

## “Análisis de Eficiencia y Tolerancia a Fallos en Arquitecturas Multiprocesador para Aplicaciones de Procesamiento de Datos”

Jorge R. Osio<sup>1,2</sup>, Diego Montezanti<sup>1,4</sup>, Eduardo Kunysz<sup>1</sup>, Diego Encinas<sup>1,4</sup> Daniel Martin Morales<sup>1,3</sup>

<sup>1</sup>Instituto de Ingeniería y Agronomía - UNAJ

<sup>2</sup>UIDET CeTAD –Fac. de Ingeniería - UNLP

<sup>3</sup>Laboratorio LINES - Universidad Tecnológica Nacional FRLP

<sup>4</sup>Instituto de Investigación en Informática LIDI - Fac. de Informática – UNLP

{josio, dmontezanti, ekunysz, martin.morales}@unaj.edu.ar

### Resumen

Dentro de la línea de investigación que se está desarrollando, existen dos enfoques diferentes. Por un lado se está trabajando sobre la implementación de algoritmos de procesamiento de imágenes sobre dispositivos reconfigurables. El objetivo es utilizar una combinación de diferentes técnicas de concurrencia y paralelismo para tener en cuenta aspectos comunes de dichos algoritmos, y así mejorar la eficiencia en el procesamiento de imágenes médicas. Por otra parte, debido a que el procesamiento en paralelo requiere de la implementación de sistemas de múltiples procesadores en dispositivos reconfigurables, resulta útil incorporar al proyecto el desarrollo de metodologías que permitan tolerar fallos transitorios que son característicos de las arquitecturas multicore, y que afectan especialmente la ejecución de aplicaciones paralelas de cómputo intensivo.

**Palabras clave:** *arquitecturas paralelas, procesamiento de imágenes, tolerancia a fallos, sistemas multicores, dispositivos reconfigurables.*

### Contexto

La línea de Investigación descrita en este trabajo forma parte del Proyecto de Investigación Científico-Tecnológico “Tecnologías de la información y las comunicaciones mediante IoT para la solución de problemas en el medio socio productivo”, que se desarrolla en la Universidad Nacional Arturo Jauretche (UNAJ) y es continuación del proyecto presentado en la edición anterior de WICC [7].

El proyecto cuenta además con financiamiento en el marco del programa “Universidad, Diseño y Desarrollo Productivo” del Ministerio de Educación a través del proyecto “Sistema de eficiencia energética” y está a la espera de su Aprobación para obtener financiamiento de la convocatoria “UNAJ Investiga 2017”.

Parte de las líneas de investigación desarrolladas se encuentran enmarcadas en los convenios de colaboración en Actividades de Investigación firmados por la UNAJ con la UIDET-CeTAD.

## 1. Introducción

En los últimos años, se ha buscado expandir el concepto del procesamiento paralelo basado en computadoras con multicores, llevándolo hacia la utilización de plataformas más específicas. Para obtener una mayor eficiencia en el procesamiento, los fabricantes de computadoras de altas prestaciones han introducido unidades basadas en FPGAs (*Field-Programmable Gate Array*) en su diseño como soporte para el cómputo [1-3].

Si bien el estudio de sistemas paralelos con multiprocesadores es un campo bien desarrollado, la utilización de múltiples cores en sistemas reconfigurables es un terreno que tiene múltiples posibilidades de exploración [4]. En el presente proyecto se exploran mejoras en la implementación de los algoritmos mediante procesamiento paralelo mediante procesadores microblaze y concurrente en VHDL (lenguaje de descripción de hardware) y la tolerancia a fallos en entornos multicore.

### **Plataformas FPGAs para procesamiento paralelo**

La implementación del paralelismo en plataformas FPGAs consiste en el uso de procesadores embebidos para ejecutar aplicaciones y en la utilización de las características que provee la lógica programable para manejar las porciones de código que se ejecutan concurrentemente [4].

La facilidad de implementar procesadores embebidos en forma rápida, junto con la posibilidad de proveer concurrencia mediante la programación en hardware, permiten combinar las FPGAs con el paralelismo obtenido

mediante sistemas multicore para alcanzar una alta eficiencia, que es un gran desafío en la búsqueda de optimizar el procesamiento de imágenes médicas. La mejora se aplica al conjunto de algoritmos basados en operadores de ventana, donde una porción de los algoritmos es común a todos posibilitando la concurrencia en VHDL, mientras que la otra parte del algoritmo se implementa en software y se ejecuta en paralelo dentro del sistema multicore.

Con las mejoras constantes que aporta la evolución de la tecnología sobre las FPGAs pueden lograrse diseños de gran magnitud, a tal punto que la tendencia actual es implementar microprocesadores de propósito general, conjuntamente con todo el hardware de propósito específico que requiere la aplicación, dentro de una FPGA.

### **Tolerancia a Fallos en sistemas multicore**

El aumento en la escala de integración, con el objetivo de mejorar las prestaciones en los procesadores actuales, sumado al crecimiento de los sistemas de cómputo, han producido que la confiabilidad se haya vuelto un aspecto relevante. En particular, la creciente vulnerabilidad a los fallos transitorios se ha vuelto crítica, a causa de la capacidad de estos fallos de corromper los resultados de las aplicaciones.

El impacto de los fallos transitorios se vuelve más notorio en el contexto del HPC, donde el escenario típicamente consiste en una supercomputadora con cientos o miles de procesadores que cooperan para ejecutar aplicaciones paralelas, debido a que el Tiempo Medio Entre Fallos (MTBF) del sistema disminuye cuanto mayor es la cantidad de procesadores. La incidencia de los fallos transitorios es aún mayor en el caso de

aplicaciones intensivas en cómputo y de gran duración, que manejan elevados volúmenes de datos [11].

El alto costo (en términos temporales y de utilización de recursos) que implica volver a lanzar una ejecución desde el comienzo, en caso de que un fallo transitorio produzca la finalización de la aplicación con resultados incorrectos, justifica la necesidad de desarrollar estrategias específicas para mejorar la confiabilidad y robustez en sistemas de HPC. En particular, resulta crucial poder detectar los fallos llamados silenciosos, que alteran los resultados de las aplicaciones pero que no son interceptados por el sistema operativo ni ninguna otra capa de software del sistema, por lo que no causan la finalización abrupta de la ejecución.

### **Antecedentes del Grupo de Trabajo**

Los antecedentes del grupo de trabajo parten de la investigación acerca de procesamiento de imágenes sobre diferentes arquitecturas paralelas [4-6] y de la investigación de la Tolerancia a fallos en sistemas multicore [8-9].

Las líneas de investigación se enmarcan fundamentalmente en la búsqueda de técnicas híbridas de procesamiento paralelo [10], sobre diferentes arquitecturas, que optimicen las prestaciones en sistemas de adquisición y procesamiento de imágenes médicas. En particular, resulta de interés su incorporación y aplicación en el Hospital “El Cruce” de Florencio Varela. Adicionalmente, se pretenden encausar temas de estudios de posgrado actuales y futuros de docentes y profesionales surgidos de la UNAJ dentro de estas líneas de investigación. El director del proyecto participa en proyectos de investigación desde el año 2005 en la Facultad de Ingeniería de la UNLP como

colaborador, y a partir de 2010 como investigador. Además dirige proyectos de investigación desde el año 2015 en la UNAJ. Actualmente, el grupo de investigación enfoca su trabajo en los temas: Procesamiento Digital de Imágenes Médicas sobre plataformas FPGA, Procesamiento de Imágenes en arquitecturas FPGA multiprocesador y Tolerancia a Fallos en Sistemas de Cómputo de Altas Prestaciones, pertenecientes a estudios de Maestrías y Doctorados que se realizan en el marco de acuerdos de colaboración entre el Laboratorio CeTAD, el Instituto III-LIDI y la UNAJ, a través de los respectivos proyectos de investigación.

Como resultado del trabajo realizado recientemente en la UNAJ se han publicado artículos en diferentes congresos nacionales e internacionales de Ingeniería Informática y Ciencias de la Computación [3-8]

## **2. Líneas de Investigación y Desarrollo**

El grupo de investigación que se ha constituido recientemente en la UNAJ es multidisciplinario, y sus miembros cuentan con experiencia en sistemas multiprocesador, sistemas embebidos, procesamiento de imágenes y tolerancia a fallos.

### **Temas de Estudio e Investigación**

- Implementación de un sistema multiprocesador en Dispositivos Lógicos Programables (FPGAs).
- Análisis y determinación de la eficiencia obtenida en el procesamiento de imágenes mediante la aplicación de técnicas de concurrencia y paralelismo.

- Implementación de una metodología distribuida basada en software para proveer tolerancia a fallos transitorios, específicamente en entornos de clusters de multicores donde se ejecutan aplicaciones paralelas científicas de paso de mensajes.

### 3. Resultados Obtenidos/Esperados

#### Investigación experimental

Hasta el momento se han obtenido los siguientes resultados en relación a los objetivos principales:

- En cuando al procesamiento de imágenes sobre dispositivos FPGAs, se han implementados varios algoritmos en sistemas multicore con buenos resultados en el procesamiento paralelo. En donde uno de cores se encarga de recibir y distribuir los datos para el posterior procesamiento.
- En cuanto a la tolerancia a fallos transitorios, se ha implementado una estrategia de detección basada en el monitoreo de las comunicaciones, y se ha diseñado un mecanismo de recuperación automático basado en el almacenamiento de un conjunto de *checkpoints* distribuidos de capa de sistema.

Para el año en curso, se esperan alcanzar importantes resultados en el área de concurrencia y cómputo paralelo, posibilitados por la utilización de sistemas basados en FPGAs. Con esta arquitectura, constituida por varios procesadores implementados en una misma FPGA, se espera lograr mayor eficiencia mediante la combinación de técnicas de paralelismo y la concurrencia lograda en VHDL. Para las pruebas se

emplearán algoritmos de procesamiento de imágenes basados en operadores de ventana, debido a que una parte del procesamiento es común para todos, por lo que puede implementarse de forma concurrente mediante un co-procesador; mientras que el resto puede paralelizarse entre todos los procesadores. La implementación de diferentes algoritmos permitirá medir la performance en la ejecución alcanzada con el sistema multiprocesador, y sacar conclusiones a partir de las características de cada algoritmo.

En cuanto a la tolerancia a fallos, se ha propuesto una metodología distribuida basada en replicación de software, diseñada específicamente para aplicaciones paralelas científicas de paso de mensajes, capaz de protegerlas de fallos transitorios que producirían ejecuciones incorrectas [9]. Bajo la premisa de que, en este tipo de aplicaciones, la mayor parte de los datos relevantes para el resultado son transmitidos entre procesos, la estrategia de detección se basa en validar los contenidos de los mensajes que se van a enviar y comparar los resultados finales, obteniendo un compromiso entre un alto nivel de cobertura frente a fallos y la introducción de un bajo *overhead* temporal y sobrecarga de operaciones, debido a que no se realiza trabajo para detectar fallos que no afectan a los resultados. Por otra parte, previene la propagación de la corrupción de datos a otros procesos, manteniéndola aislada en el contexto del proceso que ha resultado directamente afectado. De esta manera, se mejora la confiabilidad del sistema y se disminuye el tiempo luego del cual se puede relanzar la aplicación, lo cual es especialmente útil en ejecuciones prolongadas que conllevan un alto coste. Por otra parte, aprovecha el beneficio del

sistema la redundancia intrínseca de hardware de las arquitecturas multicore.

Para tolerar completamente los fallos transitorios, la propuesta se basa en restaurar el sistema a un estado consistente previo a su ocurrencia. Para ello, se debe integrar la estrategia de detección con mecanismos de protección para fallos permanentes. Se están estudiando las opciones de incorporar recuperación basada en múltiples *checkpoints* coordinados en capa de sistema (que proporcionan cobertura en el caso de que un *checkpoint* resulte afectado por un fallo, haciendo imposible la recuperación), o la utilización de un único *checkpoint* no coordinado de capa de aplicación, que puede ser verificado para asegurar que no contiene errores.

El uso de estas estrategias posibilitaría prescindir de la utilización de redundancia triple con votación para detectar y recuperar de fallos transitorios. Además, como estos fallos no requieren reconfiguración del sistema, la recuperación puede realizarse mediante re-ejecución en el mismo core en el que ocurrió el fallo.

#### 4. Formación de Recursos Humanos

Dentro de la temática de la línea de I+D, todos los miembros del proyecto participan en el dictado de asignaturas de la carrera de Ingeniería Informática de la UNAJ.

En este proyecto existe cooperación a nivel nacional. Hay dos investigadores realizando Doctorados y dos realizando Maestrías en temas relacionados con simulación de sistemas multiprocesador, sistemas embebidos, software embebido, sistemas multicore y tolerancia a fallos en HPC.

Adicionalmente, se cuenta con la colaboración de estudiantes avanzados.

#### 5. Bibliografía

- [1] O. Mencer, K. Tsoi, S. Cramer, T. Todman, W. Luk, Ming Wong and P. Leong, "CUBE: a 512-FPGA Cluster", Dept. of Computing, Imperial College London, Dept. of Computer Science and Engineering The Chinese University of Hong Kong. (2009)
- [2] Keith Underwood, "FPGAs vs. CPUs: Trends in Peak Floating-Point Performance", Sandia National Laboratories. (2011)
- [3] D. Encinas, E. Kunysz, Szymanowski, Morales M., "Performance de arquitecturas multiprocesador: técnicas de simulación y plataformas reconfigurables", Instituto de Ingeniería, UNAJ, (2014)
- [4] J. Osio, J. Salvatore, E. Kunysz, V. Guarepi, M. Morales, "Análisis de Eficiencia en Arquitecturas Multiprocesador para Aplicaciones de Transmisión y Procesamiento de Datos", UNER, Ciudad de Concordia, WICC 2016.
- [5] J. Osio, D. Montezanti, M. Morales, "Análisis de Eficiencia en Sistemas Paralelos", Ushuaia, Tierra del Fuego, WICC 2014
- [6] E. Kunysz, J. Rapallini, J. Osio, "Sistema de cómputo reconfigurable de alta performance (Proyecto HPRC)", 3ras Jornadas ITE - 2015 -Facultad de Ingeniería – UNLP
- [7] J. Osio, J. Salvatore, E. Kunysz, D. Montezanti, D. Alonso, V. Guarepi, M. Morales, "Análisis de Eficiencia en Arquitecturas Multiprocesador para Aplicaciones de Transmisión y Procesamiento de Datos", ITBA, CABA, WICC 2017
- [8] D. Montezanti, A. De Giusti, M. Naiouf, J. Villamayor, D. Rexachs, E. Luque, "A Methodology for Soft Errors Detection and Automatic Recovery", in Proceedings of the 15th International Conference on High Performance Computing & Simulation (HPCS). ISBN: 978-1-5386-3250-5/17. IEEE, 2017, pp. 434

- [9] D. Montezanti, E. Frati, D. Rexachs, E. Luque, M. Naiouf, and A. De Giusti, "SMCV: a methodology for detecting transient faults in multicore clusters," *CLEI Electronic Journal*, vol. 15, no. 3, pp. 5–5, 2012.
- [10] Grama A, Gupta A, Karypis G, Kumar V. "Introduction to parallel computing". Pearson Addison Wesley, 2003.
- [11] F. Cappello, A. Geist, W. Gropp, S. Kale, B. Kramer, and M. Snir, "Toward exascale resilience: 2014 update," *Supercomputing frontiers and innovations*, vol. 1, no. 1, pp. 5–28, 2014.