行政院國家科學委員會補助專題研究計畫成果報告

漸進式以交換技術爲基礎之網路架構

計畫類別： 個別型計畫 整合型計畫

計畫編號：NSC 90-2213-E-002-144-

執行期間： 90 年 8 月 1 日至 91 年 7 月 31 日

計畫主持人：劉邦鋒

計畫參與人員： 林藝芳，張承新，郭奕亨，胡維君．

本成果報告包括以下應繳交之附件：

赴國外出差或研習心得報告一份

赴大陸地區出差或研習心得報告一份

出席國際學術會議心得報告及發表之論文各一份

國際合作研究計畫國外研究報告書一份

執行單位：台灣大學資訊工程系

中 華 民 國 91 年 10 月 31 日

行政院國家科學委員會專題研究計畫成果報告

計畫編號：NSC 90-2213-E-002-144

執行期限：90 年 8 月 1 日至 91 年 7 月 31 日

主持人：劉邦鋒 台灣大學資訊工程系

計畫參與人員：林藝芳 中正大學資訊工程系

一、中文摘要

以交換技術為基礎之路由方法已經成為多處理機交換資訊的重要方法，但此種技術仍有其缺點。例如當一傳輸無法前進時，其使用中之網路頻寬及緩衝區接無法被其他傳輸所使用。為了能有效解決此問題，同時能保持低傳輸時間及不規則網路所特有之可擴充性，本計畫提出一簡單網路架構—漸進式三角網格架構。此種網路架構可支援無死結路由及無衝突路由方法。首先我們證明在漸進式三角網格架構，任何最短路徑陸由方法皆不會產生死結，因此漸進式三角網格架構極適宜在調適性路由方法中擔任備用路徑。其次，我們也證明我們可以將漸進式三角網格架構中之處理器加以排序成一環狀順序，使兩個在此環狀順序中互不重疊的傳輸在網路中互不衝突。此種性質在無衝突路由方法的實做上有極大的幫助。我們初步的實驗結果也證實三角網格架構了所提供的路由方法比傳統的 up-down 路由方法能提供更好的傳輸效能。

關鍵詞：網路架構，無死結路由方法，無衝突路由方法，蟲洞路由方法。

Abstract:

Wormhole switching has become the most widely used switching technique for multicomputers. However, the main drawback of wormhole switching is that blocked messages remain in the network, prohibiting other messages from using the occupied links and buffers. To address the deadlock problem without compromising communication latency and the incremental expansion capability that irregular networks can offer, we propose a simple topology called Incremental Triangular Mesh (ITM) for switch-based networks. ITM is highly scalable, allows incremental expansion of systems, has guaranteed deadlock freedom, and can support contention-free multicast. First, we show that on a ITM, any shortest path routing method will not deadlock, therefore it is ideal to be used as the escape paths in adaptive routing networks. Secondly, we show that it is possible to arrange the nodes of an ITM in a circular order so that two messages from independent parts of the circular order will not interfere with each other, and we can define a circular order for every ITM that has this contention-free property. This is extremely useful for implementing contention-free multicast and other collective communication operations. Our initial experimental results demonstrate that ITM provides better throughput than up-down routing.

Keywords: Network topology, deadlock free routing, contention free routing, wormhole routing.

二、研就成果

We describe the properties of incremental triangular mesh (ITM). The first property is that ITM guarantees freedom from deadlock for any shortest path routing. This property allows ITM to provide maximum bandwidth without the risk of a deadlock. The second property is that we can partition a special subset of ITM so that the messages traveling in different partitions will not interfere with one another. Kesavan et. al. have shown that for some irregular graphs this contention-free ordering does not exist. We show that ITM, which can be very irregular, does provide this ordering. The next two sections describe these two properties in details.

Before we introduce the concept of incremental triangular mesh (ITM) we need to define an operation called AddNode. Let G'= (V', E') be an undirected graph and e' is in E'. To add a node v into G' at edge e'= (x, y) means that we add v into V' and connect v to the two endpoints of e'. The edge e' is called the corresponding edge of v. Formally we have the following.

AddNode(G', v; (x, y)) = (V' ∪ v; E' ∪ {(v; x), (v; y)})

The incremental triangular mesh (ITM) is defined recursively as follows. First the clique of three nodes is an ITM. A graph G is an ITM if and only if there exists another ITM (denoted by G') of n –1 nodes such that G = AddNode (G', v, e'), where e' is the corresponding edge of the newly added node v.

Lemma 1 An ITM G = (V; E) has the following properties.

- G is planar.
- G has 2 |V | - 3 edges.
- For all cycles C in a graph G, there exists an edge e in E such that e connects two nodes in C that are of distance 2.

Theorem 2 Any routing discipline that takes the shortest path is deadlock-free in ITM.

Deadlock can be avoided by providing some escape paths without restricting routing [5]. ITM's deadlock-free property and incremental expansion capability make it a suitable choice for building the escape paths. Since we would like to use ITM as the escape path for the two layer routing approach in [5], we would like to know what kind of graph has ITM as its subgraph, so that part of the links can be used as ITM edges. The following theorem answers this question.

Theorem 3 A graph G = (V; E) has an ITM subgraph that contains all the nodes in V if and only if:

- G is Hamiltonian.
- There exists a Hamiltonian cycle for which G is totally triangulated. A graph G is totally triangulated for a Hamiltonian cycle C if and only if when the vertices of G are around a circle according to the order they appear in C, no edge can be added without intersecting an edge of G.

It is quite likely for a real world network topology to contain an ITM subgraph since both the construction of ITM and switching network are incremental. We only need to be certain that when we add a node, there are at least two neighbors of this newly added node that are already connected. Then we can split those ITM edges into two virtual channels, one for each layer, and deadlock-free routing can be easily and effectively achieved.

This section describes the contention-free property of ITM. We assume that each link in the network is bi-directional and two messages are contention-free as long as they do not go through the same link in the same direction. We also emphasize that each added node of the ITM will have a unique corresponding edge. We further classify the edges of ITM into interior and exterior edges. First all three edges in the kernel three-node-clique are marked as exterior edges.

When a node v is added into an ITM G = (V; E), it can only use an exterior edge (x; y) as its corresponding edge, and then (x; y) becomes an interior edge, and (v; x) and (v; y) are added into AddNode (G; v; (x; y)) as two new exterior edges. It is easy to see from Figure 3 that the exterior edges of an ITM G form a \boundary" around G. The key properties of ITM are summarized in the following lemma.

Lemma 4 For any ITM G = (V; E), we have the following properties.

- An n node ITM has n exterior edges, n-3 interior edges and, n – 2 facets.
- The exterior edges of G form a simple cycle C
- Every node in V is in C

In switch-based network routing it is desirable to have an ordering among all the nodes in a network such that two messages involving processors from different sectors of this ordering do not interfere with each other. That is, suppose we can define a total order < among processors such that when w < x < y < z, then any message-passing between w and x will not interfere with those between y and z. Using this property we can design simple contention-free recursive algorithms for broadcast and multicast, i.e. the source processor first sends the message to the processor in the middle of the list, and repeats the process on the two sub-lists. Despite that this property can be obtained for some regular graphs, it is shown in [17] that there exists irregular graphs for which an ordering is impossible. Nevertheless, we will show that for ITM we can define a similar order that has this nice non-interfering property, despite the irregularity of ITM. We now define a circular order among the nodes in an ITM G = (V, E). From Lemma 4 we know that all the nodes in G form a simple cycle in G. We then enumerate all the nodes in G clockwise or counterclockwise to define a circular order. We also define w < x < y if a node w appears before another node x, which appears before the other node y in the circular order.

Consider two messages m and m 0 . The message m goes from w to x, and m 0 goes from y to z. The two messages m and m 0 are independent if and only if w < x < y < z in the circular order defined earlier. In the following theorem we show that all the shortest paths of two independent messages will not share a communication link in the same direction. That is, a shortest path going from w to y will not share a directed edge with any shortest path going from y to z.

Theorem 5 Two independent messages will not traverse the same edge in the same direction in an ITM under any shortest path routing discipline.