



TECNOLOGÍA

Bits, Ciencia y Sociedad



sep 27 **Ciencia Computacional: manejando datos científicos masivos**

 Posteado a las 27 de Septiembre de 2011 - 14:34 | [0](#) comentario [Editar entrada](#)



 Por **Benjamín Bustos**, profesor del Departamento de Ciencias de la Computación, FCFM, U. de Chile.


Durante el mes de agosto tuve la oportunidad de participar en el "Pucón Symposium 2011: Advanced Mathematical Tools for Frontier Astronomy and other Massive Data-driven Sciences", un simposio multidisciplinario organizado por Eduardo Vera, Gerente de Innovación del **Centro de Modelamiento Matemático (CMM)**. Este simposio, que reunió a astrónomos, matemáticos, gente del área de biofísica y del área de Ciencia de la Computación, tuvo como objetivo **discutir los problemas asociados y técnicas modernas para el manejo de grandes volúmenes de datos astronómicos (y de datos científicos en general) que se están produciendo actualmente, y que van a seguir creciendo rápidamente en el corto plazo.**

Los problemas que desean resolver los astrónomos son, por ejemplo, encontrar supernovas, clasificar estrellas y galaxias, estudiar la variabilidad en la luminosidad de las estrellas, detectar fenómenos anómalos, etc., sobre datos que pueden poseer errores ("artefactos") que deben ser descartados. Todo esto se realiza analizando datos complejos, por ejemplo imágenes del cielo nocturno (en distintas frecuencias: luz visible, infrarrojo, rayos X, etc.) o series temporales con mediciones de algún fenómeno astronómico específico. Adicionalmente, una de las grandes dificultades que se avecina es la enorme cantidad de datos que es necesario analizar, lo que se traduce en problemas de transmisión, almacenamiento y procesamiento de volúmenes gigantes de datos, que son algunos de los típicos problemas que a los que investigamos en el área de Information Management nos encanta estudiar y resolver.

Por ejemplo, Henry Brandt de la IBM Systems & Technology Group mostró que **la estimación del crecimiento de la generación de datos (en general, no sólo astronómicos) para los próximos diez años era de un 4.400%, y que para 2013 se esperaba que el tráfico IP anual sea de medio Zettabyte** (un Zettabyte equivale a 1.024 Exabytes, que equivalen a más de un millón de Petabytes, que equivalen a más de mil millones de Terabytes; actualmente los discos duros más grandes pueden almacenar unos pocos Terabytes). Para el caso particular de los datos astronómicos, George Djorgovski del California Institute of Technology (Caltech) mostraba que en el mediano plazo (1 - 5 cinco años) los equipos astronómicos van a producir aproximadamente 1 Tb de datos diarios (~300 Tb de datos anuales), y que en diez años esta cantidad va a aumentar en treinta veces. Transferir dicho volumen de datos de un servidor a otro ya es técnicamente un problema complejo, analizar esa cantidad de datos y extraer información relevante en un tiempo "razonable" es todo un desafío!

Otra de las preguntas que se planteó durante el simposio fue, **¿cómo se puede traducir esta "explosión" de datos científicos en una "explosión" de descubrimientos científicos?** De acuerdo a Bruce Elmegreen del IBM T.J. Watson Research Center, para lograr esto es importante mejorar el background en Ciencia de la Computación (CS) de los estudiantes de Ciencias: **formar expertos en "computación científica" o "ciencia computacional"** que entiendan el problema en su área particular de la ciencia (astronomía, física, biología, etc.) y que además dominen las técnicas computacionales necesarias para poder analizar datos científicos masivos (data-driven science). Dentro de las materias de CS que se consideraron indispensables en una carrera de ciencia computacional están: Algoritmos, Minería de Datos en Archivos de Datos Masivos, Bases de Datos para "Large Data", Computación Paralela, Computación de Alta Performance, Visualización de Datos Complejos (multidimensionales), Simulación, e Ingeniería de Software para científicos. Lo mismo puede verse en el sentido inverso: **que haya un experto en Ciencia de la Computación interesado en resolver problemas de otra área de las ciencias, lo que hace que las oportunidades para los egresados de Computación se vean muy prometedoras.**

Tags: astronomía, benjamín bustos, biofísica, computación, matemáticas


[PERMALINK](#)

[BLOGS QUE CITAN ESTE POSTEO](#)


perfil del autor



El blog Bits, Ciencia y

Sociedad de la sección de

Tecnología de Terra es un

espacio donde académicos

del Departamento de Ciencias

de la Computación de la

Universidad de Chile hablarán

de la Tecnología y su impacto

político y social en nuestro país".

Aquí escribirán semanalmente José Miguel

Piquer, Claudio Gutiérrez, Juan Álvarez, Tomás

Barros, Nancy Hirschfeld, Benjamin Bustos,

Alejandro Hevia, Pablo Barceló y Cecilia

Bastarrica."

posteos

 VER: [MÁS RECIENTES](#) [MÁS COMENTADOS](#)

Ciencia Computacional: manejando datos científicos masivos

27 de Septiembre de 2011 - 14:34

Las competencias estudiantiles que llegaron para quedarse

21 de Septiembre de 2011 - 8:13

Egoísmo versus altruismo: ¿quiénes son los intransigentes?

13 de Septiembre de 2011 - 9:25

Locos por el lucro

30 de Agosto de 2011 - 10:54

¿Puede ser la programación entretenida para a niñas y niños?

23 de Agosto de 2011 - 10:48