

Paralelización de Algoritmos. Aplicaciones numéricas y no numéricas.

Ing. A. De Giusti¹, Lic. Marcelo Naiouf²

*Laboratorio de Investigación y Desarrollo en Informática³
Departamento de Informática - Facultad de Ciencias Exactas
Universidad Nacional de La Plata*

Descripción

La evolución del procesamiento hacia el paralelismo ha sido evidente, prácticamente desde el inicio mismo de las computadoras digitales. Los ejes que han impulsado los temas de concurrencia en software y multiprocesamiento en hardware son múltiples, pero podemos mencionar dos:

- La necesidad de reducir los tiempos de procesamiento de grandes volúmenes de datos (problemas matemáticos, modelos, grandes bases de datos, imágenes, sistemas expertos, biotecnología, etc.).
- El procesamiento de información (datos, señales) en tiempo real para la toma de decisiones tanto en ambientes administrativos como industriales (robótica, industria militar, sistemas multimediales en tiempo real, georeferenciación, reconocimiento de patrones, etc.).

Esta evolución conduce a un gran esfuerzo por transformar el procesamiento secuencial en paralelo, buscando reducir los tiempos de ejecución de procesos y de respuesta a eventos del mundo real. También en este aspecto podemos encontrar dos ejes en la respuesta tecnológica:

- El crecimiento de la potencia de cómputo, dado en la evolución de la tecnología de los componentes y en las arquitecturas de procesamiento (supercomputadoras, hipercubos de procesadores homogéneos, grandes redes de procesadores no-homogéneos, procesadores de imágenes, de audio, etc.).
- La transformación y creación de algoritmos que exploten al máximo la concurrencia implícita en el problema a resolver, de modo de distribuir el procesamiento minimizando el tiempo total de respuesta. Naturalmente esta transformación, también debe adaptarse a la arquitectura física de soporte.

En esta línea de investigación aplicada, vinculada con los proyectos "Procesamiento Concurrente y Paralelo" y "Sistemas de Tiempo Real" del LIDI se trabaja especialmente en la migración de algoritmos secuenciales a paralelos sobre diferentes modelos de arquitecturas, evaluando herramientas de software y parámetros de calidad algorítmica.

¹ Investigador Principal CONICET. Profesor Titular Dedicación Exclusiva. Dpto. de Informática. Facultad de Ciencias Exactas. Universidad Nacional de La Plata. E-mail: degiusti@info.unlp.edu.ar

² Profesor Adjunto Dedicación Exclusiva. Dpto. de Informática, Facultad de Ciencias Exactas. Universidad Nacional de La Plata. E-mail: mnaiouf@info.unlp.edu.ar

³ Calle 50 y 115 1er Piso, (1900) La Plata, Argentina, TE/Fax +(54)(221)422-7707. <http://lidi.info.unlp.edu.ar>

Objetivos de la línea de investigación

En este proyecto se investiga sobre la especificación de procesos concurrentes y paralelos; la transformación de algoritmos secuenciales en paralelos y la optimización de los mismos, así como la implementación de soluciones de procesamiento masivo de datos mediante arquitecturas multiprocesador, evaluando su eficiencia.

Interesa especialmente la aplicación de estas investigaciones al procesamiento de señales (voz e imágenes) tanto para su tratamiento en tiempo real en áreas tales como reconocimiento de patrones o robótica, como para su transmisión a distancia.

Existe un trabajo conjunto en algunos temas con otros grupos del país (CeTAD UNLP, UNS, UN Comahue) y del exterior (Universidad de la República, Universidad Autónoma de Barcelona, Universidad de New Mexico).

Los modelos de arquitectura presentes en el Laboratorio de Procesamiento Paralelo que se utilizan en la investigación experimental son tres:

- Hipercubo de 32 procesadores homogéneos (transputers t805 con 4 Mb de RAM local) que puede conectarse con estaciones con 1, 2 y 4 transputers.
- Red de procesadores heterogéneos (local y remota) que incluye PCs y workstations, en las que se trabaja emulando una arquitectura paralela mediante PVM y MPI.
- Arquitectura multiprocesador dedicada (CISC) basada en procesadores especiales para señales. (DSPs Motorola 56000 y TMS 340).

El soporte de software utilizado hasta el momento ha sido C paralelo (tanto para transputers como para DSPs), OCCAM, ADA y diferentes lenguajes ejecutables sobre las arquitecturas no homogéneas y vinculables con PVM y MPI (en particular Fortran y Java).

Se ha trabajado en la transformación de diferentes clases de algoritmos:

- Algoritmos numéricos puros (operaciones con vectores y matrices, ordenación, operaciones matriciales sobre grafos) con carga balanceada y desbalanceada y con distribución de procesos estática y dinámica.
- Algoritmos de tratamiento de caracteres, en particular compresión y descompresión sin pérdida con algoritmos estáticos y adaptivos.
- Algoritmos de tratamiento de imágenes (compresión/descompresión con pérdida) y de reconocimiento de patrones en paralelo (off-line y on-line).
- Algoritmos de scheduling de datos y procesos en sistemas físicamente distribuidos.

En todos los casos la metodología consiste en analizar las soluciones secuenciales, estudiar alternativas de paralelización, analizar su adaptabilidad a los modelos de arquitectura mencionados anteriormente y estudiar la implementación y performance de los algoritmos paralelos.

Resultados obtenidos

En el proyecto se ha trabajado en la clase de problemas mencionado anteriormente, teniendo diversos resultados (publicaciones, Tesinas de Grado, Tesis de Magister, Tesis de Doctorado) y entre las experiencias concretas realizadas puede mencionarse:

- Ordenación de datos numéricos, producto de matrices, solución de sistemas de ecuaciones y búsqueda de camino mínimo en grafos. Realizados sobre red heterogénea con PVM y transputers.
- Compresión y descompresión paralela sin pérdida, estática y adaptiva con algoritmos clásicos (Huffman, LZW, Run Length) sobre una red heterogénea con PVM.
- Compresión y descompresión paralela con pérdida, estática y adaptiva con diferentes algoritmos (JPEG, Wavelets, Fractales) sobre una red heterogénea con PVM.
- Algoritmos de separación de curvas en el plano sobre redes, transputers y DSPs.
- Paralelización de algoritmos de reconocimiento estadístico de patrones sobre redes y transputers. Extensión al tratamiento en tiempo real para productos industriales que se clasifican.
- Actualmente se está trabajando en seguimiento paralelo de trayectorias para aplicaciones en robótica y en tratamiento paralelo de video en tiempo real para su transmisión a distancia.
También se trabaja en algoritmos paralelos para clasificación de imágenes por texturas.(transputers).

Bibliografia

- [Akl89] Akl S., "The Design and Analysis of Parallel Algorithms", Prentice-Hall, Inc., 1989.
- [Akl97] Akl S., "Parallel Computation. Models and Methods", Prentice-Hall, Inc., 1997.
- [Amd67] Amdahl G., "On the Validity of the Single-Processor Approach to Achieving Large-SCALE Computing Capabilities", Proc. Of the AFIPS Conference, 1967, pp. 483-485
- [Bri95] Brinch Hansen, P., "Studies in computational science: Parallel Programming Paradigms", Prentice-Hall, Inc., 1995.
- [Bub97] Bubak, Funika, Moscinski, "Performance Analysis of Parallel Applications under Message Passing Environments", www.icsr.agh.edu.pl/publications/html/perf_full/, 1997
- [Cha88] Chandi K. M., Misra J., "Parallel Program Design. A Foundation", Addison Wesley, 1988.
- [Cof92] M. Coffin, "Parallel programming- A new approach", Prentice Hall, Englewood Cliffs, 1992.
- [Col96] Colbrook A., Brewer E., Dellarocas C., Weihl W., "Algorithms for Search Trees on Message-Passing Architectures", IEEE Transactions on Parallel and Distributed Systems, Vol. 7, No. 2: February 1996, pp. 97-108
- [CSA90] "Transputer Architecture", Computer System Architects, 1990.
- [Day94] Day K., Tripathi A., "A Comparative Study of Topological Properties of Hypercubes and Star Graphs" IEEE Trans. Parallel and Distributed Systems, vol. 5, no. 1, pp. 31-38, Jan. 1994
- [Don94] Dongarra J., et al, "PVM: Parallel Virtual Machine, A User's Guide and Tutorial for Networked Parallel Computing", MIT Press, 1994.
- [Fly72] Flynn M. J., "Some Computer Organizations and Their Effectiveness", IEEE Transactions on Computers, C-21, No 9, September, 1972.
- [Gla95] A. Glassner, "Principles of Digital Image Synthesis", Morgan Kaufmann, 1995.
- [Gon92] Rafael C. Gonz ales, Richard E. Woods, "Digital Image Processing", Addison-Wesley Publishing Comp., 1992.
- [Gup93] Gupta A., Kumar V., "Performance properties of large scale parallel systems", Journal of Parallel and Distributed Computing, November 1993.
- [Gus88] Gustafson J., "Reevaluating Amdahl's Law", Comm. Of the ACM, Volume 31 (1988), pp. 532-533.
- [Hoa85] Hoare C., "Communicating Sequential Processes", Englewood Cliffs, Prentice-Hall, 1985.
- [Hwa93] Hwang K., "Advanced Computer Architecture: Parallelism, Scalability, Programability", McGraw-Hill, 1993.
- [Kar92] Karonis N., "Timing Parallel Programs That Use Message Passing", Journal of Parallel and Distributed Computing, 14, 1992.
- [Kum94] Kumar V., Grama A., Gupta A., Karypis G., "Introduction to Parallel Computing. Design and Analysis of Algorithms", Benjamin/Cummings, 1994.
- [Kun91] Kung H. T., Sansom R., Schlick S., Steenkiste P., Arnould M., Bitz F. J., Christianson F., Cooper E. C., Menzilcioglu O., Ombres D., and Zill B., "Network-Based Multicomputers: An Emerging Parallel Architecture" Proc. Supercomputing '91, pp. 664-673. IEEE, Albuquerque, Nov. 1991.
- [Law92] H. Lawson, "Parallel processing in industrial real time applications", Prentice Hall 1992.
- [Lei92] Leighton F. T., "Introduction to Parallel Algorithms and Architectures: Arrays, Trees, Hypercubes", Morgan Kaufmann Publishers, 1992.
- [Log94] "Transputer Toolset", Logical System, 1994.
- [Luq93] Luque E, Senar M, Hernandez P, Franco D, Heymann E, Moure J, "Distributed Kernel for a Transputer Based Computer", Transputers Applications and Systems'93. Vol 2, IOS Press, 1993.
- [Luq95] Luque E, Ripoll A, Cort es A, Margalef T, "A Distributed Diffusion Method for Dynamic Load Balancing on Parallel Computers", Proceedings of the EUROMICRO Workshop on Parallel and Distributed Processing, IEEE Computer Society, Jan. 1995.
- [Mil98] Miller R., Stout Q. F., "Algorithmic Techniques for Networks of Processors", CRC Handbook of Algorithms and Theory of Computation, M. J. Atallah, ed, 1998.
- [Mor94] Morse F., "Practical Parallel Computing", AP Professional, 1994.
- [Nig95] Nigam M., Sahni S., "Sorting n^2 Numbers on $n \times n$ Meshes", IEEE Transactions on Parallel and Distributed Systems, Vol. 6, No. 12: December 1995, pp. 1221-1225
- [Sim97] Sima D, Fountain T, Kacsuk P, "Advanced Computer Architectures. A Design Space Approach", Addison Wesley Longman Limited, 1997.

- [Ste96] Steenkiste P., "Network-Based Multicomputers: A Practical Supercomputer Architecture", IEEE Transactions on Parallel and Distributed Systems, Vol. 7, No. 8, August 1996, pp. 861-875
- [Sun90] Sun X., Ni L. M., "Another view of parallel speedup", Supercomputing '90 Proceedings, pp 324-333, 1990.
- [Sun95] Sun X., Zhu J., "Performance Considerations of Shared Virtual Memory Machines", Transactions on Parallel and Distributed Systems, Vol 6, No. 11: November 1995, pp. 1185-1194.
- [Tin98] Tinetti F., De Giusti A., "Procesamiento Paralelo. Conceptos de Arquitectura y Algoritmos", Editorial Exacta, 1998.
- [Zom96] Zomaya A. (ed), "Parallel Computing. Paradigms and Applications", International Thomson Computer Press, 1996.