

Arquitecturas Multiprocesador en Computación de Alto Desempeño: Software, Métricas, Modelos y Aplicaciones

De Giusti Armando ^{1,2}, Tinetti Fernando ^{1,3}, Naiouf Marcelo¹, Chichizola Franco¹, De Giusti Laura^{1,3}, Villagarcía Horacio^{1,3}, Montezanti Diego¹, Encinas Diego¹, Pousa Adrián¹, Rodriguez Ismael¹, Rodriguez Eguren Sebastián¹, Iglesias Luciano¹, Paniego Juan Manuel¹, Pi Puig Martín¹, Dell'Oso Matías¹, Mendez Mariano^{1,4}

¹Instituto de Investigación en Informática LIDI (III-LIDI)
Facultad de Informática – Universidad Nacional de La Plata
50 y 115, La Plata, Buenos Aires

Comisión de Investigaciones Científicas de la Pcia. de Buenos Aires (CIC)
526 e/ 10 y 11 La Plata Buenos Aires

² CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas

³CIC - Comisión de Investigaciones Científicas de la Provincia de Buenos Aires

⁴ Facultad de Ingeniería - Universidad de Buenos Aires

{degiusti, fernando, mnaiouf, francoch, ldgiusti, hvw, dmontezanti, dencinas, apousa, ismael, seguren, li, mmendez, jmpaniego, mpipuig, mdelloso}@lidi.info.unlp.edu.ar, marianomendez@gmail.com

Con la colaboración en la dirección de Tesis de Posgrado de la Universidad Autónoma de Barcelona (España) y la Universidad Complutense de Madrid (España).

RESUMEN

Caracterizar las arquitecturas multiprocesador distribuidas enfocadas especialmente a cluster y cloud computing, con énfasis en las que utilizan procesadores de múltiples núcleos (multicores, GPUs y Xeon Phi), con el objetivo de modelizarlas, estudiar su escalabilidad, analizar y predecir performance de aplicaciones paralelas, estudiar el consumo energético y su impacto en la performance así como desarrollar esquemas para detección y tolerancia a fallas en las mismas.

Profundizar el estudio de arquitecturas basadas en GPUs y su comparación con clusters de multicores, así como el empleo

computadoras de alta performance. Iniciar investigación experimental con arquitecturas paralelas basadas en FPGAs. En particular estudiar performance en Clusters “híbridos”. Analizar y desarrollar software de base para clusters, tratando de optimizar el rendimiento.

Investigar arquitecturas multicore asimétricas, desarrollar algoritmos de planificación en el software de sistema operativo para permitir la optimización del

rendimiento y consumo energético en aplicaciones de propósito general.

Estudiar clases de aplicaciones inteligentes en tiempo real, en particular el trabajo colaborativo de robots conectados a un cloud y procesamiento de Big Data.

Es de hacer notar que este proyecto se coordina con otros proyectos en curso en el III-LIDI, relacionados con Computación de Alto Desempeño, Algoritmos Paralelos, Sistemas Distribuidos y Sistemas de Tiempo Real.

Palabras claves: *Sistemas Paralelos – Multicore – GPU – FPGAs - Cluster y Cloud Computing - Cluster híbridos - Performance y eficiencia energética - Tolerancia a fallas –*

Cloud Robotics de Modelos de Hardware.

CONTEXTO

Esta línea de Investigación está dentro del proyecto 11/F018: “Arquitecturas multiprocesador en HPC: Software de base, Métricas y Aplicaciones” acreditado por el Ministerio de Educación y de proyectos específicos apoyados por organismos nacionales e internacionales. También del proyecto “Computación de Alto Desempeño,

Minería de Datos y Aplicaciones de interés social en la Provincia de Buenos Aires” financiado por la CIC PBA dentro de la convocatoria a Proyectos de Innovación y Transferencia en Areas Prioritarias de la Pcia. de Buenos Aires (PIT-AP-BA).

El III-LIDI forma parte del Sistema Nacional de Cómputo de Alto Desempeño (SNCAD) del MINCYT y en esta línea de I/D hay cooperación con varias Universidades de Argentina, de América Latina y Europa en proyectos con financiación nacional e internacional.

En la Facultad de Informática de la UNLP (a partir del equipo del proyecto) se han incorporado asignaturas optativas en la currícula de grado de las carreras de Licenciatura en Informática, Licenciatura en Sistemas e Ingeniería en Computación relacionadas con Cloud Computing, Programación sobre GPGPUs y procesamiento de Big Data. Además, la Facultad aprobó y financia el proyecto “Transformación de Algoritmos para Nuevas Arquitecturas Multiprocesador”.

Se participa en iniciativas como el Programa IberoTIC de intercambio de Profesores y Alumnos de Doctorado en el área de Informática, así como el desarrollo de la Maestría y Especialización en Computación de Altas Prestaciones, acreditadas por CONEAU.

Por último, se tiene financiamiento de Telefónica de Argentina en Becas de grado y posgrado y se ha tenido el apoyo de diferentes empresas (IBM, Microsoft, Telecom, INTEL, AMAZON AWS) en las temáticas de Cloud Computing y Big Data.

1. INTRODUCCIÓN

La investigación en Paralelismo (a partir de arquitecturas multiprocesador distribuidas o concentradas en supercomputadoras) es una de las líneas de mayor desarrollo en la Ciencia Informática actual [GRA03]. La utilización de clusters, multiclusters, grids y clouds, soportadas por redes de diferentes características y topologías se ha generalizado, tanto para el desarrollo de algoritmos paralelos orientados a HPC como

para el manejo de aplicaciones distribuidas y/o servicios WEB concurrentes [GRA03][MCC12].

El cambio tecnológico, fundamentalmente a partir de los procesadores multicore, ha impuesto la necesidad de investigar en paradigmas "híbridos", en los cuales coexisten esquemas de memoria compartida con mensajes [LEI12]. Asimismo la utilización de aceleradores (GPU, FPGA, Xeon Phi) presenta una alternativa para alcanzar un alto speedup en determinadas clases de aplicaciones [KIN09][SIN12]. Debe notarse que el modelo de programación orientado a estas arquitecturas cambia sensiblemente y la optimización de código paralelo requiere nuevos recursos.

Este cambio también se refleja en la aparición de los procesadores multicore asimétricos (AMPs) que integran en un mismo chip diversos tipos de cores con distintas características (frecuencia, microarquitectura o consumo), pero con el mismo repertorio de instrucciones. Los cuales ofrecen un mayor rendimiento por watt y unidad de área que los multicores simétricos [SAE10][ANN12].

Por su lado, las arquitecturas tipo "Cloud" se presentan como una evolución natural del concepto de *Clusters* y *Grids*, integrando grandes conjuntos de recursos virtuales (hardware, plataformas de desarrollo y/o servicios), fácilmente accesibles y utilizables por usuarios distribuidos, vía WEB [VAQ09]. Estos recursos pueden ser dinámicamente reconfigurados para adaptarse a una carga variable, permitiendo optimizar su uso. Al enfocarnos en Cloud Computing aparecen problemas clásicos de la Ciencia Informática, extendidos para este nuevo modelo de arquitectura: planificación, virtualización, asignación dinámica de recursos, migración de datos y procesos [ARD09][VAZ09]. En el proyecto se ha abierto una línea específica dedicada a Cloud Computing y su uso para aplicaciones de *Big Data* o en aquellas que requieren centralizar el accionar de "robots" distribuidos en tiempo real (Cloud Robotics) [MAY13].

Los avances tecnológicos y la búsqueda continua de mayor eficiencia hacen necesario investigar diferentes componentes de las arquitecturas. En particular, las aplicaciones científicas con un uso intensivo de datos utilizan software de E/S paralelo para acceder a archivos. Contar con herramientas que permitan predecir el comportamiento de este tipo de aplicaciones en HPC es de gran utilidad para los desarrolladores. Para esto se puede utilizar el modelado basado en agentes y simulación (Agent-Based Modeling and Simulation, ABMS) [MAC06].

La importancia que ha adquirido el uso del paralelismo para aplicaciones científicas de gran duración, requiere la necesidad de estudiar los problemas de detección y tolerancia a fallos en arquitecturas paralelas, debido al alto costo de relanzar la ejecución desde el comienzo en caso de resultados incorrectos. Esto se debe lograr tratando de minimizar el overhead temporal y de aprovechar la redundancia de recursos de hardware que caracteriza a estas arquitecturas [GOL09][FIA11]. El manejo de fallos es una preocupación creciente en HPC, se esperan crecimientos en las tasas de errores, mayores latencias de detección y elevadas cantidades de fallos silenciosos con capacidad de corromper los resultados de las aplicaciones. Por ello, se han desarrollado estrategias de detección y recuperación de fallos transitorios basadas en replicación de software, detectando divergencias en las comunicaciones entre réplicas para evitar que la corrupción se propague a otros procesos, restringiendo así la latencia de detección [MON14][MON15]. De esta forma se permiten obtener ejecuciones fiables con resultados correctos o conducir al sistema a una parada segura. La recuperación puede lograrse mediante múltiples checkpoints de nivel de sistema o de un único checkpoint de capa de aplicación.

Desde otro punto de vista, interesan también los problemas que significan integración de redes de sensores con modelos del mundo real (por ej. modelos meteorológicos, hídricos o de terreno) para prevención de emergencias [SIN06][GAU16]. En esta línea,

el eje del proyecto sigue estando en la problemática del paralelismo combinado con sistemas de tiempo real, pudiendo contribuir a proyectos multidisciplinarios, en particular por temas de emergencias hídricas, exploración de recursos naturales y temas de atención sanitaria y evacuación de edificios en situaciones de emergencia.

2. LINEAS DE INVESTIGACIÓN Y DESARROLLO

- Estudio y caracterización de arquitecturas paralelas: clusters, grids, clouds, aceleradores (GPU, FPGA, Xeon Phi) e híbridos.
- Desarrollo de algoritmos de planificación de procesos orientado a procesadores asimétricos para optimizar el rendimiento general. Análisis en los diferentes niveles: sistema operativo, compiladores, técnicas de programación.
- Desarrollo de aplicaciones concretas (numéricas y no numéricas) sobre diferentes máquinas paralelas utilizando técnicas de optimización adecuadas a cada arquitectura.
- Desarrollo de técnicas de tolerancia a fallas en sistemas paralelos y distribuidos, lo cual supone una mejora en el aprovechamiento de la redundancia de recursos que no resultan eficientemente utilizadas en dichas arquitecturas.
- Desarrollo de herramientas para la transformación de código heredado, buscando su optimización sobre arquitecturas paralelas.
- Integración de métricas de rendimiento computacional y energético. Predicción de performance de aplicaciones paralelas.
- Cloud Computing. Software de base. Desarrollo de aplicaciones de HPC (principalmente de big data).
- Sistemas inteligentes distribuidos de tiempo real aprovechando la potencia de cómputo del Cloud (Cloud Robotics).
- Utilización de ABMS para desarrollar un modelo de Entrada/Salida en HPC que permita predecir cómo cambios realizados en los diferentes componentes del mismo

afectan a la funcionalidad y el rendimiento del sistema.

3. RESULTADOS OBTENIDOS/ESPERADOS

- Estudiar modelos complejos, que integren redes de sensores en tiempo real y cómputo paralelo. Estrategias de predicción de catástrofes (inundaciones, incendios por ejemplo) se basan en estos modelos con alta capacidad de procesamiento y monitoreo de señales en tiempo real [GAU16].
- Se han desarrollado diferentes aplicaciones adaptadas para diferentes arquitecturas "híbridas" (que combinan clusters, multicores y aceleradores), y analizado/comparado el rendimiento obtenido [RUC16][POU15][MON16].
- Se está trabajando en técnicas de recuperación a partir de múltiples checkpoints de nivel de sistema, que sirvan para garantizar la correcta finalización de aplicaciones científicas sobre sistemas de HPC, que resultan afectadas por la ocurrencia de fallas transitorias externas y aleatorias, integrando esta solución con las herramientas de detección desarrolladas previamente [MON14][MON15].
- Se han desarrollado nuevos planificadores de tareas para multicores asimétricos sobre diferentes sistemas operativos con el objetivo de maximizar el rendimiento y minimizar el consumo de energía [SAE15].
- Desarrollo de un modelo de la Entrada/Salida en HPC por medio de ABMS (Agent-Based Modeling and Simulation) que permita predecir cómo cambios realizados en los diferentes componentes del modelo afectan a la funcionalidad y el rendimiento del sistema [ENC15].
- Desarrollo de aplicaciones vinculadas con "Big Data", especialmente para resolver en Cloud Computing [BAS15]. (en relación con los otros proyectos del III- LIDI).
- Optimización de algoritmos paralelos para controlar el comportamiento de múltiples robots que trabajan colaborativamente, considerando la distribución de su capacidad de procesamiento "local" y la coordinación con la potencia de cómputo y capacidad de almacenamiento (datos y conocimiento) de un Cloud.
- Actualización y modernización de código fuente de Sistemas Heredados (Legacy Systems) de Cómputo Científico a través de la aplicación de un proceso de desarrollo iterativo e incremental dirigido por transformaciones de código fuente, apoyado fuertemente en las herramientas de desarrollo. Dichas transformaciones se implementan para ser aplicadas automáticamente en un entorno integrado de desarrollo [MEN14][TIN15].
- Trabajar en la implementación de transformaciones que ayuden a la paralelización del código fuente, así como también en herramientas de análisis estático de Código fuente [TIN13][TIN15].
- Adaptar las técnicas de scheduling y mapeo de procesos a procesadores de acuerdo a los objetivos actuales (en particular los relacionados con el consumo), considerando la migración dinámica de datos y procesos en función de rendimiento y consumo [GRA03][DEG10]. Se debe incluir la utilización de los registros de hardware de los procesadores para la toma de diferentes decisiones en tiempo de ejecución.

4. FORMACIÓN DE RECURSOS HUMANOS

En cooperación con Universidades iberoamericanas se ha implementado la Maestría en Cómputo de Altas Prestaciones y se continúa dictando la Especialización en Cómputo de altas Prestaciones y Tecnología GRID. Asimismo se tiene un importante número de doctorandos (del país y del exterior) realizando el Doctorado en Ciencias Informáticas de la UNLP.

Se han organizado las IV Jornadas de Cloud Computing & Big Data (JCC&BD), y se realizarán en junio de 2017 las V JCC&BD, integrando una Escuela con cursos de Posgrado relacionados con la temática de las líneas de investigación presentadas.

Existe cooperación a nivel nacional e internacional y dentro de la temática del proyecto se espera alcanzar 5 Tesis de Doctorado y 6 Tesis de Maestría en los próximos 3 años, en el país. Al menos tener 3 Doctorandos en el exterior/mixtos en el mismo período.

En 2016 se aprobaron 2 Tesis Doctorales [RUC16][MEN16] y otra está entregada para su evaluación. También se aprobaron 3 trabajos de Especialista, 1 Tesis de Magister y 2 Tesinas de grado.

5. BIBLIOGRAFÍA

- [ANN12] Annamalai A., Rodrigues R., Koren I., Kundu S., "Dynamic Thread Scheduling in Asymmetric Multicores to Maximize Performance-per-Watt," 2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum, pp. 964-971, 2012 IEEE 26th International Parallel and Distributed Processing Symposium Workshops & PhD Forum, 2012.
- [ARD09] Ardisson L., Goy A., Petrone G., Segnan M. "From Service Clouds to User-centric Personal Clouds". 2009 IEEE Second International Conference on Cloud Computing.
- [BAS15] Basgall M. J., Aquino G., Lanzarini L., Naiouf M. "Un Enfoque Dinámico para la Detección de Relaciones entre Tópicos en Textos provenientes de Redes Sociales". III Jornadas de Cloud Computing y Big Data. Facultad de Informática. UNLP. 2015.
- [DEG10] De Giusti L., Chichizola F., Naiouf M., De Giusti A.E., Luque E. "Automatic Mapping Tasks to Cores - Evaluating AMTHA Algorithm in Multicore Architectures". IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 2, No 1, March 2010. ISSN (Online): 1694-0784. ISSN (Print): 1694-0814. Págs. 1-6.
- [ENC15] D. Encinas et al., "Modeling I/O System in HPC: An ABMS Approach". The Seventh International Conference on Advances in System Simulation (SIMUL), ISBN: 978-1-61208-442-8, 2015
- [FIA11] Fialho L. "Fault Tolerance configuration for uncoordinated checkpoints". Ph.D. Thesis, Universidad Autónoma de Barcelona, Barcelona, España, Julio 2011.
- [GAU16] Adriana Gaudiani, Emilio Luque, Pablo García, Mariano Re, Marcelo Naiouf, Armando De Giusti. "How a Computational Method Can Help to Improve the Quality of River Flood Prediction by Simulation". Advances and New Trends in Environmental and Energy Informatics (part V). ISBN 978-3-319-23455-7. Pp337-351. 2016.
- [GOL09] Golander A., Weiss S., Ronen R. "Synchronizing Redundant Cores in a Dynamic DMR Multicore Architecture". IEEE Transactions on Circuits and Systems II: Express Briefs Volume 56, Issue 6, 474-478. 2009.
- [GRA03] Grama A, Gupta A, Karypis G, Kumar V. "Introduction to parallel computing". Second Edition. Pearson Addison Wesley, 2003.
- [KIN09] Kindratenko, V.V.; Enos, J.J.; Guochun Shi; Showerman, M.T.; Arnold, G.W.; Stone, J.E.; Phillips, J.C.; Wen-Mei Hwu, "GPU clusters for high-performance computing," Cluster Computing and Workshops, 2009. CLUSTER '09. IEEE International Conference on, vol., no., pp.1,8, Aug. 31 2009-Sept. 4 2009
- [LEI12] Leibovich F., Chichizola F., De Giusti L., Naiouf M., Tirado Fernández F., De Giusti A. "Programación híbrida en clusters de multicore. Análisis del impacto de la jerarquía de memoria". XII Workshop de Procesamiento Distribuido y Paralelo. CACIC2012. ISBN: 978987-1648-34-4. Pág. 306-315. 2012.
- [MAC06] C. Macal, M. North, "Tutorial on agent-based modeling and simulation part 2: how to model with agents", in: Proceedings of the Winter Simulation Conference, 2006.
- [MAY13] Mayer-Schönberger V., Cukier K., "Big Data: A revolution that will transform

- how we live, work and think” Houghton Mifflin Harcourt 2013.
- [MCC12] McCool M., Robison A., Reinders J. “Structured Parallel Programming” Elsevier-Morgan Kaufmann, 2012.
- [MEN14] Méndez M., Tinetti F. “Integrating Software Metrics for Fortran Legacy into an IDE”. XI Workshop Ingeniería de Software – CACIC 2014. ISBN 978-987-3806-05-6. Pág. 771-780. San Justo, Buenos Aires, Argentina. Octubre 2014.
- [MEN16] Méndez M. “Aplicaciones de Cómputo Científico: Mantenimiento del Software Heredado”. Tesis de Doctorado en Ciencias Informáticas (Facultad de Informática – UNLP). 2016
- [MON14] Diego Montezanti, Enzo Rucci, Dolores Rexachs, Emilio Luque, Marcelo Naiouf, Armando De Giusti. “A tool for detecting transient faults in execution of parallel scientific applications on multicore clusters”. *Journal of Computer Science & Technology (JCS&T)*, Vol. 14, Nro. 1 (Abril 2014). Páginas: 32 a 38. ISSN: 1666-6038.
- [MON15] Diego Montezanti, Dolores Rexachs, Enzo Rucci, Emilio Luque, Marcelo Naiouf, Armando De Giusti. “Caracterización de una estrategia de detección de fallos transitorios en HPC”. XXI CACIC 2015 (Congreso de Ciencias de la Computación) - XV Workshop de Procesamiento Paralelo y Distribuido (WPDP). ISBN: 978-987-3724-37-4.
- [MON16] E. Montes de Oca, L. C. De Giusti, F. Chichizola, A. E. De Giusti, M. Naiouf. “Análisis de uso de un algoritmo de balanceo de carga estático en un Cluster Multi-GPU Heterogéneo”. *Proceedings del XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016) – Workshop de Procesamiento Distribuido y Paralelo*. Octubre 2016. Pp 159-168.
- [POU15] Adrian Pousa, Victoria Sanz, Armando De Giusti. Comparación de rendimiento de algoritmos de cómputo intensivo y de acceso intensivo a memoria sobre arquitecturas multicore. Aplicación al algoritmo de criptografía AES. CACIC 2015.
- [RUC16] Rucci, Enzo. “Evaluación de rendimiento y eficiencia energética en sistemas heterogéneos para bioinformática”. Tesis de Doctorado en Ciencias Informáticas (Facultad de Informática – UNLP). 2016.
- [SAE10] Sáenz J. C. "Planificación de Procesos en Sistemas Multicore Asimétricos". Ph.D. Thesis, Universidad Complutense de Madrid, España, 2010.
- [SAE15] Juan Carlos Saez, Adrian Pousa, Daniel Chaver, Fernando Castro, Manuel Prieto Matias. ACFS: A Completely Fair Scheduler for Asymmetric Single-ISA Multicore Systems. ACM SAC 2015 (The 30TH ACM/SIGAPP Symposium on applied computing). 2015.
- [SIN06] Singh B. D. "Influence of Digital Elevation Models Derived from Remote Sensing on SpatioTemporal Modelling of Hydrologic and Erosion Processes". Ph.D. Thesis. pp 370. Cranfield University Siloe England. 2006.
- [SIN12] Sinha, R.; Prakash, A.; Patel, H.D., "Parallel simulation of mixed-abstraction SystemC models on GPUs and multicore CPUs," *Design Automation Conference (ASP-DAC)*, 2012 17th Asia and South Pacific, pp.455,460. 2012.
- [TIN13] F. G. Tinetti, M. Méndez, A. De Giusti. “Restructuring Fortran legacy applications for parallel computing in multiprocessors”. *The Journal of Supercomputing*, May 2013, Volume 64, Issue 2, ISSN 0920-8542, DOI 10.1007/s11227-012-0863-x, pp 638-659.
- [TIN15] M. Mendez, F. G. Tinetti. “Integrating Software Metrics for Fortran Legacy into an IDE” *Computer Science & Technology Series – XX Argentine Congress of Computer Science, Selected Papers.*, ISBN: 978-987-1985-71-5, Red UNCI, págs. 126-134, 2015.
- [VAQ09] Vaquero L.M. et al. "A Break in the Clouds: Towards a Cloud Definition". *ACM SIGCOMM Computer Communication Review*, vol. 39, num. 1, páginas 50-55, ISSN 0146-4833. Enero 2009.
- [VAZ09] Vázquez Blanco C., Huedo E., Montero R. S., Llorente I. M. "Elastic Management of Cluster-based Services in the Cloud". *Proceedings* pp. 19-24, *ACM Digital Library* 2009. ISBN 978-1-60558-564-2.