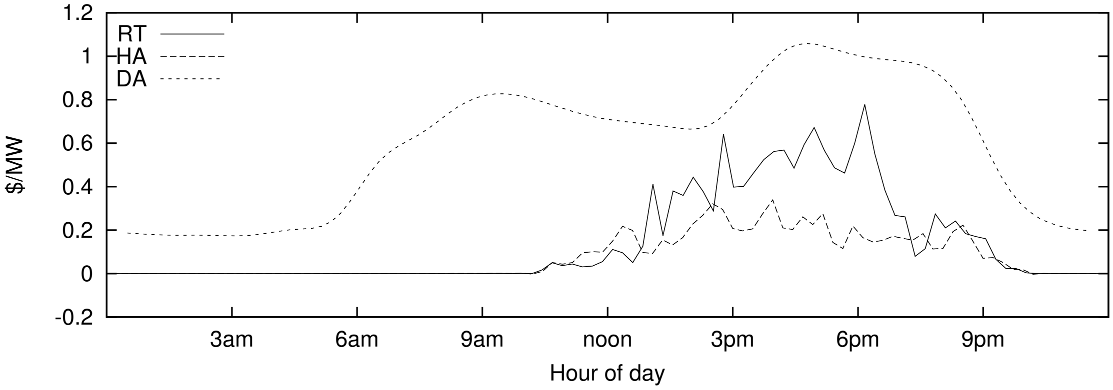
Very limited information is available regarding the frequency with which contingency reserves are called upon by power grid operators [122, 143]. In 2002 the NYISO activated some fraction of its contingency reserves 239 times, on average for under 11 minutes. ISO-NE activated contingency reserves 19 times in 2005, averaging less than



(a) 10 minute spinning



(b) 10 minute non-synchronous



(c) 30 minute operating

Figure 5.5: Availability payments for contingency reserves in the real-time, hour-ahead, and day-ahead markets. Note that the range of the Y-axis in each figure is independent.

11 minutes in each instance, and CAISO used reserves 26 times in 2005 for an average of under 9 minutes in each case. Most reserve activations lasted approximately 10 minutes. Grid operators are primarily concerned with ensuring adequate supplies of each ancillary service to meet reliability guarantees, using optimization tools to dynamically choose specific facilities to call upon to reduce their power consumption. The likelihood that a particular reserve is activated may vary dependent on some technical parameters and its location in the power grid.

The “10 minutes” or “30 minutes” in the names of NYISO services and the labels “spinning” and “non-synchronous” are primarily indicators of the time period in which resources must be able to deliver their full commitment and the latency before they begin to adapt. Insufficient information was available to estimate the extent of differences in how frequently or for what duration these different services were activated, although it seems likely that reserves with faster response time requirements will be activated more often, with the hope that problems are resolved before slower resources would be needed.

5.1.3 Other Reliability-Related Services

The market-traded ancillary services mentioned in the previous section are not the only means by which the NYISO provides compensation for actions taken to maintain reliable power. In addition to permitting large energy consumers to participate in its ancillary services markets, the NYISO also has several demand-response programs. This section outlines two such programs: emergency demand-response and day-ahead demand-response.

Emergency demand-response is called upon only when other reserve categories are likely to be inadequate. Participants generally receive day-ahead requests to reduce demand rather than acting in response to failures as contingency reserves do. NYISO historical data [22] suggests that those offering emergency demand-response may be called upon only 0.1–0.2% of each year, with periods of a year or more in which a facility might not be called upon at all. Unlike market-traded contingency reserves, organizations providing emergency demand-response receive no compensation for their availability. They are instead paid premium rates for energy reductions when called upon. $500/MWh or the real-time electrical price, if greater, is received for each MW by which consumption is reduced. Payment is provided for a minimum of 4 hours, even if the actual time for which reductions were requested is shorter.

Day-ahead demand-response allows organizations to bid reductions into day-ahead energy markets to reduce the cost of energy since energy pricing is generally set based upon the marginal cost of the most expensive unit of energy produced.

5.1.4 Data Centres as Ancillary Service Providers

Data centre characteristics make them well suited to provide ancillary services. Computing equipment has become increasingly energy proportional in recent years enabling rapid adjustment of power consumption. Data centres typically have both UPSes for local energy storage and backup generators serving as a local source of power. As high-speed, low-latency networks enable data centres to supply services to relatively distant customers, data centres can be placed in strategic locations where stability problems are most common. Such networks also make it possible to migrate workload from one data centre to another or redirect users to an alternate facility. The ability to defer execution of some workloads to a later time further increases flexibility. With ancillary service prices peaking at approximately the same time of day as power costs, providing ancillary services may enable data centres to reduce the cost of providing service at peak times.

Power costs account for much of the total cost of ownership of computing equipment and this has caused data centre operators to invest significantly in reducing their power costs and environmental footprint. As Section 5.1.1 noted, data centres are permitted to provide ancillary services in many power grids and the Federal Energy Regulatory Commission has asserted that the use of faster-responding resources for regulating reserves would increase efficiency, lead to lower power prices, and enable power grids to operate with lower levels of reserves. Power grid operators may also be able to integrate renewable energy into the grid more rapidly. Through participation in ancillary service markets, data centres can reduce their power costs and have a greater impact on renewable energy integration than by acting alone.

5.2 Simulator Overview

This section provides an overview of a discrete-event simulator constructed to evaluate the extent to which providing ancillary services may impact net energy costs and workload wait times. Section 5.2.1 outlines the modules making up the simulator. How servers and workload are represented is then discussed in Section 5.2.2. This provides context for Section 5.3 which outlines how specific services were implemented and how equipment like backup generators might be employed for better results. Further details of the simulator are available in Appendix A, Section A.4.

5.2.1 Modules

The simulator was comprised of several different types of modules. Here each of these are explained, beginning with a group of modules that supply pricing information. The module delivering workload information is described next, followed by an outline of the ancillary event generator module used to simulate the behaviour of a grid operator calling upon resources. Finally a cluster controller module is described which receives from the other modules and simulates the operations of a server cluster. The relationships between the different modules are summarized in Figure 5.6.

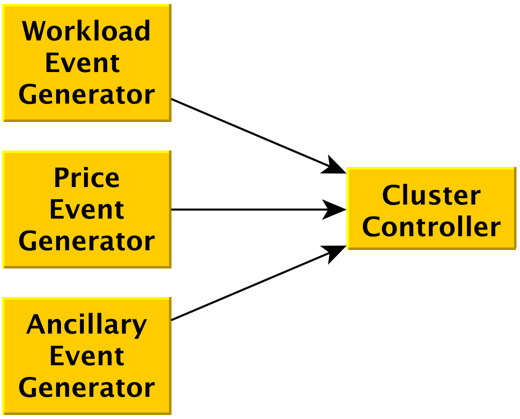


Figure 5.6: Simulator modules.

Pricing Modules

This family of modules provided pricing information for NYISO energy and ancillary service markets, supporting all NYISO regions as well as day-ahead, hour-ahead, and real-time markets.

In all simulations, price information for the appropriate region and market type was fed to the cluster controller. For data centres providing all but emergency demand-response, the current ancillary service market price for the type of service being provided was also supplied. Since emergency demand-response compensation is based on the current real-time price of power under NYISO rules and regulation service pricing calculations also necessitate real-time energy prices, the real-time energy price was also separately supplied to the controller when these services were provided.

Workload Event Generator

This module supplied the cluster controller with details of when tasks were originally executed on the system being simulated. Traces from the Parallel Workloads Archive [190] detailing the wait time of each task prior to being allocated resources were used. An event was sent to the cluster controller each time a task began executing on the original system, containing information about its wait time on the original system, the number of processors it requested, and its actual execution time.

Ancillary Event Generator

Excepting the always-active regulation service, the ancillary services considered here only require action when called upon by the grid operator. An ancillary event generator module was used to generate events simulating the grid operator's occasional activation of the ancillary service as well as informing it when it could return to its normal behaviour.

Non-regulation ancillary services were activated randomly throughout the simulations, with each activation causing a message to be sent to the data centre from the grid operator informing them of the need to reduce consumption. Once the data centre was no longer required to reduce load upon the grid, another message was sent by the grid operator indicating that normal operations could resume. Intervals between ancillary service activations were generated using an exponential distribution calibrated such that the ancillary service was active approximately a specified percentage of the time. Due to the limited information available about reserve activation patterns, a simplified model was used in which each type of reserve was activated for a fixed length of time whenever called upon. 10-minute spinning, 10-minute non-synchronous, and 30-minute operating reserves were activated for 10 minute periods. Emergency demand-response was activated for 4 hours at a time.

Cluster Controller

This module was the core of the simulator, receiving event information from the other modules and translating it into simulated cluster behaviour. It also calculated net energy costs and the duration tasks were queued awaiting resource or suspended during ancillary service activations. The behaviour of this controller is outlined in more detail throughout the remainder of this section as well as in Section 5.3.

5.2.2 Server and Task Representation

The number of processors presently assigned to tasks was the primary information maintained about the state of servers in the simulated cluster. The simplifying assumption was made that only those processors assigned to a job submission consumed power. Other researchers have reported [104] that power consumption could be reduced by over 90% in their servers in less than 10 seconds by shifting them to a low-power state. Heuristics could also be used to completely power down systems. Thus with this assumption the vast majority of the system's power consumption range could be addressed. This assumption also avoids negative impacts from tasks at the beginning or end of the workload log potentially not being fully recorded and from scheduled maintenance.

Upon receipt of a request from the grid operator to reduce power any presently executing tasks were suspended. The simplifying assumption was made that servers required 10 seconds of power consumption to transition to or from a state in which they would be able to preserve their current memory contents and device state while requiring no further power. Upon notification from the grid operator that the system could resume normal operation, previously active servers required 10 seconds to resume operation. New tasks that had arrived when the servers were suspended were also launched then, as long as sufficient processors were available to meet each task's requirements.

When notified by the workload generator module that a task had begun execution on the original system, this task was added to a queue of unfinished tasks. Information stored about the task included the number of processors it required, its actual execution time, and the amount of time that it had waited on the original system before execution. This information was then used at the simulation's end to calculate the average amount of time that each task had been required to wait. Each time during the simulation run that an event was received, incremental energy costs were tallied based on the number of processors in use and the power price. Any offsetting revenue received from providing ancillary services was also tabulated at this time, with the details of how this was done outlined in the next section.

5.3 Implementation of Ancillary Services

This section outlines how the amount of ancillary services supplied by data centres in each instance was evaluated, and how compensation for these services was determined. Section 5.3.1 first discusses regulation service, Section 5.3.2 focuses on contingency reserves, and Section 5.3.3 looks at emergency demand-response. Finally, Section 5.3.4 addresses how backup generators might be used to increase ancillary service revenue.

5.3.1 Regulation Service

As described in Section 5.1, regulation services are constantly active, required to match their energy consumption or production to the level requested by the grid operator every few seconds. Power consumers like data centres are paid the ancillary market rate for each MW of regulation capacity provided but must pay the real-time market rate for power consumed when providing the service.

For the purposes of the simulations, during time periods that a data centre wished to operate at $x\%$ utilization, where $x < 50\%$ of capacity, it was assumed to provide regulation capacity corresponding to $2x$ its normal power draw. Since the expected output requested by the grid operator averages half the capacity offered, such a facility could expect on average to be told to consume $x\%$ of electricity. During times in which a data centre wished to operate at over 50% utilization, it would purchase a portion of its energy consumption on the regular electrical markets, offering only a fraction of its capacity — $2 \times (1 - x)$ — for providing regulation service. For example, a data centre wishing to average power draw of 75% of capacity would purchase 50% of capacity on the energy markets and provide regulation capacity corresponding to the remaining 50% of capacity, resulting in power consumption averaging 75%.

Facilities providing regulation service may receive payments relatively close to the fees paid for the corresponding energy used. During any time period where the price for providing regulation service is at least 50% of the real-time cost for electricity and the system's desired utilization point is less than 50%, revenue from offering the full system capacity for regulation service would be expected to receive greater payment from regulation services than paid for energy over the long term under assumptions here. This applies even if nothing more than busy-waiting is done. A supercomputer might use busy-waiting or execution of workloads like SETI@home to boost power consumption as needed to enable it to provide more regulation service capacity. Doing so could increase its average level of power consumption, thereby potentially reducing wait times that might result from the system reducing consumption as part of its obligation to provide regulation service.

5.3.2 10-Minute Spinning, 10-Minute Non-Synchronous, and 30-Minute Operating Reserves

Until called upon by the grid operator to reduce power, the simulated cluster operated as normal. During this time it paid for energy according to energy market rates but received a payment equal to the amount by which it would be able to reduce power upon request multiplied by the market rate for the ancillary service provided. When requested to reduce power by the grid operator, the cluster transitioned active servers to a sleep state. If it had bid generator capacity into the ancillary market to provide energy it would be paid the real-time market rate for power production at this time, but no energy payments were received for reductions in power consumption. When informed by the grid operator that normal operation could be resumed, the servers were returned from their sleep state back to a normal operating state.

5.3.3 Emergency Demand-Response

Data centres providing emergency demand-response operated as normal until called upon by the grid operator to reduce demand. In the NYISO, facilities providing emergency demand-response pay for power as normal in the energy markets, and are not paid merely for their registration in the program. When called upon, the simulated clusters transitioned active servers to a sleep state. Payments corresponding to the maximum of \$500/MWh or the current real-time energy price were received by the cluster corresponding to the amount by which it reduced its expected energy consumption during this time. When informed that normal operation could resume, servers were transitioned back to normal operation.

5.3.4 Utilizing Data Centre Backup Generators

Data centres typically possess backup generators to provide emergency power in case of grid failure. By powering up generators in response to a request from the power grid operator rather than powering down servers, these generators could be used by a data centre delivering ancillary services to eliminate the need for it to shut down servers when that is not desirable. Ancillary services could thus be delivered without changes to cluster scheduling being required.

The option to bid data centre backup generator capacity into the ancillary services market was evaluated for non-regulation ancillary services. Here it was assumed that the cost of operating local generators was equal to the market price of electricity, and that facilities had generation capacity equal to their maximum power draw. Real-world data centres may instead have an $N + 1$ or $2N$ backup generator configuration depending on the criticality of their systems.

Facilities providing these services increased the capacity bid into the ancillary service market by an amount corresponding to a percentage of their maximum power draw. They would then supply this energy from their local backup generators. Local stored energy in UPSes could be used to immediately feed energy back into the grid upon request, allowing them to supply ancillary services with tighter response time requirements.

Computing equipment could still be safely shut down even when bidding data centre backup generator capacity into the ancillary services market. Stored energy in UPSes could be used when servers are being powered down or services permitted longer response time requirements could be provided. A facility's bidding behaviour could be made dependent on the priority of the workload then being executed in order to ensure that critical tasks still executed in time but allowing greater ancillary revenues when only low priority tasks are available for execution.

Given the relatively small fraction of time that contingency reserves are activated, using data centre backup generators to provide ancillary services may not require them to operate much more often than the test runs already performed to verify their operational ability. Technical faults like generator failures may also excuse the data centre from having to pay penalty fees.

5.4 Experimental Plan

Experiments were performed to evaluate how providing various types of services impacted energy costs and the wait times experienced by tasks run on the system. Ancillary services evaluated included: regulation, contingency reserves (10 minute spinning reserves, 10 minute non-synchronous reserves, 30 minute operating reserves), and emergency demand-response. Regulating reserves are continuously active, and given the lack of information known about their operational patterns no attempt was made to estimate how these services affected average wait time. Giving the limited data available and the challenges involved in incorporating additional renewable energy into the grid, experiments were performed in which the contingency reserves were activated approximately 0.5% of the time. Sensitivity analysis was performed on this assumption. Similarly the assumption was made that emergency demand-response was active 0.1–0.2% of the time and sensitivity analysis was also performed regarding this assumption. Further experiments were performed to evaluate the impact of using backup generator capacity to increase ancillary service revenue. Generators may be more likely to be used to eliminate any delays introduced by providing ancillary services, but could be used to instead increase revenue during the execution of low-priority workloads.

Data used in experiments was taken from the following locations. Both real-time and day-ahead market data from a subset of the NYISO's pricing regions was used: West, Central, New York City, and Dunwoodie. Table 5.1 shows workloads from the Parallel Workloads Archive used for the experiments described in this chapter. Start times were adjusted to fit the workloads into each of the years 2006–2010, matching the workload's local time of day, day of week, and approximate day of year.

Table 5.1: Workload logs used for evaluating the cost savings potential of providing ancillary services.

Log	**% Utilization**	**Length**
ANL-Intrepid	59.6%	8 months
DAS2-fs0	14.6%	12 months
RICC	87.2%	5 months

5.5 Results

This section outlines the results of the simulations conducted. First, the impact of providing regulation service or contingency reserves on the net energy costs of a system and the wait times experienced by tasks was evaluated for ancillary services sold in both day-ahead and real-time markets. Sensitivity analysis was performed on the assumption that contingency reserves are activated 0.5% of the time. How incorporating generator capacity into contingency reserve bids affected net energy costs was then examined. Finally, the impact of providing emergency demand-response on net energy costs and wait times is evaluated and sensitivity analysis is performed on the assumption that this service is activated 0.1–0.2% of the time. Net cost and average wait times were normalized for each test by dividing the cost and average wait using figures from an otherwise equivalent test in which no ancillary service was provided. No information regarding changes in wait times was reported for regulation service since the ability of the grid operator to set fine-grained consumption patterns makes it difficult to analyze their short term impact.

Figure 5.7 shows the impact on net energy costs for facilities bidding in day-ahead markets to offer regulating or contingency reserves. Results presented in the figure assume that data centre backup generators are unused and that contingency reserves were activated 0.5% of the time. The diagram shows an average across the different pricing regions and the different years evaluated of the normalized net cost for these workloads. Providing contingency reserves was found to result in approximately 1–12% cost savings, with greatest savings from the 10-minute spinning service. Much greater savings were possible if providing regulation service, possibly reducing power costs below zero as regulation service revenue could exceed energy costs.

Figure 5.8 shows the impact on the average amount of time tasks spent waiting prior to execution if different assumptions were made about the percentage of time that these services were activated. It shows the impact if facilities providing the 10-minute spinning service, 10-minute non-synchronous service, or 30-minute operating reserve in day-ahead markets were activated 0–1% of the time. Due to the payment

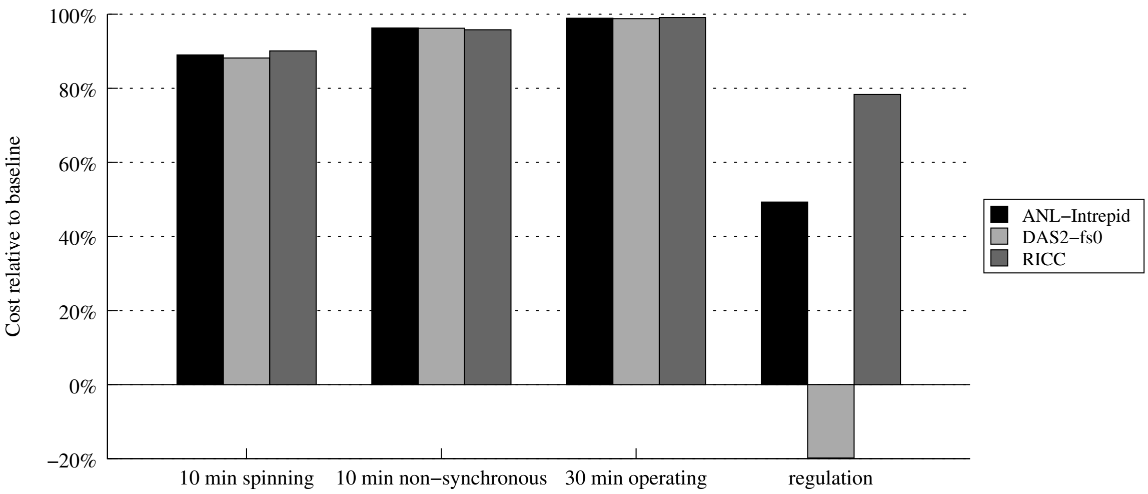


Figure 5.7: Relative cost when providing ancillary services in day-ahead markets.

structure previously outlined for contingency reserves, varying the assumptions in the range 0–1% had little impact on energy costs. As a result of scarce information, differences in the activation rates for the different contingency reserves are unknown but, as mentioned previously, the most expensive classes of reserves might be expected to be activated most often.

As shown in Figure 5.7, significant cost savings are possible when providing many of these services with minimal impact upon delays. For example, facilities providing 10-minute spinning service executing the ANL-Intrepid workload experienced an average 11.0% reduction in net energy costs with an average 2.1% increased wait time. Net energy costs might be reduced near or even below zero for low-utilization computer systems providing regulation service as was demonstrated in the results with the DAS2-fs0 workload which averaged 14.6% utilization.

Bidding capacity in real-time markets as opposed to day-ahead markets reduced cost savings. Average regulation service prices differed little between day-ahead and real-time markets, but the other reserves faced a significant decrease in availability payments. A comparative analysis of availability payments showed decreases of 38%

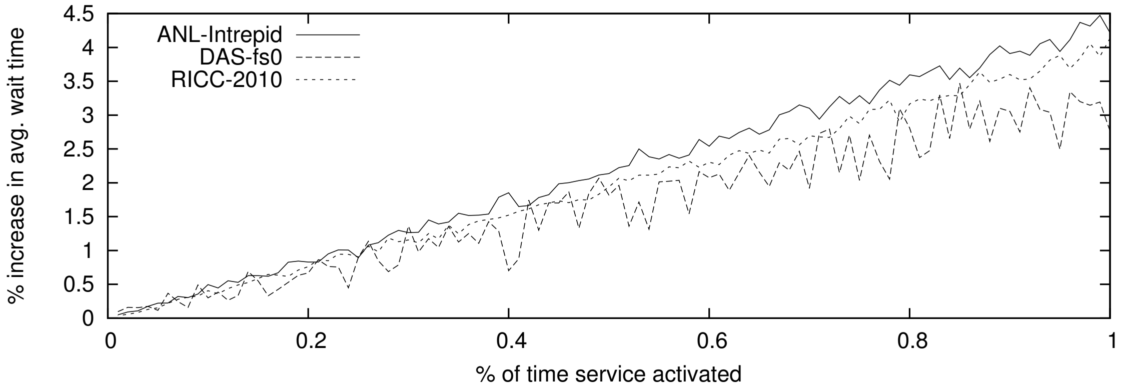


Figure 5.8: Effect of changing the percent of time the contingency reserves are activated on the average amount of time tasks spent waiting.

for 10-minute spinning service, 50% for 10-minute non-synchronous service, and 73% for 30-minute operating services which are reflected in Figure 5.9.

Bidding generation capacity into the ancillary services market significantly decreased net energy costs, increasing linearly with the amount of data centre backup generator capacity bid. For the ANL-Intrepid workload, cost savings rose from 11.0% to 29.0% when generator capacity equal to their maximum power draw was bid to provide additional volumes of 10-minute spinning service. The 10-minute non-synchronous service similarly saw energy cost savings rise from 3.7% to 9.8% and the 30-minute operating reserve saw an increase from 1.1% to 3.0% savings. The largest decreases in net energy costs were seen with the lowest utilization workload but significant decreases were seen in all cases.

Emergency demand-response is activated when normal ancillary services are inadequate to preserve reliability guarantees. Historically, emergency demand-response has been activated only 0.1–0.2% of the time. Figures 5.10 and 5.11 show, respectively, the impact upon net energy costs and delays if the percentage of time emergency demand-response is activated varies between 0–0.5% during simulations. Revenue from emergency demand-response and its impact on wait time vary more than for

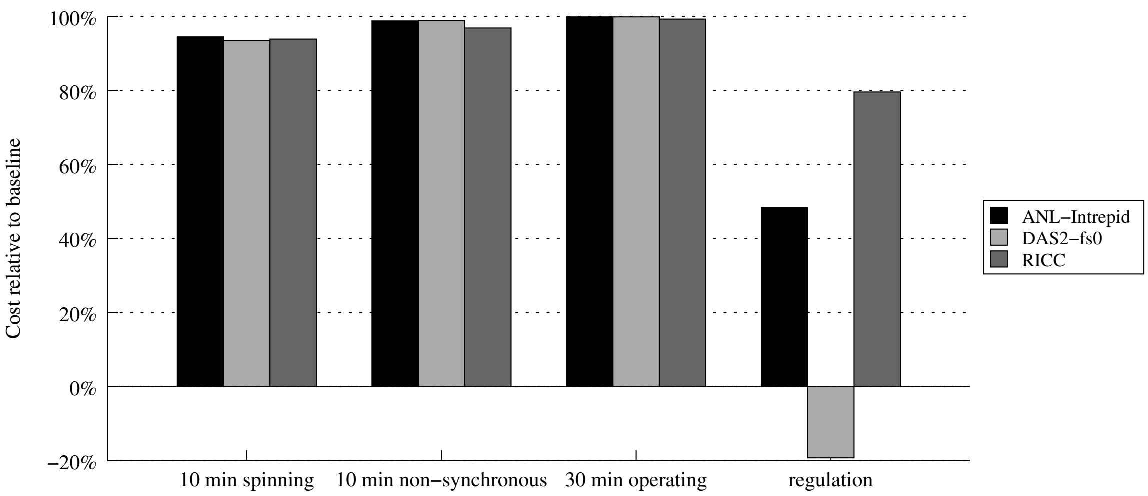


Figure 5.9: Relative cost when providing ancillary services in real-time markets.

other services. Revenue is more volatile since emergency demand-response is activated less frequently and paid when activated versus the availability payments paid for other services. Given that activation time is concentrated in far fewer and much longer periods, this also has a more volatile impact upon wait time. The lowest utilization DAS-fs0 workload saw both the greatest volatility and greatest percentage-wise change in wait time. This is to be expected since tasks executed on a high-utilization cluster might expect to have longer average wait times in normal circumstances. When a disruption occurs, it also has a larger impact on waiting tasks than a contingency reserve activation.

The differing impact of providing contingency reserves and emergency response is highlighted in Figure 5.12, which plots increases in wait times when executing the RICC workload against the percentage of time particular services are activated. Each point shows the result of a run of the simulator varying only the frequency with which the service is activated and the random number generator seed. The simulator runs for contingency reserves produced consistent increases in wait time with only a small spread, whereas the average performance of emergency demand-response in

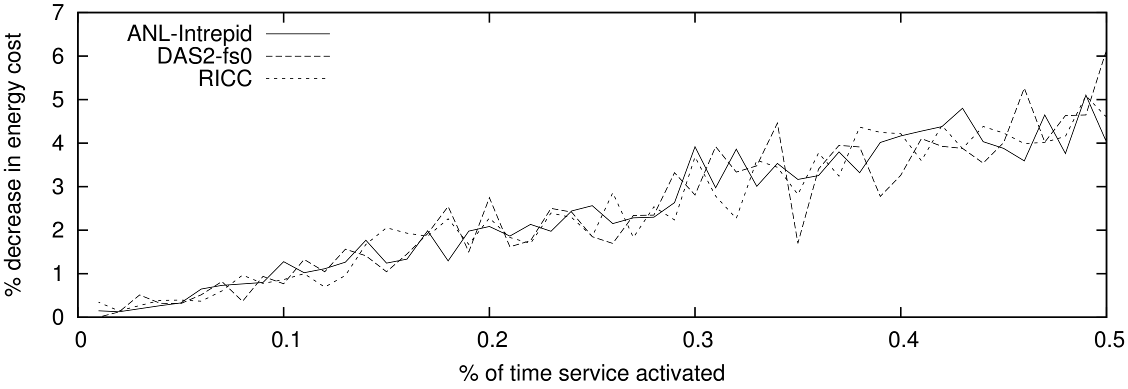


Figure 5.10: Effect of the change in percentage of time the emergency demand-response program is activated on relative energy cost.

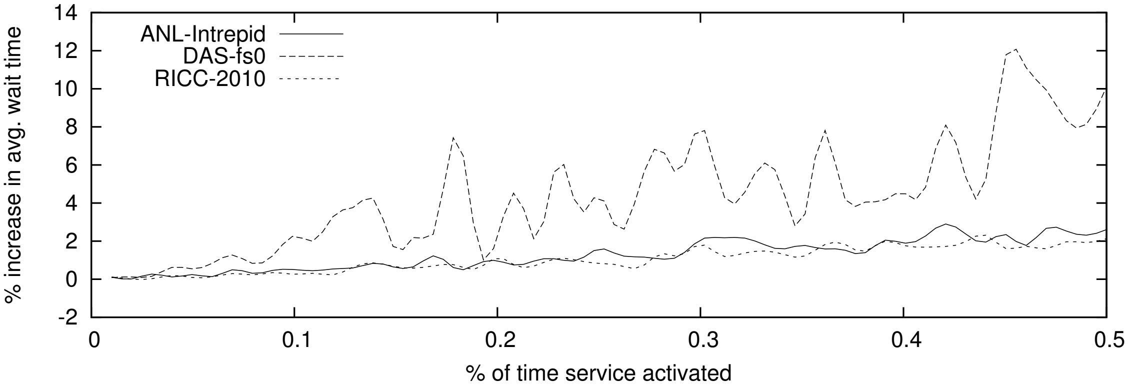
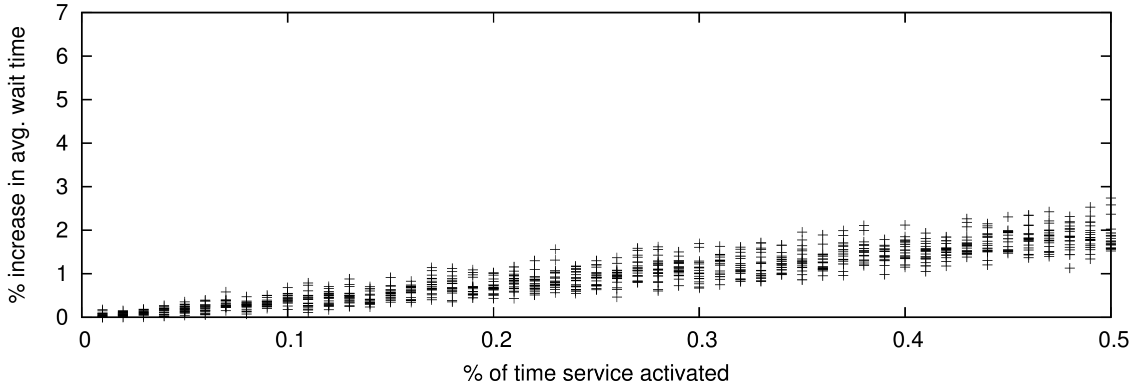


Figure 5.11: Effect of the change in percentage of time the emergency demand-response program is activated on relative wait time

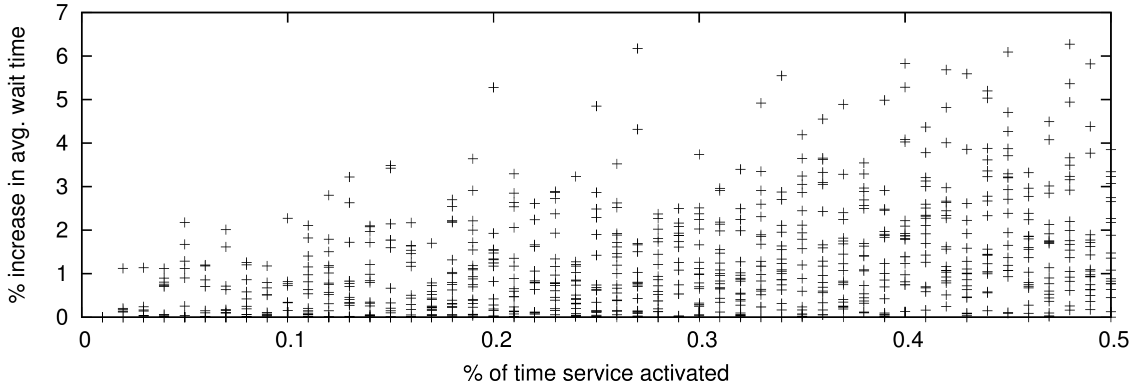
simulator runs is heavily influenced by outliers in which emergency demand-response was activated significantly more than in the other trial runs.

5.6 Versus a Follow-the-Renewables Approach

This section explores limitations in optimizing demand in response to local renewable energy production, exploring why providing ancillary services (or a price-responsive approach to power utilization) may be a more effective approach. This expands upon



(a) Contingency Reserves



(b) Emergency demand-response

Figure 5.12: Change in wait time in the RICC workload when contingency reserves and emergency demand-response are each activated with specified probability.

Chapter 4, which demonstrated that although wind turbines usually have capacity factors below 40%, they typically produce some power more than just 40% of the time. Data in this section is derived from the National Renewable Energy Laboratory's Eastern Wind Integration Dataset [179], focusing on 66 candidate wind farms for New York, from where the power market data in this study came. The proposed wind-farms ranged in capacity from 99.6–1085.5 MW, with 90% of proposed projects having capacity below 260 MW.

One argument in favour of trying to optimize power use based upon local energy production is that doing so reduces transmission loss. However, even when local wind levels are comparatively low it remains likely that some energy is currently being produced. Figure 5.13 shows for each of the sample wind farms the fraction of a constant power demand that they would be able to fulfill prior to requiring energy delivered from another more distant site. Even if data centres did not account for the output of a local wind farm in scheduling, sufficient energy would likely be produced locally to power them a majority of the time even where large data centres are concerned.

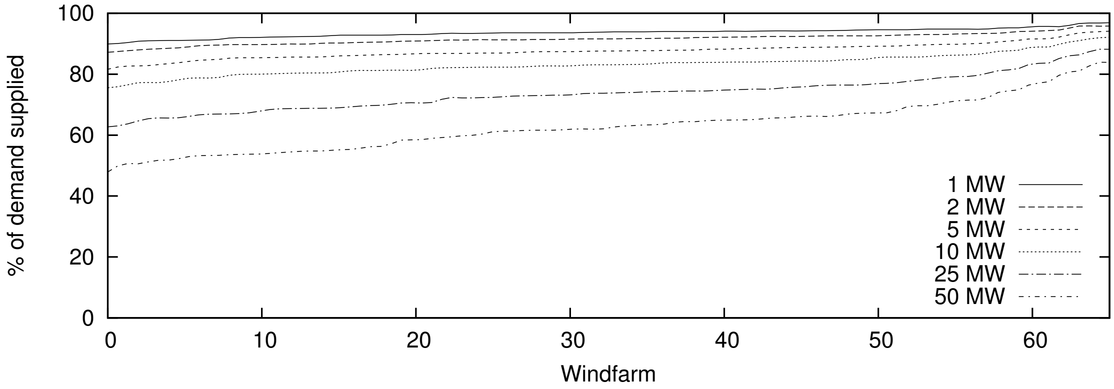


Figure 5.13: The percentage of a constant power load that the proposed New York wind farms can sustain.

Another issue that a follow-the-renewables policy ignores is changing demand in the power grid. Figure 5.14 outlines the daily change in wind generation between 6:50am and 7:00am, a time in which demand is typically increasing as people are waking up or preparing for work. Though demand is generally increasing at this time, wind generation can vary significantly upward or downward in the interval. A naive follow-the-renewables approach failing to account for human activity may thus increase grid instability at times.

Wind farm output across larger geographical areas is generally somewhat more stable, though there is still a significant correlation between local wind turbine output and the total output of all the proposed wind farms in New York. This can be seen in Figure 5.15, which looks at average, minimum and maximum wind generation across

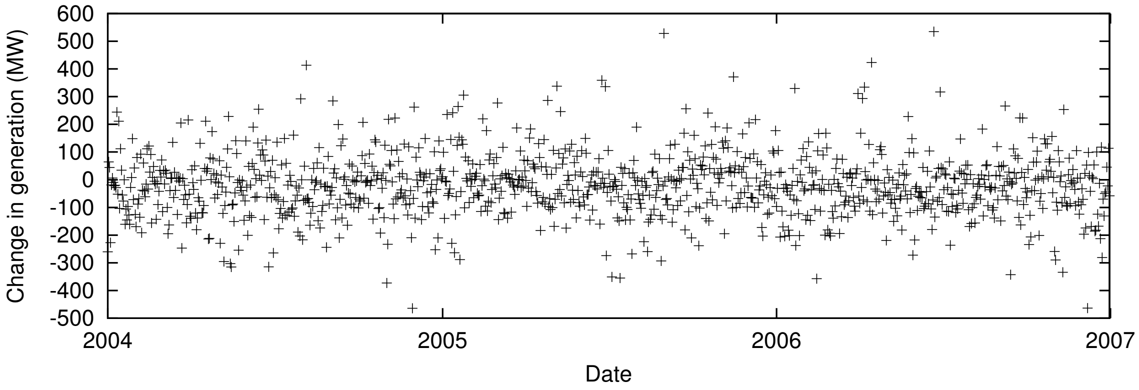


Figure 5.14: The daily change in average wind generation between two 10-minute periods at a fixed time of day.

all the proposed New York wind farms when local power production is at a particular fraction of the installed capacity. This figure demonstrates that although there is a correlation between local wind energy and that produced throughout New York state, the local output at any given point in time may differ from the wider view. As additional long-haul wind transmission lines are constructed, wind energy production is likely to become even more stable.

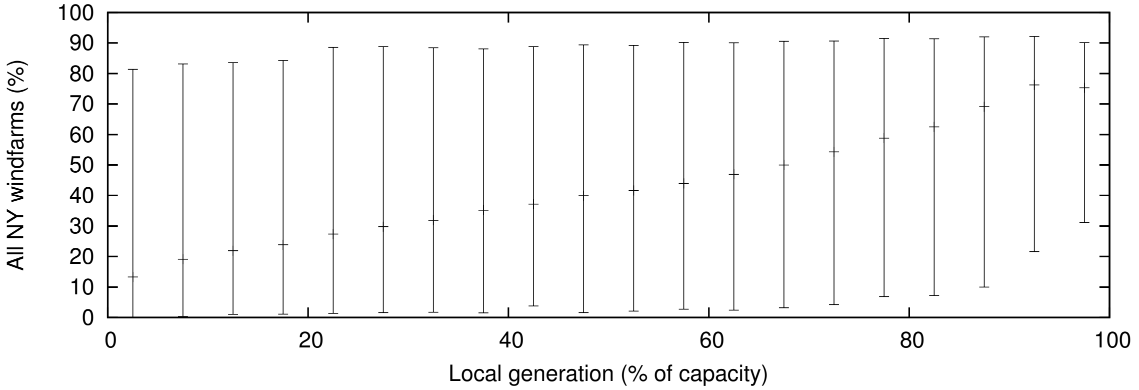


Figure 5.15: Average, minimum and maximum output across New York wind farms when an example wind farm is producing power at given levels.

5.7 Summary

Data centres may be almost ideal candidates to deliver the types of services that power grid operators have developed over the years to ensure the adequacy and reliability of the power grid as well as to provide the new and revised services intended to address the integration of more and more volatile renewables. The ability of data centres to rapidly adjust their power consumption, to defer workload to a future time, to migrate workload to a different facility, and to use local energy storage or local energy production gives them high flexibility to adjust power consumption in response to requests from the power grid operator. Acting under the direction of the grid operator may be more effective than acting independently to local renewables since the output of renewables is somewhat more stable when averaged over greater geographical areas. Providing ancillary services like regulation also enables fossil fuel generators that would otherwise provide these services to operate more efficiently. Data centres providing ancillary services may thus be more effective than a follow-the-renewables computing model.

This chapter provided a review of the different ancillary services used in power grids, discussing compensation for providing them and discussing future directions likely to be taken by these services. Simulations were performed using data from the NYISO, evaluating potential cost savings and the impact on wait time if data centres were to provide some of these ancillary services. The simulation results suggested that significant cost savings were possible – up to 12% in the case of the NYISO's 10-minute spinning service at an average cost of only 2% additional wait time. The tighter the requirements upon the ancillary service and the more control ceded to the grid operator, the greater potential savings there were. If a system was able to provide regulation service rather than the 10-minute spinning contingency reserve, savings far in excess of 12% were seen in simulation results.

Through participation in ancillary service markets, data centres stand to reduce both their energy costs and the negative impacts seen from cycling power plants. This may also enable power grids to incorporate more renewable energy more quickly.

Chapter 6

Multiple Data Centres

> I have yet to see any problem, however complicated, which, when looked at in the right way did not become still more complicated.
>
> Poul Anderson, “The Ghost in the Machine”

Chapters 4 and 5 focused on optimizing the operations of a single data centre but, in practice, large-scale computing platforms often utilize multiple data centres. This chapter outlines the reasons behind this in Section 6.1. As was outlined in Chapter 1, organizations are increasingly looking to cloud computing services in order to fulfill their computing requirements, and Section 6.2 provides an overview of the basic services common in current cloud platforms and outlines how these cope with multiple data centres. One aspect of the GreenStar Network project introduced earlier in this dissertation was the establishment of a network of renewable-energy-powered computing facilities across Canada and joined by partners from around the world. This platform used virtual machine migration to adapt to changes in the availability of renewable energy at each participating site, enabling the platform to operate using only renewable energy. Virtual machines are also one of the core services offered by current cloud platforms. Testing was conducted using the GreenStar Network platform to examine the effectiveness of migration strategies for virtual machines and migration’s impact on energy consumption. This analysis is presented in Section 6.3. Knowledge of the costs of virtual machine migration can be combined with knowledge of other types of services clouds offer to better estimate how synchronization between data centres impacts power consumption and thus enable more effective use of multiple facilities. Section 6.4 applies the lessons learned in Section 6.3 about virtual machine migration to those services that were outlined in Section 6.2. How to determine when to perform migrations and the appropriate type of migration to perform are then addressed in Section 6.5. Summary remarks in Section 6.6 then close the chapter. I created the test engine which automatically created virtual machine images, launched virtual machines and workloads, set up monitoring of the tests, and performed the specified number of migrations before performing cleanup activities. I was also responsible for the initial analysis of the experimental results and produced the figures presented herein.

6.1 The Need

Whether considering web applications, high-performance computing, or other types of workloads, the servers involved are often distributed across multiple data centres. Reasons for this vary somewhat from one type of workload to another but there are also similarities. This section outlines reasons for using a set of data centres rather than a single, centralized facility, separating these into technical, environmental and power grid, and legal aspects in Sections 6.1.1–6.1.3 respectively.

6.1.1 Technical

Reducing latency using a nearby data centre is particularly important for web workloads and was one of the constraints considered by early work into directing requests to different data centres based on power cost [202]. That early research focused on a content-distribution network (CDN) caching content in data centres around the world. CDNs have become more widely used in recent years, and an estimated 98% of Internet traffic is suitable for caching [141]. Whether using a CDN to distribute content or installing servers to run fully-featured applications in numerous data centres, such actions reduce latency and lower congestion in the core network. Using multiple data centres also provides redundancy, potentially enabling an application to remain usable despite the failure of a data centre.

6.1.2 Environmental and Power Grid

Chapter 4 examined how renewable energy availability and energy prices change, with this variation providing an incentive to shift processing when and where power is most plentiful and lowest cost. Differences in local weather conditions affecting cooling system efficiency may accentuate such differences and further incentivize organizations to use multiple facilities.

Even where centralization is feasible, as with delay-tolerant workloads like in Chapters 4 and 5, using multiple data centres rather than a single facility offers some advantages. The amount of power consumed by servers differs not only based on whether or not they are processing workload but what type of workload is being executed [115]. If a delay-tolerant workload is distributed across two data centres in different time zones, power costs can be reduced by executing high-power-consumption tasks in whichever location has lowest cost power and executing low-power jobs in the higher cost data centre, keeping utilization high but power costs lower. Researchers have also addressed workload distribution across data centres to reduce transmission system overload [175], and this may also enable data centres to more effectively provide the ancillary services discussed in the previous chapter.

6.1.3 Legal

Privacy considerations and other legal and regulatory constraints may restrict where customers are able to host their services and data, forcing clouds to operate from multiple data centres. Amazon Web Services [43] operates a separate “GovCloud” compliant with requirements sometimes imposed on US organizations to enable more users to move to its cloud infrastructure. Organizations may be compelled by legislation to avoid hosting data where security or privacy protection is weak, e.g., to prevent exposure of data to the US government as a result of the PATRIOT act [191].

High-performance computing infrastructure is also often geographically distributed. Examples include WestGrid [247] which is composed of numerous scientific computing installations in Western Canada and the LHC Computing Grid [149] using facilities across the globe to perform data analysis and simulations related to the Large Hadron Collider. Organizations may prefer to contribute computational resources rather than money, and grants funding high performance computing infrastructure often require that funds be spent locally.

6.2 Current Cloud Structure

The potential for rapid scalability and lower cost by sharing resources with other customers and only paying for the portion used have drawn many to cloud computing services. This section provides an overview of key services offered by the largest commercial cloud providers: Amazon Web Services [43], the Google Cloud Platform [106], Rackspace Cloud [204], and Windows Azure [249]. Since the Rackspace Cloud is based on OpenStack [187], a platform Rackspace founded together with NASA, this evaluation is believed to also generalize to scientific workloads though such workloads are not yet frequently executed on commercial clouds and may remain on dedicated high-performance computing platforms even if cloud tools are adopted [24].

Cloud models are sometimes divided into the following three categories [169]: software-as-a-service (SaaS), platform-as-a-service (PaaS), and infrastructure-as-a-service (IaaS). In the SaaS model, end-users access software applications managed and run by the cloud provider. As behind-the-scenes resource allocation decisions are not transparent to end-users, SaaS is not discussed further in this chapter beyond advising those evaluating SaaS providers to consider their efficiency. PaaS is a second type of service model enabling developers to build applications using cloud-provider-supported programming languages, libraries, and other services such as database access. The infrastructure underlying PaaS clouds is typically somewhat more visible to developers and those operating applications on them. The remainder of this chapter discusses PaaS but primarily focuses on IaaS, the third type of cloud service model. IaaS involves the least abstraction, providing lower-level access to hardware generally in the form of virtual machines along with other low-level services such as storage and networking.

The remainder of this section focuses on several of the services commonly used to build applications in cloud environments. Section 6.2.1 first addresses storage in the cloud. Virtual machines that handle the primary processing needs in current clouds are then discussed in Section 6.2.2. Various types of database services are available in modern clouds, and Section 6.2.3 provides an overview of these. A brief overview of other services is then given in Section 6.2.4.

6.2.1 Storage

Each of the four cloud providers examined offers some form of IaaS storage service. Amazon's Simple Storage Service (S3), Google's Cloud Storage, Rackspace Cloud Files, and Windows Azure Storage each offer scalable storage with "unlimited" space on a pay-per-use basis. Each service replicates data to several servers for redundancy. Amazon and Google expressly note that data is replicated to multiple facilities, while replication may be more limited for other cloud providers.

Not all stored data is as important or need be as readily accessible as others. Amazon and Google offer reduced redundancy storage in which data is replicated to fewer nodes, and Azure differentiates between locally-replicated and geographically-replicated storage. Amazon also offers a storage service where users are charged a significantly lower price in exchange for tolerating a delay averaging several hours before being able to access data requested for retrieval.

Amazon, Azure, and Rackspace also enable clients to use content distribution networks. As outlined in Section 6.1.1, CDNs replicate content around the world, enabling remote clients to access resources with low latency while easing congestion at central facilities. Amazon and Azure offer CDN services at added cost whereas Rackspace provides CDN support in its basic storage service.

Additional storage options are available to virtual machines being executed in the cloud environments and it is these virtual machines that are described next.

6.2.2 Virtual Machines

Virtual machines are commonly used to satisfy computing requirements in clouds. Amazon, Azure, Google, and Rackspace all offer a variety of virtual machine types. Azure, Google, and Rackspace offer virtual machine types with varied amounts of CPU power, RAM, and disk space, but Amazon also offers high-memory instances, instances with access to GPUs, and instances with solid-state local storage for high-I/O workloads.

Amazon virtual machines are ephemeral, subject to termination at any time, with disk contents lost unless Amazon's Elastic Block Storage providing network-accessible disk is used. Azure and Rackspace virtual machines are geared towards longer lifespans and are not removed until terminated by the user even if the underlying physical machine is restarted. Network disk is also available in the Rackspace and Google environments though.

In none of these environments is virtual machine migration well supported or easily available to users. Rackspace offers a VM replication service where a virtual machine's data is replicated to an alternate facility on a regular basis for disaster recovery purposes. Guidance is also available from Rackspace as to how to migrate the contents of server images between its UK and US facilities, but doing so is based on manual copying of data and the creation of a new virtual machine rather than the migration of an existing one.

6.2.3 Databases

Amazon's Relational Database Service, Google Cloud SQL, and the Rackspace Cloud Database service are cloud-provider-managed instances of traditional SQL database servers run inside a virtualized container where users are charged for allocated resources as virtual machines. Currently Amazon's SQL database service offers the broadest feature set, supporting the creation of replicas for scaling read access and also supporting synchronous replication to another facility. Azure provides a SQL database service, charging per hour based on database size, in which each database is served by three replicas for high availability. SQL Azure Federation enables the rows in database tables to be split across multiple machines, a process known as sharding. Those using Azure are also able to license Microsoft SQL Server on an hourly basis for use in virtual machines. For data-analytics purposes larger databases using SQL or SQL-like syntax are available with examples including Amazon RedShift and Google BigQuery.

According to Brewer's Conjecture or the CAP theorem [102], distributed services can maintain at most two of the following three properties: consistency, availability, and partition tolerance — continued operation in the event of a network partition. Traditional SQL databases have in recent years been augmented by various other large-scale database management services [213], and a new family of NoSQL databases [228], which grapple with the challenges of managing large-scale data. Amazon's SimpleDB service, for example, weakens consistency guarantees, allowing developers to choose eventual rather than strict consistency to maintain availability and partition tolerance. Such a consistency model can also significantly reduce the amount of data needing to be synchronized. Sacrificing certain types of database operations, particularly joins, can also enable databases to operate effectively at large scale. Microsoft's SQL Azure supports sharding through federations, but the software does not support join operations for such databases.

6.2.4 Other Types of Services

The storage services, virtual machines, and database services described in the previous sections are not the only cloud services available. This section reviews several other services of note.

PaaS options enable developers to create applications written in various languages while leaving much of the behind-the-scenes management to the cloud providers. Users are often charged based on the underlying IaaS resources used to execute their request, with Amazon's Elastic Beanstalk charging for the virtual machine instances and other resources provisioned to host applications but not charging for Elastic Beanstalk itself. PaaS versions of software like MapReduce [76] can also be used to distribute workload on a cloud. Users of Amazon's Elastic Map Reduce service are able to utilize resources at greatly reduced cost compared to what they would pay for provisioning similarly equipped virtual machines. Azure's Media Services and Amazon's Elastic Transcoding also enable users to transform media files, paying per unit of data processed or minute of video produced, again providing an alternative to managing virtual machines.

6.3 Cost of Migration

Having outlined current cloud services and how they approach the challenges of computing using multiple machines in the previous section, this section now addresses challenges in migrating workload from one machine to another. An obstacle to effectively using multiple data centres is understanding the ability of virtual machines to be migrated and the impact that migration has. File storage is often already replicated across data centres for redundancy and content distribution networks are already provided by cloud operators. Cloud databases also typically either relax consistency guarantees or execute in virtualized environments. A better understanding of virtual machine migration would thus fill a knowledge gap in how to use multiple data centres effectively, addressing both the commonly used virtual machines and the traditional database services now being executed in cloud environments.

As outlined in Chapter 2 others have studied factors limiting virtual machine migration [41], identifying as key the bandwidth available for migration and, for pre-copy live migration, the rate at which pages of memory are overwritten and the conditions determining the end of the migration's pre-copy phase. Added energy consumption from migration has received only a little attention [188] and this is thus the focus here. Networks presently exhibit very poor energy proportionality as discussed in Chapter 2, with some equipment consuming more power at idle than under full load, and thus this section focuses on server power consumption.

This section proceeds as follows. The test environment is first outlined in Section 6.3.1. Sections 6.3.2 and 6.3.3 follow with, respectively, details of the test procedure used and workloads executed during tests. Experimental results are then shown in Section 6.3.4.

6.3.1 Experimental Setup

Tests were primarily performed using three physical machines: one located at the GreenLight facility at UCSD in San Diego, California, another housed at Cybera in Calgary, Alberta and a third machine at the RackForce data centre in Kelowna, British Columbia. Data transfer between the machines used a 1 Gb/s lightpath dedicated to the GreenStar project. To compare LAN versus WAN performance, a second machine at UCSD, configured identically to the first, was used for some tests.

The GreenLight machines contained two Intel Xeon X5650 CPUs (totalling 12 cores + hyper-threading) as well as 24 GB of memory. The Cybera and RackForce machines each contained two Intel E5530 processors (totalling 16 cores + hyper-threading) alongside 48 GB of RAM. System power was managed by a Dell Active Power Controller.

Power monitors at the GreenLight facility supplied one power sample per minute. This was adequate for facility monitoring but made it difficult to estimate the power consumed by shorter operations since many VM migrations last less than a minute. Instantaneous power samples might also not be representative of average power during the time period. Power monitoring at RackForce produced finer-grained power readings, with a measurement made every three seconds. Tests were run in both directions between RackForce and UCSD as well as between RackForce and Cybera. The low sample rate of power monitoring in the GreenLight facility — the only facility in which two machines were available to perform migration testing — resulted in the focus of testing being WAN rather than addressing both WAN and LAN migration.

Each machine ran Ubuntu 10.04.3 LTS. Version 0.12.3 of the KVM hypervisor and version 0.8.7 of libvirt were used.

Due to the RAM available, hard disks were almost entirely idle during the time in which migrations were being performed, as was verified through additional tests. The destination physical machine used SSHFS to access files stored on the source physical machine, but disk activity was minimal after workloads were launched.

Networks were idle when migrations were performed, excepting a small amount of background traffic.

6.3.2 Test Procedure

Testing was carried out by a script that sequentially executed test cases described in an input file. Each test was executed in the following manner:

1. Creation of virtual machine images
2. Launch of virtual machines, workloads, and (optionally) ping processes to monitor the response time of VMs
3. Establishment of baseline power readings.
4. Migration of one or more virtual machines. This was done in series, waiting until a virtual machine had finished migrating before beginning the next migration or moving to the next step.
5. Establishment of new baseline power readings.
6. Destruction of virtual machine and deletion of their disk images.

6.3.3 Workloads

Various workloads were evaluated for suitability for use during tests, focusing on those with predictable usage of resources. Test workloads were developed using the STREAM memory benchmark and the IOzone filesystem benchmark. The Gamut workload generator was also used along with CPU-intensive random number generation. Gamut and random number generation were run concurrently in some tests. In the remainder of this section these workloads are described and issues that were encountered are noted.

The *STREAM benchmark* [10] was executed in fully automatic mode. Since KVM terminates the pre-copy phase of pre-copy live migrations when doing so is likely to result in downtime below a maximum threshold, attempts to live migrate VMs executing this workload were unsuccessful but resulted in high volumes of network activity for extended periods if the migration process was not manually terminated.

The *IOzone filesystem benchmark* [8] was also used. Shortly after the end of live migrations, the VM's filesystem became inaccessible. Migrations performed using a suspend-resume process completed successfully however. Additional tests suggest that TCP timeouts during the migration process can cause filesystems to be remounted as read-only devices.

The *Gamut workload generator* [176] was built to generate predictable workloads and was used in the tests described in this chapter to evaluate how rates of I/O activity affected migration. Gamut was set to allocate most of the VM's memory and then write to it sequentially and iteratively at a fixed, test-case-specified speed. Gamut enables repeatable tests to be run and its performance was verified as migration times grew exponentially as network I/O rates neared the throttle rate of 32 MB/s set in the KVM source code. This can be seen in Figure 6.1, which plots migration times versus I/O write rate for LAN and WAN examples and suggests that KVM's migration throttling will adversely affect migrations on low-latency, high-speed networks. Others [126] had noted that some migrations using KVM failed to complete when running certain benchmarks, but the gamut workload demonstrated that even a simple workload could cause such problems.

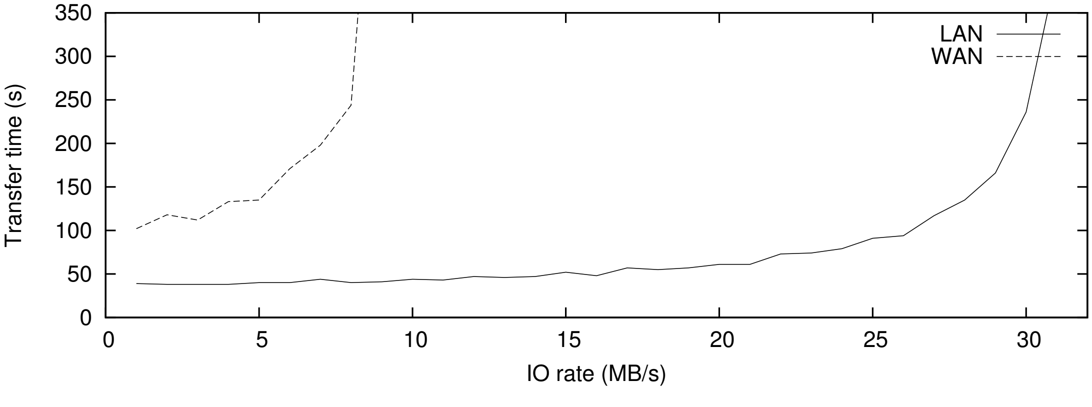


Figure 6.1: LAN and WAN migration times.

The remaining type of workload used for the experiments outlined here was *pseudo-random number generation*. This program executed a loop repeatedly calling drand48 to generate a CPU-bound workload with minimal memory activity. Tests of this sort were run both alone and concurrently with gamut.

6.3.4 Results

This section outlines the results of experiments run to examine how power consumption is affected by VM migration. The impact of live versus non-live migration, the migration of different types of workloads, and changes in the transport type used for migration are discussed in sequence.

Live Versus Non-Live Migration

Figure 6.2 shows the power consumption during two test cases in which 12 VMs executing a random number generator were migrated sequentially. The difference between the tests is that one used live migration whereas the other used suspend/resume migration. Various times were flagged during the tests to enable their stages to be observed. Several of these markers are shown as vertical bars in the figures in the remainder of this section. Vertical bars indicate the beginning of an idle period during which a baseline power reading was established, the beginning of each migration, and the beginning and end of the idle period at the end of all migrations. The added flags make clear the cause of the spike seen at the beginning of tests, attributable to the creation of VM images and starting of the VMs.

Power readings show little change in consumption regardless of whether live or non-live migration is chosen. Live migration in this instance completed slightly more quickly than non-live migration. For the last virtual machine migrated, non-live migration required added power beyond what was needed for a live migration. A similar pattern is seen when comparing a 1-VM live migration to a 1-VM non-live migration. With processor cores locked at the same frequency [11], migration may be able to use some capacity of an active system without a notable increase in system

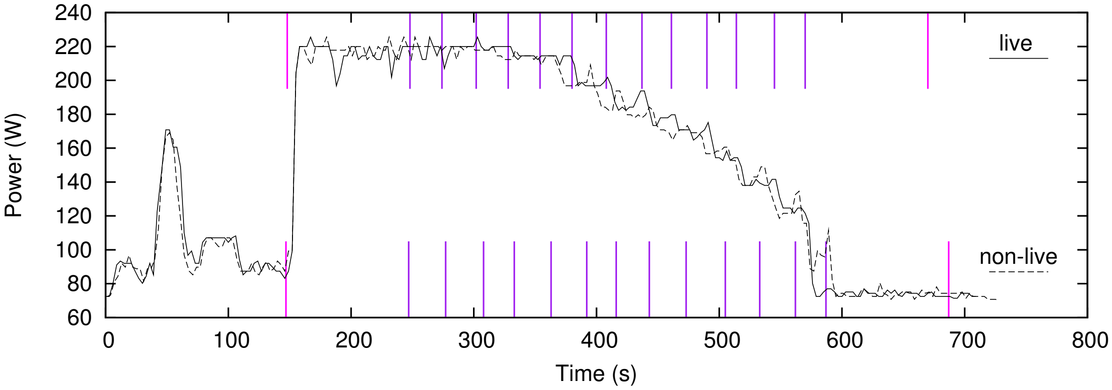


Figure 6.2: Power consumption during the live migration of 12 VMs executing a random number generator.

power consumption due to a process raising the CPU frequency of multiple cores. Non-live migration may not be faster than live migration for workloads involving little memory utilization, may increase power consumption, and also greatly increases the downtime of each virtual machine during migration.

The migration of multiple VMs can be time-consuming and thus requires a lengthy time to reduce a physical machine's power consumption. For clouds facing a significant drop in energy availability, neither standard live migration nor non-live migrations may be adequate. Power consumption can be reduced more quickly by suspending multiple VMs when starting the migration process rather than waiting until each VM's turn to be migrated arrives.

Type of Workload Migrated

The amount of additional energy required to migrate a virtual machine varies based on the workload executed on it. In Figure 6.3, for example, power consumption rises a small amount during the migration of one virtual machine writing to RAM at 4 MB/s while executing a random number generator. A more substantial increase in power consumption occurred if the VM was not performing a CPU-intensive workload. This may also be due to all active cores in a processor being forced to operate at the same frequency in the test systems.

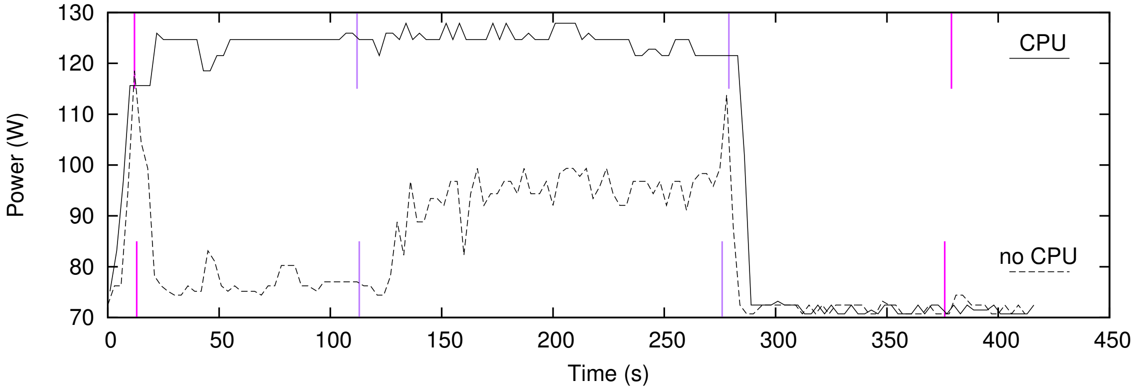


Figure 6.3: Power consumption when migrating 1 of 1 running VMs writing random numbers to RAM at 4 MB/s.

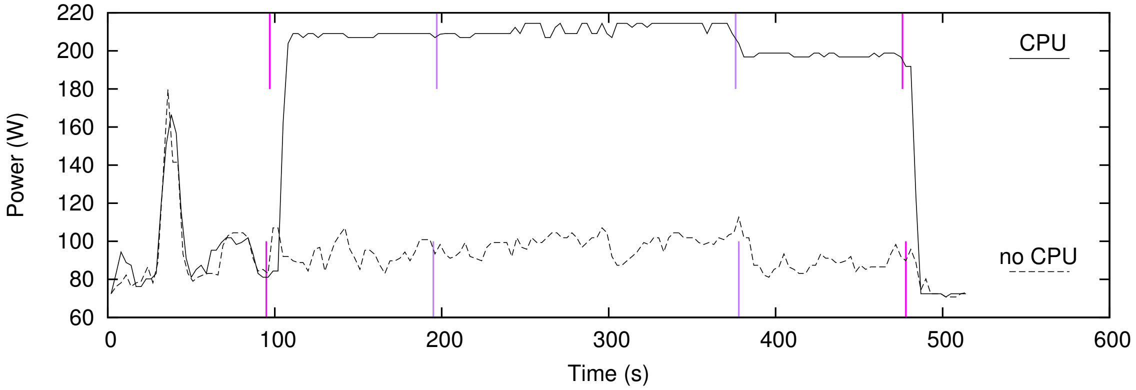


Figure 6.4: Power consumption when migrating 1 of 8 running VMs writing random numbers to RAM at 4 MB/s.

The tests in the previous paragraph address a node running just one virtual machine at the beginning of the test, but when multiple virtual machines were executed on the node the effect is not as pronounced. This can be seen in Figure 6.4 which shows the power consumption of a machine where 1 of 8 virtual machines was being migrated, with each VM executing the same workload as in the single-VM tests previously described.

Transport Type

Whether or not migration-related network traffic was encrypted had little impact on power consumption. This is shown by comparing unencrypted migrations to migrations encrypted with SSH for virtual machines writing to RAM at 5 MB/s with and without an added random number generator workload in Figures 6.5 and 6.6 respectively. Differences in power consumption might become apparent on high bandwidth links if KVM's throttling of VM migration traffic were to be lifted.

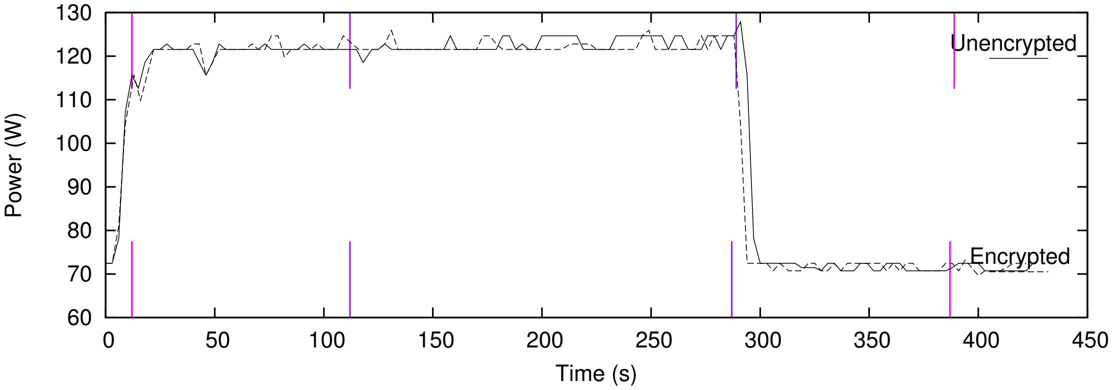


Figure 6.5: Migrating a virtual machine executing a CPU-bound random number generator and writing to RAM at 5 MB/s.

6.4 Application to Current Clouds

Combining an understanding of virtual machines and how virtual machine migration affects power consumption with an understanding of how other cloud services handle data distribution enables a broader understanding of the impact of multiple data centres and how to use them effectively.

Cloud storage services already perform data replication and enable users of cloud services to get many of the benefits of multiple data centres with no added effort.

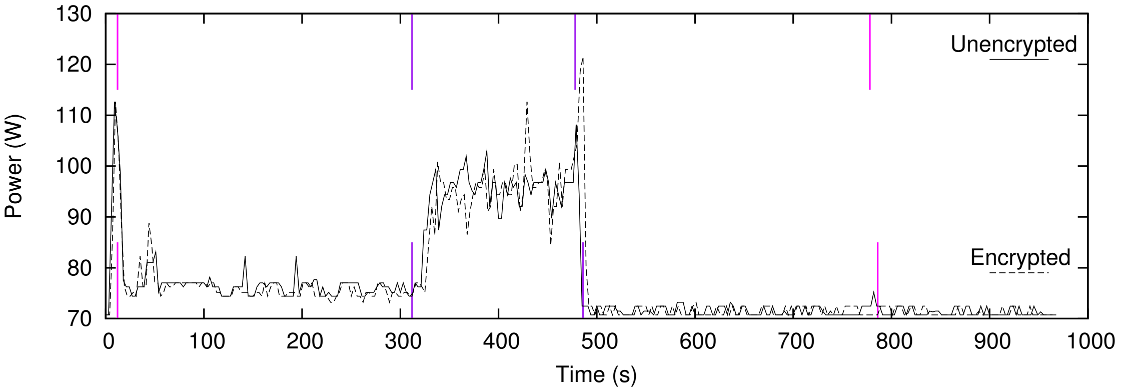


Figure 6.6: Migrating a virtual machine writing to RAM at 5 MB/s.

The content distribution network services of public cloud providers further reduce the need to explicitly manage data replication over multiple data centres.

With data already replicated to multiple facilities, if an application can be relaunched using data already available at a remote site, doing so is likely to be a more effective way to take advantage of multiple data centres than the use of either live or non-live migration. The added overhead when migrations are performed varies based not only on the workload executed in the virtual machine to be migrated but also based on the activity of other users of the machine and on the network links involved. As shown in the test results, migration overhead may at times be largely undetectable but on other occasions may add approximately 20–30W of power consumption. Due to the interaction between the VM being migrated, the migration process, and other workloads sharing the physical machine, information about the likely overhead of migrations would in most cases likely only be available to the cloud provider rather than an end-user wishing to initiate a migration.

Alternative migration strategies would be needed to allow virtual machine migration to scale. Despite having a near-dedicated 1 Gbps link for the experiments in this chapter, migrations often required significant time to complete and the migration process failed to terminate in some instances. Eliminating the throttle imposed by KVM on migration activity and choosing different stop conditions for the pre-copy phase of live migrations could reduce these problems but they are still likely to make migration infeasible at data-centre-scale even if additional network capacity were added when renewable energy is likely to be available locally [250] in order to accommodate migration activity.

Various ways exist to minimize the data transfer required to migrate applications. Migrations can be made more feasible using data deduplication [253] and compression [134] techniques as others have outlined. Shutting down a virtual machine at one facility and the creation of a new one at a different facility may often be an effective alternative to migration in many cases as previously noted. As of the time of writing, Rackspace mandates that cloud servers must be built upon Rackspace-configured images, increasing commonality between disk images. This commonality may enable migrations to be scaled more effectively. Still greater commonality may be achieved using PaaS web hosting services or PaaS versions of distributed computing software like Hadoop. Doing so increases commonality since similarly configured instances are likely to be used by many customers. They also enable computation to be routed to the data rather than a user requesting a virtual machine and then copying stored data to it.

Though web and application servers often contain little state information and are managed by load balancers which account for new and disappearing servers, the underlying database services are generally more challenging to distribute across multiple data centres. The overhead of running a database using Amazon's multi-availability-zone database service would likely be similar to but slightly lower than that of a constantly running live migration since it similarly involves duplicating state information across facilities, though only for a single application. Research into migrating cloud databases has sometimes explicitly recognized the similarity to virtual machine migration techniques [56, 83]. Since weakening consistency guarantees reduces the overhead of state replication, this can be used to more effectively utilize multiple facilities. However a lack of strong consistency may be difficult for some applications to cope with and others may need to be restructured to utilize such services. The replacement of SQL databases with NoSQL alternatives, as outlined in Section 6.2.3, would further improve the ability to effectively use multiple data centres.

6.5 When and how to migrate

It is difficult to develop a single equation to determine when virtual machines should be migrated. As can be seen by examining the related work presented in Chapter 2, research into virtual machine migration has typically been experimental or focused on more limited objectives such as estimating the amount of time that a particular migration would likely take. This section outlines different aspects to be taken into account in developing a comprehensive model to assess when virtual machine migration should be performed. Factors to consider can be grouped into the following categories: workload characteristics and behaviour, hardware properties, the objective for which migration is being performed, and information availability constraints.

The properties and quality of service requirements of both the workload to be migrated and other workloads sharing the same hardware matter when assessing whether to migrate a virtual machine. Workloads often exhibit time-varying behaviour and depending on their current activity may be possible to migrate in a more or a less efficient manner. Workload parameters, particularly the rate at which pages of memory are overwritten, impact the length of the migration period and others such as CPU utilization affect added power during the migration. Similar to how others have developed APIs enabling them to signal how they can be efficiently executed [196, 246], an API could be developed to enable applications to signal periods in which they can be migrated more efficiently. Whether applications can be suspended to be resumed at a later time, or, instead of being migrated, have the local instance terminated after launch of a new virtual machine at the destination site using a replicated system image also affect whether or not migration is likely to be a useful strategy. If a workload is nearing completion migrating it might not pay off, and if a workload has just commenced execution it may be more effective to restart it on a remote machine rather than migrating it. Some workloads might also be better executed using PaaS cloud services rather than executed in a separate virtual machine. It is not only the properties of the workload in the virtual machine that is a candidate for migration that matter, but other activity on the same hardware also has a similar impact.

The properties of the source and destination machines upon which the virtual machine that is a candidate for migration is being executed also matter. Whether a server's processor cores can scale independently can affect power consumption, and the time and energy required to change power states also has an impact. Similar effects apply to server RAM. Whether processor cores are bound to specific virtual machines can also affect power consumption.

Information availability also affects the ability to make decisions about which workloads to migrate and when to do so. Information about how other workloads executed on the same machine are likely to impact migration overhead is likely available only to the cloud operator. If virtual machine disk images and other data used by the virtual machine have been replicated elsewhere, the specific location of these replicas also may only be available to the cloud operator. Network capacity limitations are one of the primary constraints to large-scale migration, but information about a data centre's network utilization is again also likely restricted to those operating it.

Migration decisions are also affected by the type of objective that migration is intended to achieve. Whether the objective is minimizing energy consumption, reducing the financial cost of power, or implementing a follow-the-renewables approach to computing affects which migration decisions are best. With the financial cost of power differing significantly over time in a single location, and also varying from one location to another, an approach to migration that achieves lowest financial cost may consume significantly more energy than an approach aiming to minimize energy consumption. Likewise, performing migrations to locations with more renewable energy available, as part of a follow-the-renewables approach, also may not minimize energy consumption.

6.6 Summary

This chapter outlined the need for cloud providers and developers to distribute workload and applications across multiple data centres, rather than using a single, centralized facility, and provided guidance as to how to do so effectively. Technical considerations, the reduction of environmental and financial costs, and legal constraints all prompt use of multiple data centres to deliver services. A survey of the different types of cloud services currently offered outlined storage, virtual machines availability, and database services in addition to a few others. Tests conducted on the GreenStar Network testbed were then outlined, with their results illustrating how different factors impacted the energy overhead of migration. Migration-related activity caused power consumption to vary from undetectable levels through to adding approximately 20–30W of extra power consumption. The added power consumption varied based on the type of work being executed inside the virtual machine being migrated, other activity on the physical machine, and parameters of the migration itself. Some migrations might continue indefinitely without success due to the stop conditions affecting live migrations with the KVM hypervisor. The added overhead of migrating virtual machines relative to launching new ones using a copy of the disk image suggests that those using cloud services would be wise to consider alternatives to migration when using multiple data centres. Where possible, the use of PaaS options for hosting web sites or providing other needed services may enable overhead to be further reduced. Weakening consistency guarantees in database services is another avenue to improving the ability to operate using multiple data centres, though applications might need to be modified to accommodate this.

Chapter 7

Conclusions

Computing is a significant and still-rapidly-growing source of carbon emissions, and power costs account for a large portion of the total cost of ownership of computing equipment. Reducing this environmental footprint and decreasing the financial costs of data centres has been the focus of this dissertation. Some have estimated that increased use of computing may lead to a net reduction in emissions. The goal of lowering emissions may be more achievable through further growth in computing capacity. This makes efficient computing an all the more important goal. To this end, a greenhouse gas reduction protocol for the computing sector was developed. Further reductions in financial and environmental cost may be possible by altering power consumption patterns. The potential benefits of altering data centre usage patterns based on energy's financial cost, carbon intensity, local renewable energy availability, and commands from power grid operators were each evaluated. Real-world services are typically not centralized but delivered using multiple data centres. The challenges of doing so were outlined and some means for doing this in an energy efficient and environmentally friendly manner were explored. In this chapter, Section 7.1 highlights contributions made and Section 7.2 outlines opportunities for future work.

7.1 Review of Contributions

The first ISO-14064-2-compliant protocol for evaluating the emission reductions resulting from computing-related efficiency projects was discussed in Chapter 3, with the development of this protocol being a project I was closely involved in. The protocol provided guidance for estimating the effectiveness of efficiency improvement projects, enabling organizations to better assess project potential. Compliance with the ISO-14064-2 standard lends credibility to corporate sustainability reports and opens new revenue opportunities in carbon markets. Unlike the Carbon Usage Effectiveness metric's focus on operational emissions, the protocol accounts for emissions throughout the full project lifecycle and, unlike the lifecycle assessment standard released by the European Telecommunications Standards Institute at approximately the same time [23], was focused on evaluating the impact of specific potential emission reductions projects. A project relocating GeoChronos [100] from a carbon-intensive data-centre in Alberta to a low-emission facility in Kelowna, British Columbia tested and validated the protocol's procedures. I was involved throughout the development of the protocol, being a member of the Carbon Assessment Group that developed the protocol and a member of the small group that applied the protocol to a test project. Though also closely involved with many other aspects of the protocol's development and testing, I was particularly involved with comparing the power consumption of baseline and project platforms, performing risk analysis for the project, and investigating how to calculate PUE for mixed-used facilities.

Chapter 4 addressed the potential to reduce power costs by deferring high-performance computing workloads rather than changing the distribution of web requests across data centres as others had generally considered. Three measures of the environmental impact of power consumption at any given time were evaluated to assess the best target for optimization efforts. Estimates of potential cost-savings were derived using a time-step tool. Further evaluation with a discrete-event simulator was used to assess the impact of separating high and low-priority tasks. Adaptation to changes in power grid carbon intensity showed relatively little potential to reduce environmental costs, but adaptation to varying power market prices and the availability of local renewable energy both showed significant potential. Financial cost savings averaged 10–50% with oversupply problems at times reducing costs further. Data centres might increase the fraction of energy used from local renewable sources by 10–80% for data centres paired with a similar amount of wind turbine capacity. To avoid excessive delays it may be necessary to permit a significant amount of power from the grid to be used in any given time period. Further analysis was performed to evaluate how power costs might differ for low and high-priority tasks. High-priority tasks were immediately allocated resources in simulations whereas low-priority tasks might be deferred if the price of power exceeded a threshold. As power prices follow a similar diurnal pattern to the submission of tasks, the average power price paid by tasks exceeds the average power price. By deferring the execution of low priority tasks to a later point in time cost savings of 25–50% were seen over a cost-agnostic approach to execution. Simulations also evaluated the impact of shifting workload to a data centre located in a different time zone, showing the potential to reduce power costs for high-priority tasks in most workloads. I was responsible for creating the time-step tool and simulator used for the experiments reported in Chapter 4, as well as responsible for conducting the experiments themselves.

The ability of data centres to rapidly adjust power consumption and deal with sudden interruptions in power provides additional opportunity for cost savings, but prior work had not examined the impact on data centre power costs and wait time if participating in power grid programs for managing volatility. Chapter 5 addressed ancillary services, the mechanisms used by power grids to address both normal variation in supply and demand as well as to respond to failure. An overview of the different types of ancillary services was supplied along with an evaluation of the compensation available to entities providing them. Entities providing regulation services receive instructions from the power grid operator every few seconds indicating how much power they should consume or produce until next instructed. This type of ancillary service receives the highest level of compensation, being paid at a rate close to the price of the energy consumed. Contingency reserves, on the other hand, are used to compensate for the failure of power plants or transmission lines. Entities providing contingency reserves need only act infrequently and thus receive lower compensation. Simulations were conducted to assess the impact on data centre power costs and workload wait times if providing ancillary services. Cost savings of up to 12% were seen in contingency reserve simulations in which the wait times for tasks to complete increased by an average of up to 2%. This increased wait time might be eliminated using backup generators rather than powering down servers. Significantly greater savings were possible if ceding more control over operations and providing regulation service, with ancillary service revenue exceeding power costs in some simulations.

It is not only renewable energy availability that is important but also the behaviour of other users of the power grid. The local energy storage typically possessed by data centres, the speed with which servers can adjust their power consumption, the ability of some workloads to be migrated to other data centres and potential to defer other workloads to future points in time all appear to make data centres prime candidates for addressing variability and volatility in the supply of and demand for power. Data centres providing ancillary services may be able to do so in a more-efficient fashion than the power plants typically providing such service, since these power plants must typically operate at reduced efficiency. The rapid response times possible by data centres may also increase revenue opportunities, as the US Federal Energy Regulatory Commission has ordered US-based power grid operators to compensate rapid-responding reserves more highly. I was responsible for the price analysis of NYISO ancillary service markets, the creation of the simulator used and execution of the test cases. Hamidreza Zareipour helped resolve a number of questions regarding market rules for ancillary services.

Real world services are typically delivered using a number of different data centres rather than a single facility, but the energy implications of migration between facilities have not been well understood. Improved understanding of the energy implications of migration, particularly of virtual machines, was the focus of Chapter 6. The reasons for using multiple data centres were outlined, and an outline of cloud computing services and how these address multiple data centres was provided.

Virtual machine migration experiments were run to determine how to effectively distribute and reallocate workload among a set of data centres. Added power consumption from virtual machine migration varies from largely undetectable levels through to increasing system power consumption approximately 25–30W dependent on the workload being executed in the virtual machine being migrated and other activity on the host physical machine. Some types of workloads migrate relatively quickly, whereas other migrations may never terminate due to the manner in which hypervisors transition from one phase of a migration to the next. The network capacity needed to migrate even a single virtual machine severely limits the WAN migrations that can be completed in a short amount of time, restricting large-scale use of migration.

Users of infrastructure-as-a-service virtual machines would be advised to take advantage of other types of cloud services offered as they may enable multiple data centres to be used more effectively. Files in cloud storage systems are often already replicated across multiple data centres, and may be further distributed using content distribution networks. Content distribution can thus be using existing cloud services rather than reimplementing it. For applications requiring little persistent state, such as web servers or application servers depending upon a database layer for state preservation, creating new virtual machines from replicated images and then shutting down the original instance reduces the need for migration. The use of platform-as-a-service versions of application, web, or database servers should also be considered as this can further reduce the need to migrate data. Where state requires preservation, a willingness to weaken consistency guarantees below the strict consistency mandated by virtual machines and traditional relational databases provides another option to reduce the overhead of using multiple data centres. Such an approach is often taken in current cloud database services.

I was responsible for performing the actual test cases and authored the test engine that created and destroyed virtual machine images and the virtual machines themselves and also executed the test cases examined. Andrey Mirtchovski assisted with resolving issues with some of the candidate workloads and aided in collecting monitoring information.

Reducing financial and environmental costs requires careful consideration of all aspects of computing and their relation to the environment and the power grid. The full lifecycle of computer systems must be considered to achieve low total cost of ownership and low environmental costs. This requires not only selecting efficient components but also that they be used effectively, and this may lead to data centres and the equipment in them being used in non-conventional ways. Data centres may be able to reduce the average price they pay for power by shifting workload to time periods in which power is available at low cost. Dynamic power pricing already involves significant variation over the course of a typical day, primarily as a result of changing demand. As more renewable energy is integrated into power grids, pricing is likely to be more influenced by renewable energy availability. Data centres may also be able to directly shift their demand to times in which local or global metrics indicate more renewable energy is available. Due to the ability of servers in data centres to rapidly adjust their power consumption, data centres may also be able to profit from delivering ancillary services which address short-term variation in power supply and demand or make the power grid more resilient to failures. Using local energy storage or backup generators may make data centre delivery of ancillary services more practical and potentially more profitable. Different types of workloads have different quality of service requirements. Execution of some workloads can be deferred whereas other non-deferrable workloads may require that the workload be distributed across multiple data centres to react to local shifts in power availability. Understanding the migratability of workloads along with the resulting energy overhead enables multiple data centres to be managed more efficiently and effectively. Much work remains to be done, however, to fully address the environmental and financial costs of computing.

7.2 Future Work

The field of "green computing", which addresses the environmental costs of computing, is still young. This section highlights several avenues for further exploration.

The greenhouse gas reduction protocol presented in Chapter 3 covered a broad range of potential applications, but other scenarios remain unaddressed. Data centre adaptation on the basis of local renewable energy availability, the power price, and the estimated carbon intensity, or to provide ancillary services, are not yet accounted for in the protocol. It would also be worth investigating the footprint of approaches like "cycle-stealing" or the use of spot instances available at below-normal prices when excess cloud capacity exists. A better understanding of the impact of the high fraction of financial and environmental costs coming from the operational phase of the equipment's lifecycle would also be beneficial. How should this affect the frequency with which servers are replaced by new equipment? If operating in a region in which the availability of renewable energy fluctuates significantly over time might it make sense to acquire more equipment and operate it a smaller percentage of the time?

When looking to reduce power costs by providing ancillary services or otherwise altering energy usage, a better understanding of how individual jobs are affected and how variations in quality of service requirements affect power consumption would be beneficial. An ability to estimate how power costs, carbon emissions, and wait time are likely to vary for high-performance computing tasks based on the type of workload, the region or regions in which it might be executed, and the quality of service demanded would be beneficial. This could form the basis for pricing models introducing more-dynamic pricing into utility computing environments. A better understanding of ancillary service activation patterns would be useful both for deriving cost estimates as well as to better understand their impact on quality of service. Since the lowest average prices in energy markets appear to be found in real time markets [202], but payment for ancillary services is typically highest in day-ahead markets, better workload prediction is important to determine appropriate bidding strategies. Recent research has examined how to use local energy storage to safely underprovision data centre power infrastructure. Understanding how local energy storage might instead enable data centres to offer ancillary services more efficiently, without compromising the quality of service of the workload being processed, should also be explored. Other researchers have attempted to reduce maximum power draw to reduce the capacity costs that may account for a relatively large fraction of power bills, but approaches like incremental-cost-based transmission pricing [161], the use of interruptible transmission services [4], and possible deductions for duplication avoidance or customer-owned equipment [39] may enable such costs to be greatly reduced for data centres located near renewable energy sources. Such opportunities should be further explored. Assessment of the impact of simultaneously providing ancillary services in data centres connected to different power grids in different regions would also be useful.

Real-world systems are often distributed over multiple data centres and the energy implications of this deserve further consideration. Tests were run with virtual machines using the KVM hypervisor to evaluate the energy overhead of migration. Conclusions were drawn as to the likely overhead of other services during migration but further testing using other hypervisors and other cloud services should be conducted to verify their performance. The impact of data compression [134] and deduplication [253] techniques on migration overhead should also be explored further. Determining when migration of particular workloads can be performed with low overhead can reduce replication costs. Incorporating information about the migratability of each application being executed might play a role in future service pricing along with information about its deferrability. As networks become more energy proportional and green routing techniques are more widely used, the energy costs and carbon emissions associated with the network should be evaluated more closely.

Appendix A

Additional Tool Information

This appendix provides additional detail and examples of how the tools used to conduct the experiments presented in Chapters 4 and 5 function. Section A.1 first outlines the testing and validation process performed. The optimization tool described in Section 4.2 is then discussed in Section A.2. Chapter 4 also introduced a simulator in Section 4.3 and more information on this tool is supplied in Section A.3. The ancillary service simulator in Chapter 5 is another tool used in the research presented in this dissertation and this appendix closes in Section A.4 with examples of its operation.

A.1 Validation

Testing was performed throughout the creation of each of the tools developed to verify that they were functioning correctly. This involved stepping through simulations to verify that any modified code was behaving correctly and initially flagging each time newly added code was executed to verify that it wasn't executed unexpectedly. In addition, tests were run in which the progress of individual tasks was tracked to verify that they were being handled correctly. Further investigation was conducted of any counter-intuitive results to ensure that they were not due to software bugs.

A.2 Optimization tool

This tool was used to evaluate the extent to which optimizing energy use might enable the financial costs or carbon emissions associated with grid power use to be reduced or to allow the use of local renewable energy to be maximized. Section A.2.1 illustrates how cost-based optimization in the simulator functioned. This mode accepted cost information in the form of the financial cost info from power markets or the carbon cost represented by the power grid's estimated carbon intensity in each time interval. Section A.2.2 then documents how the optimization tool used information about the output of local renewable energy sources to estimate bounds on the ability to maximize its use.

A.2.1 Financial costs and carbon intensity

The optimization tool evaluated cost data and workload data on a period by period basis. Cost data might be in the form of power prices or an estimate of the average carbon intensity in an interval. Period length was required to be a multiple of the length of the interval at which prices changed in the power cost data supplied, and periods were aligned such that the start of each period coincided with the beginning of a new pricing interval. Here a scenario involving a 2-hour long period with prices changing at 30-minute intervals is described. In this example an 8 node, homogeneous cluster is assumed.

All tasks submitted during each period are aggregated to calculate a total demand for it. If a 1 node task requiring 2 hours of execution time, a 2 node task requiring 1 hour and a 4 node task requiring 1 hour of execution time arrived during a period they would form an aggregate demand of 9 node-hours.

Before processing any demand from the present period, remaining demand from previous periods is first executed. For this example, no demand was recorded for the previous period, but 5 node-hours of queued demand is assumed to have remained from the period before that and will be the first workload processed in this period.

When attempting to optimize based on the cost of power, price data for the period was sorted for low to high cost. In non-adaptive mode price data was randomly reordered. In this example it is assumed that during the interval in question per-MWh power costs over the 2 hour period were as follows: \$37.50, \$50.00, \$37.50,

\$25.00. In adaptive mode these would then be reordered in the following manner: \$25.00, \$37.50, \$37.50, \$50.00. For simplicity, clusters were assumed to consume 1 MW at full power because, as results were always reported relative to a non-adaptive case, a system's power consumption would always be factored out in the end. Here a price threshold of \$45/MWh above which tasks would not be executed was assumed.

A cluster containing 8 nodes can perform 4 node-hours of work in a single 30-minute period. Since 5 node-hours of backlog remained to be executed, in adaptive mode 4 node-hours would be executed at a cost of \$25/MWh, leaving 1 node-hour from 2 periods back to be executed along with 9 node-hours of demand submitted in the present period. In the next two pricing periods, a total of 8 node-hours of demand would be completed at a cost of \$37.50/MWh, leaving 1 node-hour of demand to be executed in the final price interval in the period. However, the price threshold of \$45/MWh previously assumed would result in this being deferred to a future period.

A.2.2 Renewable energy availability

Like the previous example, this section also assumes an 8 node cluster being evaluated and a two hour period length. However, this time the period is split into twelve 10-minute intervals. Wind generation capacity greater than the data centre's maximum power demand is assumed. Each of the following samples indicates wind generation during a 10-minute interval in the 2-hour period with, e.g., 1.2 indicating that during the interval wind generation averaged 120% of the data centre's maximum power consumption:

$$0.8, 1.0, 1.2, 1.4, 1.2, 1.0, 0.8, 0.6, 0.4, 0.2, 0.0, 0.0$$

Again a data centre maximum consumption of 1 MW was assumed as it was factored out in reported results by comparing it against a non-adaptive scenario. The fraction of local renewable energy available during the time period (RA) was calculated as follows, where $gen(p)$ represents wind generation relative to the cluster's maximum demand during the time period p:

$$
\begin{aligned}
RA &= \frac{\sum_{p \in periods} min(gen(p), 1)}{|periods|} \\
&= \Big(min(0.8, 1) + min(1.0, 1) + min(1.2, 1) + min(1.4, 1) + min(1.2, 1) \\
&\quad + min(1.0, 1) + min(0.8, 1) + min(0.6, 1) + min(0.4, 1) + min(0.2, 1) \\
&\quad + min(0.0, 1) + min(0.0, 1) \Big) \div 12 \\
&= 7.8 \div 12 = 0.65
\end{aligned}
$$

The non-adaptive mode was thus estimated to consume 65% renewable power while processing workload. Prior to needing to consume grid power or suspending execution until the next time period, in adaptive mode it was assumed that $0.65 \times 16 = 10.4$ node-hours of workload could be executed in the period. Assuming the same workload as in the previous section, of the 14 node-hours queued 10.4 could be executed using renewable energy, leaving 3.6 hours to either be deferred or processed with grid power. Assuming that the optimization tool was configured to allow grid power to be used for up to 20% of any time period, an additional $0.2 \times 16 = 3.2$ node-hours could then be processed, leaving 0.2 node-hours for the following period. Of those hours executed, a total of $\frac{10.4}{10.4+3.2} \approx 76.5\%$ were recorded as executed using renewable energy, though in this case some of the workload was deferred to a future time and not yet accounted for. Total renewable energy use and total grid power use were tracked across all periods evaluated by the optimization tool to calculate final results.

A.3 Hybrid scheduling simulator

The simulator introduced in Section 4.3 represented high-priority, interactive jobs individually and used a fluid approximation for the remainder of the workload. When power costs were estimated to be too high, interactive jobs continued execution unimpeded but any lower-priority tasks would be suspended until the cost of power fell below a threshold set based on the recent history of power prices in the region.

Beginning at time t, the following sequence outlines the simulator's behaviour through a number of events. As in the previous examples a homogeneous cluster containing 8 nodes is assumed. Each time a low-priority task arrived it was converted to a fluid approximation with, e.g., a single node task requiring two hours of execution time converted to a volume of 2 node-hours. It is with 8 node-hours of queued low-priority tasks and in a time of comparatively low power prices that this example begins:

> [**t**]: At this time when all 8 nodes were processing low-priority tasks, a high-priority, interactive task requiring 2 nodes for 1 hour arrived. This resulted in 2 nodes being reallocated to this task and the number of nodes used for processing low-priority tasks reduced to 6.
>
> [**t+0:10:00**]: An event arrives indicating a change in the price of power. An evaluation comparing it to the 75^{th} percentile of the previous week's power prices is conducted, and the time is determined to be too expensive for the low-priority jobs to continue. They are suspended while the high-priority task continues to execute using 2 processors.
>
> [**t+0:40:00**]: An event arrives indicating another change in the price of power. This time a comparison to a recent history of prices shows the hour to be relatively low-priced and the execution of non-interactive tasks

is resumed on 6 nodes.

[**t+1:00:00**]: The high priority task completes. As a result, the nodes that it was using are reallocated to the execution of low-priority tasks since the price of power remains low.

A.4 Ancillary service simulator

This section provides examples illustrating how the ancillary service simulator in Chapter 5 functioned. It begins in Section A.4.1 with an example of how contingency reserves and emergency demand-response were handled in simulations. The simulator treated emergency demand-response similarly to contingency reserves, though emergency demand-response is activated for much longer periods when called upon, generally given day-ahead notification of the need to reduce power consumption, and paid when supplying energy reductions, rather than paid for being available like the contingency reserves. Following discussion of contingency reserves, Section A.4.2 then outlines how regulation service revenue was calculated.

A.4.1 Contingency Reserves and Emergency Demand-Response

Organizations providing contingency reserves commit to reducing power consumption upon receiving such a request from the power grid operator. The following set of sequence of events illustrates how an 8 node cluster executing two tasks at time t would respond to a request from the power grid operator to reduce power consumption:

[**t**]: The power grid operator's request arrives and the system starts to suspend the tasks. Progress in executing the tasks is halted but the system continues to consume $\frac{1}{4}$ power consumption since, under the assumptions made, consuming power for a duration of 10 seconds enabled the servers

to transition in and out of a state in which their memory contents were preserved without requiring further power.

[**t+0:00:10**]: The suspension of servers is completed and the system ceases consuming power.

[**t+0:05:00**]: An event arrives indicating that a task requiring 4 nodes had begun to execute on the original system. As the system simulated is in a suspended state, this task is queued for later execution rather than immediately started.

[**t+0:10:10**]: After reducing power consumption for 10 minutes the system begins to restore servers to an operational system.

[**t+0:10:20**]: The tasks that had been running at time t were resumed and are now making forward progress again. As there is space available the task arriving at time t+0:05:00 was launched.

Systems providing contingency reserves are paid based on their availability to reduce demand rather than for their actual reduction in demand. Thus prior to time t, for as long as the system had been running tasks on 2 of its 8 nodes, it would be paid for being able to reduce $\frac{1}{4}$ its maximum power consumption on request, and from time t+0:10:20 until the next change in system state it would be paid for being available to reduce $\frac{3}{4}$ of its maximum consumption on request. When simulating emergency demand-response, systems were assumed to be compensated corresponding to the fraction of the cluster that would have been utilized in the period had a request to reduce demand not been received and, per the program rules, would be paid the greater of \$500/MWh and the real-time electrical price during the period.

In the event that tasks had arrived such that the system's 8 nodes would be inadequate to execute all of them in parallel when reawakened, the tasks which had arrived earliest were given priority over nodes. Backfilling was allowed for the remaining tasks, though backfilled tasks might be suspended partway through their execution if, due to the completion of some other task, a task which would have started before it gained priority over the nodes which had been allocated to the backfilled tasks.

Traces used in the experiments run all contained information about how long each task had been queued prior to execution on the original system. If a task was delayed due to the activation of contingency reserves the additional delay prior to the task being completed would be added to its wait time on the original system to calculate the modified wait time.

A.4.2 Regulating Reserves

Providing regulating reserves requires a higher commitment than providing contingency reserves, with organizations providing regulation service receiving commands from the power grid operator every few seconds. Managing a system providing regulation service is in practice likely to be much more complex than providing contingency reserves due to the much greater interaction with the power grid operator. However, although at present some general trends are known, such as that a facility providing n MW of regulation capacity on average consuming $\frac{n}{2}$ MW of power, the specific behaviour required of a facility providing this service at a given time is unknown. As described in Chapter 5 various factors influence the commands sent by the power grid operator to a specific facility providing regulation service. Due to this information gap, unlike for contingency reserves, the simulator made no attempt to calculate the impact upon workload wait time by providing regulation service. It instead attempted only to estimate revenue possible from providing this service.

Up to three different market prices might need to be tracked to calculate the cost impact of regulation capacity in simulations. The regulation service in the appropriate location and for the appropriate market type needed to be tracked. The

NYISO operates ancillary service markets in a smaller number of regions than it operates energy markets in, but the simulator determined the appropriate ancillary service market based upon the energy market involved. Whether the data centre was participating in the day-ahead market, the hour-ahead market, or the real-time market was also important. Power consumed aside from any consumption related to providing ancillary services might be done in the day-ahead or hour-ahead markets, but regulation-related power consumption was always purchased at real-time market rates.

Consider two time intervals. During time interval A two of the cluster's eight nodes were in use, and during time interval B six nodes were in use. During interval A the system was assumed to have bid to provide ancillary services corresponding to $\frac{1}{2}$ of its capacity. With power consumption when providing ancillary services expected to average half the ancillary service capacity offered, during the time interval the system would consume $\frac{1}{2} \div 2 = \frac{1}{4}$ of its maximum power consumption, equal to the power required to operate two nodes on average. This power consumption would be charged at real-time rates. During time interval B, regulation capacity was calculated using the following equation, earlier introduced in Section 5.3.1, since a limit of 100% of the cluster's maximum power use was placed on energy consumption (though in practice this might be exceeded by directing some energy to local storage):

$$2 \times (1 - x), \text{where } x \text{ was the fraction of the cluster's nodes in use}$$

Based on this formula and $\frac{3}{4}$ of the cluster in use, the regulation capacity provided, as a fraction of the system's total power use was calculated as follows:

$$2 \times (1 - \frac{3}{4}) = 2 \times \frac{1}{4} = 0.5$$

Average energy consumption in this period related to the regulation capacity provided would average half the 0.5 figure, resulting in a corresponding purchase of $\frac{1}{4}$ of the power at real-time rates. The remaining required power, corresponding to half the system's power use, was assumed to have taken place in the energy markets, perhaps on a day-ahead or hour-ahead basis rather than at real-time rates.

Bibliography

[1] Air emissions. `http://www.epa.gov/cleanenergy/energy-and-you/affect/air-emissions.html`.

[2] Amazon Glacier. `http://aws.amazon.com/glacier/`.

[3] Boost C++ libraries. `http://www.boost.org/`.

[4] Demand Opportunity Service - frequently asked questions. `http://www.aeso.ca/transmission/15236.html`.

[5] Green Globes NC certified buildings. `http://www.thegbi.org/assets/case_study/Green-Globes-NC-Certified-Buildings.pdf`. Accessed June 15, 2012.

[6] History of ENERGY STAR. `http://www.energystar.gov/index.cfm?c=about.ab_history`.

[7] Iceland: the coolest location for data centres. `http://www.invest.is/investment-opportunities/data-centers-in-iceland/`.

[8] IOzone filesystem benchmark. `http://www.iozone.org/`.

[9] Measuring and improving our energy use. `http://www.google.com/about/datacenters/inside/efficiency/power-usage.html`. Accessed June 14, 2012.

[10] STREAM: sustainable memory bandwidth in high performance computers. `http://www.cs.virginia.edu/stream/`.

[11] Intel Turbo Boost Technology in Intel core microarchitecture (Nehalem) based processors. White paper, Intel Corporation, Nov. 2008.

[12] Reducing data center cost with an air economizer. Technical report, Intel Corporation, Aug. 2008.

[13] Evaluating the carbon reducing impacts of ICT: An assessment methodology. Technical report, Global e-Sustainability Initiative, Sept. 2010.

[14] How ENERGY STAR's portfolio manager rates data centers. `http://www.vertatique.com/epa-energy-star039s-portfolio-manager-works-data-centers`, 2010.

[15] New York Independent System Operator (NYISO). 2010 ISO/RTO Metrics Report Appendix G, Federal Energy Regulatory Commission, 2010.

[16] Recommendations for measuring and reporting overall data center efficiency. Technical report, The Green Grid, July 2010.

[17] SAP power benchmarks specification. `http://www.sap.com/solutions/benchmark/pdf/SpecServerPowerBM.pdf`, Apr. 2010.

[18] TPC-Energy specification. Standard Specification Version 1.2.0, Transaction Processing Performance Council, June 2010.

[19] Accounting and billing manual. Manual 14, New York Independent System Operator, May 2011.

[20] Ancillary services manual. Manual 2, New York Independent System Operator, Oct. 2011.

[21] Cloud computing the IT solution for the 21st century. Technical report, Carbon Disclosure Project, 2011.

[22] Historic EDRP and SCR activation information. Technical report, New York Independent System Operator, July 2011.

[23] Life Cycle Assessment (LCA) of ICT equipment, networks and services: General methodology and common requirements. ETSI TS 103 199 V1.1.1, European Telecommunications Standards Institute, Nov. 2011.

[24] Magellan report on cloud computing for science. Technical report, U.S. Department of Energy Office of Advanced Scientific Computing Research, Dec. 2011.

[25] Order granting petition. Docket No. EL11-44-000, Federal Energy Regulatory Commission, 2011.

[26] Recommendations for measuring and reporting overall data center efficiency version 2 - measuring PUE for data centers. Technical report, The Green Grid, May 2011.

[27] Third-party provision of ancillary services; accounting and financial reporting for new electric storage technologies. Docket Nos. RM11-24-000 & AD10-13-000, Federal Energy Regulatory Commission, June 2011.

[28] The afternoon dip. `http://www.renewablesinternational.net/the-afternoon-dip/150/537/33320/`, Mar. 2012.

[29] eBay sets bold new vision for powering commerce with clean energy. Press release, eBay Inc., June 2012.

[30] Electronics Disposal Efficiency (EDE): an IT recycling metric for enterprises and data centers. Technical report, The Green Grid, 2012.

[31] Gartner says worldwide cloud services market to surpass $109 billion in 2012. HPC in the Cloud, Sept. 2012.

[32] Global investor survey on climate change. Technical report, Ceres, July 2012.

[33] Harmonizing global metrics for data center energy efficiency. Technical report, The Green Grid, Oct. 2012.

[34] ICT greenhouse gas reduction project protocol: Quantification and reporting. Protocol Version 1, Canadian Standards Associations, 2012.

[35] LEED 2012 changes by rating systems. Technical report, U.S. Green Building Council, 2012.

[36] Worldwide and regional public IT cloud services 2012–2016 forecast. Market Analysis Doc # 236552, IDC, Aug. 2012.

[37] D. Abts, M. R. Marty, P. M. Wells, P. Klausler, and H. Liu. Energy proportional datacenter networks. *SIGARCH Comput. Archit. News*, 38(3):338–347, June 2010.

[38] Cloud computing and sustainability: The environmental benefits of moving to the cloud. Technical report, Accenture and WSP Environment & Energy, Nov. 2010.

[39] AESO 2009 Tariff / Approved Rate Schedules and Riders. Technical report, Alberta Electrical System Operator, Oct. 2009.

[40] Phase two wind integration. Discussion paper, Alberta Electrical System Operator, Dec. 2010.

[41] S. Akoush, R. Sohan, A. Rice, A. W. Moore, and A. Hopper. Predicting the performance of virtual machine migration. In *2010 IEEE International Symposium on Modeling, Analysis Simulation of Computer and Telecommunication Systems*, pages 37–46, Aug. 2010.

[42] Alberta industries comply with pivotal climate change legislation. News release, Province of Alberta, Apr. 2008.

[43] Amazon Web Services. `http://aws.amazon.com/`.

[44] P. Anderson, G. Backhouse, D. Curtis, S. Redding, and D. Wallom. Low carbon computing: a view to 2050 and beyond. Technology & standards watch (techwatch), JISC, Nov. 2009.

[45] C. Ansbergs, A. Gordon, and E. Young. Understanding and engaging utility incentive programs. White Paper 38, The Green Grid, Mar. 2011.

[46] M. Arlitt, C. Bash, S. Blagodurov, Y. Chen, T. Christian, D. Gmach, C. Hyser, N. Kumari, Z. Liu, M. Marwah, A. McReynolds, C. Patel, A. Shah, Z. Wang, and R. Zhou. Towards the design and operation of net-zero energy data centers. In *Proceedings of ITHERM 2012*, 2012.

[47] M. Armbrust, A. Fox, R. Griffith, A. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia. Above the clouds: A Berkeley view of cloud computing. Technical Report UCB/EECS-2009-28, University of California at Berkeley, 2009.

[48] ASHRAE Technical Committee 9.9. ASHRAE environmental guidelines for datacom equipment. http://tc99.ashraetcs.org/documents/ASHRAE_Extended_Environmental_Envelope_Final_Aug_1_2008.pdf, 2008.

[49] D. Azevedo, C. Belady, and J. Pouchet. Water Usage Effectiveness (WUE): A green grid data center sustainability metric. White Paper 35, The Green Grid, Mar. 2011.

[50] D. Azevedo, M. Patterson, J. Pouchet, and R. Tipley. Carbon Usage Effectiveness (CUE): A green grid data center sustainability metric. White Paper 32, The Green Grid, Dec. 2010.

[51] P.-F. Bach. Germany faces a growing risk of distasterous power blackouts. Business Insider, May 2012.

[52] K. Bailey, L. Ceze, M. Greenwood, and H. M. Levy. Operating system implications of fast, cheap, non-volatile memory. In *Proceedings of the 13th USENIX Workshop on Hot Topics in Operating Systems*, May 2011.

[53] J. Baliga, R. W. A. Ayre, K. Hinton, and R. S. Tucker. Green cloud computing: Balancing energy in processing, storage and transport. *Proceedings of the IEEE*, 99(1):1–19, Jan. 2011.

[54] W. E. Bardsley. Note on the pumped storage potential of the Onslow-Manorburn depression, New Zealand. *Journal of Hydrology (NZ)*, 44(2):131–135, 2005.

[55] P. Barham, B. Dragovic, K. Fraser, S. Hand, T. Harris, A. Ho, R. Neugebauer, I. Pratt, and A. Warfield. Xen and the art of virtualization. In *Proceedings of the 19th ACM Symposium on Operating Systems Principles*, pages 164–177, New York, NY, USA, 2003. ACM.

[56] S. Barker, Y. Chi, H. J. Moon, H. Hacigümüş, and P. Shenoy. "Cut me some slack": latency-aware live migration for databases. In *Proceedings of the 15th International Conference on Extending Database Technology*, pages 432–443, New York, NY, USA, 2012. ACM.

[57] L. Barrios and A. Rodrïguez. Behavioural and environmental correlates of soaring-bird mortality at on-shore wind turbines. *Journal of Applied Ecology*, 41(1):72–81, 2004.

[58] L. A. Barroso and U. Holzle. The case for energy-proportional computing. *Computer*, 40(12):33–37, 2007.

[59] C. Bash, C. Patel, and R. Sharma. Dynamic thermal management of air cooled data centers. Technical Report HPL-2006-11, HP Laboratories, 2006.

[60] C. Belady. Thinking off the grid: Independence for today's data centers via data plants? `http://bit.ly/WuoRw4`, Apr. 2012.

[61] J. Bell and T. Weis. Greening the grid: Powering Alberta's future with renewable energy. Technical report, Pembina Institute, Apr. 2009.

[62] F. Bellosa. The benefits of event: driven energy accounting in power-sensitive systems. In *EW 9: Proceedings of the 9th workshop on ACM SIGOPS European workshop*, pages 37–42, New York, NY, USA, 2000. ACM.

[63] A. Beloglazov, R. Buyya, Y. C. Lee, and A. Zomaya. A taxonomy and survey of energy-efficient data centers and cloud computing systems. *Advances in Computers*, 82:47–111, 2011.

[64] How less became more... wind, power and unintended consequences in the Colorado energy market. Technical report, BenTeck Energy, Apr. 2010.

[65] A. Berl, E. Gelenbe, M. Di Girolamo, G. Giuliani, H. De Meer, M. Q. Dang, and K. Pentikousis. Energy-Efficient Cloud Computing. *The Computer Journal*, 2009.

[66] A. Bianzino, C. Chaudet, D. Rossi, and J. Roturier. A survey of green networking research. *IEEE Communications Surveys & Tutorials*, 14(1):3–20, 2012.

[67] I. F. Bitterlin. Exploring the concept of comparative PUE. Technical report, Prism Power, 2010.

[68] P. Bohrer, E. N. Elnozahy, T. Keller, M. Kistler, C. Lefurgy, C. McDowell, and R. Rajamony. The case for power management in web servers. *Power aware computing*, pages 261–289, 2002.

[69] BPA's Interim Environmental Redispatch and Negative Pricing Policies. Technical report, Bonneville Power Authority, May 2011.

[70] T. D. Boucher, D. M. Auslander, C. Bash, C. C. Federspiel, and C. Patel. Viability of dynamic cooling control in a data center environment. In *ITHERM '04. The Ninth Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems, 2004.*, pages 593–600 Vol.1, June 2004.

[71] J. Burge, P. Ranganathan, and J. L. Wiener. Cost-aware scheduling for heterogeneous enterprise machines (CASH'EM). Technical Report HPL-2007-63, HP Laboratories, Apr. 2007.

[72] J. M. Carrasco, L. G. Franquelo, J. T. Bialasiewicz, E. Galvan, R. C. P. Guisado, M. A. M. Prats, J. I. Leon, and N. Moreno-Alfonso. Power-electronic systems for the grid integration of renewable energy sources: A survey. *IEEE Transactions on Industrial Electronics*, 53(4):1002–1016, June 2006.

[73] J. Chase, D. C. Anderson, P. N. Thakar, A. Vahdat, and R. P. Doyle. Managing energy and server resources in hosting centers. In *Proceedings of the 18th ACM Symposium on Operating Systems Principles*, pages 103–116, New York, NY, USA, 2001. ACM.

[74] Y. Chen, A. Das, W. Qin, A. Sivasubramaniam, Q. Wang, and N. Gautam. Managing server energy and operational costs in hosting centers. In *Proceedings of the 2005 ACM International Conference on Measurement and Modeling of Computer Systems*, pages 303–314, New York, NY, USA, 2005. ACM.

[75] C. Clark, K. Fraser, S. Hand, J. G. Hansen, E. Jul, C. Limpach, I. Pratt, and A. Warfield. Live migration of virtual machines. In *Proceedings of the 2nd Symposium on Networked Systems Design & Implementation*, pages 273–286, Berkeley, CA, USA, 2005. USENIX Association.

[76] J. Dean and S. Ghemawat. MapReduce: simplified data processing on large clusters. In *Proceedings of the 6th Symposium on Operating Systems Design and Implementation*, Dec. 2004.

[77] Digital Realty Trust. US Campos survey results. `http://s3.amazonaws.com/files.digitalrealtytrust/Reports/DLR_US_2012_FINAL.pdf`, Jan. 2012.

[78] F. Douglis, P. Krishnan, and B. N. Bershad. Adaptive disk spin-down policies for mobile computers. In *Proceedings of the 2nd Symposium on Mobile and Location-Independent Computing*, pages 121–137, Berkeley, CA, USA, 1995. USENIX Association.

[79] D. L. Eager, E. D. Lazowska, and J. Zahorjan. The limited performance benefits of migrating active processes for load sharing. *SIGMETRICS Performance Evaluation Review*, 16(1):63–72, 1988.

[80] Amazon Elastic Compute Cloud. `http://aws.amazon.com/ec2/`.

[81] N. El-Sayed, I. Stefanovici, G. Amvrosladis, A. Hwang, and B. Schroeder. Temperature management in data centers: Why some (might) like it hot. In *Proceedings of SIGMETRICS 2012*, 2012.

[82] Elgg - open source social networking engine. `http://elgg.org/`.

[83] A. J. Elmore, S. Das, D. Agrawal, and A. El Abbadi. Zephyr: live migration in shared nothing databases for elastic cloud platforms. In *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*, SIGMOD '11, pages 301–312, New York, NY, USA, 2011. ACM.

[84] E. N. Elnozahy, M. Kistler, and R. Rajamony. *Energy-Efficient Server Clusters*, volume 2325/2003 of *Lecture Notes in Computer Science*, pages 179–197. Springer Berlin / Heidelberg, 2003.

[85] Enterprise server and data center energy efficiency initiatives. `http://www.energystar.gov/index.cfm?c=prod_development.server_efficiency`.

[86] EnerNOC. http://www.enernoc.com.

[87] Environment Canada. Electricity intensity tables. `http://www.ec.gc.ca/ges-ghg/default.asp?lang=En&n=EAF0E96A-1`.

[88] Environment Canada. Canadian weather energy and engineering datasets. Technical report, National Climate Data and Information Archive, 2010.

[89] Report to Congress on server and data center energy efficiency. Public Law 109-431, U.S. Environmental Protection Agency, ENERGY STAR Program, Aug. 2007.

[90] Advance notice of real-time prices: How is this handled in other competitive wholesale electricity markets? ERCOT DSWG Meeting, Oct. 2010.

[91] X. Fan, W.-D. Weber, and L. A. Barroso. Power provisioning for a warehouse-sized computer. In *Proceedings of the 34th Annual International Symposium on Computer Architecture*, pages 13–23, New York, NY, USA, 2007. ACM.

[92] Frequency regulation compensation in the organized wholesale power markets. Docket Nos. RM11-7-000 & AD10-11-000, Federal Energy Regulatory Commission, Feb. 2011.

[93] W. Felter, K. Rajamani, T. Keller, and C. Rusu. A performance-conserving approach for reducing peak power consumption in server systems. In *Proceedings of the 19th Annual International Conference on Supercomputing*, pages 293–302, New York, NY, USA, 2005. ACM.

[94] F. C. Figueiredo and P. C. Flynn. Using diurnal power price to configure pumped storage. *IEEE Transactions on Energy Conversion*, 21(3):804 –809, Sept. 2006.

[95] S. Figuerola, M. Lemay, V. Reijs, M. Savoie, and B. St. Arnaud. Converged optical network infrastructures in support of future internet and grid services using IaaS to reduce GHG emissions. *Journal of Lightwave Technology*, 27(12):1941–1946, June 2009.

[96] J. Flinn and M. Satyanarayanan. Energy-aware adaptation for mobile applications. In *Proceedings of the 17th ACM Symposium on Operating Systems Principles*, pages 48–63, New York, NY, USA, 1999. ACM.

[97] S. Fok. Data center air management report. Technical report, Pacific Gas and Electric Company, Dec. 2006.

[98] E. Frachtenberg, D. Lee, M. Magarelli, V. Mulay, and J. Park. Thermal design in the open compute datacenter. In *2012 13th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems*, pages 530 –538, 2012.

[99] V. W. Freeh, D. K. Lowenthal, F. Pan, N. Kappiah, R. Springer, B. L. Rountree, and M. E. Femal. Analyzing the energy-time trade-off in high-performance computing applications. *IEEE Transactions on Parallel and Distributed Systems*, 18(6):835–848, June 2007.

[100] J. A. Gamon, C. Kiddle, R. Curry, N. Markatchev, G. Zonta-Pastorello, Jr., B. Rivard, G. A. Sanchez-Azofeifa, R. Simmonds, and T. Tan. GeoChronos: An On-line Collaborative Platform for Earth Observation Scientists. *AGU Fall Meeting Abstracts*, Dec. 2009.

[101] Green building initiative. `http://www.thegbi.org/`.

[102] S. Gilbert and N. Lynch. Brewer's conjecture and the feasibility of consistent, available, partition-tolerant web services. *SIGACT News*, 33(2):51–59, June 2002.

[103] D. Gmach, J. Rolia, C. Bash, Y. Chen, T. Christian, and A. Shah. Capacity planning and power management to exploit sustainable energy. In *Proceedings of the 6th International Conference on Network and Services Management (CNSM 2010)*, Oct. 2010.

[104] I. Goiri, R. Beauchea, K. Le, T. D. Nguyen, M. E. Haque, J. Guitart, J. Torres, and R. Bianchini. Greenslot: scheduling energy consumption in green datacenters. In *Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis*, SC '11, pages 20:1–20:11, New York, NY, USA, 2011. ACM.

[105] I. Goiri, K. Le, T. D. Nguyen, J. Guitart, J. Torres, and R. Bianchini. GreenHadoop: leveraging green energy in data-processing frameworks. In *Proceedings of the 7th ACM European conference on Computer Systems*, EuroSys '12, pages 57–70, New York, NY, USA, 2012. ACM.

[106] Google cloud platform. `https://cloud.google.com/`.

[107] S. Govindan, A. Sivasubramaniam, and B. Urgaonkar. Benefits and limitations of tapping into stored energy for datacenters. In *Proceedings of the 38th International Symposium on Computer Architecture*, June 2011.

[108] S. Govindan, D. Wang, A. Sivasubramaniam, and B. Urgaonkar. Leveraging stored energy for handling power emergencies in aggressively provisioned datacenters. *SIGARCH Comput. Archit. News*, 40(1):75–86, Mar. 2012.

[109] The Green500 list. `http://www.green500.org`.

[110] The Green Grid. `http://www.thegreengrid.org/`.

[111] Make IT green: Cloud computing and its contribution to climate change. Technical report, GreenPeace, Mar. 2010.

[112] GreenStar network: Building a zero-carbon network. `http://www.greenstarnetwork.com/`.

[113] Y. Guo, Z. Ding, Y. Fang, and D. Wu. Cutting down electricity cost in internet data centers by using energy storage. In *2011 IEEE Global Telecommunications Conference (GLOBECOM 2011)*, pages 1 –5, Dec. 2011.

[114] S. Gurumurthi, A. Sivasubramaniam, M. Kandemir, and H. Franke. DRPM: dynamic speed control for power management in server class disks. In *Proceedings of the 30th Annual International Symposium on Computer Architecture*, pages 169–179, June 2003.

[115] D. Hackenberg, R. Schöne, D. Molka, M. Müller, and A. Knüpfer. Quantifying power consumption variations of HPC systems using SPEC MPI benchmarks. *Computer Science - Research and Development*, 25:155–163, 2010. 10.1007/s00450-010-0118-0.

[116] C. Hannemann, V. Carey, A. Shah, and C. Patel. Lifetime exergy consumption of an enterprise server. In *Proceedings of the 2008 IEEE International Symposium on Electronics and the Environment*, pages 1–5, May 2008.

[117] T. Heath, B. Diniz, E. V. Carrera, W. Meira, Jr., and R. Bianchini. Energy conservation in heterogeneous server clusters. In *Proceedings of the 10th ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming*, pages 186–195, New York, NY, USA, 2005.

[118] D. P. Helmbold, D. D. E. Long, and B. Sherrod. A dynamic disk spin-down technique for mobile computing. In *Proceedings of the 2nd Annual International Conference on Mobile Computing and Networking*, pages 130–142, New York, NY, USA, 1996.

[119] A. Heydari and P. Sabounchi. Refrigeration assisted spot cooling of a high heat density data center. In *The Ninth Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems, 2004. ITHERM '04.*, pages 601–606 Vol.1, June 2004.

[120] M. R. Hines and K. Gopalan. Post-copy based live virtual machine migration using adaptive pre-paging and dynamic self-ballooning. In *Proceedings of the 2009 ACM SIGPLAN/SIGOPS international conference on Virtual execution environments*, pages 51–60, New York, NY, USA, 2009.

[121] T. Hirofuchi, H. Nakada, S. Itoh, and S. Sekiguchi. Enabling instantaneous relocation of virtual machines with a lightweight VMM extension. In *10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing*, pages 73–83, May 2010.

[122] E. Hirst and B. Kirby. Ancillary-service details: operating reserves. Technical Report ORNL/CON–452, Oak Ridge National Laboratory, Nov. 1997.

[123] H. Hlavacs, G. Da Costa, and J. M. Pierson. Energy consumption of residential and professional switches. In *International Conference on Computational Science and Engineering, 2009*, volume 1, pages 240 –246, Aug. 2009.

[124] D. Huang, D. Ye, Q. He, J. Chen, and K. Ye. Virt-LM: a benchmark for live migration of virtual machine. In *Proceeding of the second joint WOSP/SIPEW international conference on Performance engineering*, pages 307–316, New York, NY, USA, 2011.

[125] F. Huntowski, A. Patterson, and M. Schnitzer. Negative electricity prices and the production tax credit: Why wind producers can pay us to take their power and why that is a bad thing. Technical report, The NorthBridge Group, Sept. 2012.

[126] K. Z. Ibrahim, S. Hofmeyr, C. Iancu, and E. Roman. Optimized pre-copy live migration for memory intensive applications. In *Proceedings of 2011 International Conference for High Performance Computing, Networking, Storage and Analysis*, pages 40:1–40:11, New York, NY, USA, 2011.

[127] Integrating renewable generation: SE-91 presentation of design principles. Technical report, Independent Electrical System Operator, Dec. 2010.

[128] Y. Inoue. Climate change and ICT standardization. Kyoto Symposium on ICT & Climate Change, Apr. 2008.

[129] International Organization for Standardization. ISO-14064-2, Greenhouse gases. Part 2. Specification with guidance at the project level for quantification, monitoring and reporting of greenhouse gas emission reductions or removal enhancements. Technical report, International Organization for Standardization, Geneva, Switzerland, 2006.

[130] International Organization for Standardization. ISO-14064-3, Greenhouse gases – Part 3: Specification with guidance for the validation and verification of greenhouse gas assertions. Technical report, International Organization for Standardization, Geneva, Switzerland, 2006.

[131] A. Janjani and H. Zareipour. An overview of the operation of the Alberta electricity market. In *Proceedings of the 2010 Canadian IEEE Conference on Electrical and Computer Engineering*, May 2010.

[132] R. Jansen and P. R. Brenner. Energy efficient virtual machine allocation in the cloud. In *2011 International Green Computing Conference and Workshops (IGCC)*, pages 1–8, July 2011.

[133] W. S. Jevons. *The Coal Question: An Enquiry Concerning the Progress of the Nation, and the Probable Exhaustion of Our Coal-mines.* Macmillan & Co. London, 1865.

[134] H. Jin, L. Deng, S. Wu, X. Shi, and X. Pan. Live virtual machine migration with adaptive, memory compression. In *IEEE International Conference on Cluster Computing and Workshops, 2009*, pages 1–10, Aug 31–sept. 4 2009.

[135] E. Johnson. Goodbye to carbon neutral: Getting biomass footprints right. *Environmental Impact Assessment Review*, 29(3):165 – 168, 2009.

[136] L. Johnson and C. Hope. The social cost of carbon in U.S. regulatory impact analyses: an introduction and critique. *Journal of Environmental Studies and Sciences*, pages 1–17, 2012.

[137] M. P. Johnson, A. Bar-Noy, O. Liu, and Y. Feng. Energy peak shaving with local storage. *Sustainable Computing: Informatics and Systems*, 1(3):177 – 188, 2011.

[138] S. Kamil, J. Shalf, and E. Strohmaier. Power efficiency in high performance computing. In *Proceedings of the IEEE International Symposium on Parallel and Distributed Processing*, pages 1–8, Apr. 2008.

[139] W. Katzenstein and J. Apt. Air emissions due to wind and solar power. *Environmental Science & Technology*, 43(2):253–258, 2009.

[140] W. Kempton and J. Tomic. Vehicle-to-grid power fundamentals: Calculating capacity and net revenue. *Journal of Power Sources*, 144(1):268 – 279, 2005.

[141] M. Kende. Internet global growth: lessons for the future. Technical report, Analysys Mason, Sept. 2012.

[142] C. A. King. New York state electricity markets. In *Market Design Technical Conference*. Federal Energy Regulatory Commission, 2002.

[143] B. Kirby. Load response fundamentally matches power system reliability requirements. In *IEEE Power Engineering Society General Meeting, 2007*, pages 1–6, June 2007.

[144] P. Komor. Wind and solar electricity: Challenges and opportunities. Technical report, Pew Center on Global Climate Change, 2009.

[145] V. Kontorinis, L. E. Zhang, B. Aksanli, J. Sampson, H. Homayoun, E. Pettis, D. M. Tullsen, and T. S. Rosing. Managing distributed UPS energy for effective power capping in data centers. In *2012 39th Annual International Symposium on Computer Architecture (ISCA)*, pages 488 –499, June 2012.

[146] J. G. Koomey. Growth in data center electricity use 2005 to 2010. Technical report, Analytics Press at the request of the New York Times, Oakland, CA, Aug. 2011.

[147] A. Krioukov, C. Goebel, S. Alspaugh, Y. Chen, D. E. Culler, and R. Katz. Integrating renewable energy using data analytics systems: Challenges and opportunities. *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering*, 2011.

[148] B. Krishnan, H. Amur, A. Gavrilovska, and K. Schwan. VM power metering: feasibility and challenges. *SIGMETRICS Perform. Eval. Rev.*, 38(3):56–60, Jan. 2011.

[149] M. Lamanna. The LHC computing grid project at CERN. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 534(1-2):1–6, Nov. 2004.

[150] C. Lange, D. Kosiankowski, C. Gerlach, F. J. Westphal, and A. Gladisch. Energy consumption of telecommunication networks. In *35th European Conference on Optical Communication, 2009. ECOC '09.*, pages 1 –2, Sept. 2009.

[151] B. Lawson and E. Smirni. Power-aware resource allocation in high-end systems via online simulation. In *Proceedings of the 19th Annual International Conference on Supercomputing*, pages 229–238, New York, NY, USA, 2005. ACM.

[152] M. L. Lazarewicz. Status of flywheel storage: Operation of first frequency regulation plants. In *Electrical Energy Storage Applications and Technologies*, Oct. 2011.

[153] M. L. Lazarewicz and T. M. Ryan. Integration of flywheel-based energy storage for frequency regulation in deregulated markets. In *2010 IEEE Power and Energy Society General Meeting*, pages 1 –6, July 2010.

[154] K. Le, R. Bianchini, M. Martonosi, and T. D. Nguyen. Cost- and energy-aware load distribution across data centers. In *Proceedings of the Workshop on Power Aware Computing and Systems*, Oct. 2009.

[155] K. Le, J. Zhang, J. Meng, R. Bianchini, T. D. Nguyen, and Y. Jaluria. Reducing electricity cost through virtual machine placement in high performance computing clouds. In *Proceedings of Super Computing (SC11)*, Nov. 2011.

[156] A. Lebeck, X. Fan, H. Zeng, and C. S. Ellis. Power aware page allocation. In *Proceedings of the 9th International Conference on Architectural Support for*

Programming Languages and Operating Systems, pages 105–116, New York, NY, USA, 2000.

[157] V. W. Lee, C. Kim, J. Chhugani, M. Deisher, D. Kim, A. D. Nguyen, N. Satish, M. Smelyanskiy, S. Chennupaty, P. Hammarlund, R. Singhal, and P. Dubey. Debunking the 100x GPU vs. CPU myth: an evaluation of throughput computing on CPU and GPU. *SIGARCH Comput. Archit. News*, 38(3):451–460, 2010.

[158] C. Lefurgy, X. Wang, and M. Ware. Power capping: a prelude to power shifting. *Cluster Computing*, 11(2):183–195, June 2008.

[159] M. Lenzen. Life cycle energy and greenhouse gas emissions of nuclear energy: A review. *Energy Conversion and Management*, 49(8):2178–2199, Aug. 2008.

[160] I. M. Leslie, D. Mcauley, R. Black, T. Roscoe, P. Barham, D. Evers, R. Fairbairns, and E. Hyden. The design and implementation of an operating system to support distributed multimedia applications. *IEEE Journal on Selected Areas in Communications*, 14(7):1280–1297, Sept. 1996.

[161] F. Li and D. L. Tolley. Long-run incremental cost pricing based on unused capacity. *IEEE Transactions on Power Systems*, 22(4):1683 –1689, Nov. 2007.

[162] B. Liebau and W. J. Strobele. The shortcomings of the architecture of the German electricity market. *The Electricity Journal*, 24(5):21 – 28, 2011.

[163] M. Lin, Z. Liu, A. Wierman, and L. Andrew. Online algorithms for geographical load balancing. In *Proceedings of the Third International Green Computing Conference (IGCC'12)*, June 2012.

[164] Z. Liu, Y. Chen, C. Bash, A. Wierman, D. Gmach, Z. Wang, M. Marwah, and C. Hyser. Renewable and cooling aware workload management for sustainable

data centers. In *ACM SIGMETRICS*, June 2012.

[165] N. Markatchev, R. Curry, C. Kiddle, A. Mirchovski, R. Simmonds, and T. Tan. A cloud-based interactive application service. In *Proceedings of the 5th IEEE International Conference on e-Science*, Dec. 2009.

[166] B. Marsh, F. Douglis, and P. Krishnan. Flash memory file caching for mobile computers. In *Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences, 1994*, pages 451 –460, Jan. 1994.

[167] S. Marston, Z. Li, S. Bandyopadhyay, J. Zhang, and A. Ghalsasi. Cloud computing: The business perspective. *Decision Support Systems*, 51(1):176–189, 2011.

[168] D. Meisner, B. T. Gold, and T. F. Wenisch. PowerNap: eliminating server idle power. *SIGPLAN Not.*, 44(3):205–216, 2009.

[169] P. Mell and T. Grance. The NIST definition of cloud computing. Special Publication 800-145, National Institute of Standards and Technology, Sept. 2011.

[170] M. Merritt. Alberta, B.C., share power relationship. Edmonton Journal, May 2008.

[171] R. Miller. Apple: Greenpeace's cloud math is busted. `http://www.datacenterknowledge.com/archives/2012/04/17/apple-greenpeaces-cloud-math-is-busted/`, Apr. 2012.

[172] R. Miller. Google embraces thermal storage in Taiwan. `http://www.datacenterknowledge.com/archives/2012/04/03/google-embraces-thermal-storage-in-taiwan/`, Apr. 2012.

[173] M. Milligan and K. Porter. Determining the capacity value of wind: An updated survey of methods and implementation. *Windpower '08*, June 2008.

[174] J. D. Mitchell-Jackson. Energy needs in an internet economy: A closer look at data centers. Master's thesis, University of California, Berkeley, May 2001.

[175] A. Mohsenian-Rad and A. Leon-Garcia. Coordination of cloud computing and smart power grids. In *2010 First IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 368–372, Oct. 2010.

[176] J. Moore, J. Chase, K. Farkas, and P. Ranganathan. Data center workload monitoring, analysis, and emulation. In *Eighth Workshop on Computer Architecture Evaluation using Commercial Workloads*, Feb. 2005.

[177] J. Moore, J. Chase, P. Ranganathan, and R. Sharma. Making scheduling "cool": temperature-aware workload placement in data centers. In *Proceedings of the USENIX Annual Technical Conference*, Berkeley, CA, USA, Mar. 2005. USENIX Association.

[178] B. Mulliken. Delivering green power to the people - nationwide. *Green Energy News*, 14(30), Oct. 2009.

[179] National Renewable Energy Laboratory (NREL). Wind integration datasets. `http://www.nrel.gov/wind/integrationdatasets/`.

[180] M. Nelson, B.-H. Lim, and G. Hutchins. Fast transparent migration for virtual machines. In *Proceedings of the USENIX Annual Technical Conference*, page 25, Berkeley, CA, USA, 2005.

[181] Flexibility requirements and metrics for variable generation: Implications for system planning studies. Special report, North American Electrical Reliability Corporation, Aug. 2010.

[182] Ancillary services and balancing authority area solutions to integrate variable generation. Technical report, North American Electrical Reliability Corporation, Mar. 2011.

[183] Operating practices, procedures, and tools. IVGTF task 2.4 report, North American Electrical Reliability Corporation, Mar. 2011.

[184] R. Neugebauer and D. Mcauley. Energy is just another resource: energy accounting and energy pricing in the Nemesis OS. In *Proceedings of the 8th Workshop on Hot Topics in Operating Systems*, pages 67–72, May 2001.

[185] Accommodating high levels of variable generation. Technical report, North American Electric Reliability Corporation, Apr. 2009.

[186] Greener and smarter: ICTs, the environment, and climate change. Technical report, Organization for Economic Cooperation and Development, Sept. 2010.

[187] OpenStack. `http://www.openstack.org/`.

[188] A.-C. Orgerie, L. Lefevre, and J.-P. Gelas. Demystifying energy consumption in grids and clouds. In *2010 International Green Computing Conference*, pages 335–342, Aug. 2010.

[189] S. Osman, D. Subhraveti, G. Su, and J. Nieh. The design and implementation of Zap: a system for migrating computing environments. *ACM SIGOPS Operating Systems Review*, 36(SI):361–376, 2002.

[190] Parallel Workloads Archive. `http://www.cs.huji.ac.il/labs/parallel/workload/`.

[191] S. Pearson and A. Benameur. Privacy, security and trust issues arising from cloud computing. In *2010 IEEE Second International Conference on Cloud Computing Technology and Science (CloudCom)*, pages 693 –702, 2010.

[192] M. V. Pereira, L. A. Barroso, and J. Rosenblatt. Supply adequacy in the Brazilian power market. In *Proceedings of the 2004 IEEE Power Engineering Society General Meeting*, pages 1016–1021, June 2004.

[193] T. Pering, T. Burd, and R. Brodersen. The simulation and evaluation of dynamic voltage scaling algorithms. In *Proceedings of the 1998 International Symposium on Low Power Electronics and Design*, pages 76–81, New York, NY, USA, 1998.

[194] R. Peters and L. O'Malley. Storing renewable power. Technical report, Pembina Institute, June 2008.

[195] J. Pfeifenberger and A. Hajos. Demand response review. Technical report, The Brattle Group, Mar. 2011.

[196] E. Pinheiro, R. Bianchini, E. Carrera, and T. Heath. Load balancing and unbalancing for power and performance in cluster-based systems. Technical Report DCS-TR-440, Rutgers University, 2001.

[197] E. Pinheiro, R. Bianchini, and C. Dubnicki. Exploiting redundancy to conserve energy in storage systems. *SIGMETRICS Perform. Eval. Rev.*, 34(1):15–26, 2006.

[198] E. Pinheiro, W.-D. Weber, and L. A. Barroso. Failure trends in a large disk drive population. In *5th USENIX Conference on File and Storage Technologies (FAST 2007)*, pages 17–29, 2007.

[199] M. Poess and R. O. Nambiar. Energy cost, the key challenge of today's data centers: a power consumption analysis of TPC-C results. *Proc. VLDB Endow.*, 1(2):1229–1240, Aug. 2008.

[200] D. Pratt. Demand side ancillary services. `http://www.nyiso.com/public/webdocs/services/market_training/workshops_courses/other/demand_side_ancilsvcs.pdf`, Nov. 2008.

[201] A. Qouneh, C. Li, and T. Li. A quantitative analysis of cooling power in container-based data centers. In *2011 IEEE International Symposium on Workload Characterization (IISWC)*, pages 61 –71, nov. 2011.

[202] A. Qureshi, R. Weber, H. Balakrishnan, and J. Guttag. Cutting the electric bill for Internet-scale systems. In *ACM SIGCOMM*, 2009.

[203] Rackforce. `http://www.rackforce.com/`.

[204] Rackspace cloud. `http://www.rackspace.com/cloud/`.

[205] J. Rambo and Y. Joshi. Modeling of data center airflow and heat transfer: State of the art and future trends. *Distributed and Parallel Databases*, 21(2):193–225, June 2007.

[206] P. Ranganathan, P. Leech, D. Irwin, and J. Chase. Ensemble-level power management for dense blade servers. *ACM SIGARCH Computer Architecture News*, 34(2):66–77, 2006.

[207] N. Rasmussen. Re-examining the suitability of the raised floor for data center applications, revision 1. APC, 2007.

[208] N. Rasmussen and J. Spitaels. A quantitative comparison of high efficiency AC vs. DC power distribution for data centers. White Paper 23 (revision 3), Schneider Electric, 2012.

[209] S. Rivoire, M. A. Shah, P. Ranganathan, and C. Kozyrakis. Joulesort: a balanced energy-efficiency benchmark. In *Proceedings of the 2007 ACM SIGMOD*

international conference on Management of data, SIGMOD '07, pages 365–376, New York, NY, USA, 2007.

[210] D. Rosenberg, R. A. Bodaly, and P. J. Usher. Environmental and social impacts of large scale hydroelectric development: who is listening? *Global Environmental Change*, 5(2):127 – 148, 1995.

[211] R. Rudervall, J. P. Charpentier, and R. Sharma. High voltage direct current (HVDC) transmission systems technology review paper. In *Energy Week 2000*, Mar. 2000.

[212] S. Ruth. Green IT more than a three percent solution? *IEEE Internet Computing*, 13(4):74 –78, July-Aug. 2009.

[213] S. Sakr, A. Liu, D. M. Batista, and M. Alomari. A survey of large scale data management approaches in cloud environments. *IEEE Communications Surveys Tutorials*, 13(3):311 –336, 2011.

[214] D. Salomonsson, L. Soder, and A. Sannino. An adaptive control system for a DC microgrid for data centers. In *Conference Record of the 2007 IEEE Industry Applications Conference, 2007*, pages 2414 –2421, Sept. 2007.

[215] H. D. Saunders. The Khazzoom-Brookes Postulate and neoclassical growth. *The Energy Journal*, 13(4):131–148, Jan. 1992.

[216] R. Sharma, T. Christian, M. Arlitt, C. Bash, and C. Patel. Design of farm waste-driven supply side infrastructure for data centers. In *Proceedings of ASME 2010 4th International Conference on Energy Sustainability*, May 2010.

[217] Smart 2020: Enabling the low carbon economy in the information age. Technical report, The Climate Group, 2008.

[218] S. C. Smith, P. K. Sen, and B. Kroposki. Advancement of energy storage devices and applications in electrical power system. In *2008 IEEE Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century*, pages 1–8, July 2008.

[219] Report on cloud computing to the OSG steering committee. Technical report, Standard Performance Evaluation Corporation, Apr. 2012.

[220] Server Efficiency Rating Tool. `http://www.spec.org/sert/`.

[221] SPECpower_ssj2008 results for Fujitsu PRIMERGY RX300 S7 (Intel Xeon E5-2660). `http://www.spec.org/power_ssj2008/results/res2012q1/power_ssj2008-20120305-00429.html`.

[222] SPECpower_ssj2008 results for Huawei Technologies Co., Ltd RH2288 V2. `http://www.spec.org/power_ssj2008/results/res2012q1/power_ssj2008-20120306-00439.html`.

[223] SPECpower_ssj2008. `http://www.spec.org/power_ssj2008/`.

[224] Specvirt_sc2010. `http://www.spec.org/virt_sc2010/`.

[225] K. Srinivasan, S. Yuuw, and T. J. Adelmeyer. Dynamic VM migration: assessing its risks & rewards using a benchmark. *SIGSOFT Softw. Eng. Notes*, 36(5):42–42, Sept. 2011.

[226] C. Stewart and K. Shen. Some joules are more precious than others: Managing renewable energy in the datacenter. In *Proceedings of the Workshop on Power Aware Computing and Systems*, 2009.

[227] J. Stoess, C. Lang, and F. Bellosa. Energy management for hypervisor-based virtual machines. In *Proceedings of the USENIX Annual Technical Conference*, pages 1–14, Berkeley, CA, USA, 2007. USENIX Association.

[228] C. Strauch. NoSQL databases. `http://www.christof-strauch.de/nosqldbs.pdf`.

[229] N. Tanaka. The oil market: A longer-term perspective. `http://www.iea.org/speech/2008/tanaka/g8_economy.pdf`.

[230] J. W. Tester et al. The future of geothermal energy - impact of enhanced geothermal systems (EGS) on the United States in the 21st century. Technical report, Idaho National Laboratory, 2006.

[231] Top 500 supercomputers. `http://top500.org/`.

[232] F. Travostino, P. Daspit, L. Gommans, C. Jog, C. de Laat, J. Mambretti, I. Monga, B. van Oudenaarde, S. Raghunath, and P. Yonghui Wang. Seamless live migration of virtual machines over the MAN/WAN. *Future Generation Computer Systems*, 22(8):901–907, Oct. 2006.

[233] W. Tschudi, O. VanGeet, J. Cooley, and D. Azevedo. ERE: A metric for measuring the benefit of reuse energy from a data center. White Paper 29, The Green Grid, Sept. 2010.

[234] W. Tschudi and C. Williams. Power supply options for data centers. Technical Report LBNL-63339, Lawrence Berkeley National Laboratory, July 2007.

[235] R. Urgaonkar, B. Urgaonkar, M. J. Neely, and A. Sivasubramanian. Optimal power cost management using stored energy in data centers. In *SIGMETRICS*, pages 221–232, June 2011.

[236] U.S. Department of Energy Solar Energy Technologies Program. Photovoltaic basics. http://www1.eere.energy.gov/solar/pv_basics.html.

[237] G. Varsamopoulos, Z. Abbasi, and S. K. S. Gupta. Trends and effects of energy proportionality on server provisioning in data centers. In *2010 International Conference on High Performance Computing (HiPC)*, pages 1–11, Dec. 2010.

[238] V. Venkatachalam and M. Franz. Power reduction techniques for microprocessor systems. *ACM Comput. Surv.*, 37(3):195–237, Sept. 2005.

[239] A. Verma, P. Ahuja, and A. Neogi. pMapper: power and migration cost aware application placement in virtualized systems. In *Proceedings of the 9th ACM/IFIP/USENIX International Conference on Middleware*, pages 243–264, New York, NY, USA, 2008.

[240] K. V. Vishwanath, A. Greenberg, and D. A. Reed. Modular data centers: how to design them? In *Proceedings of the 1st ACM workshop on Large-Scale System and Application Performance*, pages 3–10, New York, NY, USA, 2009. ACM.

[241] D. Wang, C. Ren, A. Sivasubramaniam, B. Urgaonkar, and H. Fathy. Energy storage in datacenters: what, where, and how much? In *Proceedings of the 12th ACM SIGMETRICS/PERFORMANCE joint international conference on Measurement and Modeling of Computer Systems*, SIGMETRICS '12, pages 187–198, New York, NY, USA, 2012.

[242] P. Wang, H. Zareipour, and W. D. Rosehart. Characteristics of the prices of operating reserves and regulation services in competitive electricity markets. *Energy Policy*, 39(6):3210 – 3221, 2011.

[243] X. Wang and M. Chen. Cluster-level feedback power control for performance optimization. In *IEEE 14th International Symposium on High Performance Computer Architecture*, pages 101–110, Feb. 2008.

[244] Y. Wang, X. Wang, and Y. Zhang. Leveraging thermal storage to cut the

electricity bill for datacenter cooling. In *Proceedings of the 4th Workshop on Power-Aware Computing and Systems*, pages 8:1–8:5, New York, NY, USA, 2011.

[245] M. Weiser, B. Welch, A. Demers, and S. Shenker. Scheduling for reduced CPU energy. In *Proceedings of the 1st USENIX conference on Operating Systems Design and Implementation*, page 2, Berkeley, CA, USA, 1994.

[246] A. Weissel, B. Beutel, and F. Bellosa. Cooperative I/O: a novel I/O semantics for energy-aware applications. *ACM SIGOPS Operating Systems Review*, 36(SI):117–129, 2002.

[247] WestGrid. `http://www.westgrid.ca/`.

[248] E. J. Wilson. *Managing the Risks of Geologic Carbon Sequestration: A Regulatory and Legal Analysis*. PhD thesis, Purdue University, Oct. 2004.

[249] Windows Azure. `https://www.windowsazure.com/`.

[250] J. Wu, J. Zhang, and M. Savoie. Lightpath scheduling and routing for green data centres. *Telecommunication Systems*, pages 1–14. 10.1007/s11235-011-9644-8.

[251] D. Xu and X. Liu. Geographic trough filling for internet datacenters. In *2012 Proceedings IEEE INFOCOM*, pages 2881 –2885, Mar. 2012.

[252] K. Ye, D. Huang, X. Jiang, H. Chen, and S. Wu. Virtual machine based energy-efficient data center architecture for cloud computing: A performance perspective. In *Proceedings of the 2010 IEEE/ACM Int'l Conference on Green Computing and Communications & Int'l Conference on Cyber, Physical and Social Computing*, pages 171–178, Washington, DC, USA, 2010.

[253] X. Zhang, Z. Huo, J. Ma, and D. Meng. Exploiting data deduplication to accelerate live virtual machine migration. In *2010 IEEE International Conference on Cluster Computing*, pages 88–96, Sept. 2010.

[254] Y. Zhang, P. Chowdhury, M. Tornatore, and B. Mukherjee. Energy efficiency in telecom optical networks. *IEEE Communications Surveys Tutorials*, 12(4):441 –458, 2010.