

中央氣象局全球模式之平行計算處理及在新超級電腦上之 測試

馮欽賜¹ 李長華²

¹中央氣象局科技研究中心

²台灣富士通公司

摘 要

中央氣象局從兩年前新引進日本富士通 VPP 系列超級電腦，它不僅具備有強大向量化運算能力的處理器，而且兼具分散記憶體多重處理器的平行電腦架構。爲了能夠充分發揮新型態超級電腦的最佳效能，在該公司技術人員的參與協助之下，目前我們擁有一版利用標準 MPI (Message Passing Interface) 語言爲修改基礎，使程式平行化的新版中央氣象局全球模式。在這篇報告中，將簡要描述此一全球模式處理平行計算的基本概念及架構，並且說明它在中央氣象局新超級電腦上測試所呈現的平行化效率和各種相關的統計結果。

一、前言

爲了提昇預報品質，中央氣象局天氣預報作業模式在提高解析度及希望使用較詳細複雜的物理參數化的需求之下，對電腦運算能力的要求也就與日俱增。爲了滿足這方面的需求，從 1998 下半年中央氣象局開始購置新型的超級電腦作爲日常數值天氣預報模式的研究和作業機器，以逐漸取代原本所使用的克雷 (Cray) YMP 電腦系統。新引進的超級電腦乃屬於日本富士通公司製造的 VPP 系列超級電腦，它的特色在於不僅具備有強大向量化運算能力的處理器，而且兼具分散記憶體多重處理器的平行電腦架構。由於平行化的硬體架構與之前克雷電腦系統中共享記憶體的特徵截然不同，這使得如何修改作業預報模式的程式構造，使其能夠充分發揮新型態超級電腦的最佳效能，成爲電腦環境轉換過程非常重要的環節，而這也形成中央氣象局數值預報模式研究人員新的挑戰。所幸，經由日本富士通公司技術人員的參與協助之下，目前我們擁有一版利用標準 MPI (Message Passing Interface) 語言爲修改基礎，使程式平行化的新版中

央氣象局全球模式。在這篇報告中，我們將簡要描述此一全球模式處理平行計算的基本概念及架構，並且說明它在中央氣象局新超級電腦上測試所呈現的平行化效率和各種相關的統計結果。

二、全球模式的平行化

中央氣象局全球模式是典型的波譜動力模式並且具備完整的物理過程，詳細的說明可參考 Liou et al. (1997)。進行計算平行化的設計，最主要的是確定各個運算過程所涉及資料空間分布的相關性及可分割性。以物理部份的運算而言，顯而易見地陣列資料僅在垂直方向的訊息會被同時採用而有相關性。但關於動力部份的運算，則資料在不同的階段會出現不同的相關性條件，在進行傅立葉轉換 (Fourier transform) 時同一緯度上的資料訊息具備計算上的相關性，但進行勒壤得轉換 (Legendre transform) 時則換成同一經度上的資料訊息具備計算上的相關性。由於動力部份這種特性，使其平行化過程的考量遠比物理部份來得複雜許多。基於計算平行化的原則在於完整保留資料相關的維向部份，分割其它維向，全球模式的分割作

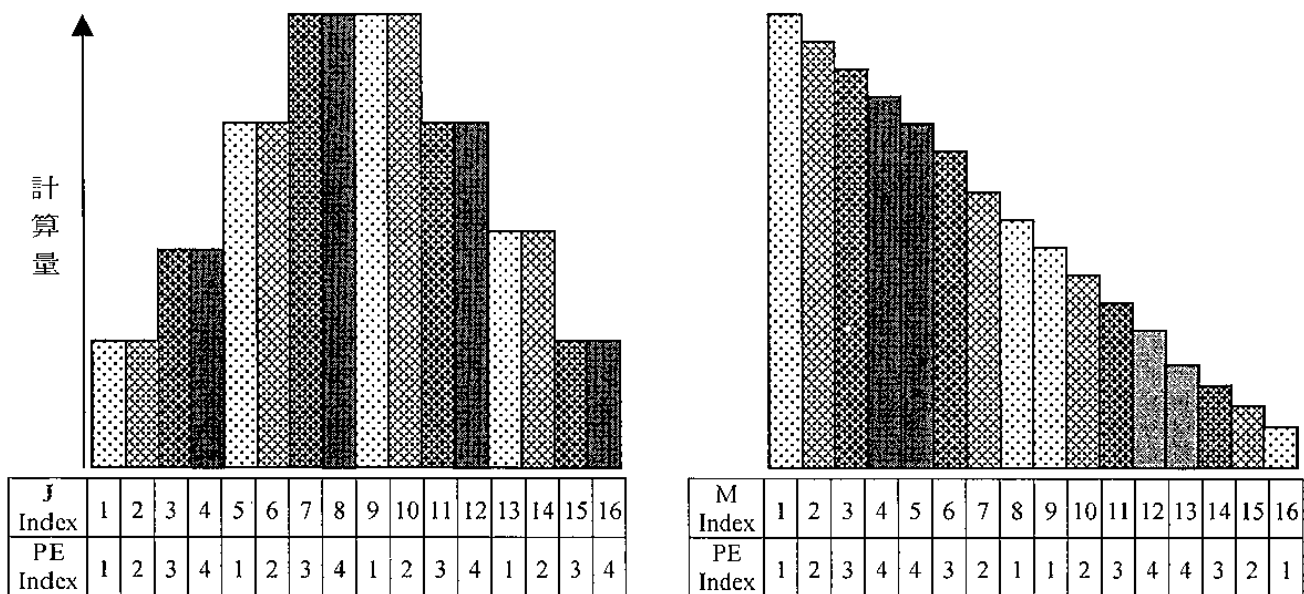


圖 1：全球模式針對經向格點及緯向波數所採取的切割設計。直方格長度代表對應之計算量，表格所列的是被分割緯向與各處理器之間分配對應關係。J 代表經向格點，M 代表緯向波數，PE 代表處理器代號，上圖是以 4 顆處理器舉例。

法上，凡屬格點空間的運算性質以及動力的傅立葉轉換階段都是選擇在經向上進行資料切割，但進入動力的勒壤得轉換階段以及屬於波譜空間的運算性質則在緯向波數上進行資料切割。目前模式僅作單一維向的分割設計，使平行化中所能採用的處理器個數上限有較大的限制，因此它較傾向適合於單一處理器本身即具有強大運算功能的平行電腦系統，例如現今之 VPP 系列超級電腦。

如同前述，動力結構方面牽涉到兩種不同階段的資料相關性，而且後一階段資料互相關連的維向卻是前一階段被切割的維向，因此由某一階段進入另一階段之前需進行資料的重組，把相關性的資料匯集，並且同時把即將被分割的部份進行分散。本質上，資料的重組在完成類似陣列的轉置 (Transposition) 過程，而資料陣列的轉置牽涉各個處理器之間資料的傳遞和交換，這部份的工作完全透過 MPI 介面提供的副程式加以處理。

此外，為了達到計算平行化的最佳效果，一個相當重要的關鍵在於是否做到處理器工作負荷的均勻化，而這需針對運算特性小心地讓每個處理器分配到相同的計算量。首先注意的問題便是關於波譜方法涉及的三角截斷空間，這使得緯向波數上的切割需考慮計算量伴隨波數增加而線性遞減的特性。其次，中央

氣象局全球模式為了降低長波輻射計算所耗費的龐大時間，刻意讓長波輻射部份在同一緯度上有跳點運算的設計，使進行長波輻射計算的格點數目隨緯度提高而能逐漸減少。因此，在經向上的切割則需額外考慮這方面的問題。圖 1 便是目前全球模式針對經向格點及緯向波數所採取的切割設計，它利用交錯分發的方式，使計算工作相當平衡地分配至各個處理器上。值得注意的是在經向格點及緯向波數方面分發序列在雙週期的差異。

三、新級超電腦上的測試結果

目前中央氣象局擁有富士通 VPP300 型及 VPP5000 型兩種超級電腦系統，前者共有 16 個處理器，後者則有 15 個處理器，單以每一處理器具有的浮點運算之理論最佳績效值而言，VPP5000 型可達 9.6GFLOPS 是 VPP300 型的四倍。由於 VPP5000 是日常天氣預報模式的作業機器，因此在這篇報告所討論的測試結果都在 VPP5000 上進行。測試過程共使用三種不同解析度條件，分別是 T79L18 (240x120x18)、T120L18(360x180x18)以及 T159L30 (480x240x30)，而且執行效能的統計都是根據不含輸出之七天預報的模式執行表現。

圖 2 呈現的是全球模式使用不同數目處理器所產生的加速效率，若越接近幾何級數遞增 (即圖中理想

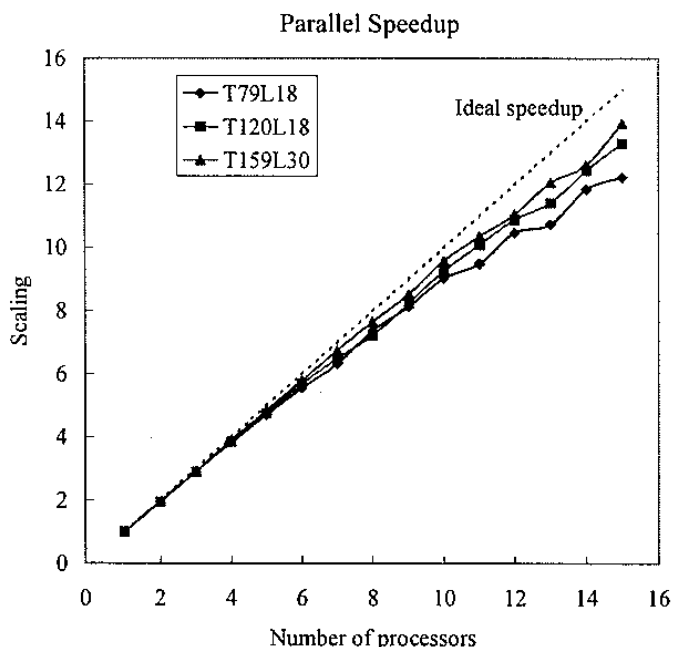


圖 2：全球模式使用不同數目處理器所產生的加速效率。圖中虛線代表最理想的加速結果。

線) 表示模式整體平行化的效果越佳。圖中的結果顯示當模式的解析度提高平行化效率越顯著，當使用 15 顆處理器時 T159L30 可實質達到 13.7 顆的加速效益。能達到 90% 以上加速率，顯示模式伴隨平行化過程耗費在資料於處理器間交換傳遞的負荷上並不明顯，另一方面也表示各處理器間的工作負擔被調整相當平均。圖 3 是統計全球模式在不同數目處理器下完成七

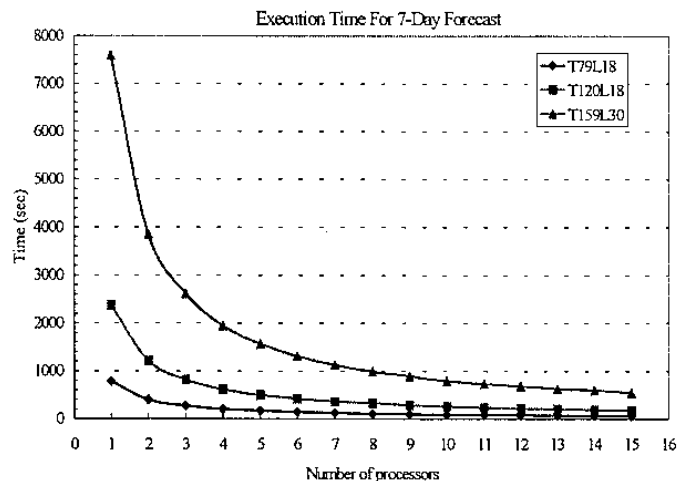


圖 3：全球模式在不同數目處理器下完成七天預報所花的時間。

天預報所花的時間，其遞減率與圖 2 是對應的。當使用 15 顆處理器時 T79L18、T120L18 和 T159L30 分別只需 69 秒、187 秒和 563 秒。若日常作業要求全球模式 7 天預報只需在 60 分鐘內完成而言，3 顆 VPP5000 處理器便可符合執行 T159L30 的需求。

由於 VPP 系列單一處理器本身具有強大的計算能力，如何妥善設計程式結構運用其向量化的功能和硬體的便利以發揮最佳的效能，亦提高模式整體執行效率很重要的關鍵。檢視整體全球模式使用單一處理

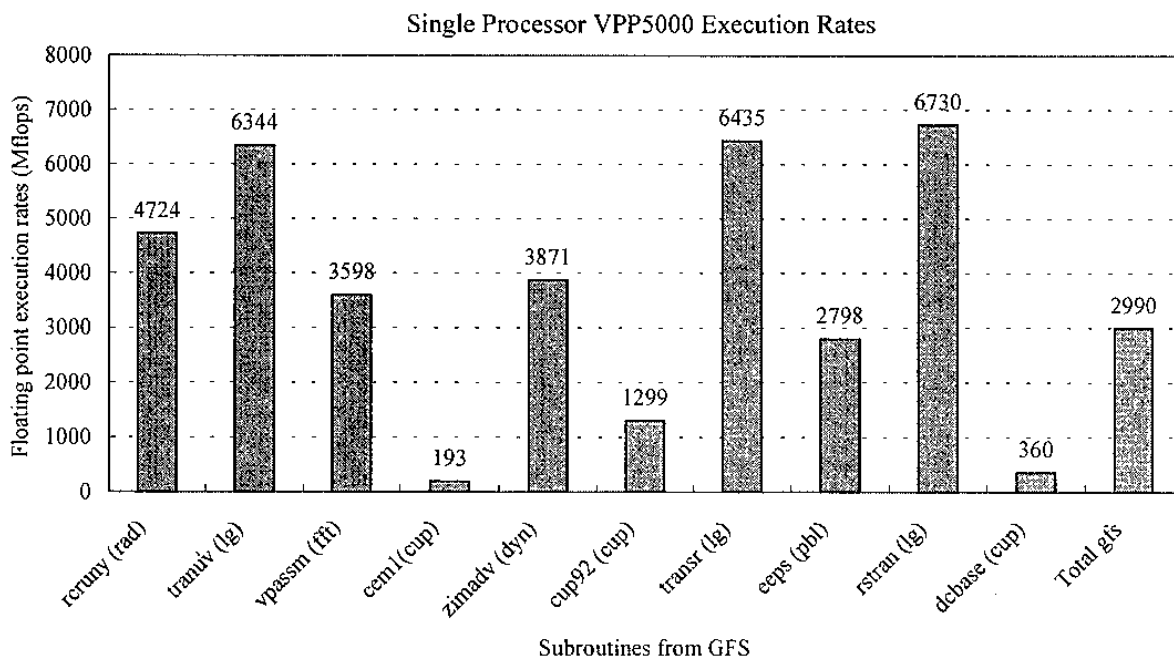


圖 4：T159L30 全球模式前十名最耗時副程式所表現的浮點運算值。其中所屬作用分別為 rad (radiation)、lg (Legentre transform)、fft (Fourier transform)、cup (Cumulus)、dyn (Dynamics)、pbl (Boundary layer)。

器所表現的實質浮點運算數據，顯示三種解析度下分別各達到 2185MFLOPS (T79L18)、2600MFLOPS (T120L18) 與 2990MFLOPS (T159L30) 的結果。當模式解析度提高而增加迴路運算的向量化長度，可以更有效地發揮電腦的效能。以 T159L30 表現的實質浮點運算值而言，約達到理論最佳值的 31%。根據歐洲展期預報中心 (ECMWF) 在 VPP300 型之處理器上 (實際是 VPP700 型，但中央處理器與 VPP300 型完全相同) 測試他們 T213L30 全球模式的報告 (Dent et al. 1997)，結果指出其實質浮點運算值約達到理論最佳值的 35%。雖然模式的結構彼此可能有很大的差異，但若按此數據評估目前中央氣象局全球模式的浮點運算效能應屬中上水準。

圖 4 進一步分析 T159L30 模式中前十名最耗時副程式所表現的浮點運算值，這十個副程式佔據幾乎 60% 的整體執行時間。圖中顯示這些程式的浮點運算效率差異甚鉅，高低之間差別達數百倍之多。其中，效率表現最突出的部份都屬於動力模組的範圍，其浮點運算可達理論最佳值的 60% 以上 (例如，tranuv, transr, rstran)，富士通公司技術人員對動力模組中運算迴路的重新調整是獲致如此效能的重要因素。至於表現較不理想的數個程式 (例如，cem1、cup92 和 dcbase) 則屬於積雲參數化的部份，這現象當然與積雲參數化計算過程遭遇較多的條件判斷，以及緯度上發生積雲之網格不均勻降低向量化效益有密切關係。不過，其浮點運算效率過低的現象，也有可能與程式的撰寫方法不良有某些程度關連，這方面的改善仍待進一步的測試。

四、結語

相對原本克雷電腦上作業的全球模式，此版以 MPI 為基礎的平行化全球模式，內部最大的差異主要在動力部份，差異包括為提升平行化下向量計算效能對迴路運算結構的重新調整以及為資料轉置所加入的 MPI 副程式。由於 MPI 副程式僅出現在傳立葉轉換和勒壤得轉換之間的階段，它相當清楚獨立於程式的某小部份，因此仍能維持程式相當理想的可讀性。從目前測試結果顯示，模式的平行化效果相當不錯，其

次浮點運算的效能在動力部份也相當優異，但相對上一些物理部份的浮點運算則仍有待進一步加強改善，這方面正進行之中。

致謝

感謝富士通公司曾經參與中央氣象局全球模式修改工作的所有日本技術人員。

參考文獻

- Liou et al., 1997: The second-generation global forecast system at the Central Weather Bureau in Taiwan. *Wea. Forecasting*, 12, 653-663.
- Dent et al., 1997: Use of the Fujitsu VPP700 for weather forecastinh at ECMWF. *Fujitsu Scientific & Technical Journal*, 33, 88-100.