

第三章 Grid 工作排程分析模型設計

本研究參考Lee, H. M., Lee, T. Y., and Hsu (2006)網格架構的運作，分成監督節點與執行節點，分別執行監督節點工作排程分析模型和執行節點工作排程分析模型。它們的模型描述如下如圖 3-1所示：

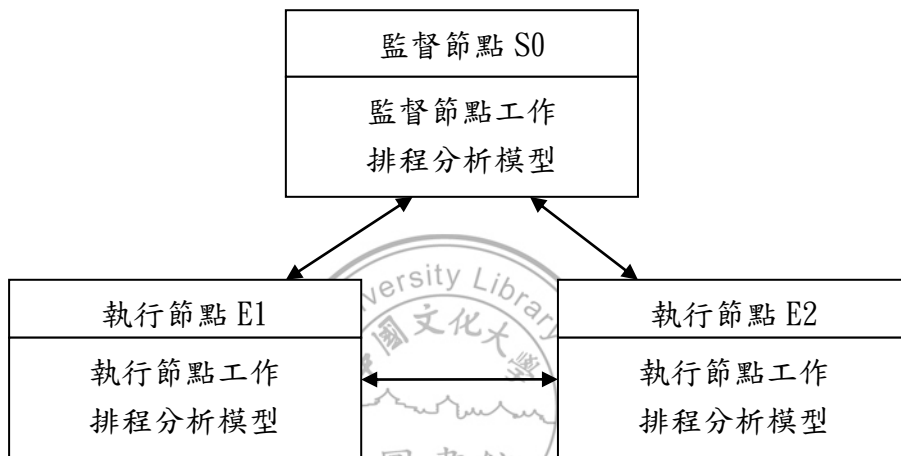


圖 3-1 網格節點工作排程模型架構圖

因本研究主要重點在網格執行節點，所以本章首先粗略介紹監督節點工作排程分析模型，其次則較為詳細的介紹執行節點工作排程分析模型，而後是一些有關的資料庫與表格，最後探討工作排程分析設計。

第一節 監督節點工作排程分析模型

監督節點工作排程模型分析如下：

一、功能

監督節點的功能描述如下：

- (一)處理新的工作
- (二)處理工作排程
- (三)處理移轉的工作
- (四)接收移轉工作的完成
- (五)蒐集工作資訊
- (六)蒐集節點資訊
- (七)更新 Grid 節點資訊和日誌裡的總資訊

二、模組

監督節點工作排程分析模型的模組如下：

- (一)信息接收模組
- (二)工作排程模組
- (三)工作移轉模組
- (四)工作完成模組
- (五)收集資訊模組
- (六)信息發送模組



第二節 執行節點工作排程分析模型

執行節點工作排程模型分析如下：

一、功能

執行節點的功能描述如下：

- (一)處理新的工作
- (二)處理新的排程
- (三)發送執行節點的資訊到監督節點
- (四)發送需要移轉的工作到監督節點
- (五)接收由監督節點傳來已完成的工作
- (六)接收由監督節點傳來移轉的工作

- (七)發送已完成的移轉工作到監督節點
- (八)發送執行節點之工作狀況給監督節點
- (九)發送執行節點目前之設備狀況給監督節點
- (十)完成本地工作

二、模組

執行節點工作排程分析模型的模組如下：

(一)接收模組

用為接收工作。

(二)工作排程模組

用為置入排程演算法。

(三)工作移轉模組

用為將工作移轉至別的節點。

(四)工作完成模組

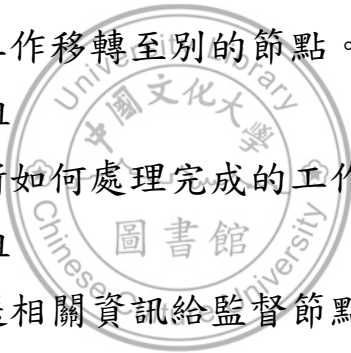
用為判斷如何處理完成的工作。

(五)發送資訊模組

用為發送相關資訊給監督節點。

(六)信息發送模組

用為發送工作及工作資訊。



工作處理模型圖如圖 3-2 所示：

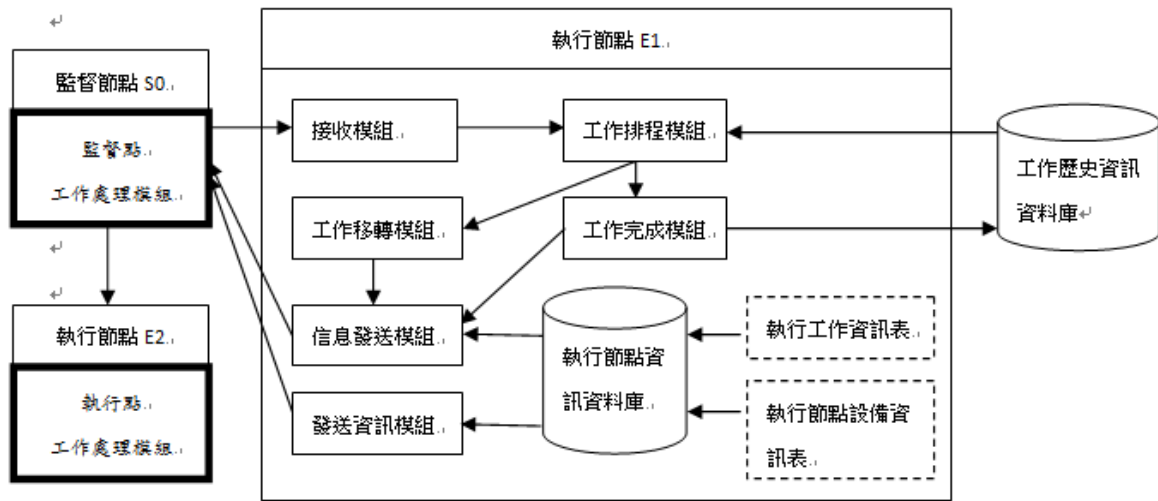


圖 3-2 執行點工作處理模型

三、模組工作說明

執行節點模組工作說明如下：

(一)接收模組

- (A)接收由監督節點傳來已完成的工作。
- (B)接收由監督節點傳來要執行的工作。
- (C)接收本地新工作。

(二)工作排程模組

依照排程演算法處理新的工作排程。

(三)工作移轉模組

當本執行節點工作忙碌時，將工作移轉給監督節點執行。

(四)工作完成模組

- (A)如是由監督點移轉來已完成的工作，則傳回監督點。
- (B)如是本執行節點已完成的工作，則更新工作歷史資訊。

(五)發送資訊模組

發送執行節點的資訊到監督節點，內容包括執行節點

之工作狀況給監督節點及執行節點目前之設備狀況給監督節點。

(六)信息發送模組

- (A)當工作忙碌時，發送工作到監督節點執行。
- (B)發送已完成的移轉工作到監督節點。
- (C)送出本執行節點已完成的工作資訊。

第三節 資料庫與表格

一、工作歷史資訊資料庫(job information history database, JIHDB)
建立每個 Grid 節點完成之工作資料，如表 3-1 所示：

表 3-1 工作資訊歷史資料庫

工作名稱	檔案大小(KB)	Memory 使用量(MB)	CPU 完成使用量 (Sec)	來回時間(Sec)	工作分類 等級
Job1	100,000	20	58	1200	E
Job2	50,000	6	312	500	B
Job3	39,000	1	805	1020	B
Job4	75,120	4	70	340	D
Job5	15,075	10	156	180	A
JobD	50,000	3	840	1789	C
JobE	20,000	8	1235	1300	A

(一)說明

(A)工作名稱：

工作之索引名稱，可供執行工作資訊表作搜尋。

(B)CPU 完成使用量(Sec)：

工作處理完所使用的CPU時間，可供工作排程來

計算CPU使用率。

(C)來回時間(Sec)：

是從工作讀入到完成所經過時間，可供工作排程來計算CPU使用率。

(D)工作分類等級：

CPU 使用率(%)等級分五級：

A (81-100)

B (61-80)

C (41-60)

D (21-40)

E (0-20)

(二)工作分類等級可由(CPU 完成使用量/(來回時間+1)*100)取得。

(三)每次工作正常結束時，數據好者取代之

二、執行工作資訊表(run job information table, RJIT)

當有一項新工作進入執行點時，一筆新資料就會在執行工作資訊表中建立出來，其資料可由工作歷史資訊資料庫供給，如表3-2所示：

表 3-2 執行工作資訊表

工作名稱	檔案大小 (KB)	CPU 完成使用量 (Sec)	來回時間 (Sec)	Memory 使用量 (MB)	執行	目前 CPU Time (Sec)	目前 來回時間 (Sec)	工作 分類 等級	工作 傳送	接收 工作
JobE	20,000	1235	1300	8	R	1234	1288	A	0	0
Job5	15,075	156	180	10	R	85	97	A	0	0
Job3	39,000	805	1020	1	R	710	903	B	0	0
JobF				4	R	133	270	C	0	0
Job4	75,120	70	340	4	R	28	132	D	0	0

(一)說明

(A)CPU 完成使用量(Sec)：

為已過CPU處理時間的歷史資訊。

(B)執行：

表示目前工作執行的狀態。

(C)工作傳送：

顯示該工作將被移轉傳送的目標。若本地節點工作忙碌需要移轉工作，則設為1；初值設為0，也就是本地可執行的工作。

(D)接收工作：

顯示該工作從何被傳送過來。若為監督點移轉過來要執行的工作，則設為1；初值設為0，也就是本地工作。

(二)當找不到工作名稱時，則以Memory 使用量 4 工作分類等級 C 處理，如 JobF。

三、執行節點設備資訊表

建置一表格於執行節點資訊資料庫將目前設備狀況記錄起來。如表3-3所示：

表 3-3 執行節點設備資訊表

節點名稱	CPU 等級	Memory Size(MB)	硬碟容量(GB)
GRID2	Level2	1024	450

四、執行節點資訊資料庫：

執行節點從執行工作資訊表和執行節點設備資訊表蒐集資訊並傳送Grid資訊給執行節點工作移轉模組，提供監督節點目前執行節點的狀況來予以分配資源。如表3-4所示：

表 3-4 執行節點資訊資料庫

節點名稱	工作次數	CPU 完成 使用量(Sec)	運行次數	完成次數	執行節點 設備等級
GRID2	50	64326	5	45	L1

(一)說明

(A)工作次數：

該節點總共工作的次數。

(B)運行次數：

該節點目前正在運行工作的次數。

(C)完成次數：

該節點目前完成工作的次數。

(D)執行節點設備等級：

包括CPU等級、Memory Size及硬碟容量之整體設備等級；有L1、L2、L3、L4、L5。

第四節 執行節點工作排程分析設計

工作排程設計有以下幾項條件為前提：

1. 電腦是單一 CPU 系統。
2. 系統將提供可用的 CM 值(可用記憶體)。
3. 工作載入連續記憶體。
4. 工作排程為人為操作。

為了達到系統績效最佳化，CPU bound 和 I/O bound 工作需要同時執行，執行的工作數目越大越好。電腦系統為了達到目的必須保持工作資訊的歷史。用工作資訊來建立工作歷史資訊資料庫，可在系統日誌檔案中或手動產生。當新工作進入系統，資訊是由工作歷史資訊資料庫取得，可利用來排程。本排程分初始之工作排程設計、新工作進入之工作排程設計、工作結束之工作排程設計及工作

排程設計的調整來介紹。

(一)初始之工作排程設計

1. 將讀入之工作，在執行工作資訊表中，首先照工作分類等級升冪，其次照 Memory 使用量升冪，最後照來回時間升冪排列。
2. 記錄每個工作分類等級之工作數。
3. 依排入工作分類等級 A、E、B、D、C 的順序依序排入執行一個工作，這是為了 CPU bound 和 I/O bound 可以取得平衡，也是本研究主要所提出的排程方法；然後每等級一個工作排入執行就換下一個等級，記錄每個工作分類等級之執行數。
 - a. 當前工作，剩餘記憶體空間足夠時，則排入執行。
 - b. 當前工作，剩餘記憶體空間不夠時，則找下一個等級的工作排入執行，一直到無法排入執行為止。

(二)新工作進入排程設計

1. 由工作歷史資訊資料庫取得工作資訊，若是找不到其資訊，則以 Memory 使用量 4、工作分類等級 C 處理。
2. 照初始之工作排程的設計來排列。
3. 此工作分類等級之工作數加 1，若是排入執行時，執行數也加 1。

(三)工作結束排程設計

1. 此工作分類等級工作數及執行數各減 1。
2. 看同等級之下一個工作是否可排入執行。
 - a. 剩餘記憶體空間足夠時，則排入執行。
 - b. 剩餘記憶體空間不夠時，則找下一個等級的工作，再看是否可排入執行，一直到都無法排入執行為止。
3. 其工作新資訊可與工作歷史資訊(JHDB)對照，以決定是新

增或更新。

(四)排程設計的調整

系統執行一段時間後，每個工作分類等級之執行數可能不同，就需要調整；當某一工作結束時，將工作分類等級執行數最少者，優先排入執行。

