Free Lunch: Exploiting Renewable Energy For Computing

Sherif Akoush, Ripduman Sohan, Andrew Rice, Andrew W. Moore and Andy Hopper

Computer Laboratory, University of Cambridge

`firstname.lastname@cl.cam.ac.uk`

Abstract

This paper argues for "Free Lunch", a computation architecture that exploits otherwise wasted renewable energy by (*i*) colocating datacentres with these remote energy sources, (*ii*) connecting them over a dedicated network, and (*iii*) providing a software framework that supports the seamless execution and migration of virtual machines in the platform according to power availability. This work motivates and outlines the architecture and demonstrates its viability with a case study. Additionally, we discuss the major technical challenges facing the successful deployment of Free Lunch and the limiting factors inherent in its design.

1 Introduction

There is currently strong support for mitigating both individual and organisational impact on the environment by reducing fossil fuel consumption and, by extension, decreasing the associated carbon footprint. Our dependency on information technology in our daily lives has led to computing infrastructure becoming a significant consumer of energy. A study commissioned in Japan in 2006 showed that digital infrastructure used 4% of total electricity production [1]. Similarly in the USA and UK it accounts for 3% [2] and 10% [3] of countrywide energy generation respectively.

As information technology is deployed more widely (and in ever increasing domains), it is likely that the total energy consumed by this sector will continue to increase despite gains achieved by technological enhancements. The total energy consumed by information technology in Japan for example, grew by 20% over the past five years [1] even though there has been marked improvement in computing and communication efficiency over that period.

In anticipation of this growth, our industry is beginning to explore renewable energy as an alternative energy source to fossil fuels in the hope of mitigating our ecological footprint [4]. For example, British Telecom, which consumed 0.7% of the total UK energy in 2007, has announced plans to develop wind farms to generate a quarter of its electricity requirement by 2016.

However, it is difficult to reliably and efficiently exploit renewable energy due to its unpredictability, variability and remote location. Recent studies have shown that incorporating renewable energy into the electricity grid is an expensive and inefficient process [5]. A significant amount of infrastructure investment is necessary to achieve the required levels of scale that result in a financially viable system [6]. The electricity grid requires upgrading in order to collect, store and distribute this type of energy. Additionally, renewable energy is typically located in remote areas leading to significant (up to 15%) attenuation losses when it is transmitted over long distances.

We have embraced the challenge of reducing individual and organisational energy consumption and are working towards an optimal digital infrastructure, as part of our research into computing for the Future of the Planet [7].[a] In this paper we outline "Free Lunch", an architecture that enables us to exploit otherwise wasted renewable energy without requiring the expensive investment necessary to deploy these resources at scale. Our design alleviates many of the disadvantages associated with incorporating renewable energy resources into the electrical grid. We locate datacentres close to renewable energy sites and link them together with a dedicated communications network. We migrate workloads seamlessly to locations that have enough power for them to continue execution. Our proposal is considered a free lunch because we make use of remote renewable energy that would otherwise be unutilised. Additionally, the energy consumed by computations in our architecture may be offset against an equal amount that would have been obtained from traditional (fossil fuel) sources.

In the remainder of this paper we describe our architecture in detail (Section 2), describe a case study that provides a broad overview on the viability of this system (Section 3), highlight the pertinent technical challenges (Section 4), and finally outline the limitations of our design (Section 5).

2 Architecture

The premise of our argument is analogous to the observation that it is only worth moving data to a remote computing facility if the computation requires more than 100,000 CPU cycles per byte of data [8]. Similarly, we believe that it is only worth moving computation and data to sources of renewable energy if it is economically cheaper than transmitting power to datacentres. As we have discussed previously, transmitting intermittent power over long distances requires significant infrastructure investment and is not considered optimal.

---

[a]http://www.youtube.com/watch?v=LN4H6vk1xYA
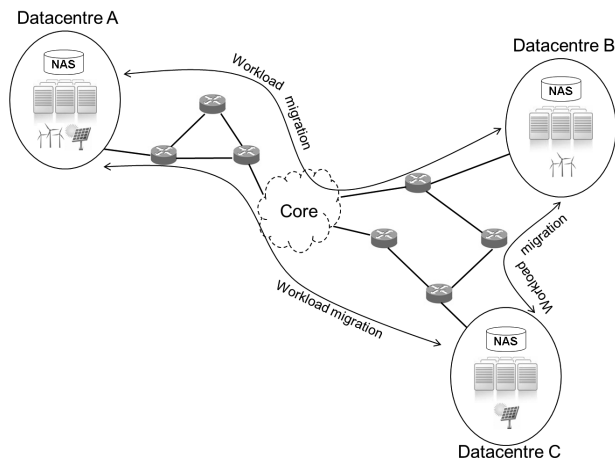
Figure 1: Free Lunch Infrastructure

Figure 1 illustrates the basic architecture. Our design is composed of a number of independent datacentres colocated with sources of renewable energy. These datacentres are interconnected through dedicated links. Compute jobs are submitted to the architecture where they are executed. Moreover, the system automatically migrates workloads between datacentres depending on energy availability. Free Lunch differs from traditional models in the following dimensions:

(i) Fluctuating Node Availability: Unlike traditional datacentre models, this architecture does not employ redundant or backup energy sources (e.g. generators). Moreover, due to their expense and impact on the environment we assume a minimal amount of backup energy sources exist on-site [6]. Consequently, server uptime is correlated to energy availability at a given location. The high intermittency of renewable energy [6] therefore results in a system with variable node availability.

(ii) Workload Mobility: Fluctuating node availability means that the continuity of executing tasks is insecure. Before servers are shutdown, tasks running on them are migrated to other servers on the same or different datacentres. Tasks are oblivious to migration and continue to run with minimal interruption. However if there is not enough capacity elsewhere in the system tasks are paused and restarted when energy resources are available.

(iii) Dedicated Network: We assume sites are linked by a dedicated high-speed and low-latency network. Because we have complete control over the physical connection we are able to accurately schedule task migration between sites without congestion or throughput concerns.

2.1 Enabling Technologies

Free Lunch is dependent on the following technologies:

(i) Virtualisation: Virtualisation platforms (such as Xen and VMware) enable three key functions in this architecture. Firstly, virtualisation provides abstraction and encap-

sulation functionality. The task (augmented with operating system, libraries and data) is packaged as a virtual machine (VM) that is loosely coupled with hardware thus simplifying storage and management. Secondly, virtualisation facilitates consolidation, which leads to improved hardware utilisation. Finally, current platforms provide migration functionality (such as XenMotion and VMotion) enabling the seamless mobility of VMs between physical hosts.

(ii) Cloud Computing Infrastructure: Cloud computing platforms enable an extensible software control plane that provides VM consolidation, scheduling, migration and load balancing functionalities. We intend to modify this platform for the purposes of our design.

(iii) Fast Networks: Developments in networking have delivered high-bandwidth low-latency networks with the capacity to render our architecture practical. Advances in control plane technology such as OpenFlow [9] permit the creation and deployment of non-standard networking architectures. Additionally, the area of wavelength-division multiplexing (WDM) optical IP networks which promise a low-energy transport system is becoming increasingly mainstream [10].

(iv) Scalable Modular Datacentres: The recent modular design of datacentres includes machines, networking, power distribution and cooling in one single unit. Consequently, it is simple and cheap to deploy (and later on add or remove) computational resources at renewable energy sites. A typical modular datacentre has a cost of approximately $600,000 which is 1% of the cost of building a traditional one.[b]

(v) Energy Harvesting Equipment: As the fundamental technology in renewable energy harvesting equipment continues to improve and economies of scale are achieved, it is likely that power efficiency will increase while the manufacturing cost of these devices decreases [11]. This sustained growth renders our architecture feasible and attractive.

3 Case Study

The viability of the Free Lunch architecture is dependent on harnessing enough energy to carry out a useful amount of computation without consuming a significant portion of it for workload migration between datacentres.[c] In this section we present a simple and approximate case study to illustrate broadly (*i*) the magnitude of energy that can be extracted from renewable energy resources and (*ii*) the overhead (in terms of energy and VM interruption) of migrations in the architecture. We stress that it is by no means an exact model, however, it is sufficient to provide a useful insight on the viability of our design. We focus on wind and solar energy in our study.

---

[b]Details available at http://dlc.sun.com/pdf/820-5806-10/820-5806-10.pdf

[c]Additionally, we have to minimise the idle time of servers so that their embodied energy is not wasted.

3.1 Energy Availability

We chose two locations that we believe have superior wind and solar power characteristics: near the Red Sea and the South West of Australia. Both sites are close to the equator and have excellent wind speeds all year round [12]. These sites are picked to complement each other: (*i*) their time zones are almost twelve hours apart which maximises the aggregate amount of solar energy over a 24 hour period and (*ii*) each location is in a different hemisphere which means that summer and winter are interchanged. With the recent push from industry to operate datacentres at higher temperatures, the choice of these two locations could be practically possible from a climate perspective.

We assume 10,000 m^2 of solar cells will be installed with an average efficiency of 10% [13]. Moreover, each location has one General Electric 1.5 MW wind turbine.

We estimate the wind energy generated by extrapolating wind speed from the weather stations' readings [14] at ground level to 80 m (the height of the wind turbine). Additionally, the wind energy is modelled as a function of the power curve of the turbine.

On the other hand, solar irradiation is dependent on a variety of factors such as latitude, time, atmospheric conditions and aerosols. We calculate the amount of solar energy using a simple insolation model [15] that approximates clear-sky solar irradiation after modifying it to account for cloud cover taken from the weather stations' readings [14].

For the year 2007 the *average* amount of energy (wind and solar) that is obtained per-site is 1 MW with a sum of 2 MW across our example configuration. This is sufficient to power approximately ten fully loaded Sun Modular Datacentres with a maximum combined capacity of 2,800 machines. This value increases linearly with the installation of additional energy harvesting equipment.

3.2 Migration Overhead

There is a considerable variation in the energy production due to the seasonality of solar power and the intermittency of wind power. Free Lunch is designed to cope with these fluctuations by (*i*) pausing VMs on one site until energy is available or (*ii*) migrating VMs to other datacentres that have excess power.

3.2.1 Number of Migrations

Consider the following simple scenario: each site is loaded initially to average capacity. When power generation falls below this average on either site we determine whether there is excess capacity (i.e. the amount of power being produced is above average) on the remote site. If this is the case we assume that we can proceed with a migration, otherwise we have to pause jobs on the local site. The underlying aim of this analysis is to determine how often we can migrate jobs (and therefore keep them running in spite of power fluctuation) as opposed to stopping them.

Using the configuration outlined above, for the year 2007 our results showed that there were 615 *combined* instances where power dropped below average (we assume 150 W per machine). This means that each site would experience an energy shortage approximately once a day. Out of these instances, excess power was available on the remote site 331 times allowing workloads to be relocated. However, we have to stop tasks 284 times as there is not enough capacity elsewhere.

3.2.2 Impact on Availability

For the purpose of quantifying the impact of migrations on VM availability, we extend our simple model by assuming that there is a third datacentre (possibly powered by non-renewable energy) which will *always* have the energy capacity to accept VMs. For the 615 instances when we have a power drop, jobs can be migrated elsewhere instead of being paused.

In our previous work on measuring the migration downtime on a number of production workloads [16], we showed that this downtime is highly workload, network and machine specific. On a 10 Gbps wide-area network spanning 10,000 miles, we found out that the worst case VM downtime is 677.1 ms. This value accounts for typical communication delay (100 ms) coupled with the final stop-and-copy stage for synchronising memory pages and disk sectors.[d] Assuming 615 migrations are done over the year, a VM will experience a total of 416.4 s migration originated downtime. For a service level agreement (SLA) defined as 99.5% uptime/year, this interruption in service due to migrations represents 0.3% of the maximum allowed outage.

3.2.3 Energy Cost

We also calculated the energy cost of a migration in the architecture. We assume Cisco CRS-1 routers with an energy profile of 3 nJ/bit operating within a network topology similar to a previous study [10]. This topology has 10 core hops between sites with a redundancy/cooling/growth factor of 8. We do not consider the optical transport system in our calculations as its energy consumption is minute (one order of magnitude less) compared to the energy consumption of the routers [10].

Additionally, we assume a VM memory size of 7.5 GB (similar to an Amazon EC2 "large" instance). The 615 migrations in our example occur with uniform periodicity (i.e. every 0.5 days). Furthermore, based on results obtained from our study, we assume 20 GB of *modified* disk state between migrations.

---

[d]We expect that datacentres are using dedicated storage systems, which are faster than communication interconnects. We also assume periodic synchronisation of disk state between datacentres.

Our results indicate that the total energy consumed during a migration in this architecture (including all networking elements) is 57.5 kJ ($\approx$ 0.16 kWh). We note that this value can be considerably lowered if we simplify the network topology by reducing the number of hops and use simpler routers as is possible with dedicated links.

4 Technical Challenges

This section highlights the technical challenges facing the successful deployment of the Free Lunch architecture:

4.1 VM Migration

Live migration [17] enables us to move entire *running* VMs between physical hosts with minimal downtime. While this technology is designed for intra-datacentre movement of VMs, we will also be using it for inter-datacentre migrations.

Current migration designs employ a simple "greedy" algorithm that is based on rounds of copying of modified memory pages between hosts over a short time period (typically in the order of seconds). Termination conditions are triggered if: (*i*) state changes between successive copy iterations are smaller than a pre-determined threshold or (*ii*) the migration process has run for longer than a predefined period [17].

While this algorithm performs well within datacentres, the higher latency and lower bandwidth properties of inter-datacentre links render the termination conditions inaccurate [16] which reduce the efficiency of the migration process. In the Free Lunch architecture physical limits can also become an issue. For example, two sites located at antipodes experience a minimum propagation delay of approximately 67 ms [18].

Furthermore, while existing algorithms are optimised for migrating a single VM between physical hosts we expect that, in the common case, Free Lunch migrations will occur at the rack level which requires a redesign of the migration subsystem to accommodate the simultaneous transfer of thousands of VMs.

4.2 Storage Synchronisation

Current migration architectures assume that all hosts in the system have common access to network attached storage [17]. This is a reasonable assumption given that the architectures were designed to work within a single datacentre. However, Free Lunch spans multiple geographically disperse sites resulting in the need to synchronise VM disk state between locations.

When a VM is migrated, its disk image has to be locally available in order for it to continue running without performance degradation. Keeping disk state synchronised is fundamentally a tradeoff between (*i*) the amount of data transferred for the purposes of maintaining sector synchronisation between hosts and (*ii*) the latency overhead of transferring sectors in the final round of a VM migration.

We are currently evaluating a disk migration algorithm that strives to minimise the amount of data transferred *and* migration latency.

Memory migration algorithms (Section 4.1) are typically unsuited to synchronise disk images that are larger (1–2 orders of magnitude) in size and have lower modification rates. We are currently studying this challenge and have created a *unified* migration algorithm that works for both memory and disk; it is able to build an accurate profile of modified memory pages and disk blocks to efficiently synchronise them with the destination host.

4.3 Scheduling and Placement

The traditional problem of optimising VM placement needs to be extended to account for the intermittency of energy and the transition overhead during migrations. A successful optimisation algorithm will strive to schedule jobs to minimise energy requirements while maximising resource utilisation without violating SLAs.

The placement of disk images too is an important concern in our architecture. Upon migration, VMs require the disk image to be locally accessible to continue functioning properly. However, it is difficult to predict when migrations will be needed in response to a power shortage. It is therefore likely that in a practical deployment of the architecture, the set of datacentres that a VM may be relocated to will be limited to the set of datacentres between which the VM disk state is periodically synchronised.

While an effective migration scheduler will strive to migrate all the VMs on a physical host without violating any individual SLA this may not be possible under certain circumstances. In this case, mechanisms for defining VM (migration) priorities are required.

5 Limitations

The Free Lunch architecture is, however, bounded by the following logical and physical limitations:

5.1 Workload Suitability

The non-standard properties of the Free Lunch architecture (specifically fluctuating node availability and workload mobility) may render it unsuitable for tasks requiring high uptime guarantees. Similarly, tasks with large working sets may be unsuitable candidates for migration.

Nevertheless, there is a large number of existing applications that would be appropriate for computation in our architecture. Simple stateless CPU bound tasks (e.g. quantitative analysis), scientific computations with small data requirements (e.g. n-body simulations) and non-interactive data processing workloads (e.g. mapreduce) are all ideal.

Ultimately, it is the answer to the question "will this workload be able to meet its SLA?" that determines whether the workload is appropriate for inclusion in the Free Lunch architecture. While the definition of SLA is highly application and user specific [19], in practice we

expect that large scale batch processing computations that are unaffected by downtime will be well suited to run on this system. On the other hand, latency sensitive or interactive applications will fare poorly.

Even if the SLA can be met, it is important to consider workloads with large disk footprints (i.e. modifying a large number of disk blocks) that may be unsuitable for inclusion in the architecture due to the high time and data transfer overhead associated with their migration.

5.2 Hardware Installation and Dependability

The nature of our design requires computation resources to be colocated with renewable energy resources. Typically these sites are in remote geographic regions. Gaining access to these datacentres for the purpose of managing hardware equipment is likely to be expensive and logistically challenging [20].

Limited hardware lifetimes (about three years) and component failure are also an issue. It is imperative that datacentres in the architecture are designed to gracefully accommodate failure. Harsh environmental conditions will also adversely impact hardware reliability.

6 Conclusion

In this work we have motivated and outlined "Free Lunch", an architecture that enables the exploitation of otherwise wasted renewable energy resources for the purpose of computation. We have provided a broad overview of the system, a simple case study that demonstrates its viability, the major technical challenges facing the successful implementation of our design and its inherent limitations. It is our belief that this infrastructure will be instrumental in reducing the ecological footprint of information technology both currently and in the future.

References

[1] M. Etoh, T. Ohya, and Y. Nakayama, "Energy consumption issues on mobile network systems," in *Proc. IEEE Symposium on Applications and the Internet (SAINT'08)*, pp. 365–368, August 2008.

[2] J. G. Koomey, "Estimating Total Power Consumption By Servers In The U.S. And The World," tech. rep., Lawrence Berkeley National Laboratory, February 2007.

[3] R. Barrington, "Green Computing: ICT and Climate Change - Saint or Sinner," in *BCS Business Lecture (IT-Wales'07)*, November 2007.

[4] C. Stewart and K. Shen, "Some Joules Are More Precious Than Others: Managing Renewable Energy in the Datacenter," in *Proc. USENIX Workshop on Power Aware Computing and Systems (HotPower'09)*, (San Diego, CA, USA), October 2009.

[5] B. V. Mathiesen and H. Lund, "Comparative analyses of seven technologies to facilitate the integration of fluctuating renewable energy sources," *Journal of Renewable Power Generation, IET*, vol. 3, no. 2, pp. 190–204, 2009.

[6] M. A. Delucchi and M. Z. Jacobson, "Providing all global energy with wind, water, and solar power, Part II: Reliability, system and transmission costs, and policies," *Energy Policy*, Dec. 2010.

[7] A. Hopper and A. Rice, "Computing for the future of the planet," *Philosophical Transactions of the Royal Society*, vol. 366, no. 1881, pp. 3685–3697, 2008.

[8] G. Bell, J. Gray, and A. Szalay, "Petascale computational systems," *Computer*, vol. 39, no. 1, pp. 110–112, 2006.

[9] N. McKeown, T. Anderson, H. Balakrishnan, G. Parulkar, L. Peterson, J. Rexford, S. Shenker, and J. Turner, "Openflow: enabling innovation in campus networks," *SIGCOMM Comput. Commun. Rev.*, vol. 38, pp. 69–74, March 2008.

[10] J. Baliga, R. Ayre, K. Hinton, W. V. Sorin, and R. S. Tucker, "Energy Consumption in Optical IP Networks," *Journal of Lightwave Technology*, vol. 27, no. 13, pp. 2391–2403, 2009.

[11] R. Gross, M. Leach, and A. Bauen, "Progress in renewable energy," *Environment International*, vol. 29, no. 1, pp. 105–122, 2003.

[12] C. L. Archer and M. Z. Jacobson, "Evaluation of global wind power," *Journal of Geophysical Research*, vol. 110, 2005.

[13] D. J. MacKay, *Sustainable Energy-without the hot air*. UIT, December 2008.

[14] "MIDAS Land Surface Stations data (1853-current), British Atmospheric Data Centre." http://badc.nerc.ac.uk/data/ukmo-midas, 2006.

[15] D. R. Brooks, "Simple model of solar irradiance." http://www.srrb.noaa.gov/highlights/sunrise/gen.html, March 2006.

[16] S. Akoush, R. Sohan, A. Rice, A. W. Moore, and A. Hopper, "Predicting the performance of virtual machine migration," in *Proc. IEEE Symposium on Modeling, Analysis, and Simulation of Computer Systems (MASCOTS'10)*, vol. 0, (Los Alamitos, CA, USA), pp. 37–46, August 2010.

[17] C. Clark, K. Fraser, S. Hand, J. G. Hansen, E. Jul, C. Limpach, I. Pratt, and A. Warfield, "Live migration of virtual machines," in *Proc. USENIX Symposium on Networked Systems Design and Implementation (NSDI'05)*, (Berkeley, CA, USA), pp. 273–286, 2005.

[18] A. Wissner-Gross and C. Freer, "Relativistic statistical arbitrage," *Journal of Physical Review E*, vol. 82, pp. 056104+, November 2010.

[19] A. Rice, S. Akoush, and A. Hopper, "Failure is an option," tech. rep., Microsoft Research Technical Report MSR-TR-2008-61, June 2008.

[20] G. Werner-Allen, K. Lorincz, M. Welsh, O. Marcillo, J. Johnson, M. Ruiz, and J. Lees, "Deploying a wireless sensor network on an active volcano," *IEEE Internet Computing*, vol. 10, pp. 18–25, March 2006.