

# Evolución y Tendencias en Sistemas Paralelos para HPC

Marcelo Naiouf<sup>(1)</sup>, Franco Chichizola<sup>(1)</sup>, Laura De Giusti<sup>(1)</sup>, Enzo Rucci<sup>(1)(2)</sup>, Adrián Pousa<sup>(1)</sup>, Ismael Rodríguez<sup>(1)</sup>, Sebastián Rodríguez Eguren<sup>(1)</sup>, Erica Montes de Oca<sup>(1)</sup>, María José Basgall<sup>(1)(2)</sup>, Juan Manuel Paniago<sup>(1)</sup>, Martín Pi Puig<sup>(1)</sup>, Emmanuel Frati<sup>(3)</sup>, Javier Balladini<sup>(4)</sup>, Armando De Giusti<sup>(1)(2)</sup>

<sup>1</sup>Instituto de Investigación en Informática LIDI (III-LIDI)

Facultad de Informática – Universidad Nacional de La Plata

<sup>2</sup>CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas

<sup>3</sup>Universidad Nacional de Chilecito

<sup>4</sup>Universidad Nacional del Comahue

{mnaiouf,francoch,ldgiusti,erucci,apousa,ismael,seguren,emontesdeoca,mjbasgall,jmpaniego,mpipuig,fefrati,degusti}@lidi.info.unlp.edu.ar; [javier.balladini@gmail.com](mailto:javier.balladini@gmail.com)

## Resumen

El eje de esta línea de I/D lo constituye el estudio de tendencias actuales en las áreas de arquitecturas y algoritmos paralelos. Incluye como temas centrales:

- Arquitecturas Many-core (GPU, procesadores MIC), híbridas (diferentes combinaciones de multicores y GPUs), heterogéneas, asimétricas y FPGAs.
- Cloud Computing para HPC (especialmente para aplicaciones de Big Data) y sistemas distribuidos de tiempo real (Cloud Robotics).
- Desarrollo y evaluación de algoritmos paralelos sobre nuevas arquitecturas y su evaluación de rendimiento energético y computacional.
- Empleo de contadores de hardware, en particular en toma de decisiones en tiempo de ejecución.

**Palabras clave:** *Sistemas Paralelos. Clusters híbridos. Arquitecturas asimétricas. GPU, MIC, FPGA. Cloud Computing y cloud robotics. Performance y eficiencia energética. Contadores de hardware.*

## Contexto

Se presenta una línea de Investigación que es parte de los Proyectos 11/F018 “Arquitecturas multiprocesador en HPC: Software de Base, Métricas y Aplicaciones” y 11/F017 “Cómputo Paralelo de Altas Prestaciones. Fundamentos y Evaluación de Rendimiento en HPC. Aplicaciones a Sistemas

Inteligentes, Simulación y Tratamiento de Imágenes” del III-LIDI acreditados por el Ministerio de Educación y de proyectos específicos apoyados por organismos nacionales e internacionales. Y del proyecto “Cloud Computing y Big Data. Aplicaciones a HPC y Cloud Robotics” financiado por la Facultad de Informática de la UNLP.

En el tema hay cooperación con varias Universidades de Argentina y se está trabajando con Universidades de América Latina y Europa en proyectos financiados por CyTED, AECID y la OEI (Organización de Estados Iberoamericanos).

Por otra parte, 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) en la temática de Cloud Computing.

Se participa en iniciativas como el Programa IberoTIC de intercambio de Profesores y Alumnos de Doctorado en el área de Informática.

Asimismo el III-LIDI forma parte del Sistema Nacional de Cómputo de Alto Desempeño (SNCAD) del MINCYT.

## Introducción

Una de las áreas de creciente interés lo constituye el cómputo de altas prestaciones, en el cual el rendimiento está relacionado con dos aspectos: por un lado las arquitecturas de soporte y por el otro los algoritmos que hacen uso de las mismas.

A la aparición de arquitecturas many-core (como las GPU o los procesadores MIC), se

ha sumado el uso de FPGAs debido a su potencia de cómputo y rendimiento energético. Su combinación en sistemas HPC da lugar a plataformas híbridas con diferentes características [RUC16].

Lógicamente, esto trae aparejado una revisión de los conceptos del diseño de algoritmos paralelos (incluyendo los lenguajes mismos de programación y el software de base), así como la evaluación de las soluciones que éstos implementan. También resulta necesario investigar las estrategias de distribución de datos y procesos a fin de optimizar la performance.

Además de las evaluaciones clásicas de rendimiento prestacional como el speedup y la eficiencia, otros aspectos comienzan a ser de interés, tales como el estudio del consumo y la eficiencia energética de tales sistemas paralelos [CAS12].

Los avances en las tecnologías de virtualización y cómputo distribuido han dado origen al paradigma de Cloud Computing, que se presenta como una alternativa a los tradicionales sistemas de Clusters y Multicluster para ambientes de HPC [ROD07][BER08]. A su vez, este concepto se puede ampliar a sistemas distribuidos de tiempo real, en particular sistemas inteligentes como son los robots que pueden trabajar en paralelo utilizando su propia capacidad de procesamiento y al mismo tiempo conectándose con la potencia de un servidor en la nube (Cloud Robotics) [LOR13][GUO12][KEH15].

Una herramienta que ha comenzado a ser utilizada para la monitorización y evaluación de soluciones paralelas en tiempo de ejecución son los contadores de hardware, que (en sus diferentes variantes) permiten detectar fallas de concurrencia o estimar problemas de consumo instantáneo o decidir una migración de datos o procesos [BOR05] [SPR02].

En esta línea de I/D se trabaja sobre estos aspectos que marcan tendencias en el área.

### **GPUs y Cluster de GPUs**

La combinación de GPUs con otras plataformas paralelas como clusters y multicores, brindan un vasto conjunto de

posibilidades de investigación en arquitecturas híbridas, a partir de diferentes combinaciones a saber:

- Cluster de máquinas uncore cada una con placa GPU, lo que permite combinar herramientas de programación paralela como MPI/CUDA.

- Máquinas multicore con más de una GPU, que combinan herramientas de programación paralela como OpenMP/CUDA o Pthread/CUDA.

- Cluster de máquinas multicore cada una con placa de GPU, lo que permite combinar OpenMP/MPI/CUDA o Pthread/MPI/CUDA.

Los desafíos que se plantean son múltiples, sobre todo en lo referido a distribución de datos y procesos en tales arquitecturas híbridas a fin de optimizar el rendimiento de las soluciones.

### **MIC**

En forma reciente Intel brinda una alternativa a partir de los procesadores MIC (Many Integrated Core Architecture), permitiendo utilizar métodos y herramientas estándar de programación con altas prestaciones.

MIC combina varios cores Intel en un solo chip, y la programación puede realizarse usando código standard C, C++ y FORTRAN.

El mismo programa fuente escrito para MIC puede ser compilado y corrido en un Xeon, por lo que los modelos de programación familiares remueven barreras de entrenamiento y permiten enfocarse en el problema más que en la ingeniería del software.

Esta tendencia resulta promisoría, siendo el co-procesador Phi un ejemplo de esta arquitectura [JEF13].

### **FPGAs**

Una FPGA (Field Programmable Gate Array) es una clase de acelerador basado en circuitos integrados reconfigurables. La capacidad de adaptar sus instrucciones de acuerdo a la aplicación objetivo le permite incrementar la productividad de un sistema y mejorar el rendimiento energético para ciertos tipos de aplicaciones. En la actualidad las

principales empresas fabricantes de FPGA trabajan en nuevas herramientas de programación empleando estándares familiares para HPC. Se espera que su empleo contribuya a reducir los tradicionales tiempos y costos de programación [SEA13][XIL15].

### **Consumo energético**

Un aspecto de interés creciente en la informática actual, principalmente a partir de las plataformas con gran cantidad de procesadores, es el del consumo energético que los mismos producen.

Muchos esfuerzos están orientados a tratar el consumo como eje de I/D, como métrica de evaluación, y también a la necesidad de metodologías para medirlo.

Entre los puntos de interés pueden mencionarse:

- Caracterización energética de las instrucciones, de algoritmos complejos y de sistemas paralelos desde el punto de vista de la potencia máxima y el consumo total.
- Modelos de predicción de performance energética.
- Análisis de esquemas de distribución de procesos entre procesadores, con ajuste dinámico de la frecuencia de clock en función del consumo.
- Estudio de técnicas para reducir el consumo energético en aplicaciones de HPC de acuerdo a las arquitecturas utilizadas.
- Evaluación de eficiencia energética de diferentes algoritmos y plataformas paralelas.

### **Cloud Computing**

Cloud Computing, proporciona grandes conjuntos de recursos físicos y lógicos (como ser infraestructura, plataformas de desarrollo, almacenamiento y/o aplicaciones), fácilmente accesibles y utilizables por medio de una interfaz de administración web, con un modelo de arquitectura “virtualizada” [SHA10][XIN12]. Estos recursos son proporcionados como servicios (“as a service”) y pueden ser dinámicamente reconfigurados para adaptarse a una carga de trabajo variable (escalabilidad), logrando una mejor utilización y evitando el sobre o sub

dimensionamiento (elasticidad) [VEL09][VAQ09].

Más allá de las potenciales características y beneficios que brinda un Cloud, de por sí atractivas, es de gran interés estudiar el despliegue de entornos de ejecución para cómputo paralelo y distribuido (Clusters Virtuales), como así también realizar I/D en la portabilidad de las aplicaciones de HPC en el Cloud [DOE11][ROD11].

Los Clusters Virtuales (VC), están conformados por VMs configuradas e interconectadas virtualmente para trabajar en forma conjunta como un recurso de cómputo único e integrado. Cada una tiene asociado un S.O., recursos de almacenamiento, protocolos de comunicación, configuraciones de red y entornos de software para ejecución de algoritmos paralelos [VAZ09][HAC11].

Por otro lado, Cloud Robotics es una de las áreas más prometedoras de la investigación informática actual en la cual se cuenta con “robots” dotados de diferentes sensores y capacidades, conectados a un Cloud vía Internet. Los temas de investigación derivados son múltiples: sensores, redes de sensores e inteligencia distribuida; robótica y sistemas colaborativos de tiempo real basados en robots; aplicaciones críticas (por ej. en ciudades inteligentes o en el ámbito industrial).

### **Contadores de hardware**

Todos los procesadores actuales poseen un conjunto de registros especiales denominados contadores de hardware [SPR02]. Estos registros se pueden programar para contar el número de veces que ocurre un evento dentro del procesador durante la ejecución de una aplicación. Tales eventos pueden proveer información sobre diferentes aspectos de la ejecución de un programa (por ejemplo el número de instrucciones ejecutadas, la cantidad de fallos cache en L1, cuántas operaciones en punto flotante se ejecutaron, etc). Los procesadores actuales poseen una gran cantidad de eventos (más de 300) y la capacidad de usar hasta 11 registros simultáneamente [INT12]. El acceso a estos recursos de monitorización se puede llevar a

cabo usando diferentes herramientas en función del nivel de abstracción deseado.

Interesa plantear la utilización de contadores de hardware en dos líneas:

- Optimización de programas. La dependencia que tiene un programa paralelo sobre su arquitectura para ser altamente eficiente requiere comprender los motivos que penalizan su desempeño. Los contadores hardware permiten acceder a información precisa sobre aspectos específicos de la ejecución de los programas, ayudando al programador en la tarea de encontrar esos motivos y comparar con datos concretos los beneficios de los cambios que realiza [TIN14].
- Sintonización dinámica de aplicaciones. La información que proporciona la PMU (Unidad de Monitorización de Performance) de los procesadores puede ser utilizada para modificar el comportamiento o tomar decisiones en tiempo de ejecución. Esto permite construir algoritmos dinámicos de gran precisión, que se ajustan a los eventos que ocurren en el hardware [FRA14].

## **Líneas de Investigación, Desarrollo e Innovación**

- Arquitecturas many-core (procesadores MIC y GPU) y FPGA. Análisis de este tipo de máquinas y de técnicas para desarrollar código optimizado.
- Arquitecturas híbridas (diferentes combinaciones de clusters, multicores, manycores y FPGAs). Diseño de algoritmos paralelos sobre las mismas.
- Cloud Computing para realizar HPC. Evaluación de performance en este tipo de arquitectura. Análisis del overhead por el software de administración del Cloud.
- Sistemas inteligentes de tiempo real distribuidos (Cloud Robotics).
- Migración en vivo de VCs homogéneos y heterogéneos en las arquitecturas de Cloud Computing.
- Consumo energético en las diferentes arquitecturas paralelas, en particular en relación a los algoritmos paralelos y la configuración de la arquitectura. Modelos y

predicción de performance energética en sistemas paralelos.

- Contadores de hardware. Aplicaciones en la optimización de aplicaciones paralelas.

## **Resultados y Objetivos**

### **Investigación experimental a realizar**

- Desarrollo y evaluación de algoritmos paralelos sobre nuevas arquitecturas. Análisis de rendimiento, eficiencia energética y costo de programación.
- Empleo experimental de contadores de hardware, orientados a la detección de fallas de concurrencia [FRA15] y a la toma de decisiones sobre la frecuencia de clock de los procesadores en función del consumo.
- Análisis del overhead introducido por el sistema gestor del Cloud en un entorno de HPC para aplicaciones científicas de Big Data. Comparar el rendimiento entre Cloud y Cluster Computing.
- Analizar y comparar las técnicas de migración en vivo de VMs, con el fin de implementar migraciones de VCs homogéneos y heterogéneos en Cloud Computing.
- Realizar el desarrollo de nuevos planificadores de tareas para multicores asimétricos sobre diferentes sistemas operativos con el objetivo de maximizar el rendimiento y minimizar el consumo de energía [SAE14][SAE15].
- 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.

### **Resultados obtenidos**

- En la línea de planificación de procesos sobre multicores asimétricos se evaluaron nuevos algoritmos de planificación sobre los kernel Solaris y Linux, se analizaron las mejoras en rendimiento y consumo energético para un conjunto diverso de cargas de trabajo [SAE14][SAE15].

- Utilización combinada de cluster de multicores y cluster de GPUs [MON14][POU15].

- Se ha finalizado una tesis de doctorado focalizada en evaluación de rendimiento y eficiencia energética de diferentes plataformas heterogéneas, en la cual se desarrollaron diferentes soluciones paralelas para un problema del área bioinformática con alta demanda computacional [RUC16]. Se analizaron sus rendimientos y el consumo energético y costo de programación asociados.

### Organización de Eventos

En el año 2015 se ha organizado la III Jornada de Cloud Computing y Big Data (JCC&BD 2015) en Argentina, con participación de especialistas académicos del país y del exterior y de empresas con experiencia en Cloud Computing. En junio de 2016 se organizará la IV Jornadas de Cloud Computing y Big Data (JCC&BD 2016).

### Formación de Recursos Humanos

Dentro de la temática de la línea de I/D el último año se concluyeron: 2 tesis doctorales, 2 trabajos de Especialización y 1 tesina de grado. Al mismo tiempo se encuentran en curso 4 tesis de Doctorado en Ciencias Informáticas.

Además, se participa en el dictado de las carreras de Doctorado en Ciencias Informáticas, y Magíster y Especialización en Cómputo de Altas Prestaciones de la Facultad de Informática de la UNLP (todas acreditadas A por la CONEAU), por lo que potencialmente pueden generarse nuevas Tesis de Doctorado y Maestría y Trabajos Finales de Especialización.

Existe cooperación con grupos de otras Universidades del país y del exterior, y hay tesis de diferentes Universidades realizando su Tesis con el equipo del proyecto.

Respecto a las carreras de grado, se dictan por parte de integrantes de la línea de investigación dos materias directamente relacionadas con los temas de la misma: "Taller de Programación sobre GPUs" y

"Cloud Computing. Aplicaciones en Big Data"

### Referencias

[BER08] Bertogna, M., Grosclaude, E., Naiouf, M., De Giusti, A., Luque, E.: "Dynamic on Demand Virtual Clusters in Grids". In: 3rd Workshop on Virtualization in High-Performance Cluster and Grid Computing (VHPC 08). España. (2008).

[BOR05] S. Y. Borkar, P. Dubey, K. C. Kahn, D. J. Kuck, H. Mulder, S. S. Pawlowski, y J. R. Rattner: "Platform 2015: Intel® Processor and Platform Evolution for the Next Decade". In: Intel Corporation, White Paper, 2005.

[CAS12] Casanova B., Balladini J., De Giusti A., Suppi R., Rexachs D., Luque E.: "Mejora de la eficiencia energética en sistemas de computación de altas prestaciones". In: XII Workshop de Procesamiento Distribuido y Paralelo. CACIC 2012. ISBN: 978987-1648-34-4. Pág. 377-386. Argentina (2012).

[DOE11] Doelitzcher, F., Held, M., Sulistio, A., Reich, C. Viteras: "Virtual Cluster as a Service". In: 3rd IEEE International Conference on Cloud Computing Technology and Science. Grecia (2011).

[FED09] Alexandra Fedorova, Juan Carlos Saez, Daniel Shelepov, Manuel Prieto: "Maximizing Power Efficiency with Asymmetric Multicore Systems". In: Communications of the ACM, Vol. 52 (12), pp 48-57. 2009.

[FRA14] F. E. Frati, K. Olcoz Herrero, L. Piñuel Moreno, M. R. Naiouf, A. De Giusti: "Modelo para sintonización dinámica de aplicaciones guiada por eventos de hardware". In: XX Congreso Argentino de Ciencias de la Computación. 2014.

[FRA15] Frati E: "Software para arquitecturas basadas en procesadores de múltiples núcleos. Detección Automática de errores de concurrencia". Tesis Doctoral, Universidad Nacional de La Plata. 2015.

[GUO12] Guoqiang Hu, Wee Peng Tay, Yonggang Wen: "Cloud robotics: architecture, challenges and applications". In: Network, IEEE, vol.26, no.3, pp.21-28. 2012.

[HAC11] Hacker, T., Mahadik, K.: "Flexible Resource Allocation for Reliable Virtual Cluster Computing." In: Supercomputing Proceedings (SC11). USA (2011).

[INT12] Intel Corporation: "Intel® 64 and IA-32 Architectures Software Developer's Manual". Manual 253669-043US, 2012. Recuperado a partir de <http://www.intel.com/content/www/us/en/processors/architectures-software-developer-manuals.html>

[JEF13] Jeffers, James; Reinders, James: "Intel Xeon Phi Coprocessor High Performance Programming", Morgan Kaufmann, 2013.

[KEH15] Kehoe B., Patil S., Abbeel P., Goldberg K.: "A Survey of Research on Cloud Robotics and Automation". In: IEEE Transactions on Automation Science and Engineering (T-ASE): Special Issue on Cloud Robotics and Automation. Vol. 12, no. 2. 2015.

- [LOR13] Lorencik D., Sincak P.: "Cloud robotics: Current trends and possible use as a service". In: Applied Machine Intelligence and Informatics (SAMI), 2013 IEEE 11th International Symposium on , vol., no., pp.85-88. 2013.
- [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". XXI Congreso Argentino de Ciencias de la Computación. Argentina (2015).
- [ROD07] Rodriguez, I. P., Pousa, A., Pettoruti, J.E., Chichizola, F., Naiouf, M., De Giusti, L., De Giusti, A.: "Estudio del overhead en la migración de algoritmos paralelos de cluster y multicluster a GRID". In: Proceedings del XIII Congreso Argentino de Ciencias de la Computación. Argentina (2007).
- [ROD11] Rodriguez, I., Pettoruti, J.E., Chichizola, F., De Giusti, A.: "Despliegue de un Cloud Privado para entornos de cómputo científico". In: Proceedings del XVII Congreso Argentino de Ciencias de la Computación. Argentina (2011).
- [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.
- [SAE14] Juan Carlos Saez, Adrian Pousa, Fernando Castro, Daniel Chaver, Manuel Prieto: "Exploring the Throughput-Fairness Trade-off on Asymmetric Multicore Systems". In: Euro-Par 2014: Parallel Processing Workshops - 2nd Workshop on Runtime and Operating Systems for the Many-core Era (ROME 2014). 2014.
- [SAE15] Juan Carlos Saez, Adrian Pousa, Daniel Chaver, Fernando Castro, Manuel Prieto Matias: "ACFS: A Completely Fair Scheduler for Asymmetric Single-ISA Multicore Systems". In: ACM SAC 2015 (The 30TH ACM/SIGAPP Symposium on applied computing). 2015.
- [SEA13] Sean Settle: "High-performance Dynamic Programming on FPGAs with OpenCL". In: IEEE High Performance Extreme Computing Conference. 2013.
- [SHA10] Shafer, J.: "I/O virtualization bottlenecks in cloud computing today". In: Proceedings of the 2nd conference on I/O virtualization (VIOV10). USA (2010).
- [SPR02] B. Sprunt: "The basics of performance-monitoring hardware". In: IEEE Micro, vol. 22, nro 4, pp. 64- 71. 2002.
- [TIN14] F. G. Tinetti, S. M. Martin, F. E. Frati, M. Méndez: "Optimization and Parallelization experiences using hardware performance counters". In: Proceedings of the 4th International Supercomputing Conference. México (2013).
- [VAQ09] Vaquero, L. M., Rodero-Merino, L., Caceres, J., Lindner, M.: "A Break in the Clouds: Towards a Cloud Definition". In: ACM SIGCOMM Computer Communication Review, Volume 39, Issue 1, pp. 50--55. USA. (2009).
- [VAZ09] Vázquez Blanco, B., Huedo, E., Montero, R. S., Llorente, I. M.: "Elastic Management of Cluster-based Services in the Cloud". In: Proceedings pp.19-24, ACM Digital Library. ISBN 978-1-60558-564-2. 2009.
- [VEL09] Velte, A.T., Velte, T.J., Elsenpeter, R.: "Cloud Computing: A Practical Approach". McGraw Hill Professional. 2009.
- [XIL15] Xilinx Inc. "SDAccel Development Environment". [Online]. Disponible en <http://www.xilinx.com/products/design-tools/software-zone/sdaccel.html>
- [XIN12] Xing, Y., Zhan, Y.: "Virtualization and Cloud Computing". In: Proceedings pp.305-312, Springer Link. ISBN 978-3-642-27323-0. (2012). Morgan Kaufmann. 2013.