

Aplicaciones de cómputo intensivo con impacto social: Salud y Consumo de energía en HPC

Javier Balladini¹, Marina Morán¹, Claudia Rozas¹, Candelaria Álvarez¹,
Rafael Zurita¹, Rodrigo Cañibano¹, Pablo Bruno¹, Belén Casanova¹,
Cristina Orlandi², Armando De Giusti³, Remo Suppi⁴, Dolores Rexachs⁴, Emilio Luque⁴

¹ Facultad de Informática, Universidad Nacional del Comahue
{javier.balladini, marina, claudia.rozas, candelaria.alvarez, rafa, rodrigo.canibano}@fi.uncoma.edu.ar;
{psbruno90, mb.casanova.retamal}@gmail.com

² Hospital Francisco Lopez Lima
orlandi.mariacristina@gmail.com

³ Instituto de Investigación en Informática LIDI, Universidad Nacional de La Plata
degiusti@lidi.info.unlp.edu.ar

⁴ 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. Actualmente, las dos líneas principales de investigación que llevamos adelante son: aplicaciones para la salud con problemas de Big Data, y consumo energético de los sistemas de computación de altas prestaciones. Estas líneas de investigación se desarrollan en colaboración con una universidad nacional y otra del extranjero, y un hospital público.

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

1. Contexto

Las líneas de investigación aquí presentadas están enmarcadas 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.

Uno de los dos 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,

pertenecientes al "Hospital Francisco López Lima" de la ciudad de General Roca, provincia de Río Negro. 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 HPC, ha ganado una enorme difusión en el público general debido al impedimento para seguir aumentando la frecuencia de reloj de los procesadores; el problema se encuentra fundamentalmente 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 de la arquitectura de la mayoría de los sistemas

de cómputo, incluyendo computadoras personales de escritorio, notebooks, celulares y tablets.

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 las 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, como el sector salud. 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 y comunicación entre tareas, y complejidad de la arquitectura de las plataformas hardware. A su vez, una de las mayores problemáticas que afectan a estos sistemas de cómputo de grandes dimensiones es el elevado consumo energético y potencia eléctrica demandada.

Nuestro interés está centrado en dos líneas: aplicaciones de cómputo intensivo para la salud, y gestión del 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, las Unidades de Cuidados Intensivos (UCI) proveen cuidados continuos y rigurosos para personas adultas que están críticamente enfermas y que pueden beneficiarse de tratamiento, y dan un buen morir a los pacientes irrecuperables. En las Unidades de Vigilancia Intermedia (UVI), se atienden pacientes que no requieren de tratamientos de soporte vital pero que, por su condición, son necesarios mayores cuidados y controles de los que se reciben en una unidad de internación convencional. En ambas unidades (UCI y UVI), los datos de los pacientes involucran datos clínicos de baja frecuencia, y flujos de datos fisiológicos de alta frecuencia generado por el equipamiento médico (como monitores médicos de signos vitales, respiradores artificiales, etc.). En una UCI/UVI típica, los enfermeros completan manualmente datos en formularios, registrando datos clínicos y fisiológicos. Los datos fisiológicos se obtienen por observación de las pantallas del equipamiento médico (ubicados a los lados de las camas) a intervalos de tiempo (horas o fracciones) que definen los médicos para cada paciente. Cabe mencionar que el equipamiento médico emite alertas,

siempre limitadas a lecturas de sus propios parámetros (sin posibilidad de interconexión con otros equipos o de suministrarle información adicional para emitir alertas más inteligentes). Al final del día, los médicos analizan exhaustivamente los datos de los formularios y dan a los enfermeros las indicaciones de tratamientos a realizar a los pacientes. Todo este trabajo es llevado a cabo por enfermeros y médicos de manera manual, sin la asistencia de ningún sistema informático.

Los principales problemas que ocurren en las UCI y UVI típicas son:

1. La metodología actual de gestión de la información es muy propensa a errores humanos.

Causas:

- Podrían cometerse errores en el registro manual, que realizan los enfermeros, de datos observados en los monitores de signos vitales u otro equipamiento médico.

- El análisis realizado por los médicos, que interrelaciona una gran cantidad de variables, provenientes de enormes volúmenes de información de diferente tipo, origen y soporte (datos fisiológicos, radiografías, análisis de laboratorio, datos de observación del paciente, etc.) es una tarea sumamente compleja. Ella, sin asistencia informática alguna, es difícil de concretar sin cometer errores.

Efectos: Posibles diagnósticos erróneos o información inconsistente que implica un esfuerzo adicional del personal para detectar su origen.

2. Detección tardía del deterioro de la salud de los pacientes.

Causas: Desde que los datos se toman por los enfermeros hasta que son analizados por los médicos, puede transcurrir un tiempo prolongado.

Efectos: Incidencia negativa en la salud de los pacientes, que puede impactar en:

- Aumento del riesgo de muerte.
- Aumento de los costos económicos de los tratamientos.
- Estadías más largas, y por ende la atención de menos pacientes.

3. Pérdida de información de pacientes.

Causas: Los datos observados en el equipamiento médico es registrado en formularios a intervalos de horas (o varios minutos), perdiendo todas las muestras intermedias (por ejemplo, el monitor de signos vitales actualiza la frecuencia cardíaca una vez por segundo).

Efectos:

- Disminución de la precisión de los diagnósticos.
- Imposibilidad de registrar el conjunto completo de todos los datos medidos por el equipamiento médico para uso futuro. Disponer de un gran conjunto de datos históricos permite la extracción de nuevo conocimiento, útil para futuras investigaciones médicas y el descubrimiento de patrones para predicción de patologías. La predicción de patologías, como sepsis (una de las principales causas de muerte en UCIs), permitiría comenzar los tratamientos más temprano, reduciendo el riesgo de muerte, el costo económico de los tratamientos, y las estadías de los pacientes.

Un gran avance sería disponer de un sistema de cómputo que detecte el deterioro en la salud de los pacientes en tiempo real, basándose en múltiples parámetros de diferentes orígenes. La detección temprana del deterioro permitirá aumentar la efectividad de los tratamientos, y por consiguiente la mejora de la salud de los pacientes, la reducción del costo económico de los tratamientos, y el aumento del número de pacientes atendidos. El problema principal que debe enfrentarse para la construcción de este sistema, y que creemos viable con la aplicación de técnicas de computación paralela y distribuida, 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 [11, 3, 1, 2].

Consumo energético de sistemas de HPC

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 [12, 10].

El consumo energético es hoy en día un gran problema. Para dar una idea de la magnitud del

mismo, la máquina de mayor prestaciones de la actualidad, la máquina de EEUU Summit, demanda casi 10 MW de potencia, lo mismo que se requiere para abastecer a los hogares de una ciudad con alrededor de 135.000 habitantes (cálculo 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, y 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, y el consumo energético de los sistemas de HPC.

Aplicaciones para la salud

Esta línea está enfocada en el diseño y desarrollo de un sistema para detección automática y temprana del deterioro de la salud de pacientes en Unidades de Terapia Intensiva (UTI) y 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.

El sistema está basado en reglas clínicas. Cada regla define las condiciones, relacionando parámetros y valores, que deben cumplirse para ge-

nerar una alerta indicando posible riesgo (actual o futuro) en la salud de los pacientes. Cada paciente puede asociarse a un conjunto de reglas particulares, creadas desde el conocimiento de expertos o guías clínicas. El sistema adquiere datos desde sistemas de historia clínica electrónica y del equipamiento médico como los monitores de signos vitales. Las reglas clínicas asociadas a cada paciente son continuamente evaluadas (en tiempo real) y, si una de ellas concuerda, se emite una alerta.

La construcción del sistema involucra diferentes desafíos tecnológicos relacionados fundamentalmente a:

- Extracción de datos del equipamiento médico.
- Procesamiento en tiempo real de grandes volúmenes de datos (Big Data).

Consumo energético de sistemas de HPC

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 erróneos.
- Diseñar la infraestructura de un sistema de Big Data de tiempo real para el procesamiento de reglas médicas.
- Diseñar y desarrollar una aplicación para la interacción del sistema con médicos y enfermeros.

Los avances/resultados actualmente comprenden:

- Un análisis del estado general de las UTI del hospital Francisco Lopez Lima, y la propuesta del diseño de alto nivel del sistema [8].
- El diseño de la infraestructura del sistema de procesamiento de reglas y un prototipo, y el desarrollo de un dispositivo embebido adquirente de datos (actualmente obtiene información de la señal analógica del electrocardiograma desde un monitor médico) que envía datos a la plataforma de procesamiento por medio de una red de WiFi [4].
- Estudios para la realización de un dispositivo para obtener la saturación de oxígeno en sangre de los pacientes, y su integración con el dispositivo adquirente de datos.
- El desarrollo parcial de una aplicación que detecta complejos QRS de manera eficiente en señales de electrocardiogramas.

Consumo energético de sistemas de HPC

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) [7], 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, y en [6] expusimos un análisis de los factores que afectan el consumo energético de operaciones de checkpoint y restart en clusters. Actualmente, hemos agregado la predicción para la operación de restart (adicionalmente a la de checkpoint), y se contemplan distintas alternativas de configuraciones del sistema, relacionadas a: almacenamiento en NFS, configuración de la aplicación de checkpoint/restart (compresión/sin compresión), 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 local a la Universidad Nacional del Comahue, cuenta con un doctor en la temática (recibido en 2008). En Marzo de 2019 se ha presentado la tesis de grado “Una Infraestructura de Big Data de Tiempo Real para un Sistema de Detección del Deterioro de Pacientes en Terapia Intensiva”, en la Universidad Nacional del Comahue. En el transcurso del 2019 se espera la finalización de dos tesis de grado en temas de aplicaciones para la salud, y en 2020 una tesis doctoral en el tema de consumo energético de sistemas de HPC.

Referencias

- [1] Amara health analytics, <http://www.amarahealthanalytics.com> (accedido en marzo de 2019).
- [2] ehcos smarticu, <http://www.ehcos.com/productos/ehcos-smarticu/> (accedido en marzo de 2019).
- [3] Excel medical, <http://excel-medical.com/> (accedido en marzo de 2019).
- [4] Javier Balladini, Pablo Bruno, Rafael Zurita, and Cristina Orlandi. An automatic and early detection of the deterioration of patients in intensive and intermediate care units. *Journal of Computer Science and Technology*, 18(03):e25, Dec. 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, Marina Morán, Dolores Rexachs, and Emilio Luque. Factores que afectan el consumo energético de operaciones de checkpoint y restart en clusters. *XXIV Congreso Argentino de Ciencias de la Computación (CACIC 2018)*, 2018.
- [7] 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.
- [8] 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.
- [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] 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, Adolfo Hoisie, Dean Miron 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 Miron 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.
- [11] Carolyn McGregor. Big data in neonatal intensive care. *Computer*, 46(6):54–59, 2013.
- [12] 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.