Green Aware Datacenter Selection

Khin Swe Swe Myint[1], Thandar Thein[2]
University of Computer Studies, Yangon
khinsweswemyint@gmail.com,thandartheinn@gmail.com

Abstract

Datacenters have been the key system infrastructure for cloud computing. The demand for datacenter computing has increased significantly in recent years resulting in huge energy consumption. Renewable energy resources, such as wind and solar power are rapidly becoming generation technologies of significance in the United States and around the world. The integration of renewable energy resources is usually very challenging because of their intermittency and inter-temporal variations. The high energy footprint of datacenters leads to serious environmental issues. Energy expenditure has become a significant fraction of datacenter operating costs. The proposed system is explicitly modeled the intermittent generation of renewable energy (Wind Power Model and Solar Power Model) with respect to varying weather conditions in the geographical location of each datacenter. Renewable energy resources datacenter selection framework is proposed to reduce the environmental impact and the system takes into account the benefit from the location diversity of different types of available renewable energy resources and an efficient datacenter selection algorithm is proposed.

major concern of broader research community participating both from academia and industry in the recent years.

Large Internet companies (e.g. Google and Microsoft) have significantly improved the energy efficiency of their multi-megawatt datacenters. However, the majority of the energy consumed by datacenters is actually due to countless small and medium-sized datacenters, which are much less efficient. These facilities range from a few dozen servers housed in a machine room to several hundreds of servers housed in a large enterprise installation. These cost, infrastructure, and environmental concerns have prompted some datacenter operator to generate their own solar/wind energy of draw power directly from a nearby solar/wind farm. Green energy sources promise to mitigate the issues surrounding non-renewable generation, but their output is very susceptible to environmental changes. This limits the

1. Introduction

Internet services such as web-mail, social networks, searching, online chatting and gaming are becoming part of people's day-to-day activity. Consequently Internet based service providers, such as Amazon, Google, Yahoo, Facebook etc, are under tremendous pressure in making their systems reasonably scalable, and providing guaranteed performance with low latency and high availability. In order to achieve these objectives such services need to be deployed over dedicated datacenters containing massive number of servers, large storage and heterogeneous networking elements together with an infrastructure to distribute power and provide cooling. Ever-increasing service demand, complexity of services and growing client population drive these data centers to consume enormous amount of power incurring a major part of their operating costs.

These datacenters place a heavy burden on both the environment and energy resources. Thus efficient power management of datacenters has become a

use of green energy in time-sensitive applications. Datacenters are sensitive applications. Datacenters are a significant source of energy consumption with an estimated 2% global greenhouse gas emissions attributed to them [4]. However, the time-sensitive nature of their service-level workloads has precluded the use of green energy, as jobs might need to be stopped when the available green energy drops [16]. This trend will likely continue, as these technologies' capital costs continue to decrease (e.g., the cost of solar energy has decreased by 7-fold in the last two decades [6]) and governments continue to provide incentives for green power generation and use (e.g., federal and state incentives in New Jersey can reduce capital costs by 60% [7]). For the scenarios in which green datacenters are appropriate, we argue that they should connect to the solar/wind energy source and the electrical grid, which acts as a backup when green energy is unavailable. The major challenge with solar or wind energy is that, unlike brown energy drawn from the grid, it is not always available. For example, photovoltaic (PV) solar energy is only available during the day and the amount produced depends on the weather and the season.

In this paper, green datacenter framework is proposed to distribute the incoming service requests among geographically dispersed datacenters based on renewable power availability. It calculates the wind power model and solar power model to generate the renewable energy with respect to the varying weather conditions in the geographically location of each data center.

The rest of this paper is organized as follows. The related work is introduced in Section 2. The power consumption in datacenter is presented in Section 3. It expresses the knowledge of the brown energy and renewable power generation. The architecture of the green datacenter selection is described in Section 4 and simulation setup is illustrated in Section 5. Conclusion is discussed in Section 6.

2. Related Work

Greening data centers is becoming an increasingly important topic in operating cloud-scale datacenter networks for Internet service operators, due to the fact that datacenters has become major energy consumers and environmental concerns. To our best knowledge I.Goiri et al. [9] leveraged renewable energy to handle scientific workloads. They further integrated green awareness into Hadoop (namely GreenHadoop [10]). Krioukov et al. [5] advocated a supply-following computing paradigm for data intensive applications, and developed a green aware scheduling algorithm to maximize green usage while meeting the deadline of data processing jobs. Baris Aksanli et al. [3] developed green- aware scheduling for both on-line services and batch jobs in a single datacenter. These existing studies for a single datacenter have major drawback that didn't consider the complementary effects on green energy supply among multiple datacenters there are also other proposals on architectural integrations into datacenters [8].

Most of the works that have considered variable electricity prices have targeted request distribution across multiple datacenters in interactive Internet services [16]. The exception is recent work [14], which considers variable electricity prices in multi-datacenter high-performance computing clouds and differs from these previous efforts as it seeks to maximize green energy use, predict green energy availability and schedule batch jobs within a single datacenter.

At the scale of multiple datacenters, many studies target at interactive Internet services. Le et al. [7] dynamically scheduled online services across multiple datacenters to maximize green energy usage for Web hosting with cost budget constraints. Since interactive services have little slack, most of these workload scheduling algorithms target at maximizing instant usage of green energy while satisfying online services QoS. They do not consider dynamic workload placement, migration and discrete cooling granularity, which are physically related to the outside temperature of geographically distributed datacenters.

These previous efforts focus on dispersed datacenter. Willow [11] assumes that reductions in green energy supply affect the servers differently, and migrates the load. In contrast, Blink [15] considered managing server power states when the amount of green energy varies but the datacenter is not connected to the electrical grid. It is not realistic for datacenters to depend completely on green energy, since this may cause unbounded performance degradation. In addition, our approach for managing green energy consumption is through job scheduling, rather than load migration or server power state management. In contrast with these higher level approaches, SolarCore [5] is a multi-core power management scheme designed to exploit PV solar energy and focuses on a single server, so it is closer to the works that leverage green energy in embedded systems.

3. Power Consumption in Datacenters

Brown Energy: Brown energy is energy produced from resources that when burned is harmful to the environment. These include coal, natural gas and other fossil fuels. The burning of fossil fuels is one of the most efficient ways to produce energy, but it can also be harmful to the environment. Some side effects of burning fossil fuels include air pollution and acid rain. Brown energy is the type of energy most commonly used in American homes. Many consumers assume it is cheaper than green energy.

Datacenters often contract with their power companies to pay variable brown electricity prices, i.e. different dollar amounts per kilowatt ($/kWh) of consumed brown energy. The most common arrangement is for the datacenter to pay less for brown electricity consumed during an off-peak period than during an on-peak period. Typically, off-peak prices are in effect during the night, whereas on-peak prices apply during the day. Thus, it would be profitable for the datacenter to schedule part of its workload (e.g., maintenance or analytics tasks, activities with loose deadlines) during the night.

Renewable Energy: Green energy is energy produced using technology that does not leave long-lasting effects on the environment. There are several strategies for producing green energy from renewable resources such as wind power, solar energy, hydroelectricity and more. Solar and wind are two of the most promising green energy technologies, as they do not cause the environmental disruption of hydroelectric energy and do not have the waste storage problem of nuclear energy. Solar/wind equipment produces Direct Current (DC) electricity, which is typically converted to Alternating Current (AC) by DC/AC inverters. In this paper, the datacenter generates its own PV solar energy is assumed. Self-generation is attractive for multiple reasons, including (1) the fact that energy losses with power transformation and transmission can exceed 40%; (2) the ability to survive grid outages, which are common in some developing countries; (3) the fact that PV power scales poorly, as the cost/W does not decrease beyond 800KW (the maximum inverter size today) (4) the ability to eventually lower costs. This amortization period is substantially shorter than the typical 20-30 years lifetime of the panels. The period will be even shorter in the future, as solar costs continue to decrease at a rapid pace [6]. The increasing popularity of distributed generation and micro grids suggests that many people find self-generation attractive.

There are multiple ways to connect solar panels to a datacenter. The solar panels can be connected to batteries for storing excess energy during periods of sunlight and discharging it during other periods. The datacenter must also be connected to the electrical grid via a grid-tie device if it must be operational even when solar energy is not available. The available wind power generated from wind turbines is modeled based on the ambient wind speed.

4. Green Datacenter Architecture

In this paper, an efficient geographically distributed green datacenter selection framework is proposed and shown in Figure 1. There are two main components in this framework: power grid that supplies the renewable energy sources and SLO guaranteed green datacenter selection.

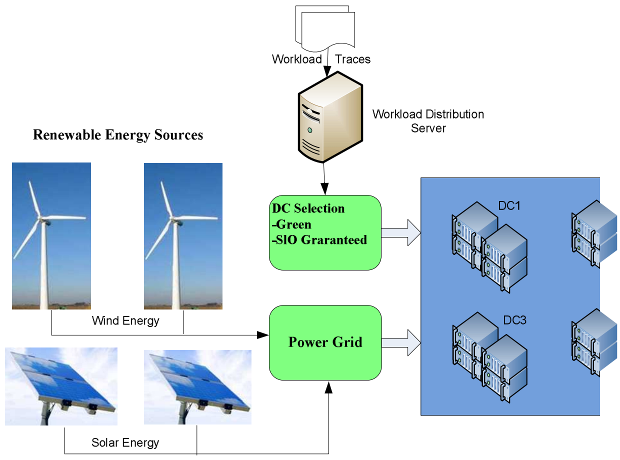

Figure 1. Geographically Dispersed Green Datacenter Architecture

Solar Power Model: Each datacenter is associated with a solar farm. We simulate solar farms with different numbers of solar panels. The default size of a solar farm is 10,000 solar panels. Under the default setting, the peak green energy supply is equal to the peak energy consumption of each datacenter. We calculate the solar energy production according to the parameter specification of BP-MSX 120 panels. Basically, given instant irradiance as input, the estimation model presented in the specification derives the amount of solar energy produced by a solar panel. The total amount of solar power is the power of each solar panel multiplied by the number of solar panels.

$$E = A * r * H * PR \qquad (1)$$

E = Energy (watt)
A = Total solar panel Area (m2)
r = solar panel yield (%)
H = Annual Average solar radiation on tilted panels
PR = Performance ratios,

Wind Power Model: The number of wind turbine installations is rapidly growing worldwide. It is expected that the United States can get 20% of its electricity from wind energy by the year 2030 [15]. It has been shown that wind power generated by wind turbines in a wind farm can be modeled as a function of the actual wind speed.

–

(2)

P = power in watts

ρ = The air density

A = The swept area of the turbine blades

V = wind speed (meters per second)

Even though the wind has that much power, it's not possible to extract it all due to losses. Typical wind turbine efficiency (Ct) is about 40%, and axial flux alternator efficiencies (Ca) are about 60% efficient. These numbers will vary depending on the designs and their specific performance curves. But, it's a good starting point.

–

(3)

P = power in watts

ρ = The air density

A = The swept area of the turbine blades

V = wind speed (meters per second)

C_t = wind turbine efficiency

C_a = alternator efficiency

Green Datacenter Selection: If a datacenter is fully occupied with the cloud services, other datacenters will be considered. This simple approach has the main problems in minimization of the environmental impact. Without the awareness of the hourly changes of the renewable energy and workload deadline, it may schedule too much workload with deadline to a datacenter with insufficient green energy. As its decision is oblivious, the environmental impact of the datacenter will increase.

5. Simulation Setup

Datacenter Parameters: In our evaluation, a large system composed of four geographically distributed datacenters for an Internet service operator (e.g., Google) is simulated. Accordingly, four different locations are assumed in the simulator, i.e., New York, California, New Jersey and New England, which are the locations whose typical meteorological year data are available in [19]. The power consumption profile of each server in the same location is assumed to be approximately the same, which is usually true when homogeneous servers and configurations are used in each datacenter.

Renewable Energy Availability: To emulate the intermittent availabilities of wind energy in the locations of different datacenters, i.e., wind power, the meteorological data from the Measurement and Instrumentation Data Center (MIDC) [18] of the National Renewable Energy Laboratory is used. A variety of meteorological data, including irradiances and wind speed, is covered in those records from the MIDC. Moreover, prior studies have shown that the data from the MIDC is sufficiently accurate. In particular, the meteorological data from the four stations, e.g., Sun Spot One, Loyola Marymount University (LMU) Rotating Shadowband Radiometer, New Jersey and La Ola Lanai is used, since they have consistent time periods with available meteorological data, beginning from December 1st, 2012 to December 31th, 2012. We further assume that there are 200 turbines installed in each wind farm.

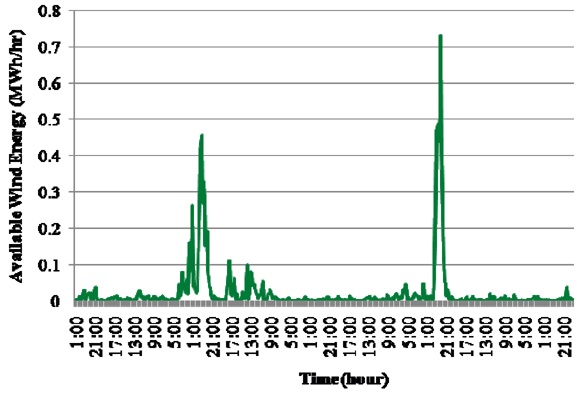

Figure 2. The Sample Wind Energy Availability of December for DC1

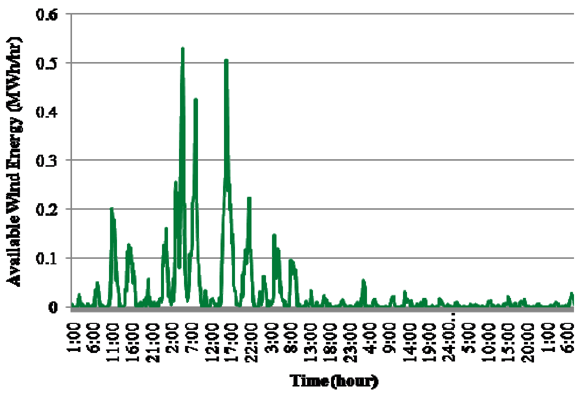

Figure 3. The Sample Wind Energy Availability of December for DC2

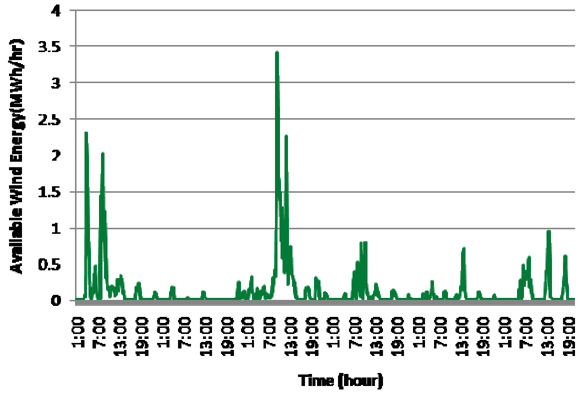

Figure 4. The Sample Wind Energy Availability of December for DC3

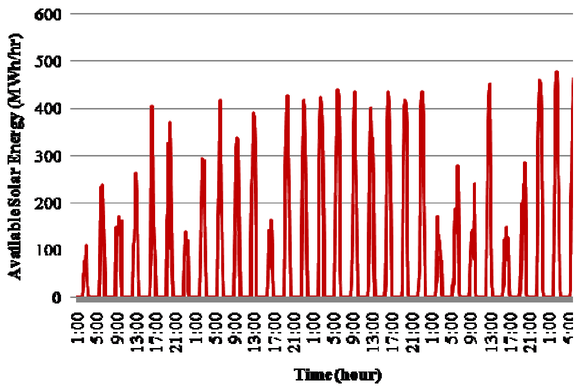

Figure 7. The Sample Solar Energy Availability of December for DC2

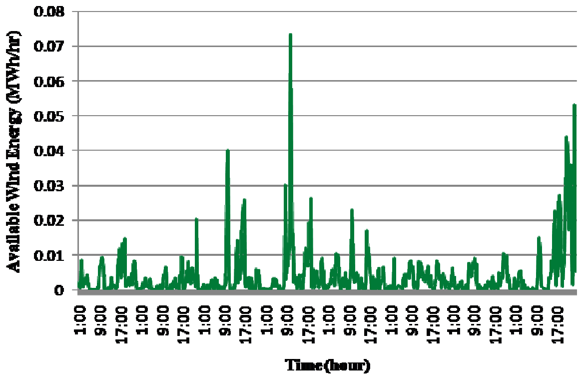

Figure 5. The Sample Wind Energy Availability of December for DC4.

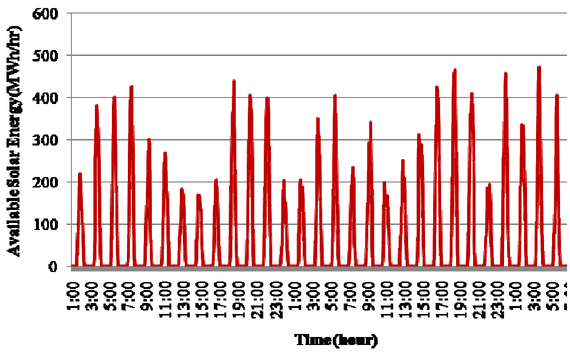

Figure 7. The Sample Solar Energy Availability of December for DC3

As shown in these four Figures 2-5, the available wind energy shows a diurnal pattern. This is due to the fact that the local weather conditions have a nearly diurnal pattern.

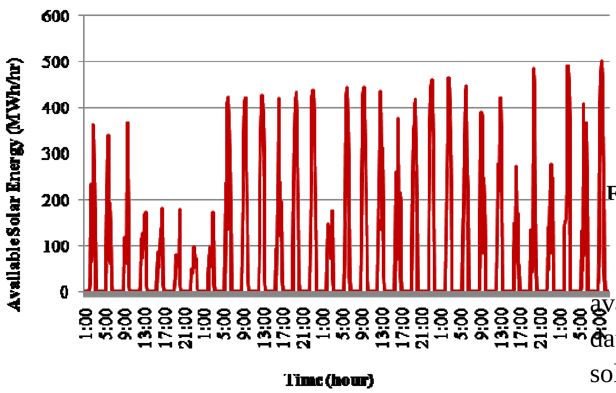

Figure 7. The Sample Solar Energy Availability of December for DC1

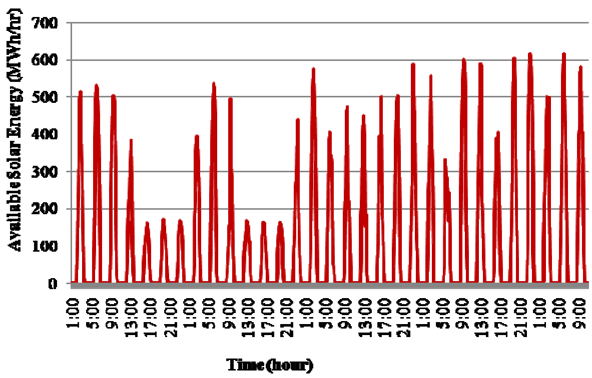

Figure 9. The Sample Solar Energy Availability of December for DC4.

The typical meteorological year data availabled [19] for solar energy is used and each datacenter have 10,000 solar panels installed in each solar plant to provide solar energy to the local power utilities of the 4 datacenters is assumed. In particular, BP-MSX 120 panels produced by British Petroleum are assumed to be used in the solar plants [1] and derived the power from its specification. Specifically, this calculation based on the power models discussed in Sections 4, as well as the varying weather conditions obtained from MIDC, the available solar energy of all the 4 datacenters throughout the entire simulated month is demonstrated in Figures 6- 9.

Green datacenter selection algorithm is presented in Table 1. The algorithm assigns a job to its suitable datacenter by considering the green energy available by wind and solar power. At the first step of algorithm, the initial datacenter for the job is assigned. Then, calculate the renewable energy for each datacenter and rank the amount of green energy by descending. If the job want to access to the cloud services then deadline of the job is checked to meet the SLA guaranteed and then the selected datacenter with the largest available amount of green energy is routed. In the proposed algorithm, an effective way is selected to direct clients across the wide area to an appropriate DC location to meet the service level objective.

Table 1

Green Datacenter Selection Algorithm

```
Algorithm: Max_Green (minimum brown)
Input:   job;
         deadline;
         DC_List;
Output:  Max_Green_DC;
1:  for DC_i ← DC_List    // Initialize the datacenters list
2:      Green_Energy_i ← calculate green energy for DC_i; // Calculate the green energy for each datacenter
3:  end for
5:  sorted_Green_DC_List ← Rank DCs; //Descending order of maximum green energy DC
6:   if (job access to cloud services) then
8:     Check job's deadline;
9:  for DC_i: sorted_DC_List
10:    if (DC_i meet the job's deadline) then
11:    begin
12:       Max_Green_DC ← DC_i;
13:     allow the job to access the Max_Green_DC;
14:     exit;
15:   end
16:end for
```

6. Conclusion

In this paper, the problem of minimizing brown energy consumption of high performance computing (HPC) workload for cloud providers that operate multiple geographically distributed datacenters with renewable energy sources have studied. The generation of renewable energy model such as wind power and solar power is explicitly described from varying weather conditions in the geographically location of each datacenter. Specifically, a green datacenter selection algorithm to meet the service level objective is also presented and simulation with real world weather trace is described.

REFERENCES

[1] A. Krioukov, C. Goebel, S. Asplaugh, Y. Chen, D. Culler, R. Katz. "Integrating Renewable Energy Using Data Analytics Systems: Challenges and Opportunities". IEEE Data Engineering Bulletin. March 2011.

[2] A. Qureshi, R. Weber, H. Balakrishnan, J. Guttag, and B. Maggs."Cutting the Electric Bill for Internet-Scale Systems", In Proceedings of SIGCOMM, , August 2009, pp 123-134.

[3] B. Aksanli, et al. "Utilizing green energy prediction to schedule mixed batch and service jobs in data centers", In Proceedings of HotPower'11, 2011,pp 51-55.

[4] B. Watson, A. Shah, M. Marwah, C. Bash, R. Sharma, C. Hoover, T. Christian, C. Patel. "Integrated Design and Management of a Sustainable Data Center". In Proceedings of IPACK2009. 2009.

[5] C. Li, W. Zhang, C. Cho, and T. Li. "SolarCore, "Solar Energy Driven Multi-core Architecture Power Management", In Proceedings of the International Symposium on High-Performance Computer Architecture, February 2011.

[6] 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, October 2009.

[7] D. Lifka, "The ANL/IBM SP Scheduling System", In Proceedings of the International Workshop on Job Scheduling Strategies for Parallel Processing, 1995.

[8] G. Sriram, et al, "Benefits and limitations of tapping into stored energy for datacenters", In Proceedings of ISCA'11, pp 341–352, 2011.

[9] I. Goiri, et al, "GreenSlot: Scheduling energy consumption in green datacenters", In Proc. of SC'11, 2011, pp 21-31.

[10] I. Goiri, et al, "GreenHadoop: Leveraging green energy in data-processing frameworks" In Proceedings of EuroSys'12, NY, 2012, pp 57–70.

[11] K. Kant, M. Murugan, and D. H. C. Du. Willow, "A Control System for Energy and Thermal Adaptive Computing", In Proceedings of the International Parallel and Distributed Processing Symposium, May 2011.

[12] K. Le, R. Bianchini, M. Martonosi, and T. D. Nguyen, "Cost-And Energy-Aware Load Distribution Across Data Centers", In Proceedings of HotPower, 2009.

[13] K. Le, O. Bilgir, R. Bianchini, M. Martonosi, and T. D. Nguyen, "Capping the Brown Energy Consumption of Internet Services at Low Cost", In Proceedings of the International Green Computing Conference, August 2010.

[14] K. Le, J. Zhang, J. Meng, Y. Jaluria, T. D. Nguyen, and R. Bianchini. "Reducing Electricity Cost Through Virtual Machine Placement in High Performance Computing Clouds", In Proceedings of Supercomputing, November 2011.
[15] N. Sharma, S. Barker, D. Irwin, and P. Shenoy. Blink, "Managing Server Clusters on Intermittent Power", In Proceeding of the 16th International Conference on Architectural Support for Programming Languages and Operating Systems, March 2011.
[16] Z. Liu, M. Lin, A. Wierman, S. Low, and L. Andrew, "Greening Geographical Load Balancing", In Proceedings of the International Conference on Measurement and Modeling of Computer Systems, June 2011.
[17] US Environmental Protection Agency. Report to Congress on Server and Data Center Energy Efficiency, August 2007.
[18] Measurement and Instrumentation Data Center. http://www.nrel.gov/midc/.
[19] http://rredc.nrel.gov/solar/old_data/nsrdb/1991-2005/tmy3.