

# **Simulación paramétrica paralela. Aplicación a modelos de predicción de inundaciones.**

Adriana A. Gaudiani<sup>1</sup>, Armando De Giusti<sup>2</sup>, Emilio Luque<sup>3</sup>, Marcelo Naiouf<sup>2</sup>

<sup>1</sup> Instituto de Ciencias, Universidad Nacional de General Sarmiento, Bs. As., Argentina - agaudi@ungs.edu.ar

<sup>2</sup> Instituto de Investigación en Informática LIDI (III-LIDI), Universidad Nacional de La Plata, Bs. As., Argentina – {degiusti, mnaiouf}@lidi.info.unlp.edu.ar

<sup>3</sup> Depto. de Arquitectura de Computadores y Sistemas Operativos, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona) España- Emilio.Luque@uab.es

## **RESUMEN**

El modelado y la simulación de inundaciones provocadas por el desborde de ríos brinda sistemas computacionales para el estudio y la predicción de estos fenómenos naturales, con el objetivo de pronosticar su comportamiento. Estos sistemas necesitan tomar gran cantidad de datos de entrada para aumentar su precisión, como también deben generar múltiples escenarios para cubrir todas las situaciones de riesgo. Por esto, son de cómputo intensivo y pueden tomar días de procesamiento hasta lograr resultados. A este problema se le suma la falta de certeza en los valores de los datos de entrada del proceso.

Mediante la programación paralela y los avances en cómputo de alto rendimiento en clusters de computadoras, se pretende atenuar el problema de la incertidumbre de los datos de entrada y optimizar el proceso de predicción mediante la simulación de múltiples escenarios.

Con este trabajo se pretende desarrollar una metodología para optimizar la predicción de inundaciones provocadas por el desborde de ríos, en principio de llanuras o planicies, y en particular en la Cuenca del Río Salado o en el Paraná Medio.

## **MARCO TEÓRICO**

Muchos investigadores afirman que el cómputo se ha convertido en el tercer método para hacer investigación detrás de la teoría y la experimentación para las ciencias y la ingeniería. Aunque no existe un total acuerdo sobre el lugar destinado al cómputo científico junto a las otras dos partes de esta tríada, es innegable que los métodos de cómputo son una herramienta indispensable en la mayoría de la disciplinas.

El cómputo científico, utilizando los métodos de las ciencias de la computación, hace posible el estudio de problemas que son demasiado complejos para ser tratados de manera analítica, o de aquellos muy costosos o peligrosos para estudiarlos de manera natural mediante experimentación.

Habitualmente los problemas del mundo real son sistemas muy complejos para ser evaluados mediante modelos analíticos, y requieren de una simulación numérica para su estudio. La simulación con computadoras provee de un recurso para imitar el comportamiento de sistemas complejos del mundo real. Como afirman Law y Kelton, “en una simulación usamos una computadora para evaluar el modelo numéricamente y los datos son reunidos para estimar sus verdaderas características” [14].

Los modelos de simulación son una abstracción de la realidad: se basan en una representación matemática en tiempo y espacio del sistema real, que existe o podría existir. Todos los modelos pueden comparar las salidas alternativas antes de seleccionar

un curso de acción y brindan un entorno virtual dinámico de experimentación o de formación.

El modelo es una representación más simple del problema, y uno de sus propósitos es predecir los efectos de variaciones en el sistema. Un buen modelo debe equilibrar realismo y simplicidad. La simulación del sistema es la operación del modelo. Este puede reconfigurarse y volver a experimentar, algo que puede no ser posible en el sistema que representa [2].

La simulación de problemas categorizados como de gran tamaño y complejidad constituyen un desafío debido a la gran cantidad de disciplinas involucradas y al grado de dificultad que presenta su modelización. Una definición ampliamente aceptada acerca de este tipo de problemas es la que brinda The Office of Science and Technology Policy: “*Grand Challenges are... fundamental problems in science and engineering, with potentially broad social, political, and scientific impact, that could be advanced by applying high performance computing resources*”. La simulación de inundaciones provocadas por ríos de planicie está clasificada como uno de estos problemas.

El estado del arte en modelos computacionales nos dice que un buen modelo de un sistema complejo usualmente es una herramienta cara para la ciencia. Día a día crece la demanda de recursos de hardware, mucho tiempo de procesamiento y mucha memoria para seguir la evolución del sistema o para mejorar las salidas gráficas. Actualmente los sistemas de cómputo que ofrecen un alto rendimiento son cada vez más accesibles, algo en parte posible gracias a los clusters de PC's y la computación grid[5] [6][17][18].

No se debe perder de vista que los modelos utilizados en simulaciones de la ciencia y la ingeniería nunca ofrecen un modelo perfecto de la naturaleza, y sólo son un subconjunto de la realidad. Siempre los experimentos y las observaciones de los expertos en cada área serán imprescindibles como puntos de referencia para comprender los fenómenos naturales.

## MOTIVACION

Las inundaciones de ríos de planicie son un fenómeno natural que ocurre habitualmente y con el rápido cambio del clima global se espera que acontecimientos extremos más severos y más frecuentes ocurran en el futuro cercano. El control de las inundaciones es una preocupación primordial, junto con la necesidad de supervisar las inundaciones en situaciones de crisis.

Por estas razones, los modelos de simulación basados en este fenómeno físico son necesarios para permitir una comprensión mejor de los procesos implicados en él. Modelos globales o locales y toda la información necesaria sobre el entorno que se pretende estudiar deberán integrarse para lograr un sistema de simulación que pueda predecir eficientemente la evolución en tiempo y espacio del escurrimiento del agua durante la inundación.

En particular, estos fenómenos agravados por el fenómeno El Niño, tuvieron graves consecuencias en el litoral argentino como sucedió en el año 1997 en el río Paraná, y como sucedió en el 2003 con la trágica crecida del cauce del Salado.<sup>1</sup>

El país cuenta con varios modelos de predicción, pero muy poco sobre simuladores de inundaciones [7][8][9]. Existen abundantes datos de las crecidas del Paraná y en menor cantidad respecto al Salado. El trabajo de Giampieri *et.al.* [13] presenta una interesante reconstrucción numérica del desastre provocado por la cuenca del Salado en el 2003. Esta situación representa un desafío al brindar una metodología de optimización y el

---

<sup>1</sup> <http://ultra31.unl.edu.ar/noti/noticia.php?idnoticia=1480>

desarrollo de un DSS que permitan predecir de la manera más certera posible y tomar decisiones anticipadas sobre estos eventos.

Por otro lado, los avances en Computer Science y en High Performance Computing son muy valiosos y pueden ponerse al servicio de este desarrollo para incrementar la eficiencia del DSS y así lograr la predicción de manera confiable y rápida [5] [6].

## **OBJETIVOS**

Este trabajo busca proveer de una metodología y una herramienta que permita buenas predicciones y ayudar en la prevención de los desastres que provocan estos fenómenos naturales.

Se pretende desarrollar una metodología para la predicción de inundaciones provocadas por el desborde de ríos, en principio de llanuras o planicies y, en particular en la Cuenca del Río Salado o en el Paraná Medio.

Se espera brindar las herramientas necesarias para atenuar el problema de la incertidumbre de los datos de entrada y optimizar el proceso de predicción mediante métodos estadísticos basados en experimentación factorial.

Con el desarrollo de un framework de experimentación se espera estudiar y evaluar la posibilidad de implementar un sistema de soporte para la toma de decisiones ante el desarrollo de este tipo de eventos.

Mediante el aporte de la programación paralela sobre clusters de PC's (y eventualmente distribuida sobre grid) se pretende ir algunos pasos más allá de la metodología clásica provista por los modelos de predicción de inundaciones, aprovechando el paralelismo inherente en algunas técnicas de optimización. Gracias a los avances en High Performance Computing se busca estudiar su aplicación en este trabajo para intentar proveer al simulador potencia de cómputo que mejore el proceso de predicción.

## **DESARROLLO**

En la primera etapa se hará un intenso estudio de los modelos numéricos de uso actual buscando evaluar su confiabilidad al ser utilizados en sistemas de simulación y predicción de inundaciones.

El desarrollo del framework se hará reuniendo optimización y simulación manteniendo como Solver o Modelo Numérico del problema hidrodinámico alguno de los utilizados actualmente y de dominio público, como son RMA2 y HEC-RAS.<sup>2</sup>

La predicción a corto plazo se logra aumentando la cantidad de parámetros de entrada, pero, debido a la inevitable falta de precisión que se produce en dichos valores, es necesario implementar técnicas de optimización con el fin de definirlos con la máxima exactitud. Una manera de lograrlo es mediante la programación paralela tratando con sus técnicas de posibilitar una mayor certeza en las predicciones.

Algunas de las técnicas a investigar para su utilización son algoritmos genéticos, máximo gradiente, búsqueda Taboo, entre otras posibles. Se generará una función de ajuste para evaluar el grado de predicción lograda.

Debido a la necesidad de cubrir los riesgos posibles que se generarían ante una situación real de desborde del río, se contemplan múltiples escenarios. Cada escenario es el resultante de la combinación de los factores que determinan el comportamiento del

---

<sup>2</sup> <http://www.hec.usace.army.mil/>  
[http://smig.usgs.gov/cgi-bin/SMIC/model\\_home\\_pages/model\\_home?selection=rma2](http://smig.usgs.gov/cgi-bin/SMIC/model_home_pages/model_home?selection=rma2)

sistema. Se desarrollará una metodología adecuada utilizando experimentación factorial utilizando para su implementación técnicas de cómputo de altas prestaciones.

La gran cantidad de datos consumidos por este tipo de simulación y los métodos utilizados en la optimización, como por ejemplo la basada en experimentación factorial, requieren una gran potencia de cómputo pues es necesario ejecutar un gran número de simulaciones. Debido a esto, se buscará mejorar la velocidad de convergencia de la optimización utilizando un esquema de programación paralela del tipo Master-Worker o peer-to-peer implementado sobre un cluster de PC's, una colección de clusters de PC's o sobre un grid. Un procesador puede calcular cada combinación de parámetros y enviarlos a los Workers para que desarrollen la simulación, construyendo un mapa de riesgo o impacto ambiental.

La mejora en el rendimiento general del sistema mediante los aportes logrados con los avances en High Performance Computing hace factible el análisis de escenarios hipotéticos, como los vinculados al cambio climático.

Por otro lado, se estudiará la distribución del área de datos durante el proceso de simulación aprovechando los aportes del cómputo distribuido y su impacto en la mejora del tiempo de procesamiento.

### **Respecto a la relación establecida con Instituciones Académicas y grupos de investigación:**

- Este trabajo se construye en base a la relación de trabajo establecida con investigadores del Departamento de Arquitectura de Computadores y Sistemas Operativos (DACSO-CAOS) de la Universidad Autónoma de Barcelona - España, donde se llevan adelante investigaciones sobre simulación de propagación y predicción de incendios forestales, y sobre aplicaciones paralelas en ciencia computacional, entre otras. Especialmente con el grupo dirigido por el Dr. Emilio Luque, con quienes se ha participado en trabajos de predicción y sintonización de aplicaciones en entornos multi-cluster [3].

- El trabajo es desarrollado en el Instituto de Investigación en Informática LIDI (III-LIDI) de la Universidad Nacional de La Plata, bajo la dirección del Ing. De Giusti y del Dr. Naiouf. El avance del mismo se ve favorecido por la relación establecida entre el Instituto y el Departamento DACSO-CAOS de la UAB, respecto a sus actividades de investigación.

- Desde el III-LIDI se ha establecido contacto con el Dr. Vionnet de la Fac. de Ingeniería y de Ciencias Hídricas de la Universidad Nacional del Litoral, lo que ha permitido relevar los programas que utilizan en la simulación y predicción de inundaciones quedando abierta la posibilidad de cooperación entre ambas Instituciones.

### **AGRADECIMIENTO**

Agradecemos los comentarios de la Dra. Dolores Rexachs, los cuales ayudaron a mejorar esta presentación.

### **REFERENCIAS BIBLIOGRÁFICAS**

- [1] Anna Morajko, Paola Caymes-Scutari, Tomàs Margalef, Emilio Luque: *Automatic Tuning of Data Distribution Using Factoring in Master/Worker Applications*. International Conference on Computational Science (2) 2005: 132-139
- [2] Anu Maria: *Introduction to modeling and simulation*. Proceedings Winter Simulation Conference of the 1997.

- [3] Argollo, Gaudiani, Rexachs, Luque. *Tuning application in a multi-cluster environment*. Springer - Lecture Notes in Computer Science, Volume 4128/2006, p. 78-88 (2006)
- [4] Baker Abdalhaq, Ana Cortés, Tomàs Margalef, Emilio Luque: *Enhancing wildland fire prediction on cluster systems applying evolutionary optimization techniques*. Future Generation Comp. Syst. 21(1): 61-67 (2005)
- [5] Baker Abdalhaq, Ana Cortés, Tomàs Margalef, Emilio Luque: *Accelerating Wildland Fire Prediction on Cluster Systems*. International Conference on Computational Science 2004: 220-227
- [6] Baker Abdalhaq, Germán Bianchini, Ana Cortés, Tomàs Margalef, Emilio Luque: *Improving Wildland Fire Prediction on MPI Clusters*. PVM/MPI 2003: 520-528
- [7] Clause, Dalponte, Rinaldi, Vénere, Cazenave, Varni, Vives. *Simulación de inundaciones en llanuras aplicación a la cuenca del arroyo Santa Satalina – Azul*”, Mecánica Computacional Vol. XXIII, G.Buscaglia, E.Dari, O.Zamonsky (Eds.) Bariloche, Argentina, Noviembre 2004.
- [8] Clause, Euillades, Vénere,, Vives, *Modelos de elevación de precisión y su importancia para la simulación de inundaciones*. Congreso Nacional del Agua, Córdoba (2002).
- [9] Clause, Vénere, A computational environment for water flow along floodplains. International Journal on Computational Fluid Dynamics, Vol. 16, pp. 327-330 (2002).
- [10] Germán Bianchini, Ana Cortés, Tomàs Margalef and Emilio Luque. *S2F2M - Statistical System for Forest Fire Management*. Departament d'Informàtica, E.T.S.E, Universitat Autònoma de Barcelona, 08193-Bellaterra (Barcelona), Spain. 2005.
- [11] Germán Bianchini, Ana Cortés, Tomás Margalef, Emilio Luque and Emilio Chuvieco: *Wildland Fire Propagation Danger Maps Based on Factorial Experimentation*. ITEE 2005: 173-185
- [12] Gene Golub and James M. Ortega. *Scientific Computing, An Introduccion with Parallel Computing*. Academic Press – 1993.
- [13] Giampieri R.C., P.A. Tassi, L.B. Rodríguez, y C.A. Vionnet. *Simulación Numérica de la Crecida Extraordinaria del Río Salado en la Ciudad de Santa Fe*. Revista Internacional de Desastres Naturales, Accidentes e Infraestructura Civil, Puerto Rico. ISSN 1535-0088. Vol. 4, No.1, Mayo 2004. 85-102.
- [14] Law, Kelton. *Simulation Modeling and Analisis*. McGraw-Hill International Series, Third Ed. ISBN 0-07-116537-1. 2000.
- [15] Pablo Jacovski, *Computadoras, Modelización Matemática y Ciencia Experimental, Mecánica Computacional*. Vol. XXIII. G.Buscaglia, E.Dari, O.Zamonsky (Eds.) Bariloche, Argentina, Noviembre 2004.
- [16] Sixto Rios Insua, Bielza Lozoya, Alfonso Caballero: *Fundamentos de los Sistemas de Ayuda a la decisión*. RaMa 2002 ISBN 84-7897-494-6
- [17] A. Chervenak, I. Foster, C. Kesselman, C. Salisbury, S. Tuecke, “The data Grid: Towards and Architecture for the Distributed Mannagement and Analysis of Large Scientific data Sets”. Journal of Network and Computer Applications, 2001,pp.187-200.
- [18] Ian Foster, Carl Kesselman. “The Grid 2: Blueprint for a New Computing Infrastructure” (The Morgan Kaufman Series in Computer Architecture and Desing). Morgan Kaufmann; 2<sup>nd</sup> edition, 2003.