

EXTENSIÓN FUNCIONAL DE CLUSIM: SIMULACIÓN DE APLICACIONES PARALELAS SPMD, PIPELINE Y DIVIDE/CONQUER

Claudio Marcelo Pérez Ibarra¹, Adelina García¹, Sandra Adriana Méndez², Cecilia Lasserre¹.

¹Grupo de Ingeniería de Software (GIS). Facultad de Ingeniería. Universidad Nacional de Jujuy
Avda Martiarena esq. Italia. San Salvador de Jujuy. Jujuy. CP 4600
TE 54+388+4221591 CEL 54+388+ 0388-154145307

²Departamentos de Arquitectura de Computadores y Sistemas Operativos. Escuela Técnica Superior de Ingeniería. Universidad Autónoma de Barcelona
cmperezi@gmail.com, adelinagarcia_jujuy@hotmail.com

RESUMEN

El rendimiento de un sistema paralelo depende tanto de la configuración de recursos de la computadora como del tipo de aplicación a ejecutar. Por ende, el estudio del comportamiento de aplicaciones paralelas en *clusters*, considerando diferentes configuraciones, puede establecer qué configuraciones son más adecuadas según el tipo de aplicación. El estudio de rendimiento de estos sistemas no es tarea fácil, ya que implica detener el *cluster*, reconfigurar sus recursos, analizar el comportamiento del sistema y evaluar el rendimiento; todo esto con el consecuente impacto en tiempo y costo. CluSim es un simulador de *clusters* para aplicaciones paralelas, basado en OMNeT++, que permite parametrizar la configuración de un *cluster* de modo que sea posible evaluar y predecir el impacto en el rendimiento de diferentes configuraciones para aplicaciones tipo *Master/Worker*. En este proyecto se propone extender la funcionalidad CluSim incorporando a su modelo de simulación aplicaciones paralelas tipo SPMD, *Pipeline* y *Divide/Conquer*.

Palabras claves: cómputo de altas prestaciones, clusters, simulación, paradigmas de programación paralela

CONTEXTO

Este proyecto está inserto en una línea de investigación que se viene desarrollando desde el año 2008 en dos proyectos acreditados por la Universidad Nacional de Ju-

jujuy: Aplicaciones del Cómputo de Altas Prestaciones y Sistemas de Cómputo de Altas Prestaciones con Alta Disponibilidad: Evaluación de la Performance en Diferentes Configuraciones.

El proyecto se encuentra acreditado y financiado por la Secretaria de Ciencia y Técnica y Estudios Regionales de la Universidad Nacional de Jujuy (SECTER-UNJu), y asesorado por la Mg. Sandra Méndez de la Escuela Técnica Superior de Ingeniería, Dpto. Arquitectura de Computadoras y Sistemas Operativos de la Universidad Autónoma de Barcelona.

INTRODUCCIÓN

Hoy, la ciencia, la ingeniería, la industria, la medicina, la economía, entre otras áreas, plantean problemas que por su complejidad, tamaño y/o restricciones de tiempo (resultados en tiempo real) demandan gran poder de cómputo para su resolución. Sin embargo, esta demanda supera la capacidad computacional de elementos de procesamiento simple, tales como equipos de escritorio o computadoras portátiles. No obstante, el avance de las arquitecturas multinúcleo, las redes de interconexión de alta velocidad, librerías de paso de mensaje y memoria compartida para programación de múltiples procesos han hecho posible que la Computación de Alto Rendimiento (*High Performance Computing*, HPC) trate aquellos problemas que exceden a la computación serial [Carver, 2007].

Es así que el procesamiento paralelo, im-

plementado mediante supercomputadoras, entornos *cluster*, *multicluster* y *grid*, constituye un modelo tecnológico en el que n procesadores, interconectados, resuelven problemas p veces más rápido que un único procesador [Jortberg, 1997].

Las supercomputadoras, como las listadas en el top500 (www.top500.org), están integradas por miles de procesadores o multicomputadoras que aprovechan la potencia conjunta de varios elementos de procesamiento simple. Lógicamente, la construcción y mantenimiento de equipos de semejante envergadura requiere de presupuestos millonarios, por lo que estos sistemas no están disponibles para todo el mundo. Sin embargo, arquitecturas como los *clusters* de computadoras ofrecen una alternativa económica para sistemas de alto rendimiento [Catalán, 2003].

Los *clusters* de computadoras basados en *workstation* (WS) y redes convencionales (local o remota) o COTS (*Commodity Off The Shelf*), permitieron que muchas instituciones y organizaciones construyeran y usaran computadoras paralelas, teniendo en mente, aumentar la capacidad de procesamiento al aprovechar el potencial ofrecido por un gran número de procesadores y proporcionar servicio ininterrumpido al administrar eficientemente los recursos en todo momento [Rodrigues, 2006].

A fin de lograr esto, la configuración de un *cluster* requiere no sólo de un cuidadoso diseño de su arquitectura y funcionalidad, sino también de un conocimiento detallado de la distribución de datos y del modelo de comunicación que siguen las aplicaciones a ejecutar. La distribución de datos y el modelo de comunicación determinan los patrones de cómputo y comunicación de la aplicación según el paradigma de programación (SPMD, *Divide/Conquer*, *Pipeline*, *Master/Worker*) que ésta siga [Muresano, 2008].

Por ello, en la evaluación de rendimiento de estos sistemas se debe considerar diferentes configuraciones de recursos y distintos patrones de cómputo y comunicación, basándose tanto en las características del

cluster como de la aplicación. En la práctica, realizar esta evaluación considerando distintos escenarios no siempre es posible, ya sea, porque el sistema aún no está disponible o porque el impacto práctico o económico es demasiado alto. En estos casos, una de las estrategias más recomendada es la simulación.

En la literatura existen muchos trabajos centrados en simular grandes redes y aplicaciones de HPC. Sin embargo, la mayoría de estos simuladores de redes se enfocan en arquitecturas específicas. Molero y otros [Molero *et al.*, 2000] presentan un simulador de Redes de Área de Almacenamiento (SANs) que permite trabajar tanto con trazas de tráfico reales como sintéticas, y simula fallos en enlaces y *switches*, canales virtuales, diferentes algoritmos de ruteo, etc. SIMLAB [Berenbrink *et al.*, 2001] es otro entorno de simulación para SANs que fue implementado en C++ y permite modelar discos rígidos, nodos de enrutamiento e interfaces de redes. Tiene como objetivo ayudar en el desarrollo y verificación de algoritmos distribuidos para la red PRES TO SAN (que soporta la entrega de datos en tiempo real) ayudando a decidir qué algoritmos son más adecuados y eficientes para ese tipo de red. PARSEC [Bagrodia *et al.*, 1998] es un entorno de simulación de eventos discretos, que, mediante un compilador mejorado de C++ permite simular entidades y constructores de mensajes de comunicación entre entidades. SIMCAN [Nuñez *et al.*, 2008] es un entorno de simulación para grandes redes complejas de almacenamiento que simula estas redes y sus subsistemas subyacentes (I/O, Networking, etc.).

También existen trabajos que simulan aplicaciones MPI. Por ejemplo, en [Bagrodia *et al.*, 2008] se presenta a MPI-SIM, un simulador paralelo diseñado para predecir el rendimiento de aplicaciones MPI y MPI-IO ya existentes. El objetivo de MPI-SIM es predecir el rendimiento de estos programas en función de características de arquitectura tales como número de procesadores, latencias en la comunicación, algo-

ritmos de caché, etc. En [Riesen, 2006] se presenta el prototipo de un simulador que consiste en un enfoque híbrido entre la ejecución de una aplicación paralela en modo *stand-alone* y la simulación de una red que se utiliza para el paso de mensajes de MPI. Este prototipo necesita que la aplicación simulada se ejecute en el mismo (o casi el mismo) hardware que se desea simular (por ejemplo, igual CPU a la misma velocidad, igual subsistema de memoria, etc.) También es posible encontrar entornos de simulación de redes de propósito general que permiten crear diferentes configuraciones de redes, con diferentes tipos de nodos, *switches*, topologías, protocolos, etc. Ejemplos de éstos son OMNeT++ (www.omnetpp.org) y OPNET Modeler (www.opnet.com).

CluSim es un simulador de *clusters* homogéneos para aplicaciones *Master/Worker* que permite parametrizar la configuración de recursos de un *cluster*, los patrones de cómputo y comunicación de la aplicación paralela y el tamaño del problema, de modo que sea posible evaluar y predecir el impacto en el rendimiento del sistema considerando diferentes configuraciones (número de nodos del *cluster*). Este simulador, basado en OMNeT++, fue presentado en los trabajos [Pérez Ibarra *et al.*, 2010a], [Valdiviezo *et al.*, 2010] y [Pérez Ibarra *et al.*, 2010b]. CluSim pretende servir como herramienta de soporte a la toma de decisiones para la selección de las configuraciones más adecuadas para un *cluster* que ejecuta un determinado tipo de aplicación paralela.

Considerando lo expuesto, este proyecto pretende:

1. Determinar los parámetros que caracterizan el comportamiento (cómputo/comunicación) de aplicaciones MPI tipo SPMD, *Pipeline* y *Divide/Conquer*.
2. Definir un modelo de simulación para los paradigmas de programación paralela SPMD, *Pipeline* y *Divide/Conquer* basándose en los patrones de comunicación y cómputo.

3. Incluir en el modelo de simulación de CluSim dichos tipos de aplicación.

LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

Evaluación de rendimiento de sistemas paralelo en base a la simulación de *clusters* y aplicaciones paralelas SPMD, *Pipeline* y *Divide/Conquer*.

RESULTADOS Y OBJETIVOS

El proyecto, iniciado en febrero de 2012, aún no presenta resultados. Los objetivos del proyecto son:

- Determinar los parámetros que caracterizan el comportamiento de aplicaciones tipo SPMD, *Pipeline* y *Divide/Conquer*.
- Incluir en el modelo de simulación de CluSim los tipos de aplicación indicados.

FORMACIÓN DE RECURSOS HUMANOS

El equipo de trabajo dedicado a la actual línea de investigación está integrado por:

- Pérez Ibarra, Claudio Marcelo (Director)
- García, Adelina (Codirectora)
- Méndez, Sandra Adriana (Asesora)
- Valdiviezo, Luis Marcelo
- Chosco, Víctor Hugo
- Márquez Ruíz, Pablo Ezequiel
- Nolasco, Silvia Alejandra
- Ganami, Martín

Estudios de postgrado y Tesinas de finalización de carrera de grado:

- Doctorado en curso: 1
- Trabajo final de Especialización: 1
- Tesinas en curso: 2
- Beca CIN: 1. Directora MSc. Cecilia Lasserre

REFERENCIAS

[Bagrodía *et al.*, 1998] Bagrodía, R., M. Takai, Y. Chen, X. Zeng, and J. Martin. Parsec: A parallel simulation environ-

- ment for complex systems. *IEEE Computer*, 31:77–85, 1998.
- [Bagrodia *et al.*, 2008] Bagrodia, R., E. Deelman, and T. Phan. Parallel simulation of large-scale parallel applications. *Int. J. High Perform. Comput. Appl.*, 15(1):3–12, 2001.
- [Berenbrink *et al.*, 2001] Berenbrink, P., A. Brinkmann, and C. Scheideler. Simlab - a simulation environment for storage area networks. In *Workshop on Parallel and Distributed Processing (PDP)*, pages 227–234, 2001.
- [Carver, 2007] Carver, Jeffrey C. Third international workshop on software engineering for high performance computing (hpc) applications. In *ICSE COMPANION '07: Companion to the proceedings of the 29th International Conference on Software Engineering*, page 147, Washington, DC, USA, 2007. IEEE Computer Society.
- [Catalán, 2003] Catalán i Coit, Miguel. “Nuevo Modelado de Computación Paralela con Clusters Linux”. VI Congreso HIPALinux. miKeL a.k.a.mc2. Septiembre 2003.
- [Jortberg, 1997] Jortberg, Charles A. “The Supercomputers”. Editorial Abdo & Daughters. Minneapolis. USA. 1997.
- [Molero *et al.*, 2000] Molero, X., F. Silla, V. Santonja, and J. Duato. Modeling and simulation of storage area networks. *Modeling, Analysis, and Simulation of Computer Systems, International Symposium on*, 0:307, 2000.
- [Muresano, 2008] Muresano, Ronal. “Aplicaciones Single Program Multiple Data (SPMD) en Ambientes Distribuidos”. Memoria del trabajo de investigación del “Máster en Computación de Altas Prestaciones”, presentada en la Escuela Técnica Superior de Ingeniería (Departamento de Arquitectura de Computadores y Sistemas Operativos). 2008.
- [Nuñez *et al.*, 2008] Nuñez, A., J. Fernandez, J. D. Garcia, L. Prada, and J. Carretero. Simcan: a simulator framework for computer architectures and storage networks. In *Simutools '08: Proceedings of the 1st international conference on Simulation tools and techniques for communications, networks and systems & workshops*, pages 1–8, ICST, Brussels, Belgium, Belgium, 2008. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).
- [Pérez Ibarra *et al.*, 2010a] Pérez Ibarra, C. M., L. M. Valdiviezo, N. M. Pérez Otero, H. P. Liberatori, D. Rexachs, E. Luque y C. M. Lasserre. CLUSIM: Simulador de Clusters para Aplicaciones de Cómputo de Altas prestaciones basado en OMNeT++. XVI Congreso Argentino de Ciencias de la Computación. Buenos Aires (Morón). Octubre 2010.
- [Pérez Ibarra *et al.*, 2010b] Pérez Ibarra, C. M.; C. M. Lasserre; Pérez Otero N. M.; Liberatori, H. P. y L. M Valdiviezo. “Simulador de un cluster tolerante a fallos basado en OMNeT++”. XII Workshop de Investigadores en Ciencias de la Computación. Área de Procesamiento Paralelo y Distribuido. ISBN 978-950-34-0652-6. El Calafate. Argentina. pp. 587-591. 2010. Mayo 2010.
- [Riesen, 2006] Riesen, Rolf. “Supercomputer Simulation Design Through Simulation”. Cray User Group (CUG). ISBN: 1-4244-0327-8. Barcelona. pp. 1–9. 2006.
- [Rodrigues, 2006] Rodrigues de Souza, Josesmar. “FTDR: Tolerancia a fallos, en clusters de computadoras geográficamente distribuidos, basada en Replicación de Datos”. Tesis doctoral de la Universidad Autónoma de Barcelona.
- [Valdiviezo *et al.*, 2010] Valdiviezo, L. M.; Pérez Otero, N. M.; Pérez Ibarra, C. M. y C. M. Lasserre. Caracterización de una aplicación paralela con distintas configuraciones en CluSim. *Investigaciones en Facultades de Ingeniería del NOA – 2010*. ISSN 3367-5072. Ed. EdiUNJu. S. S. de Jujuy. pp. 499-504. Noviembre 2010.