Aplicaciones de Cómputo Intensivo con Impacto Social

Javier Balladini[1], Marina Morán[1], Pablo Bruno[1], Belén Casanova[1], Claudia Rozas[1], Federico Uribe[12], Sebastián Gomez[2], Nestor Vicente[3], Cristina Orlandi[3], Armando De Giusti[4], Remo Suppi[5], Dolores Rexachs[5], Emilio Luque[5]

[1] Facultad de Informática, Universidad Nacional del Comahue
{javier.balladini, marina, claudia.rozas, federico.uribe}@fi.uncoma.edu.ar; psbruno90@gmail.com; mb.casanova.retamal@gmail.com

[2] Poder Judicial del Neuquén
sebastian.gomez@jusneuquen.gov.ar

[3] Hospital Francisco Lopez Lima
nestorvicente07@yahoo.com.ar; orlandi.mariacristina@gmail.com

[4] Instituto de Investigación en Informática LIDI, Universidad Nacional de La Plata
degiusti@lidi.info.unlp.edu.ar

[5] Departamento de Arquitectura de Computadores y Sistemas Operativos, Universidad Autónoma de Barcelona
{remo.suppi, dolores.rexachs, emilio.luque}@uab.es

Resumen

Los avances tecnológicos de los sistemas de cómputo paralelo y distribuido permiten el desarrollo de aplicaciones antes impensadas. Nuestra investigación se centra en desarrollar metodologías, modelos y soluciones informáticas para colaborar en la resolución de problemas que tengan una alta demanda computacional e impacto social. Hemos definido tres ejes de investigación: aplicaciones para la salud, aplicaciones de informática forense, y consumo energético de los sistemas de cómputo paralelo. Estas líneas de investigación se desarrollan en colaboración con una universidad nacionales y otra del extranjero, un hospital público, y un gabinete provincial de informática forense.

Palabras claves: computación de altas prestaciones, eficiencia energética, aplicaciones para la salud, aplicaciones de informática forense.

1. Contexto

La línea de investigación aquí presentada está enmarcada dentro del proyecto de investigación 04/F013 "Aplicaciones de Cómputo Intensivo con Impacto Social", financiado por la Universidad Nacional del Comahue (UNComa), con inicio el 01/01/2017 y finalización el 31/12/2020, y acreditado por el Ministerio de Educación de Argentina. Otra fuente de financiamientos es el Inter-U "Colaboración UNComa-UNLP: docencia e investigación en Sistemas Paralelos" de PROMINF.

Uno de los tres ejes centrales de nuestra investigación, las aplicaciones para la salud, se desarrolla en colaboración con la Unidad de Terapia Intensiva y la Unidad de Vigilancia Intermedia, ambos pertenecientes al "Hospital Francisco López Lima" de la ciudad de General Roca, provincia de Río Negro. El eje de aplicaciones de informática forense, se desarrolla dentro del marco de colaboración con el Gabinete de Pericias Informáticas del Poder Judicial de la provincia de Neuquén. Respecto al eje relacionado con el consumo energético de los sistemas de cómputo paralelo, se desarrolla en colaboración con el Instituto de Investigación en Informática LIDI de la Universidad Nacional de La Plata (UNLP), y el grupo de investigación HPC4EAS (High Performance Computing for Efficient Applications and Simulation) de la Universidad Autónoma de Barcelona (UAB) de España.

2. Introducción

Una de las áreas de mayor interés en la actualidad es la Computación de Altas Prestaciones (HPC, del inglés, *High Performance Computing*). La computación paralela es un tipo de computación en el que los cálculos se realizan de forma simultánea. Si bien el paralelismo ha sido empleado históricamente en la computación de altas prestaciones, ha ganado un enorme interés debido al impedimento para seguir aumentando la frecuencia de reloj de los procesadores; el problema se encuentra en el alto consumo energético y disipación del calor a altas frecuencias. Como

no se podía seguir aumentando la frecuencia para que las aplicaciones ejecuten más rápido, la solución fue incrementar la cantidad de unidades de procesamiento, dando así lugar a la aparición de procesadores multinúcleos. Desde entonces, la computación paralela se ha convertido en el paradigma dominante en la arquitectura de computadoras.

Para algunas aplicaciones, será suficiente con utilizar una plataforma comprendida por una única computadora con uno o más procesadores multinúcleos. En otros casos, podrá ser necesario el poder de cómputo de una agregación de computadoras, como por ejemplo del tipo *cluster*. La masificación de la tecnologías de cómputo paralelo hacen que ellas sean cada vez más accesibles, y se pueda pensar en el desarrollo de nuevas aplicaciones, muchas de las cuales pueden tener un fuerte impacto para el beneficio social. Sin embargo, extraer el máximo rendimiento de estas plataformas requiere utilizar técnicas específicas de programación paralela, que son más difíciles que las típicas de programación secuencial, principalmente debido a la sincronización, comunicación de tareas, y complejidad de la arquitectura de las plataformas hardware.

Nuestro interés está centrado en tres áreas temáticas: aplicaciones de cómputo intensivo para la salud, aplicaciones de cómputo intensivo de informática forense, y el consumo energético de los sistemas de HPC. A continuación se introduce cada uno de estos temas:

Aplicaciones para la salud

En los centros de salud, la Unidad de Cuidados Intensivos (UCI) atiende a pacientes cuya salud está en condiciones críticas, con riesgo de muerte. Por esta razón deben contar con asistencia médica las 24 horas del día y ser controlados en forma rigurosa; mínimamente están conectados a un monitor que mide sus signos vitales y da alertas, considerando aquellos aspectos que pueden indicar un riesgo para la salud del paciente bajo criterios aplicados según la población estándar.

El equipamiento utilizado para monitorear a los pacientes y realizar diferentes estudios ha sido desarrollado para trabajar en forma independiente, sin contemplar la posibilidad de incorporar información adicional obtenida a través de otros medios (ya sea otros estudios, datos inherentes del paciente, etc). Es decir, no se ha previsto un uso integral de toda la información del paciente, sino que los estudios se realizan en forma aislada y luego el personal médico debe realizar el análisis de los mismos considerando todas las variables conocidas. Esta metodología demora la detección de patologías. Además, debido a que el equipamiento médico generalmente no está preparado para el registro histórico de datos, el personal de la UCI toma registros manuales en intervalos de horas. Así, la mayoría de las mediciones que realiza el equipamiento médico se pierden, y esta omisión de información podría reducir la precisión del diagnóstico.

Un gran avance sería disponer de un sistema de cómputo que detecte patologías en tiempo real, basándose en múltiples parámetros de diferentes medios. La detección temprana de patologías permitirá aumentar la efectividad de los tratamientos, y por consiguiente la mejora de la salud de los pacientes y la reducción del costo económico. El problema principal que debe enfrentarse para la construcción de este sistema, y que creemos viable con la aplicación de técnicas de computación paralela, es el procesamiento en tiempo real de un gran volumen de datos generado por el equipamiento (especialmente las curvas como, por ejemplo, el electrocardiograma).

No hay muchos sistemas de este tipo, algunos de ellos se encuentran en etapa experimental inicial y otros ya llevan algunos años de investigación. La información disponible de estos sistemas es normalmente escasa por tratarse mayormente de software privativo. En la bibliografía se encuentran algunos trabajos como [13, 3, 1, 2].

Aplicaciones de informática forense

En la actualidad, una gran parte de las investigaciones judiciales involucran elementos de prueba que son potenciales fuentes de evidencia digital [11, 10]. El volumen de datos digital que debe ser sometido a análisis forense está aumentando de forma sostenida, incrementando el tiempo de procesamiento requerido para el análisis. Este aumento en el tiempo de análisis forense lleva a un crecimiento excesivo de la lista de espera de pericias en trámite e impacta negativamente en la investigación que deben llevar a cabo los organismos jurisdiccionales.

Muchas tareas periciales involucran un elevado tiempo de procesamiento: detección de imágenes sospechosas de contener pornografía, la generación de índices para realizar consultas dinámicas sobre el corpus digital, la creación de diccionarios personalizados para descifrado de contraseñas en base a la información digital contenida en los dispositivos de almacenamiento, la generación de listas de valores hash para análisis de firmas, la detección de malware o localización de archivos relevantes a la investigación, y la búsqueda de evidencia digital mediante palabras clave. En especial, la detección de imágenes con contenido pornográfico [8] es una de las tareas

que requieren mayor tiempo de procesamiento.

Las herramientas disponibles en el mercado no satisfacen los tiempos de procesamiento requeridos para estas actividades. Los avances tecnológicos y técnicas de programación de los sistemas de cómputo paralelo pueden hacer viable el desarrollo de aplicaciones para informática forense que tengan tiempos de procesamiento menores a los actuales. En muchos casos, las herramientas también carecen de capacidades de ejecución no interactiva que posibilite la automatización de tareas para ser utilizadas en entornos de cómputo paralelo y distribuido.

Consumo energético

Mientras el rendimiento de los sistemas de computación de altas prestaciones (HPC, High Performance Computing) continúa creciendo, las máquinas aumentan significativamente la cantidad de unidades de procesamiento. Este aumento en el número de componentes hace disminuir la confiabilidad y aumentar el consumo energético de un sistema de cómputo. Así, a pocos años de arribar a la era exaescala (prevista para 2020), el consumo energético se han identificado como uno de los mayores desafíos a enfrentar [14, 12].

El consumo energético es hoy en día un gran problema. Para dar una idea de la magnitud del mismo, utilizaremos de ejemplo la máquina de mayor prestaciones de la actualidad, la máquina China Sunway TaihuLight. Esta máquina demanda 15 MW de potencia, lo mismo que se requiere para abastecer a los hogares de una ciudad con alrededor de 200.000 habitantes (calculo realizado en base al consumo de un hogar en Argentina). Además del alto impacto económico, la generación de tanta energía podría tener un alto impacto medioambiental, por ejemplo, represas hidroeléctricas que modifican el ecosistema, y social, por ejemplo, la mayor fuente de energía mundial se obtiene del carbón, cuya extracción minera es altamente peligrosa.

La computación ecológica es el estudio y la práctica de la computación ambientalmente sostenible. Ella se ocupa de diferentes aspectos de los sistemas de cómputo: diseño, manufactura, eliminación, y uso. Este último aspecto, el uso ecológico, se refiere al uso de los sistemas de cómputo con conciencia ambiental. Es posible reducir el consumo de energía de los sistemas de cómputo utilizando diferentes estrategias que deben ser consideradas a nivel del software, y consisten en realizar cambios en la configuración del sistema o en las aplicaciones. Estas estrategias incluyen: explotación del paralelismo (muchos cores lentos consume menos energía que pocos cores rápidos), uso adecuado de la jerarquía de memoria, hibernación de recursos, escalado dinámico de frecuencia y tensión, rediseño de algoritmos, planificación de tareas, asignación de tareas a recursos hardware.

3. Líneas de investigación

El eje central de nuestra investigación es desarrollar metodologías, modelos y soluciones informáticas para colaborar en la resolución de problemas que tengan una alta demanda computacional e impacto social en los siguientes campos: aplicaciones para la salud, aplicaciones de informática forense, consumo energético de los sistemas de cómputo paralelo.

Aplicaciones para la salud

Esta línea está enfocada en el estudio y desarrollo de un sistema para detección temprana de patologías en Unidades de Terapia Intensiva (UTI), y que puede también abarcar a las Unidades de Vigilancia Intermedia (UVI). Nuestro objetivo está orientado a resolver el procesamiento de una gran cantidad de datos en tiempo real, proveniente de señales de equipamiento médico, utilizando técnicas de computación paralela.

Aplicaciones de informática forense

Esta línea se centra en acelerar el procesamiento de grandes volúmenes de datos, principalmente de videos e imágenes, mediante técnicas de computación paralela. En especial, interesa acelerar el proceso de detección automática de imágenes con pornografía.

Consumo energético

Nos centramos en el desarrollo de metodologías, modelos y construcción de software para administrar y gestionar el consumo de energía y prestaciones de sistemas de cómputo paralelo. Nuestros objetivos principales son:

- Predicción de energía y rendimiento. Es importante proveer a un administrador de sistema de herramientas que permitan predecir la energía y el rendimiento que producirían distintas configuraciones del sistema al ejecutar una dada aplicación paralela, y así poder seleccionar la configuración adecuada que mantenga el compromiso deseado entre tiempo de ejecución y eficiencia energética.

- Gestión energética en mecanismos de tolerancia a fallos. La tolerancia a fallos agrega

una carga de trabajo significativa al sistema de cómputo, sobre todo en sistemas que tienen enormes cantidades de unidades de procesamiento [9], haciendo necesario gestionar el consumo energético de los distintos mecanismos.

4. Resultados y objetivos

Aplicaciones para la salud

Los objetivos específicos en curso son:

- Desarrollar hardware y software para extraer datos del equipamiento médico, y transmitirlos por WiFi a la plataforma de procesamiento de la información. La extracción de datos no es trivial debido al uso de protocolos de comunicaciones propietarios que los fabricantes no dan a conocer.
- Desarrollar aplicaciones para el procesamiento eficiente de señales (como el electrocardiograma). Éstas deben ejecutar en máquinas con procesadores de propósito general (CPUs), y utilizar los recursos de cómputo de manera eficiente para reducir el tamaño de la plataforma hardware que requiere el sistema. Se incluye también la detección de anomalías en las señales para evitar la contaminación del sistema con datos erroneos.
- Diseñar la infraestructura de un sistema de Big Data que permita el almacenamiento, análisis y procesamiento en tiempo real de señales extraídas del equipamiento médico (monitores de signos vitales, respiradores, etc.) y otros datos ingresados manualmente.
- Desarrollar una aplicación para la interacción con médicos y enfermeros.

Los avances/resultados actualmente comprenden:

- Un análisis del estado general de las UTI del hospital Francisco Lopez Lima [7].
- El desarrollo de un dispositivo embebido que obtiene información de la señal del electrocardiograma de un monitor médico, a través de una salida analógica, y la transmite por WiFi para su posterior procesamiento.
- El desarrollo parcial de una aplicación que detecta complejos QRS de manera eficiente en señales de electrocardiogramas.
- Análisis parciales de rendimientos de diferentes alternativas tecnológicas de la infraestructura de Big Data.

Aplicaciones de informática forense

Se han evaluado distintas alternativas de software para detección de pornografía disponible en la literatura, y se ha seleccionado la aplicación desarrollada por Yahoo [4] debido a que posee una alta tasa de aciertos. Actualmente, se está trabajando en el análisis de la aplicación seleccionada, y en su optimización para acelerar el tiempo de procesamiento de imágenes en plataformas paralelas basadas en CPUs. Además, se espera avanzar en la ejecución distribuida de la aplicación de detección de imágenes pornográficas para aumentar la productividad en el procesamiento de una cantidad masiva de imágenes, y en el desarrollo de una aplicación que permita el lanzamiento de la aplicación para el procesamiento de un conjunto específico de imágenes y generación de reportes forenses.

Consumo energético

Tras dar los primeros pasos en la predicción de energía y rendimiento para aplicaciones SPMD construidas con el modelo de programación de paso de mensajes (MPI) [6], actualmente se está avanzando en la predicción energética para aplicaciones SPMD implementadas con un modelo de programación híbrido de paso de mensajes (MPI) y memoria compartida (OpenMP), y en una metodología de predicción mejorada que permita una mayor precisión.

Los métodos de tolerancia a fallos tienen fuerte incidencia en el consumo energético de los sistemas de HPC, y resulta de suma importancia conocer, antes de ejecutar una cierta aplicación, el impacto que pueden producir los diferentes métodos y configuraciones del mismo. En [5], presentamos una metodología para predecir el consumo energético producido por el método de checkpoint coordinado remoto. Actualmente, hemos agregado la predicción para la operación de restart (adicionalmente a la de checkpoint), y se están contemplando distintas alternativas de configuraciones del sistema, relacionadas a: almacenamiento en NFS, configuración de la aplicación de checkpoint/restart, y energéticas (estados C y P de las CPUs). A futuro, esperamos extender el trabajo a la propuesta de mecanismos de gestión de tolerancia a fallos que procuren un uso eficiente del cluster, permitiendo maximizar la productividad y minimizar el consumo energético.

5. Formación de recursos humanos

El equipo de trabajo cuenta con un integrante doctorado en la temática (año 2008). En 2018 se espera la finalización de dos tesis de grado en temas de aplicaciones para la salud, y en 2019 una tesis doctoral en el tema de consumo energético.

Referencias

[1] Amara health analytics, http://www.amarahealthanalytics.com (accedido en marzo de 2018).

[2] ehcos smarticu, http://www.ehcos.com/productos/ehcos-smarticu/ (accedido en marzo de 2018).

[3] Excel medical, http://excel-medical.com/ (accedido en marzo de 2018).

[4] Open sourcing a deep learning solution for detecting nsfw images, https://yahooeng.tumblr.com/post/151148689421/open-sourcing-a-deep-learning-solution-for (accedido en marzo de 2018).

[5] Javier Balladini, Marina Morán, Dolores Rexachs, and Emilio Luque. Metodología para predecir el consumo energético de checkpoints en sistemas de hpc. *XX Congreso Argentino de Ciencias de la Computación (CACIC 2014)*, 2014.

[6] Javier Balladini, Ronal Muresano, Remo Suppi, Dolores Rexachs, and Emilio Luque. Methodology for predicting the energy consumption of spmd application on virtualized environments. *Computer Science and Technology (JCST)*, 13(3):130–136, 2013.

[7] Javier Balladini, Claudia Rozas, Emmanuel Frati, Nestor Vicente, and Cristina Orlandi. Big data analytics in intensive care units: challenges and applicability in an argentinian hospital. *Computer Science and Technology (JCST)*, 2015.

[8] Carlos Caetano, Sandra Eliza Fontes de Avila, William Robson Schwartz, Silvio Jamil Ferzoli GuimarÃ£es, and Arnaldo de Albuquerque AraÃºjo. A mid-level video representation based on binary descriptors: A case study for pornography detection. *Neurocomputing*, 213:102–114, 2016.

[9] Franck Cappello, Al Geist, William Gropp, Sanjay Kale, Bill Kramer, and Marc Snir. Toward exascale resilience: 2014 update. *Supercomputing Frontiers and Innovations*, 1(1), 2014.

[10] Simson L. Garfinkel. Digital forensics research: The next 10 years. *Digit. Investig.*, 7:S64–S73, August 2010.

[11] Sebastián Gómez, Hernán Herrera, and Federico Uribe. Automatización y computación distribuida para laboratorios de informática forense. In *45Âº JAIIO - Jornadas Argentinas de Informática*, 2016, CABA, Argentina.

[12] Robert Lucas, James Ang, Keren Bergman, Shekhar Borkar, William Carlson, Laura Carrington, George Chiu, Robert Colwell, William Dally, Jack Dongarra, Al Geist, Rud Haring, Jeffrey Hittinger, Adolfy Hoisie, Dean Micron Klein, Peter Kogge, Richard Lethin, Vivek Sarkar, Robert Schreiber, John Shalf, Thomas Sterling, Rick Stevens, Jon Bashor, Ron Brightwell, Paul Coteus, Erik Debenedictus, Jon Hiller, K. H. Kim, Harper Langston, Richard Micron Murphy, Clayton Webster, Stefan Wild, Gary Grider, Rob Ross, Sven Leyffer, and James Laros III. Doe advanced scientific computing advisory subcommittee (ascac) report: Top ten exascale research challenges.

[13] Carolyn McGregor. Big data in neonatal intensive care. *Computer*, 46(6):54–59, 2013.

[14] John Shalf, Sudip Dosanjh, and John Morrison. Exascale computing technology challenges. In *Proceedings of the 9th International Conference on High Performance Computing for Computational Science*, VECPAR'10, pages 1–25, Berlin, Heidelberg, 2011. Springer-Verlag.