

El cielo capturado en 500 petabytes

Un ingeniero de sistemas y computación uniandino participa como coordinador de informática de la contribución de Francia al procesamiento de datos del *Large Synoptic Survey Telescope* (LSST). El 5 de abril del 2017 estuvo en Los Andes y explicó la magnitud del proyecto.

Foto: NASA/ESA/Hubble

Durante 10 años, una cámara de 1,65 metros de diámetro instalada en el Gran Telescopio para Rastreo Sinóptico (LSST, por su nombre en inglés) tomará 2000 imágenes por noche del cielo visible desde el hemisferio sur. Cada 40 segundos, el aparato de 350 toneladas cambiará de posición para capturar nuevas imágenes. Al cabo de 4 noches, el telescopio habrá barrido la totalidad del universo observable en el norte de Chile.

Allí se construyen las instalaciones del LSST, que comenzará las operaciones en el 2022. Al finalizar el 2032, será posible apreciar la evolución de cerca de 37.000 millones de cuerpos siderales como galaxias, estrellas, planetas, nebulosas; de objetos transitorios como supernovas, cometas o asteroides y cuanto sea visible. Algunos son de luminosidad muy débil y no solo del espectro visible, sino también de luz cercana a la infrarroja.

Dos egresados del Departamento de Ingeniería de Sistemas y Computación (DISC) de Los Andes trabajan en este proyecto y participaron en el Coloquio Franco-colombiano de Astrofísica que tuvo lugar el 5 de abril del 2017, en el marco del año Francia-Colombia. Son Fabio Hernández, ingeniero de investigación del Institut National de Physique Nucléaire et de Physique des Particules (In2p3) que pertenece al Centre National de la Recherche Scientifique (CNRS), en Lyon (Francia), y coordina la computación de la contribución francesa al procesamiento de datos para el LSST; y Juan Pablo Reyes, quien cursa un doctorado en la Université d'Aix Marseille. Revista Foros ISIS publica un registro de la conferencia del primero y una entrevista con el segundo (ver página 8).

Fabio Hernández, quien ha trabajado durante 25 años en investigación, impartió una conferencia sobre el proyecto del telescopio LSST que captará las fotografías del

Para qué el *Large Synoptic Survey Telescope*

El objetivo del proyecto *Large Synoptic Survey Telescope* (LSST) es realizar un reconocimiento del cielo durante diez años. Las imágenes y datos capturados servirán para que, posteriormente, los científicos tengan a su disposición los más grandes catálogos astronómicos jamás compilados. Con ellos podrán estudiar la estructura y evolución del universo y lo que hay en él, en cuatro áreas de ciencia:

- > Comprender la materia y la energía oscuras.
- > Detectar asteroides peligrosos para la tierra y estudiar el sistema solar remoto.
- > Registrar el cielo óptico transitorio.
- > Estudiar la formación y estructura de la Vía Láctea.

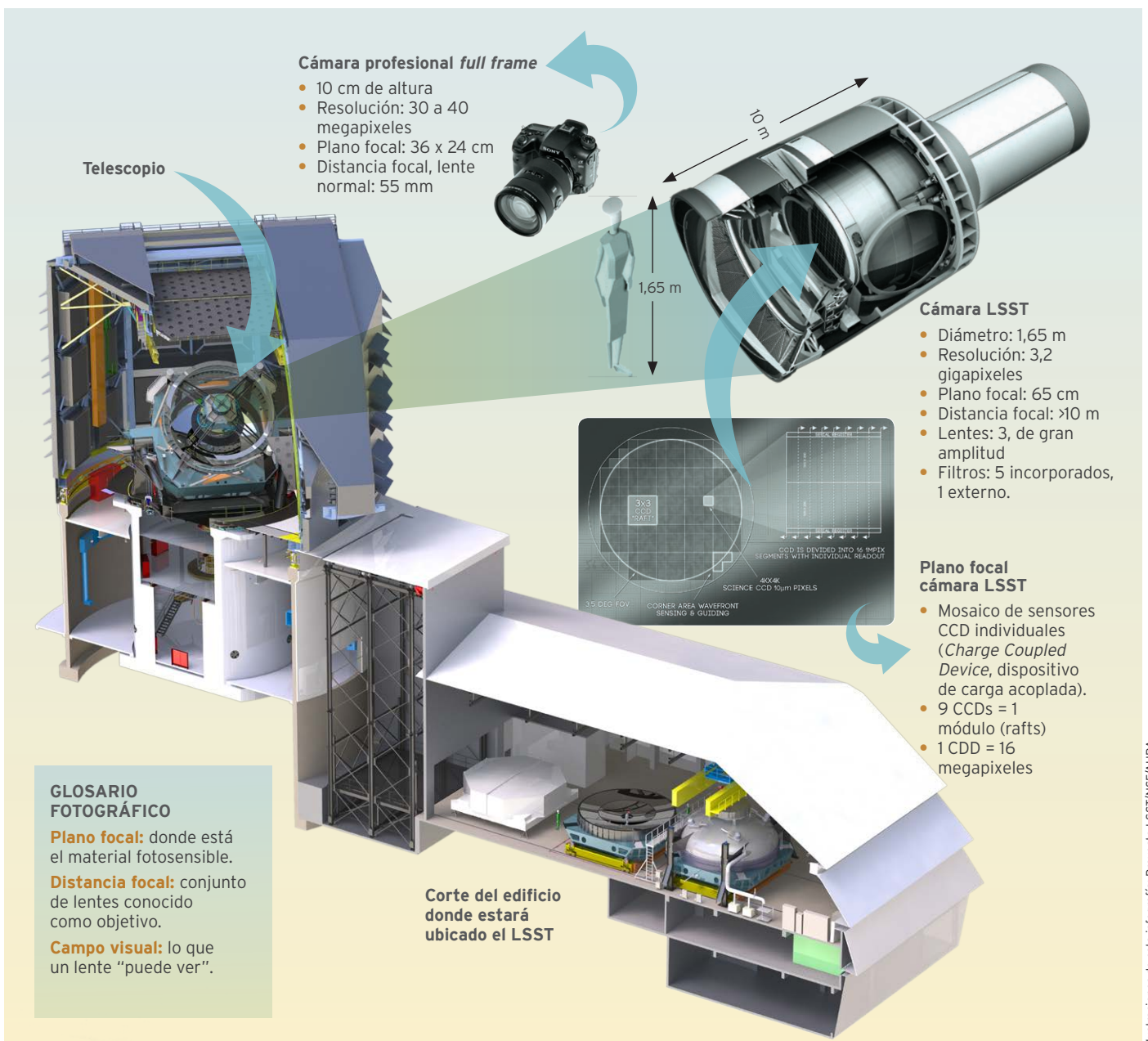
En dólares

El costo de construcción del proyecto (que comenzó en el 2014, aunque su concepción y diseño se inició hace 10 años) es de 600 millones de dólares, aproximadamente. El 20 % de esa cifra está dedicada al manejo de los datos. La cámara vale cerca de 117 millones de dólares y se calcula la operación del proyecto más o menos en 55 millones por año; 11 millones de dólares del presupuesto se han destinado a educación.

Los principales socios son Estados Unidos, Chile y Francia y la financiación proviene del National Science Foundation, del Ministerio de Energía de Estados Unidos, y de fondos conseguidos por la Corporación LSST, integrada por cerca de 40 miembros institucionales y 34 contribuyentes internacionales que representan a 23 países.

universo y sobre el trabajo de manejo de datos obtenidos en las imágenes, su procesamiento y almacenamiento para el posterior uso por los científicos.

Según Fabio Hernández, el propósito del LSST es componer el más ambicioso catálogo que se haya intentado, con las propiedades físicas de los objetos astronómicos. La cámara es un elemento clave para lograrlo y su instituto participó en el diseño del mecanismo de movimiento de los filtros, que se pueden intercambiar antes de tomar una imagen.



Ilustraciones de esta infografía: Proyecto LSST/NSF/AURA



Imagen: Proyecto LSST/NSF/AURA

Espejos

- El tamaño del espejo principal determina la cantidad de luz que captará y la distancia a la cual el telescopio puede observar objetos.
- Diámetro espejo 1: 8,4 m
- Diámetro espejo 2: 5 m
- Diámetro espejo 3: 3,5 m
- Peso: 5,9 toneladas
- Costo: 20 millones de dólares
- Fabricante: Universidad de Arizona

El manejo de datos, un asunto enorme

Noche tras noche, el proyecto requerirá una capacidad de almacenamiento de 15 *terabytes* para preservar 2000 imágenes brutas y, cada año, 4,5 *petabytes*. Un pixel será representado en 18 bits y una imagen ocupará cerca de 7 *gigabytes*. Al cabo de 10 años, se habrán producido cerca de 500 *petabytes* en imágenes brutas y derivadas y todos los

Universo por demanda

Una pieza central del *data management* será el catálogo, el punto de entrada probable para el estudio a partir de los datos. El investigador podrá preguntarle, por ejemplo, por las galaxias en una región del cielo. O por todos los objetos similares en términos de color, de longitud de onda, o por todos los cuásares con un desplazamiento hacia el rojo; o por todas las galaxias azules.

Cada año se espera obtener información sobre las observaciones y medidas de propiedades físicas de 37.000 millones de objetos, que estará disponible en computación por demanda, de tal manera que los científicos usuarios procesen los datos, los exploren y extraigan lo necesario para sus investigaciones.

catálogos de los objetos astronómicos ocuparán un poco más de 80 *petabytes*.

En el momento de su conferencia, Fabio Hernández estaba involucrado en el diseño de procesamiento de datos. “El manejo de este volumen, desde el punto de vista computacional, es uno de los asuntos más interesantes, pero todavía tenemos más preguntas que respuestas”, asegura. El *data management* tiene tres objetivos: Primero, obtener las imágenes brutas y registrarlas, transportarlas, almacenarlas y crear un archivo con todo el material recolectado durante 10 años. Este se transmitirá por fibra óptica desde el telescopio al sitio de base de La Serena y luego al centro de computación de la NCSA (Illinois). De ahí se enviará a Lyon (Francia).

El segundo objetivo se trabaja en tiempo real: consiste en tratar aquellas imágenes que revelan eventos transitorios, es decir, detectar los cambios del cielo al comparar las imágenes del mismo objeto con el fin de localizar acontecimientos como supernovas; 60 segundos después, se transmitirán alertas sobre estos hechos para que otros telescopios del mundo los estudien. Se calcula que se encontrarán cerca de 10 millones de eventos transitorios por noche.

El tercer objetivo es producir anualmente un conjunto de archivos fijos e inmutables que constituyen un *data release*: du-

Los desarrolladores del catálogo

El Stanford National Accelerator Laboratory (SLAC) está a cargo del desarrollo del *software* para el manejo del catálogo de los objetos celestes; el National Center of Supercomputing Applications (NCSA), en la Universidad de Illinois en Urbana-Champaign, se encarga del procesamiento y la producción anual del *data release*. La Universidad de Princeton y la de Washington desarrollan los algoritmos de ciencia y del tratamiento de las imágenes; el Infrared Processing and Analysis Center (IPAC), de la Universidad de Caltech, diseña el *software* de visualización y exploración de las fotografías; en Lyon, el Centre National de la Recherche Scientifique (CNRS) trabajará en el tratamiento masivo y el almacenamiento a largo plazo.



Fabio Hernández es ingeniero de investigación del Institut National de Physique Nucléaire et de Physique des Particules (In2p3).

rante cuatro meses se reprocesan todas las imágenes capturadas desde el inicio del sondeo hasta el año en curso, con los algoritmos afinados y puestos al día. Las imágenes tratadas y el catálogo producidos serán el material de trabajo de los científicos.

Para esto es necesario desarrollar *software*: para el tratamiento de imágenes, para transportarlas, para la detección de los objetos celestes, para el manejo del catálogo, para producir el *data release* y, finalmente, para publicarlos, de tal manera que los científicos los puedan usar. ■