Diseño una Estrategia de Planificación Basada en las Comunicaciones para Sistemas Heterogéneos

V. Arnau, J.M. Orduña, A. Ruiz, R. Valero , J. Duato

Resumen— **En los últimos años se han realizado muchas investigaciones sobre la planificación de tareas en sistemas heterogéneos enfocadas exclusivamente desde el punto de vista computacional. Sin embargo, una estrategia de planificación ideal debe también tener en cuenta los requisitos de comunicación de las aplicaciones y el ancho de banda que puede ofrecer la red. En este trabajo, proponemos por una parte un criterio para medir la calidad de cada asignación de recursos de red a cada aplicación paralela, de acuerdo con los requisitos de comunicación. Por otra parte, proponemos una técnica de planificación basada exclusivamente en este criterio que proporciona una asignación subóptima de procesos a procesadores de acuerdo con los requisitos de comunicación. La evaluación de los resultados muestra que el uso de esta técnica de planificación aprovecha al máximo el ancho de banda disponible en la red, mejorando sensiblemente las prestaciones de esta. Por tanto, la técnica de planificación propuesta puede ser utilizada en el diseño de estrategias de planificación basadas en las comunicaciones para aquellas situaciones donde los requisitos de comunicación generen el cuello de botella de las prestaciones del sistema.**

Palabras clave— **Redes de interconexión, sistemas heterogéneos.**

I. Introducción

En los últimos años se han realizado muchas investigaciones sobre sistemas heterogéneos enfocadas hacia el problema *NP-completo* de planificar adecuadamente diversos grupos de tareas en las máquinas que forman el sistema [3], [6], [8], [9]. No obstante, estas propuestas sólo se centran en aspectos computacionales, y no consideran el coste de la comunicación, suponiendo que en cualquier caso el subsistema de comunicación proporciona suficiente ancho de banda. Sin embargo, conforme la potencia computacional de los nuevos procesadores se incrementa, la red de interconexión puede convertirse en el cuello de botella de estos sistemas heterogéneos. Por esta razón el problema de la planificación debe ser estudiado no solo desde el punto de vista computacional, sino también desde el punto de vista de la comunicación. Dado un sistema heterogéneo (que puede estar formado por diferentes grupos de sistemas homogéneos interconectados) y dado un cierto conjunto de aplicaciones diferentes (paralelas o secuenciales) de diferentes usuarios, una estrategia de planificación ideal sería aquella que realizara la asignación de los procesos a los procesadores teniendo en cuenta tanto los requerimientos computacionales como los de comunicación de las aplicaciones ejecutadas en la máquina. El planificador elegiría entre estrategias de planificación de tareas basadas en computación o basadas en comunicaciones dependiendo de la clase de requisitos que provocaran el cuello de botella en las prestaciones del sistema.

El desarrollo de una estrategia de planificación para sistemas heterogéneos basada en las comunicaciones presenta no obstante varios problemas. Por una parte, se deben medir los requisitos de comunicación que plantean las aplicaciones. Por otra parte, se deben caracterizar de alguna forma los recursos de red existentes en el sistema. Además es necesario definir un criterio para medir la calidad de cada una de las potenciales asignaciones de recursos de la red a cada una de las aplicaciones. Utilizando este criterio como base, se debe entonces desarrollar alguna técnica de asignación de procesos a procesadores basada en los requisitos de comunicación. Por ultimo, dicha técnica debería integrarse dentro de la planificación de procesos, para poder ser aplicada cuando los requisitos de comunicación sean los que provoquen el cuello de botella en las prestaciones del sistema.

En un trabajo previo [1], hemos propuesto un modelo del coste de las comunicaciones, con el fin de caracterizar los recursos de la red para una topologías dada. En este artículo, proponemos por un lado un criterio para medir la calidad de cada asignación de recursos de la red a cada aplicación paralela, de acuerdo con los requisitos de comunicación. Por otro lado, proponemos una técnica de planificación que proporciona una asignación (sub)óptima de procesos a procesadores con respecto a los requisitos de comunicación. En este primer enfoque hemos considerado el problema bajo hipótesis simplificadas (todas las aplicaciones generan tráfico dirigido sólo a procesos del mismo grupo de procesos, sólo hay un proceso por procesador y todos los procesos generan los mismos requisitos de comunicación) para realizar un análisis rápido del comportamiento de las técnicas de planificación propuestas. El propósito es realizar una primera evaluación de las mejoras en las prestaciones de la red que se pueden alcanzar utilizando estas técnicas.

Dada la gran complejidad de medir los requisitos de comunicación de las aplicaciones ejecutadas sobre la máquina y de integrar las técnicas de asignación propuestas en la planificación de procesos, dejaremos la solución de estos problemas para futuros trabajos.

Este trabajo ha sido financiado por el proyecto coordinado CICYT TIC97–0897–C04–01

Departamento de Informática, Universidad de Valencia. Correo electrónico: `Juan.Orduna,Vicente.Arnau@uv.es`.

José Duato pertenece al DISCA, Universidad Politécnica de Valencia. Correo electrónico: `jduato@gap.upv.es`

El resto de este trabajo se organiza de la siguiente forma: en la Sección II resume el problema de la caracterización de redes irregulares y propone un modelo de coste de comunicación. En la Sección III se describen tanto el criterio propuesto para medir la calidad de una determinada asignación de recursos de la red para las aplicaciones ejecutadas sobre la máquina, como la técnica propuesta de planificación de tareas basadas en este criterio. La Sección IV presenta los resultados de la evaluación de prestaciones obtenida con el método propuesto. Finalmente, en la Sección V presentaremos las conclusiones extraídas.

II. Modelo del Coste de las Comunicaciones

En los computadores masivamente paralelos las redes de interconexión han sido tradicionalmente caracterizadas por sus propiedades topológicas, como el número de nodos, el ancho de la bisección y el diámetro. Sin embargo, estas propiedades no nos dan información a cerca de la organización de los canales, una de las cuestiones más importantes cuando caracterizamos topologías irregulares. Por lo tanto, estas propiedades no pueden ser usadas para medir el coste de la comunicación para topologías irregulares. Adicionalmente, el algoritmo de encaminamiento puede también afectar seriamente las prestaciones de topologías irregulares, pues determina la distribución del tráfico en la red. Consideremos, por ejemplo, el esquema de encaminamiento up*/down* usado en redes Autonet [7]. En este esquema, algunos caminos mínimos son prohibidos por el algoritmo de encaminamiento, mientras este tiende a sobrecargar los enlaces cercanos al nodo raíz. Como resultado, la red puede saturarse prematuramente. Por consiguiente, tanto la topología como el algoritmo de encaminamiento deben ser considerados cuando caracterizamos redes irregulares.

En un artículo previo hemos propuesto un nuevo modelo de coste de comunicación entre nodos, la *tabla de distancias equivalentes* [1], [?]. Este modelo es totalmente independiente del patrón de tráfico, y sólo depende de la topología de la red y del algoritmo de encaminamiento. La fuerte correlación del modelo con las prestaciones de la red la convierte en una base se partida para métodos de caracterización de redes que no dependan del patrón de tráfico. Adicionalmente, este modelo puede ser usado como base para una asignación eficiente de procesos a procesadores, ya que proporciona una métrica basada en distancias entre nodos.

Concretamente, nuestro modelo de coste de comunicación entre nodos propone una métrica simple, la *distancia equivalente* entre cada par de nodos (a partir de ahora referiremos un conmutador como un nodo). Esta métrica mide el coste de la comunicación entre dos nodos independientemente del patrón de tráfico [1]. Una tabla de distancias equivalentes se obtiene calculando la distancia equivalente entre cada par de nodos de la red. En particular, el método usado para calcular la distancia equivalente entre dos nodos n_i y n_j se basa en una analogía con los circuitos eléctricos. Primero, se toma en cuenta únicamente los canales pertenecientes a los caminos más cortos que unen n_i y n_j. Los canales restantes se eliminan de la red. Hay qe destacar que en este método estamos utilizando sólamente los canales permitidos por el algoritmo de encaminamiento. Entonces, asumiendo que todos los canales de la red tienen el mismo coste, cada canal es reemplazado por una resistencia de valor unidad. Finalmente, se calcula la distancia equivalente entre n_i y n_j como la resistencia equivalente entre ellos.

La aplicación del método produce una *tabla T_N de $N \times N$ distancias equivalentes* , donde N es el número de nodos de la red. En esta tabla, el elemento T_{ij} representa la distancia equivalente entre el nodo i y el nodo j. La tabla de distancias no satisface la desigualdad triangular, y por tanto no define un espacio métrico. En otras palabras, no podemos encontrar un espacio métrico en el cual representar los nodos con las distancias equivalentes entre ellos. Sin embargo, la tabla de distancias da una medida de la distancia de interconexión, y correlaciona muy bien con las prestaciones de la red, como se muestra en [1].

III. Una Técnica de Planificación Basada en las Comunicaciones

Basándonos en la tabla de distancias, proponemos primero un criterio para medir la calidad de cada asignación de los recursos de la red a las aplicaciones ejecutadas sobre la máquina. Este criterio ya ha sido anteriormente propuesto para caracterizar los recursos de la red [2]. Por otro lado, proponemos un método de búsqueda heurística para establecer la asignación óptima de procesos a procesadores para un conjunto dado de aplicaciones pertenecientes a diferentes usuarios. En general, podemos asumir que los procesos pertenecientes a la misma aplicación podrán comunicar intensivamente entre ellos, aunque no se comunicarán en absoluto con el resto de procesos de otras aplicaciones. Por lo tanto, podemos agrupar los procesos que se ejecutan en la máquina formando conjunto de grupos lógicos de procesos, donde cada grupo está formado por procesos pertenecientes a cada aplicación. Además, para poder analizar rápidamente la técnica de planificación propuesta asumiremos también que existe el mismo número de procesos que de procesadores y que todos los procesos tienen los mismos requisitos de comunicación. Con estas estas limitaciones, el algoritmo propuesto intenta proporcionar una partición de la red adaptada a cualquier conjunto de grupos lógicos existente.

A. Función de calidad

Cuando todo el tráfico generado por el conjunto de grupos lógicos es tráfico dirigido dentro de cada grupo (hemos asumido que cada aplicación pertenece a diferentes usuarios) entonces los procesos pueden ser asignados a los procesadores de forma que aprovechen al máximo los canales de comunicación existentes en la red. Sin embargo, es necesario definir una métrica del ancho de banda de las comunicaciones alcanzadas para cada una de las posibles asignaciones de procesos a procesadores, con el fin de seleccionar la mejor asignación.

Partiendo de la tabla de distancias, hemos definido dos funciones de calidad globales distintas y complementarias, la función *similitud* y la función *desimilitud*. Hay que resaltar que ambas funciones deben tener en cuenta la asignación de procesos a procesadores. Dado que hemos asumido que todo el tráfico generado por el conjunto de agrupaciones lógicas va dirigido al interior de las agrupaciones de procesos, y que existe sólo un proceso por cada procesador, la asignación de procesos a procesadores determinará los destinos de los mensajes generados por cada nodo de la red, y además el coste de las comunicaciones para una topología dada. En este sentido, denotaremos como *cluster* al subconjunto de nodos (conmutadores de la red) asignado a cada grupo lógico de procesos (por motivos de simplicidad, también hemos asumido que todos los nodos de la red son iguales y tienen el mismo número de procesadores conectados a ellos, y que todos los grupos lógicos están formados por un número de procesos que resulta ser un múltiplo entero de nodos de la red). Por tanto, una asignación dada de procesos a procesadores determinará una partición de la red formada por la unión de todos los clusters físicos.

Sea una partición P de la red, formada por M clusters $A_1, A_2, \cdots, A_M$, y sea que un cluster A_i está formado por x_i nodos $a_1, a_2, \cdots, a_{x_i}$ ($x_i < N$, donde N is el número de nodos de la red). En general, los nodos $a_1, a_2, \cdots, a_{x_i}$ formarán un cluster A_i sólo si los procesos de un grupo lógico dado son asignados a este conjunto de nodos. Bajo estas condiciones, se define la función de calidad de cluster para el cluster A_i como

$$F_{A_i} = \sum_{k=1}^{x_i-1} \sum_{j=k+1}^{x_i} T^2_{a_k a_j} \tag{1}$$

donde T_{ij} es la distancia entre el nodo i y el nodo j en la tabla de distancias. Si un cluster A_i contiene x_i nodos, entonces F_{A_i} se define como la suma de los cuadrados de todas las distancias de entre nodos de la agrupación. La función global de similitud para la partición final se define como

$$F_G = \frac{\dfrac{\sum_{i=1}^{M} F_{A_i}}{\sum_{i=1}^{M} \dfrac{x_i (x_i - 1)}{2}}}{\dfrac{\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} T^2_{ij}}{\dfrac{N (N-1)}{2}}} \tag{2}$$

donde M es el número de agrupaciones en la partición final, F_{A_i} representa la función de similitud de cluster para cada cluster A_i y el término

$$\sum_{i=1}^{M} \frac{x_i (x_i - 1)}{2} \tag{3}$$

es el número total de distancias entre nodos de dentro de cada una de cada cluster de la partición P. F_G se calcula como la suma de todos los valores F_{A_i} dividida por el número total de distancias interiores a los clusters existentes en la partición P y normalizado por el valor cuadrático medio de todas las distancias entre nodos de la red. De este modo, un valor de F_G mayor que 1 significa que la asignación final presenta un coste de las comunicaciones interiores de los clusters mayor que cuando la asignación de procesos a procesadores es aleatoria. Un valor de F_G se acerca a 0 significa que la asignación obtenida muestra un coste de las comunicaciones interiores que es muy pequeño comparado con la distancia media cuadrática de los nodos de la red.

La función de desimilitud D_{A_i} para el cluster A_i se define como

$$D_{A_i} = \sum_{k=1}^{x_i} \sum_{j=1}^{N-x_i} T^2_{a_k j} \quad \forall j \notin A_i \tag{4}$$

Es decir, D_{A_i} se define como la suma cuadrática de todas las distancias de cada uno de los nodos del cluster A_i al resto de nodos pertenecientes a otros clusters. La función global de desimilitud se define como

$$D_G = \frac{\dfrac{\sum_{i=1}^{M} D_{A_i}}{\sum_{i=1}^{M} x_i (N - x_i)}}{\dfrac{\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} T^2_{ij}}{\dfrac{N (N-1)}{2}}} \tag{5}$$

donde M representa el número de clusters en la agrupación final. D_G se calcula como la suma de todos los valores D_{A_i} dividida por el número total de distancias existentes entre nodos de distintos clusters en la partición P, y todo ello normalizado por el valor cuadrático medio de los elementos de la tabla de distancias.

F_G y D_G dan una medida de los costes de comunicación entre nodos de un mismo cluster y entre los nodos de distinttos clusters, respectivamente. Por ello, se pueden considerar como medidas inversamente relacionadas al ancho de banda de las comunicaciones interiores a los clusters y al ancho de banda de las comunicaciones entre clusters, respectivamente. Por tanto, el cociente de D_G dividido por F_G nos da la relación entre el ancho de banda interior a los clusters y el ancho de banda entre clusters. Denominaremos a este cociente *coeficiente de clustering* C_c. La técnica de planificación deberá maximizar este coeficiente a la hora de calcular una asignación de procesos a procesadores, ya que se supone que el conjunto de agrupaciones lógicas genera el 100% del tráfico con destino a procesos del mismo grupo de procesos.

B. *Técnica de planificación*

Partimos de un conjunto dado de grupos lógicos de procesos que definirá, junto con la topología de la red, un espacio de soluciones Ω consistente en todas las posibles asignaciones de procesos a procesadores . Este espacio de soluciones vendrá asociado a una función objetivo F que asigna un coste a cada solución particular (asignación de procesos a procesadores) $P \in \Omega$. Como la asignación de procesos se ha demostrado que es un problema $NP-completo$, hemos aplicado un método heurístico de búsqueda para tratar de encontrar la mejor asignación para un conjunto de grupos lógicos dado. Este método heurístico de búsqueda debe encontrar una solución particularP_0 tal que

$$F(P_0) \leq F(P) \;\; \forall \, P \in \Omega$$

Concretamente, hemos utilizado como función objetivo F para cada asignación P la función global de similitud F_G, tal y como se ha definido en la Sección III-A. Dado que esta función representa el coste de las comunicaciones interiores de los clusters, al minimizarla estamos maximizando el coeficiente de clustering C_c, es decir, la relación entre el ancho de banda interior y el ancho de banda de las comunicaciones entre clusters. De esta forma, la búsqueda heurística conduce a una asignación que proporcionará las mayores prestaciones de la red.

Hemos probado varios de los métodos heurísticos propuestos en [3], y hemos obtenido los mejores resultados para una variante del método Tabú [5]. Este método heurístico porporciona resultados iguales o mejores que otros métodos que requieren mayor coste computacional. Dada una asignación de procesos a procesadores P_i (que determinará una partición de la red con M clusters físicos $A_1, A_2, \cdots, A_M$), el método propuesto busca otra asignación P_{i+1} consistente en P_i con la permutación de dos nodos que resulte en el mayor decremento de del valor de F. Por tanto, $F(P_{i+1}) \leq F(P_i)$. Sin embargo, cuando la función objetivo entre en un mínimo local, entonces no habrá ninguna permutación que resulte en un decremento de F. Para poder continuar la búsqueda, el método Tabú establece que en este caso la siguiente iteración P_{i+1} consistirá en P_i con la permutación que resulte en el menor incremento de F. Adicionalmente, la permutación que condujo de P_i a P_{i+1} se prohibe durante un número dado de iteraciones h (de hecho el nombre del método se deriva de estos movimientos "tabú"). La búsqueda termina cuando F alcanza su valor mínimo. El problema es que no existe ninguna forma de asegurar que el valor actual de F sea realmente el valor mínimo absoluto.

En nuestro caso particular comenzamos la búsqueda Tabú con una asignación aleatoria de procesos a procesadores, y aplicamos el método Tabú hasta que hemos alcanzado al menos tres veces el mismo mínimo local, o bien hasta que la búsqueda se ha realizado durante 20 iteraciones. En este punto, el valor mínimo encontrado por el método $F(P_{MIN})$ se considera el valor mínimo, y la asignación correspondiente P_{MIN} es almacenada. Entonces se genera una nueva semilla (asignación aleatoria) $P_{i'}$, continuando la búsqueda desde otro punto de partida diferente. Después de repetir este proceso 10 veces, para redes pequeñas (hasta 16 nodos) el mínimo obtenido con este método fué el mismo valor $F(P_0)$ que el obtenido con una búsqueda exhaustiva. Para redes de mayor tamaño no pudimos realizar una búsqueda exhaustiva, debido a la alta capacidad computacional necesaria para ello. No obstante, no pudimos encontrar ningún valor de F menor con ninguno de los otros métodos heurísticos de búsqueda.

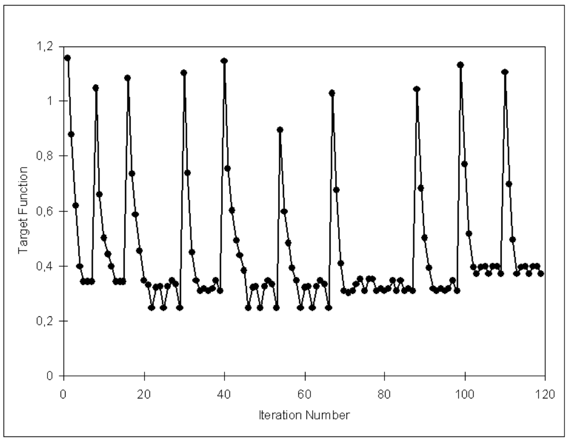

Fig. 1. Búsqueda Tabú para una red de 16 nodos

La Figura 1 muestra los valores $F(P_i)$ obtenidos en la búsqueda Tabú realizada sobre una red de 16 nodos. En esta Figura aparece en el eje de abcisas el número total de itraciones realizadas. Los valores de F en los 10 puntos de partida del método forman los 10 picos que aparecen en la Figura. En esta puede apreciarse cómo el valor de F decrece rápidamente en las primeras iteraciones siguientes a los puntos de partida. Hay que destacar que el valor mínimo de F no se alcanza desde todos los puntos de partida. En este ejemplo se alcanza únicamente desde el tercero, quinto y sexto puntos de partida. Por otra parte, en el séptimo punto de partida se alcanza 20 iteraciones sin llegar a repetir 3 veces el mismo mínimo local.

IV. Evaluación de Prestaciones

En esta sección vamos a estudiar el incremento en las prestaciones de la red que puede proporcionar la técnica de planificación propuesta, así como la correlación entre el coeficiente de clustering y las prestaciones de la red. Para ello hemos evaluado las prestaciones de varias topologías de red con asignaciones aleatorias de procesos a procesadores y también con la asignación que nos da la técnica de planificación propuesta. Hemos calculado el coeficiente de clustering para cada una de estas asignaciones, con el fin de mostrar que esta función realmente nos da a priori una medida de las prestaciones de la red que una asignación determinada puede proporcionar. Este estudio de prestaciones asume que todas las comunicaciones entre procesos existentes van dirigidas a procesos de una misma agrupación y que además todos los procesos generan la misma cantidad de tráfico.

Hemos evaluado las prestaciones de algunas redes irregulares mediante simulación. La metodología de evaluación utilizada se basa en la propuesta en [4]. Las medidas más importantes de prestaciones son la latencia y la productividad. La latencia del mensaje abarca desde que el mensaje es inyectado en la red hasta que el último flit es recibido en el nodo destino. La productividad es la máxima cantidad de información entregada por unidad de tiempo (máximo tráfico aceptado por la red). El tráfico es la tasa de recepción de flits. La latencia se mide en ciclos de reloj. El tráfico se mide en flits por nodo y por ciclo. Nuestro simulador modela la red a nivel de flit.

A. Modelo de la Red y Partición Lógica de Procesos

La red está compuesta por una serie de conmutadores. La topología de la red es irregular y ha sido generada aleatoriamente. Sin embargo, por motivos de simplicidad imponemos tres restricciones. La primera consiste en que hay exactamente 4 estaciones de trabajo conectadas a cada conmutador. La segunda consiste en que dos conmutadores vecinos están conectados por un enlace sencillo. Finalmente, todos los conmutadores de la red poseen el mismo tamaño. Asumimos 8 puertos por conmutador. Sin embargo, cada conmutador posee 4 puertos disponibles para conectar a otro conmutador. De estos 4 puertos, sólo 3 son utilizados cuando generamos la topología. El puerto sobrante se deja libre. Hemos evaluado redes con tamaños comprendidos entre 16 conmutadores (64 estaciones de trabajo) y 24 conmutadores (96 estaciones de trabajo). Para estos tamaños de red, han sido analizadas distintas topologías.

Con el fin de demostrar que la técnica de planificación propuesta es capaz de incrementar las prestaciones de la red, hemos evaluado cada red con distintas asignaciones de procesos a procesadores. Para mayor simplicidad, hemos supuesto un conjunto fijo de N procesos (donde N es también el número de estaciones de trabajo en la red) agrupados en 4 agrupaciones con $X = \frac{N}{4}$ procesos, donde X es también múltiplo de 4 (ya que estamos suponiendo un proceso por procesador). Con esta suposición aseguramos que cada grupo de procesos puede ser asignado en un número entero de conmutadores de la red). Cada proceso se supone que envía todos los mensajes generados a procesos de la misma agrupación lógica de procesos. Para cada red, hemos realizado la búsqueda Tabú hasta que ha proporcionado la mejor partición para 4 agrupaciones de $N/4$ procesos, calculando la correspondiente asignación de procesos. Adicionalmente, hemos calculado algunas asignaciones aleatorias para cada red considerada.

B. Resultados de Simulaciones y Correlaciones

$\boxed{(5,6,8,15)\ (0,1,11,12)\ (3,9,10,14)\ (2,4,7,13)}$

Fig. 2. Partición de 4 grupos obtenida para una red de 16 nodos

La Figura 2 muestra la partición de 4 agrupaciones obtenida para una red de 16 nodos (conmutadores). Dado que para este tamaño de red el conjunto de agrupaciones lógicas está formado por 4 grupos de 16 procesos, la partición ideal estaría formada por agrupaciones o clusters de 4 nodos cada uno. La técnica de planificación proporciona una partición compuesta exactamente por cuatro agrupaciones de 4 nodos. Adicionalmente, hemos calculado 9 asignaciones aleatorias.

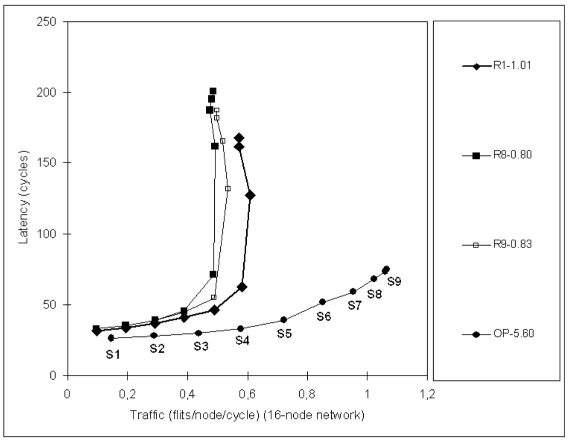

Fig. 3. Resultados de simulación para una red de 16 nodos.

La Figura 3 muestra las prestaciones de red obtenidas para la asignación proporcionada por la técnica de planificación (curva etiquetada como OP), comparadas con las prestaciones obtenidas con varias asignaciones aleatorias (curvas etiquetadas como R_i) (Aunque ejecutamos simulaciones para todas las asignaciones generadas, sólo mostramos unas cuantas por cuestión de claridad en la gráfica). Para cada una de las asignaciones, se simuló la red desde una baja carga (punto de simulación $S1$) hasta el punto de saturación (punto de simulación $S9$). El coeficiente de agrupamiento o clustering obtenido para cada asignación se muestra a la derecha de las etiquetas de cada curva. Como puede observarse, las prestaciones obtenidas con la asignación que proporciona el método propuesto es sobre un 85% mayor que las prestaciones obtenidas con cualquiera de las asignaciones aleatorias. Estos resultados muestran que la técnica de planificación propuesta puede incrementar las prestaciones de la red si el conjunto de agrupaciones lógicas está bien definido. Por otro lado, el valor del coeficiente C_c es claramente menor para las asignaciones aleatorias, mostrando que esta función de calidad está directamente relacionada con las prestaciones de la red.

Adicionalmente hemos estudiado la correlación entre el coeficiente de clustering C_c y las prestaciones

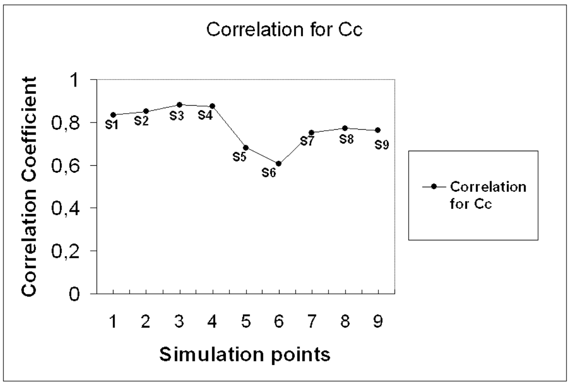

Fig. 4. Correlación del coeficiente de clustering C_c con las prestaciones de la red.

de la red. Dado que las agrupaciones lógicas de procesos están perfectamente definidas (el 100% de los mensajes van dirigidos a procesos de la misma agrupación) es de esperar que un valor más alto de C_c se corresponda con mejores prestaciones de red (tal y como muestra la Figura 3). La Figura 4 muestra la correlación entre el coeficiente C_c y las prestaciones de red obtenidas para todas las asignaciones sobre la red de 16 nodos (ver la Figura 3), desde una carga baja hasta el punto de saturación profunda (punto S9). El coeficiente de correlación es de alrededor del 85% para los puntos correspodientes a una carga baja (puntos S1 a S4). Para los puntos de simulación correspondientes a la saturación (puntos S7 a S9) el coeficiente de correlación está sobre el 75%, demostrando que este coeficiente correlaciona muy bien con las prestaciones de la red incluso bajo condiciones de saturación profunda. Los índices de correlación obtenidos para los puntos de simulación S5 y S6 no son significativos debidos a los diferentes niveles de saturación de la red para cada una de las asignaciones. Por ejemplo, para el nivel de tráfico del punto de simulación S6 la red está en saturación profunda con las asignaciones aleatorias, mientras que con la asignación obtenida mediante la técnica de planificación la red aún no ha entrado en saturación.

Aunque los resultados no se muestran aquí por cuestiones de espacio, también hemos estudiado el índice de correlación para otras redes con distintos tamaños En todos los casos el coeficiente de correlación superó el 70 % para cualquier condición de carga de la red. Estos resultados validan al coeficiente de agrupamiento como medida "a priori" de las prestaciones relativas de la red.

V. Conclusiones y Trabajo Futuro

Es este trabajo hemos propuesto un técnica de planificación consistente en la aplicación de métodos heurísticos de búsqueda utilizando un criterio basado en la tabla de distancias [1]. Esta técnica de planificación intenta ser un primer paso hacia estrategias de planificación sobre sistemas heterogéneos que tengan en cuenta tanto los requisitos de computación como los requisitos de comunicación que presenten las aplicaciones ejecutadas sobre la máquina. El objetivo es evaluar la mejora de las prestaciones de la red que una estrategia de planificación basada en comunicaciones puede ofrecer.

Hemos evaluado las prestaciones de la red para asignaciones de procesos a procesadores proporcionadas por la técnica de planificación, comparándolas con las prestaciones obtenidas con asignaciones aleatorias. La productividad de la red aumenta significativamente cuando se utiliza la asignación obtenida mediante el método propuesto. Por tanto, la técnica de planificación propuesta se puede utilizar como estrategia de planificación cuando la red sea el cuello de botella de un sistema heterogéneo.

Adicionalmente hemos estudiado la correlación entre el criterio de búsqueda propuesto y las prestaciones de la red, mostrando que cuando existe comunicación entre procesos dirigida únicamente a procesos del mismo grupo, entonces este criterio correlaciona muy bien con las prestaciones de la red, y por tanto se puede utilizar como una medida "a priori" de las prestaciones relativas de la red.

Como trabajo futuro nos planteamos el estudio de los aspectos del diseño de la estrategia de planificación que no han sido estudiados en este trabajo: la medida o estimación de los requisitos de comunicación de las aplicaciones, y la integración del método propuesto dentro de las técnicas de planificación.

Referencias

[1] V. Arnau, J.M. Orduña, A. Ruiz, J. Duato, *On the Characterization of Interconnection Networks with Irregular Topology: A New Model of Communication Cost*, Proceedings of XI IASTED International Conference on Parallel and Distributed Computing and Systems (PDCS'99), November 1999, pp.1-6.

[2] V. Arnau, J.M. Orduña, S. Moreno, R. Valero, A. Ruiz, *A Clustering Approach for Improving Network Performance in Heterogeneous Systems, Euro-Par'2000 – Parallel Processing, Springer-Verlag, 2000. Lecture Notes in Computer Science* .

[3] T. D. Braun, H. J. Siegel et al., "A Comparison Study of Static Mapping Heuristics for a class of Meta-tasks on Heterogeneous Computing Systems", in *Proceedings of* 8^{th} *IEEE Heterogeneous Computing Workshop (HCW'99)*, April 1999, pp. 15-29.

[4] J. Duato, "A new theory of deadlock-free adaptive routing in wormhole networks," *IEEE Trans. Parallel and Distributed Systems*, vol. 4, no. 12, pp. 1320–1331, December 1993.

[5] F. Glover, M. Laguna, , *The Grid: Tabu Search*, *Kluwer Academic Publishers, 1998*. ISBN: 0-7923-8187-4

[6] M. Maheswaran, S. Ali, H. Siegel, D. Hensgen and R. Freund, "Dynamic Matching and Scheduling of a Class of Independent Tasks onto Heterogenous Computing Systems", in *Proceedings of* 8^{th} *IEEE Heterogeneous Computing Workshop (HCW'99)*, April 1999, pp. 30-44.

[7] M. D. Schroeder et al., "Autonet: A high-speed, self-configuring local area network using point-to-point links," Technical Report SRC research report 59, DEC, April 1990.

[8] H. Topcuoglu, S. Hariri and M.Y. Wu, "Task Scheduling Algorithms for Heterogenous Processors", in *Proceedings of* 8^{th} *IEEE Heterogeneous Computing Workshop (HCW'99)*, April 1999, pp. 3-14.

[9] S. C. S. L. Wang, H. J. Siegel, V. P. Roychowdhury and A. A. Maciejewski, "Task Matching and Scheduling in heterogeneous Computing Environmets using a Genetic-algorithm-based Approach ", in *Int. Journal of High Speed Computing*, Vol. 7, No. 1, 1995, pp. 45-71.