

## 使用可變多工伺服器管理電源

Jia Hong Lian(連嘉宏)  
黎明技術學院資訊科技系  
chlien@mail.lit.edu.tw

Rui Yang Dong(董瑞洋)  
黎明技術學院資訊科技系  
eric-1024123@yahoo.com.tw

## 摘要

網際網路伺服器中心使用多工 server pool 提供多工服務，本論文中說明伺服器中心管理資源所使用的可變多工伺服器(Scalable Multiple Server, SMS)結構設計，此結構能夠讓使用電力和效能上達到適當的平衡，目的為提供不同的服務到不同伺服器，並且使用動態重置 server pool 提供讀取資料和改善能源效率，server pool 的分析模型皆相同。我們從實驗數據的到達率分析網頁伺服器，並用線性和指數這兩種實驗模型分析電源消耗和使用率的關係。線性模型用行列近似分析達到網頁伺服器最小功率消耗，我們比較不同工作量的量測基準證實我們所用的模型是好的，從實體機器做追蹤此模型計算的數據得到實驗結果，我們也應用了網頁伺服器中心測試 SMS 結構，最後減少了 16.9% 的電力消耗。

**關鍵詞：**電源管理、可變多工伺服器、網頁伺服器、M/M/1 模型、省電

## Abstract

An Internet server center can provides multiple service sites from a multiple sever pool. The paper depict a Scalable Multiple Server (SMS) architecture design for manage resource in a server center which can balance power consumption and performance. The purposes are to provide different servers for different service sites and servers utilizes dynamically reconfiguring the server pool offered load and increase the energy efficiency of server pool in accordance with an analytic model. We analysis web server power consumption relate to utilization from the experiment arrival data. We utilize linear model and exponential model to analysis that. The linear model utilizes queuing analysis to get minimal power consumption. We compare different work load measure for validate our model. The experimental result tracks from this model real machine calculate. We also apply web server center SMS test, and achieving reduce 16.9% energy consumption.

**Keyword:** power management、Scalable Multiple Server、web server、M/M/1 model、power saving

## 1. 緒論

網際網路發展至今已經快要二十年，從最早期的撥接網路開始，一直發展到現在智慧型手機常用到的行動網路和熱門的雲端系統，網路的便利確實改變了人們的生活型態，我們可以看到網際網路正不斷地快速成長，也影響了我們的生活，現在我們的生活與網路息息相關，其中網頁伺服器的高效能對網路發展而言影響很大，是造成網路成功的一個重要因素。

使用電源切換的節能方法在十多年前的論文中[1,2]就已經開始被討論，其中最成功的節能技術為系統層動態電源管理[3,4,5]，節能方法為關閉部分沒有用到的閒置裝置電源，以達到節能效果，而近年來也有人用此方法在做雲端系統的省電機制[6]。

伺服器資源管理的基本問題是讀取資料量大小估計問題:有多少資源需要執行?大小是否符合讀取資料量等級?這些問題會牽涉到網路流量特性，最好的近似方法為使用指數移動平均點樣本或是超過某樣本週期的平均資源使用率[3]。另外一個方法為使用靜態網頁伺服器模型，原始碼讀寫資料指令的程式內碼將源自於請求數和位元傳輸[4]。而這些研究已經成為了許多網頁的基準，像是 SPEC web 和 Web Stress[7,8]。這些研究之目的是在改進網頁伺服器的效能，無論是網頁伺服器之間的效能，基本的運算系統或是硬體 I/O 和網路 I/O 皆是。

本論文探討主機配置資源的方法，主要著重於電源管理。我們設計和執行可調整型資源管理結構--可變多工伺服器(Scalable Multiple Server, SMS)--可控制伺服器配置資源和例行程式請求伺服器選取可重構交換基礎設施。我們透過資源配置，關閉不必要的電源以達到節能。

伺服器使用 SMS 管理會有很好的節能效果，圖 1 為 SMS 結構的基礎:在圖中的 server pools 將一起分享由主機所提供的資源，server pool 會一起動作以支援共同主機的載入請求，而伺服器的資源是一般性並且是可替換的。DNS 伺服器為動態，當獲取網域名稱時會對應到此網域名稱的位址，改變路徑引導至正確的伺服器。每個主機會

將服務傳送到外部單一虛擬伺服器，電源消耗了多少和載入資料有多少與請求的可用資源有關，伺服器運算系統和 DNS 伺服器連續地監控和載入。管理者的功能:中央主機 OS 的「大腦」決定如何分配資源的配置與重置，為了達到節能，SMS 會使用網頁伺服器電源模型去管理伺服器配置資源，並設法得到最大資源但有最小電源消耗量，在第二節中將對電源管理模型詳細介紹。

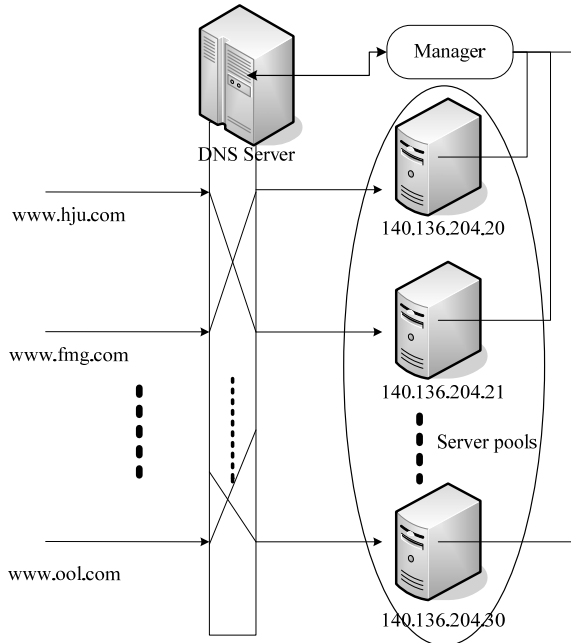


圖 1. SMS 結構圖

本篇論文共分為五小節來討論，第一節為本論文研究的概述，第二節介紹網頁伺服器的電源配置系統如何形成，此模型以實驗數據為基礎，在第三節中將說明為了達到最小功率，網頁伺服器配置模型的服務率要如何選擇，第四節為實驗結果，討論功率和使用率的關係，第五節是本論文的結論。

## 2. 伺服器電源配置系統建構

網頁伺服器可以很大、很複雜，也可以是不同種類的系統，常常會包括多種資源像電子、光學和力學等裝置，本論文使用 SMS，SMS 會配置每個服務到合適的伺服器以符合伺服器所需要的載入。在本篇論文中，我們著重在於發展系統的抽象模型給予網頁伺服器服務等級快速估算電源消耗和使用資源動態重構會如何衝擊整體的電力消耗。從服務等級電源模型看網頁伺服器模型就像一個黑箱子，我們不會對資源如何設計有興趣，而是資源會怎麼在此環境內運作，因此在黑箱子內部的模型並不是那麼的重要，研究的重點在於黑箱子內整體的電源消耗、電源應該要如何配置達到節能。

網頁伺服器可以在兩個不同電源狀態下運算，狀態轉換的示意圖如圖 2 所示。從圖中可看到有兩個圈圈，在圈圈內分別寫著 Busy, b 和 Idle, i，這分別代表網頁伺服器的狀態為忙碌和閒置，這兩種狀態會一直做轉換，狀態從忙碌轉換成閒置僅僅只有一瞬間，因此只會消耗少量的電源，此外執行效能也不會因為轉換而變差，在後面將介紹本論文網頁伺服器的狀態轉換用 M/M/1 排隊模型，用此模型計算出不同使用率的電力消耗。

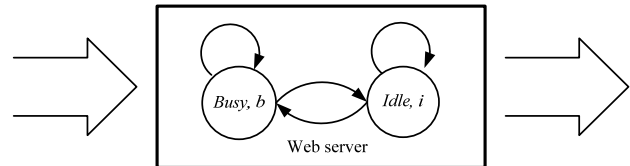


圖 2. 網頁伺服器的電源轉換狀態模型

過去，

$$pw_i = I \quad (1)$$

此處  $I$  是閒置狀態平均電力消耗，與不同結構的系統很有關聯性。

忙碌狀態不同於閒置狀態，其電力消耗的計算如下

$$pw_b = I + B \quad (2)$$

此處  $B$  是忙碌狀態額外的平均電力，和閒置狀態比較，除了  $I$  以外還要再加上  $B$  才是忙碌狀態時的電力消耗。

### 2.1 從實驗結果推論陳述電源模型

在本小節中將介紹網頁伺服器電源消耗簡單的估算方法。

我們量測電源消耗使用 ACPI 網頁伺服器系統，擁有單核心 2.4 GHz Pentium 4 處理器，266MHz 匯流排，1.0GB RAM, IBM 80GB IDE 硬碟，一個單核心技嘉乙太卡，網頁伺服器執行 Windows2000 高階伺服器和 IIS5.0 網頁伺服器系統。

本測量會導致有以下這些現象:雖然在閒置狀態時網頁伺服器沒有在執行，但這並不表示此時沒有電力消耗，電力消耗來自於部分資源中的主程式仍然在運作所造成，因此在閒置狀態時仍然會有很小的電力消耗變化，並且平均電源消耗幾乎是不變的。而根據我們觀察了伺服器的使用狀況，發現在忙碌狀態時電源消耗值的變化量較大，電源消耗率將隨時間變化。我們用三個不同機率分布的使用率 0.15, 0.5 和 0.85 來測量，結果如圖 3 所示。

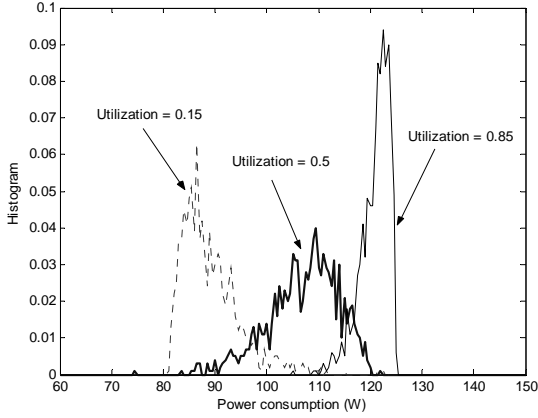


圖 3. 伺服器電源消耗的分佈圖

從圖 3 很明顯地看到不同的使用率造成的電力消耗差異大，較大的使用率會得到較大的值，並且從使用率 0.85 的圖形來看，此圖形和其他兩個使用率圖形比較起來較窄，其標準差較小，分布較集中，其他兩個使用率的圖形較寬，標準差較大，分布上會比較分散，我們會在後面的章節討論兩種模型合併的電力消耗。我們的重點放在不同使用率時伺服器的平均電源消耗，第一個模型的模擬更準確和更合適，第二個模型可用於一個簡單的管理結構。以此觀察為基礎，我們建構閒置電源消耗如下：

A. 指數模型

在此模型中使用指數函數做電源消耗量的計算，伺服器的使用率 $\rho$ 是估測忙碌狀態電源消耗量的重要參數，我們一開始先建立了忙碌狀態時電源的指數函數，忙碌狀態時電源消耗量的計算式如下：

$$pw_b = I + (B - I) \left( 1 - e^{-\frac{\rho}{\rho-1}} \right) \quad (3)$$

這裡  $I$  是在閒置狀態的平均電源消耗量， $B$  是系統在忙碌狀態 $\rho=1$ 大量載入資料時的額外電源消耗。使用率 $\rho$ 很小時將保留最多的資源，此時系統為閒置狀態，請求數最少，反之高使用率會造成很大的請求數，當電源消耗量很大時，就會造成很大的請求數。

B. 線性模型

以上的指數模型雖然在分析上很容易，但是會導致管理上更複雜，所以要有其他的模式做分析，以下是我們的第二種模型，對忙碌狀態的電源消耗是 $\rho$ 的線性函數：

$$pw_b = I + (B - I)\rho \quad (4)$$

第(4)式應用在忙碌狀態伺服器有不同使用率 $\rho$ 的電源消耗，電力系統在忙碌狀態，並且電力消耗會隨著使用率 $\rho$ 增加而增加。

2.2 M/M/1 排隊模型穩態機率

我們使用 M/M/1 排隊模型做研究，此模型簡單的講就像是一個只有單一服務窗口的郵局，大家只排成一隊列，客人進來、排隊、接受服務然後離開，用此模型可以用來管控顧客流量，系統管理者可以控制進入系統的顧客流量到一個最適合系統的情況[9]。在網頁伺服器的研究中此排隊模型只有單一伺服器提供服務，等候時間使用卜瓦松過程(Poisson process)量測，服務時間為指數分布，並且在計算上為求方便假設為無限緩衝容積，因此伺服器的請求數量需要很大，在程序上遵守一進一出，我們用此模型計算出不同使用率的電力消耗。使用率 $\rho$ 表示進入和離開之比值 $\lambda/\mu$ ，然後此狀態的機率可以寫成(5)式。我們假設達成率不影響系統的狀態，此傳輸方法可以歸納為圖 4，從圖 4 可以看到此排隊模式很有次序的遵守一進一出規則。

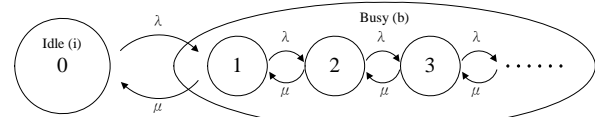


圖 4. 等候網頁伺服器處理請求狀態轉換

要成為穩態，每個轉換率在互相轉換的過程會相等，我們可以很容易得到狀態機率如下：

$$P_n = (1 - \rho)\rho^n \quad \rho < 1 \quad (5)$$

比較圖 2 和圖 4，當系統狀態為  $P_0$  電源消耗為閒置狀態時，閒置狀態平均機率

$$P_{idle} = P_0 = 1 - \rho \quad (6)$$

不論何時，系統只要一離開閒置狀態後就立即進入了忙碌狀態。因此，忙碌狀態平均機率為

$$P_{busy} = \sum_{n=1}^{\infty} P_n = 1 - P_0 = \rho \quad (7)$$

系統的電源消耗為系統忙碌和閒置狀態的電源消耗的總合，量測忙碌狀態和閒置狀態電源

$$PW = pw_i \cdot P_{idle} + pw_b \cdot P_{busy} \quad (8)$$

在常數模型中，我們得到

$$PW = I \cdot (1 - \rho) + (I + B) \cdot \rho$$

因此

$$PW = I + B \cdot \rho \quad (9)$$

這個傳統模型並不適合我們。在指數模型中我們得到

$$PW = I \cdot (1 - \rho) + \left( I + (B - I)(1 - e^{-\frac{\rho}{\rho-1}}) \right) \cdot \rho$$

因此

$$PW = I + \rho(B - I)(1 - e^{-\rho}) \quad (10)$$

所以只要知道使用率 $\rho$ 就能獲得網頁伺服器平均功率。

在線性模型中，我們得到

$$PW = I \cdot (1 - \rho) + (I + (B - I)\rho) \cdot \rho$$

因此

$$PW = I + (B - I) \cdot \rho^2 \quad (11)$$

這三種模型的準確性如圖 5 所示，從圖中可以看到指數模型和線性模型的準確度比常數模型好。

線性模型中允許使用率 $\rho < 0.65$ ，若使用率 $\rho$ 超過 0.65 則沒有意義，原因是因為高使用率的伺服器此時已接近滿載，另一方面則是因為已超過載入的最大範圍，因而請求緩衝。從圖 5 可以看出這些估測模型的準確度，指數模型得到錯誤率有 15%。

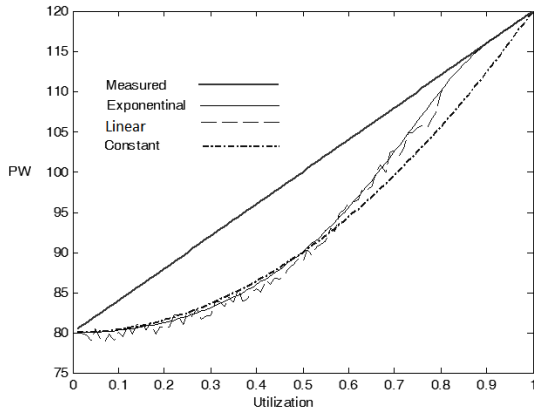


圖 5. 平均電源消耗和使用率 $\rho$ 關係圖

### 3. 選擇網頁伺服器以節省電力

選擇網頁伺服器以節省電力這也許是機器服務率為無限時的一個好方法，但是若以經濟來講這並不是最佳的管理方法，因為到達請求會是隨機的，並且前面提到，機器在閒置狀態時耗電量不為零，一樣需要花費耗電成本，因此必須要時常測量機器大小或電力消耗的服務容量。

在本節我們將定義一些符號以使用在本節中，討論電力消耗的計算和網頁伺服器的效能變化。一些觀念在前面已經很簡單的介紹過了，在前面的章節中介紹過的觀念如下：

$I_0$ : 維持網頁伺服器運算必然會消耗電力，數值和機器大小無關。

$\mu_i$ : 網頁伺服器能提供給主機的服務率。

$\mu$ : 網頁伺服器能提供給取代舊主機的新主機服務率。

$I_1$ : 主機在閒置狀態時增加的平均電力消耗。

$B_1$ : 主機在忙碌狀態的最大電力消耗。

經由觀察了不同機器的實體量測，我們可以使用  $\mu_i$  和  $\mu_b$  寫出下列的算式

$$pw_i = I_0 + \sqrt{\frac{\mu}{\mu_1}} \cdot I_1 \quad (12)$$

$$pw_b = I_0 + I_1 \cdot \sqrt{\frac{\mu}{\mu_1}} + (B_1 - I_1) \cdot \sqrt{\frac{\mu}{\mu_1}} \cdot \rho \quad (13)$$

此時把(12)和(13)代入 M/M/1 模型後得到

$$\begin{aligned} PW &= (I_0 + I_1 \cdot \sqrt{\frac{\mu}{\mu_1}})(1 - \rho) + (I_0 + I_1 \cdot \sqrt{\frac{\mu}{\mu_1}} + (B_1 - I_1) \cdot \sqrt{\frac{\mu}{\mu_1}} \cdot \rho) \cdot \rho \\ &= I_0 + I_1 \cdot \sqrt{\frac{\mu}{\mu_1}} + (B_1 - I_1) \cdot \sqrt{\frac{\mu}{\mu_1}} \cdot \rho^2 \end{aligned}$$

帶入使用率 $\rho = \lambda / \mu$ 得到

$$PW = I_0 + I_1 \cdot \sqrt{\frac{\mu}{\mu_1}} + (B_1 - I_1) \cdot \frac{\lambda^2}{\sqrt{\mu_1}} \mu^{-\frac{3}{2}} \quad (14)$$

要對(14)式服務率取得最小值，我們可以對(14)式 $\mu$ 微分，對電源函數 PW 兩邊微分後得到

$$\frac{d}{d\mu} PW = \frac{I_1}{2\sqrt{\mu_1}} \cdot \mu^{-\frac{1}{2}} - \frac{3(B_1 - I_1)\lambda^2}{2\sqrt{\mu_1}} \mu^{-\frac{5}{2}}$$

當 PW 是 0 時  $\mu$  是有最小值的，設定電力消耗是 0 解開方程式，最後得到最小值，解開上式得到最佳服務率  $\mu^*$ ：

$$\mu^* = \lambda \sqrt{\frac{3(B_1 - I_1)}{I_1}} \quad (15)$$

先前的研究結果數據顯示，當我們測量電源消耗量的時間夠久時，某些網站的網頁伺服器流量會是很穩定的，並且流量是可預測的。在一般的網站中流量是很穩定的，並且是可預期，如果網站  $i$  請求的到達率是  $\lambda(i)$ ，此時離開率我們可以選擇  $\mu^*(i)$  以達到最小電源消耗。

## 4. 實驗

### 4.1 實驗 1

從圖 6 我們可以了解此模型是如何獲得新機器的決定，圖中顯示了在不同服務率時的主機功率消耗(power consumption, PW)，主機用以下數據做計算， $\mu = 221.25$  請求數/秒， $I_0 = 20.12W$ ， $I_1 = 21.56W$  和  $B_1 = 61.03W$ 。假設 PW 為連續函數，當使用率為  $\mu_k$  時有最小值，當  $\mu < \mu_k$  時 PW 變小，反之，當  $\mu > \mu_k$  時 PW 變大。在圖 6 中虛線部分為計算結果，從實機計算獲得。此計算值為追蹤實機所產生之結果。

### 4.2 實驗 2

本小節介紹從 SMS 模型所得到的動態伺服器資源管理的實驗結果，這些實驗研究的是單一 server pool 對稱集群案例。

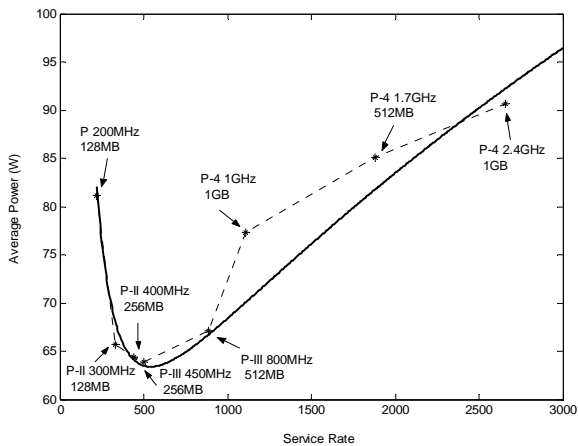


圖 6. 平均電源消耗和使用率 $\mu$ 關係圖

圖 1 說明中央資料測試是由 server pool 組成，經由 DNS 取得位址。管理者把使用層伺服器用在專門的機器上。此監控是用 800 行 C 語言所組成，編譯到管理者，除了讀取監控碼以外，他包含了 300 行 C 語言並使用第 4 節描述的 single pool 同質伺服器，在我們的實驗中，管理者會 24 小時平均地從 server pool 的所有服務之中讀取。

本論文硬體設備使用 2.8 GHz Pentium-4 系統 (Asus P4PE) 經由 Gigabit 乙太 NICs 連接 server pool 和網路，本伺服器使用 Windows 2000 advanced DNS 伺服器，在 Windows 2000 中活動目錄會是固定地固定在網域名稱，他將支援 server pool 的伺服器動態更新記錄。DNS 伺服器在標準 HTTP 1.0 Web 所傳輸的龐大資料量足以使 server pool 工作量飽和。伺服器包含一個 200 MHz Pentium 系統 (GA-586SG)，三個 400 MHz Pentium-II 系統 (Asus P2B/440BX)，三個 800 MHz Pentium-III 系統 (Asus P3B)，二個 1 GHz Pentium-4 系統 (Asus P4B)，二個 1.7 GHz Pentium-4 系統 (Asus P4B) 和二個 2.4 GHz Pentium-4 系統 (Asus P4PE)。這些伺服器連接乙太網路傳輸速度為 100 Mb/s，所有伺服器包含完整的網頁檔案備份，我們的工作量適合從記憶體產生大量的請求，也因此 CPU 速度將會非常重要。

為了顯示多工伺服器在不同負載條件下的影響，我們使用 Web Stress 企業版虛擬網頁服務負載產生器來展示多工伺服器不同載入量的影響，Web Stress 產生了高流量擠爆網頁工作量，由於是重尾分布，因此每個請求服務是容易改變的。虛擬流量的一個優點為負載產生器是一次性的，這代表請求抵達率容易受反應延遲影響，因此，我們可以透過控制產生器的數量產生虛擬負載“波”在任何振幅和週期，這實驗呈現了設計的系統在方案中的行為衍生出來的不同觀點。我們的實驗結合 Web Stress 虛擬流量工作量和實體請求追蹤，我們從輔大 (net.cc.fju.edu.tw) 追蹤了五個網頁的伺服器日誌，實驗時間總共 9 天，在這 9 天的時間共接收了 520M 請求，在圖 7 中顯示了使用

DW-6090 電力分析儀分析伺服器集區每秒的耗電量。

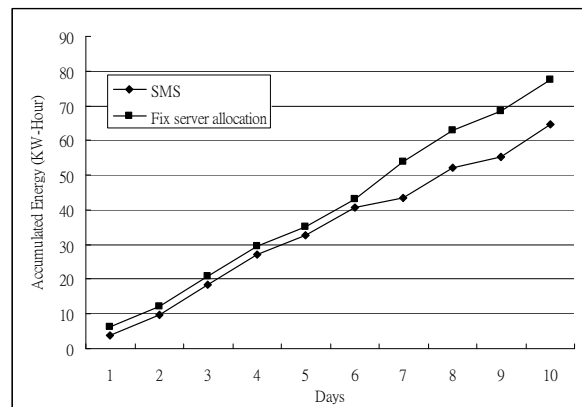


圖 7. 網頁請求累積能量追蹤

我們從圖 7 去比較 SMS 傳統固定式伺服器能量曲線，執行會造成 SMS 總能量消耗 64.48KWh，共節省了 16.9% 的電力，省電的成效不錯。在圖中反映了管理者若使用過時的舊機器，在週期性衰退的讀取中可以看到需要更多的電力消耗量。此外，從量測平均反應時間我們看到 SMS 有較短的反應時間並且滿足效能需求，反應時間低於 500ms。

## 5. 結論

在本論文中說明如何設計並且建立 SMS、中央主機資源管理系統，SMS 結構是由中央主機配置資源給 server pool，並且使用智慧型電源處理 (power-conscious approach) 以達到節能效果，主要目的是將資料中心電源管理系統合併到廣大的資源管理框架。此研究能夠使中央主機改善在資料載入時的能源效率，減少耗電量，並且動態的配對資料讀取和減少電力消耗，達到最大效率和最小功率的目標。

我們說明了如何使用可重置的網路層請求更改進來的請求路徑到動態供應的伺服器裝置，並且也說明了智慧型電源會怎麼運作，智慧型電源主要會用在僕伺服器的載入請求，我們的 SMS 網頁伺服器電源管理會自動的轉換任何閒置伺服器的電源狀態，控管中央全部的電源層。我們用指數模型、線性模型和常數模型與 M/M/1 模型結合做分析，這三種模型在不同使用率的電源消耗分析會得到不一樣的結果，線性模型和指數模型的結果比常數模型要來的好。

我們提出了新的系統模型和方法做動態伺服器服務層管理，這個模型用來敘述 SMS 資料中心測試，從樣板發展出我們的模型潛在的適合服務，提供給動態可變資源利用和 server pool 電源消耗，這個服務層的電源消耗問題為隨機的模型，最後從實驗數據得知本樣板可以在長期的網

頁運算中減少了 16.9% 的能源消耗，新的系統模型成功的達到了節能的目標。

## 6. 參考文獻

- [1] Vivek S. Pal, Mohit Area, Gaurav Banga, Michael Svendsen, Peter Drnschel, Willy Zwaenopel, and Erich Nahum. "Locality-Aware Request Distribution in Cluster-based Network Servers". *Proceedings of 8th International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS VIII)*. San Jose, California, United States, pages 205-216, Oct. 1998.
- [2] Yasushi Saito, Brian N. Bershad, and Henry M. Levy, "Manageability, Availability and Performance in Porcupine: A Highly Scalable Cluster-Based Mail Service", *Proceedings of 17th ACM Symposium on Operating Systems Principles (SOSP)*. Charleston, South Carolina, United States, pages 1-15, Dec. 1999.
- [3] Armando Fox and Steven D. Gribble and Yatin Chawathe and Eric A. Brewer and Paul Ganthier, "Cluster-based scalable network services", *Proceedings of 16th ACM Symposium on Operating System Principles (SOSP)*. Saint Malo, France, pages 78-91, Oct.1997.
- [4] Darrell C. Anderson, Jeffrey S. Chase, and Amin M. Vahdat, "Interposed Request Routing for Scalable Network Storage", *ACM Transaction on Computer Systems*, 20(1):25-48, Feb. 2002.
- [5] 卓昆輝，「虛擬主機動態分配策略之研究」，國立中正大學，碩士論文，2012
- [6] 吳書瑋，「藉由伺服器整併技術以實現雲端運算中省電機制之研究」，國立中興大學，碩士論文，2012
- [7] Paessler. Last update 2013. Webserver Stress Tool, Retrieved 2013, from <http://www.paessler.com/webstress/>
- [8] SPEC. Last update 2011. SPEC web 2009, Retrieved 2013, from <http://www.spec.org/web2009/>
- [9] 林守信，「具有優先權之 M/M/1 排隊模型的最佳門檻策略」，國立東華大學，碩士論文，2001