

Clúster de software libre del Laboratorio Nacional de Observación de la Tierra

Información del reporte:

Licencia Creative Commons



El contenido de los textos es responsabilidad de los autores y no refleja forzosamente el punto de vista de los dictaminadores, o de los miembros del Comité Editorial, o la postura del editor y la editorial de la publicación.

Para citar este reporte técnico:

Aguilar Sierra, A. (2025). Clúster de software libre del Laboratorio Nacional de Observación de la Tierra. *Cuadernos Técnicos Universitarios de la DGTIC*, 3 (2) páginas(24 - 34).

<https://doi.org/10.22201/dgtic.30618096e.2025.3.2.107>

Alejandro Aguilar Sierra

Instituto de Ciencias de la Atmósfera
y Cambio Climático

Universidad Nacional Autónoma de México

asierra@unam.mx

ORCID: 0000-0003-1018-8521

Resumen:

Se implementó un clúster de procesamiento y distribución de datos satelitales para el Laboratorio Nacional de Observación de la Tierra, con el objetivo de contar con un sistema de alta disponibilidad para brindar datos de manera robusta y continua a los usuarios. El laboratorio recibe imágenes de satélite constantemente y con ellas se generan productos útiles para estudios territoriales, meteorología y prevención de desastres. El proceso de datos, desde la recepción, la generación de productos, el almacenamiento y la distribución a usuarios finales, requiere un sistema organizado, eficiente y tolerante a fallas. El clúster de procesamiento y distribución de productos satelitales, con base en software libre de código abierto, fue implementado para cubrir estas necesidades.

Palabras clave:

Clúster heterogéneo, cómputo de alto rendimiento, tolerancia a fallas, alta disponibilidad.

Abstract

A cluster for satellite data processing and distribution was implemented for the National Laboratory for Earth Observation, to provide a high-availability system so that robust and continuous data could be offered to users. The laboratory constantly receives satellite images, which are used to generate products that are useful for territorial, meteorology and disaster prevention studies. From reception, product generation, storage, and distribution to end users, data processing operations require an organized, efficient, and fault-tolerant system. A cluster for the processing and distribution of satellite products was implemented to meet these needs, based on open-source software.

Keywords:

Heterogeneous cluster, high-performance computing, fault tolerance, high availability

1. INTRODUCCIÓN

El Laboratorio Nacional de Observación de la Tierra (LANOT), cuya sede es el Instituto de Geografía de la Universidad Nacional Autónoma de México (UNAM), fue creado en 2017 con apoyo del entonces Consejo Nacional de Ciencia y Tecnología y de un consorcio de instituciones de la UNAM y externas, con el fin de proporcionar datos e información de satélite para prevención de riesgos, estudios territoriales y meteorológicos (Aguirre, 2018). Desde entonces ha crecido el número de socios del consorcio, con instituciones tan importantes para la seguridad nacional como el Centro Nacional de Prevención de Desastres y la Secretaría de la Marina. Conjuntando información de diversas fuentes en tiempo casi real, es posible coadyuvar a generar alertas tempranas y prevenir riesgos, monitorear tormentas, actividad volcánica, incendios y posibles accidentes industriales. Además, el LANOT es un espacio para la docencia y la difusión de las ciencias, al que continuamente se incorporan estudiantes, becarios y tesis.

El objetivo del proyecto presentado en este reporte técnico es contar con un sistema confiable de alta disponibilidad (*High Availability*) tolerante a fallas (*fail over*) y prevención de denegación de servicio (*Denial of service*), para brindar datos de manera robusta y continua a los usuarios del consorcio LANOT. En este trabajo se reporta el estado actual de un proyecto activo y en desarrollo.

2. DESARROLLO TÉCNICO

2.1 INFRAESTRUCTURA Y NECESIDADES

El LANOT cuenta con tres antenas (véase la Figura 1): una con dirección fija al satélite geoestacionario GOES-East (NOAA, 2017), que abarca el hemisferio americano; una móvil que sigue a algunos satélites de órbita polar como los JPSS de la Oficina Nacional de Administración Oceánica y Atmosférica (NOAA por sus siglas en inglés) (NOAA, 2024); y una pequeña antena para el sistema GeonetCast Américas (NOAA, 2023).

Figura 1

Antenas del LANOT ubicadas en la azotea del edificio anexo del Instituto de Geografía en Ciudad Universitaria.



Nota. De izquierda a derecha, antena móvil polar, antena de GeonetCast y antena para el satélite geoestacionario GOES-East.

La cantidad de información que se recibe diariamente es considerable, más de 1 TB, y debe ser procesada, publicada y almacenada. El clúster de software libre del LANOT, llamado así porque utiliza únicamente sistemas de código abierto de alta calidad, desarrollados en instituciones como la Universidad de Wisconsin en Madison o la misma UNAM y patrocinado por instituciones como la NOAA y la Administración Nacional de Aeronáutica y el Espacio, (NASA por sus siglas en inglés), fue implementado con este propósito.

Un clúster se compone de un conjunto de servidores, llamados nodos, que unen fuerzas para ejecutar un conjunto de tareas de manera eficiente y robusta. Entre las ventajas que proporciona un clúster están las siguientes:

- Alta disponibilidad: minimiza puntos débiles por medio de redundancia y replicación.
- Alto rendimiento: acelera procesos por medio de paralelismo, concurrencia y distribución de procesos.
- Balanceo de carga: distribuye tareas de acuerdo a la demanda, de modo que haya un equilibrio entre los nodos y ninguno se sature.

Dada la variedad de tareas que se requieren para proporcionar un buen servicio a los usuarios del LANOT y el enorme volumen de productos satelitales, se consideró que un clúster era la solución apropiada. Es heterogéneo porque fue construido en diferentes etapas desde mediados de 2018, con diferentes

marcas pero al mismo tiempo, sus necesidades son muy específicas, al contrario de un clúster de uso general y multiusuario, como los descritos en la implementación de Grid UNAM (Frausto, 2024) .

Cada nodo participa en alguna de las siguientes operaciones:

- 1. Adquisición.** Proporciona la materia prima de todo el proceso: los datos crudos obtenidos en las estaciones receptoras conectadas a las antenas externas.
- 2. Procesamiento.** Esos datos crudos deben procesarse para obtener productos útiles para distintas aplicaciones a partir de ellos. Para una descripción detallada de esta operación, así como los productos y sus niveles, ver la sección 2.4.
- 3. Publicación.** Los productos ya procesados se ponen al alcance de los usuarios del LANOT por FTP y al público en general por HTTP.
- 4. Almacenamiento.** Tanto los datos crudos como los productos de mayor nivel pueden almacenarse para su consulta posterior.

Aunque son de marcas diferentes y con hardware ligeramente distinto, todos los nodos tienen memoria RAM de 250 GB, arreglos RAID de discos con capacidad total de 20 a 40 TB, procesadores Intel Xeon con múltiples núcleos y dos puertos Ethernet de 10 Gb.

2.2 ANTECEDENTES

Originalmente se contaba solamente con los servidores bajo licencia comercial de la empresa *Sea Space*. A mediados de 2018 ya estaban saturados y no podían generar las animaciones que se muestran en la página principal del LANOT (<https://www.lanot.unam.mx/>), además de que la geolocalización presentaba errores que no se podían corregir porque no se contaba con el código fuente. En 2018 se agregó un primer servidor de software libre con sistema operativo GNU/Linux con el que se empezaron a crear imágenes con geolocalización correcta y animaciones de disco completo y de Estados Unidos de América (EUA) continental (véase la Figura 3). Más tarde en ese mismo año se construyó el sistema de almacenamiento masivo (Aguilar, 2020). Con este sistema se implementó una red local con un *switch* de alta velocidad de 10Gb, donde posteriormente se armaría el clúster.

Antes de la pandemia ya se contaba con dos servidores de software libre: el principal se dedicaba al procesamiento de imágenes del satélite GOES-16 y el secundario al procesamiento de las imágenes polares, pero ambos eran capaces de asumir las tareas del otro en caso de falla o mantenimiento, con lo que ya se contaba con un método básico de redundancia y disponibilidad.

Durante la pandemia se adquirieron nuevos servidores, con lo que se tuvo el suficiente número de nodos para armar un sistema con base en clúster. Al principio se implementó un sistema sofisticado con *Pacemaker* (Pacemaker, 2024) como gestor del clúster y *GlusterFS* (Gluster, 2024) como un sistema de archivos distribuido, que funcionó en pruebas. Al poner esos servidores en producción, empezaron a fallar, y ante la urgencia de mantener el servicio activo, se tomó la decisión de simplificarlo y armar un sistema con prestaciones similares pero sin herramientas sofisticadas, programas en shell en lugar del *Pacemaker* y el clásico sistema de archivos en red NFS en lugar de *Gluster*. Este sistema con base en recursos convencionales es lo que se reporta en este trabajo.

2.3 OPERACIÓN

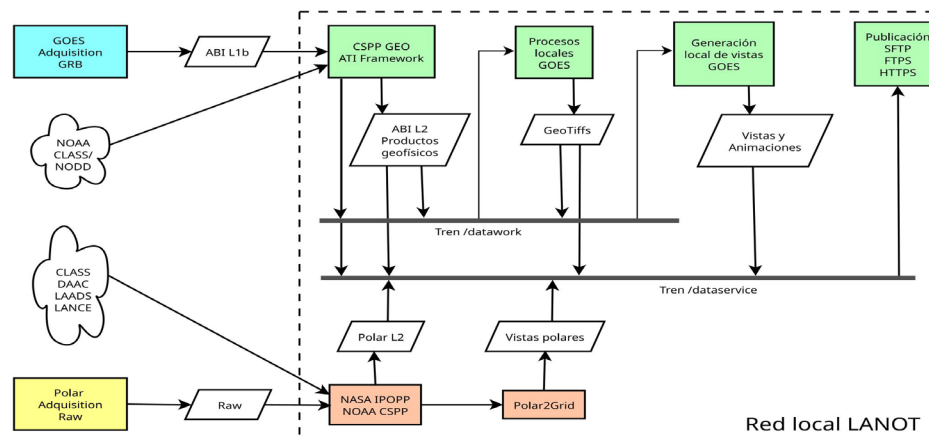
Todos los nodos usan una versión reciente del sistema operativo GNU Linux y se comunican entre sí sin necesidad de contraseña, por medio de llaves públicas con *secure shell*. Un solo usuario virtual realiza todas las operaciones por medio de un conjunto de programas y scripts ejecutables en carpetas replicadas en todos los nodos, con control de versiones *git* en el nodo principal. Los scripts están principalmente escritos en Bash y Python3 y los programas ejecutables en los lenguajes C y C++. El área de trabajo local de cada nodo es su arreglo de discos RAID con 20 a 40 TB de capacidad total.

Al realizar cualquiera de las operaciones descritas en la sección 2.1, los nodos producen nuevos datos que son almacenados. Todos tienen acceso a dos discos virtuales distribuidos, montados por NFS, llamados trenes porque surten de materia prima y transportan productos generados en los distintos nodos (véase la Figura 2). El primero de ellos se llama */datawork* porque se usa principalmente para trabajo de procesamiento de datos. En su carpeta *input* se guardan los productos de bajo nivel y conforme se generan productos de mayor nivel, se almacenan en la carpeta *output*. En el segundo tren, llamado */dataservice* pues se dedica al servicio de datos, se almacenan todos los productos que serán publicados. Los nodos de publicación ponen el contenido de */dataservice* accesible al exterior del clúster por medio de los protocolos FTPS, SFTP y HTTP.

Un primer uso de redundancia en el sistema permite reconstruir cualquiera de los trenes en caso de falla, pues todos los productos generados en los nodos también se encuentran en sus respectivos espacios de disco locales. Toda esta información almacenada en esta parte del sistema es efímera ya que se está renovando constantemente, y sólo permanece unos días o un par de semanas, a lo mucho, pero un conjunto de datos e imágenes selectos se almacenan a largo plazo en el sistema de almacenamiento masivo.

Figura 2

Diagrama de flujo de datos del LANOT



Nota. Del lado izquierdo se muestran los sistemas propietarios de adquisición conectados directamente a las antenas externas, así como proveedores externos de datos auxiliares. Dentro del rectángulo punteado se muestran todos los procesos en la red local aislada y el uso de los trenes de datos. Los únicos servidores con acceso a la red externa son los de publicación.

Toda la operación del sistema es automatizada, impulsada por eventos de recepción de imágenes, que ocurren cada 10 minutos (disco completo), 5 minutos (sector EUA continental) y 1 minuto en el caso geoestacionario, y a intervalos de horas o días en el caso polar. Todas las actividades se monitorean por medio de eventos programados a distintos horarios y cada uno genera un mensaje local de correo electrónico, lo que constituye en conjunto una bitácora del sistema.

En las siguientes secciones se detallan las operaciones definidas en la sección 2.1, con excepción de adquisición, realizada por el sistema propietario TeraScan en los servidores de las estaciones receptoras conectadas a las antenas.

2.4 PROCESAMIENTO

No se pueden usar directamente los datos crudos que se reciben de las antenas, por lo que deben ser procesados. De acuerdo con las convenciones en Percepción Remota (NASA, 2025), cada etapa de procesamiento genera productos de mayor nivel. Los de nivel 1 son usualmente datos ya legibles pero aún no muy útiles pues están en unidades de radiancia. Los de nivel 2 ya están calibrados y georreferenciados y convertidos a unidades geofísicas como albedo, temperatura o presión. Los de nivel 3 ya pasaron por un proceso de control de calidad y están mapeados en una cuadrícula uniforme y una proyección geográfica adecuada a su aplicación. Los de nivel 4 fueron combinados con modelos y datos de otros instrumentos.

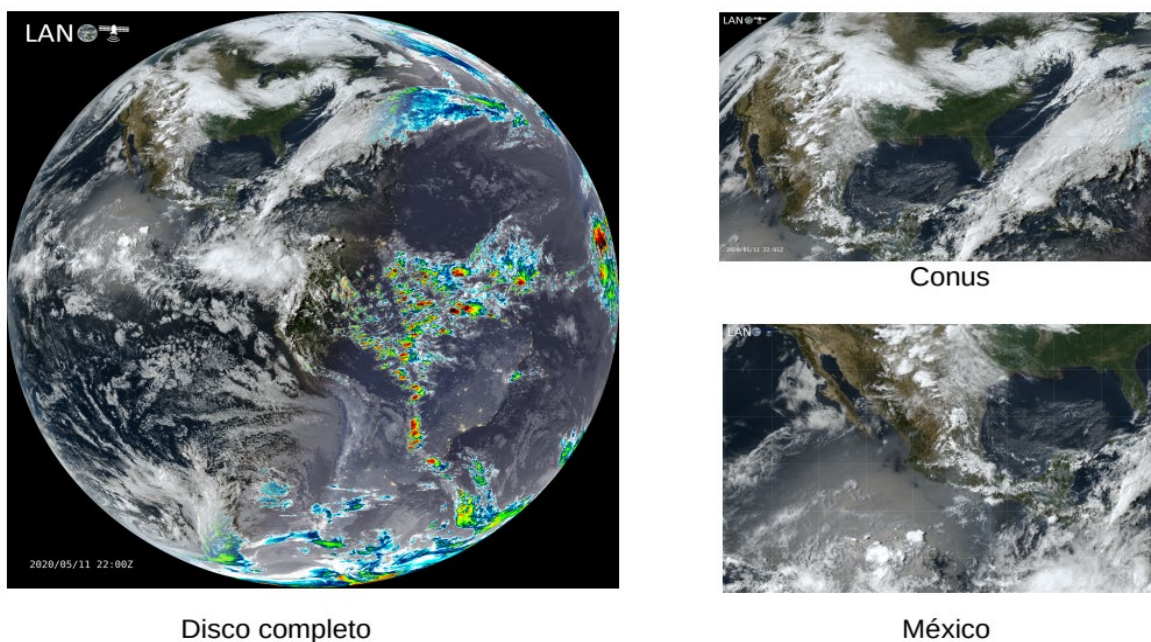
Las vistas son productos finales que ya no se pueden procesar más, exclusivos para visualización humana. En la Figura 2 se ilustra el proceso de cómo, a partir de datos crudos o de nivel 1, se generan productos de mayor nivel y vistas útiles. Una descripción más detallada del conjunto de productos que se generan en el LANOT, se encontrará en una nota técnica actualmente sometida a la revista Investigaciones Geográficas (Aguilar y Jiménez, 2025).

Los paquetes de software libre más importantes usados en el LANOT son las bibliotecas NetCDF, GDAL (Warmerdam, 2024), los Paquetes de Procesamiento Satelital de la Comunidad (CSPP GEO, 2024 para procesar imágenes geoestacionarias, CSPP LEO, 2024 para imágenes polares), desarrollados por la NOAA y la UW Madison y el Paquete Internacional de Procesamiento de Observación Planetaria (IPOP por sus siglas en inglés, detenido en 2024), desarrollado por la NASA. Dentro de lo posible, todas las tareas están configuradas para aprovechar los núcleos de los servidores, lo que confiere alto rendimiento con paralelismo local (no distribuido entre servidores).

También se desarrollaron programas originales para generar versiones de los mismos productos en el formato GeoTiff (GeoTiff, 2024) y reproyectados, que son más útiles para algunos usuarios que el original NetCDF, por ejemplo, para uso en sistemas de información geográfica, así como recortes de diferentes sectores de interés. Se generan diversas vistas, como los compuestos RGB (combinaciones de distintas bandas para cada color primario) notablemente en color verdadero, que se publican en disco completo cada 10 minutos, y en CONUS (región continental de EUA que incluye a casi todo México) cada 5 minutos, tanto en imagen fija como en animación de las últimas 24 horas, en su proyección geoestacionaria original. También se generan recortes derivados del disco completo y reproyectados de distintos sectores de interés, como los sectores A1 a A5 de la Secretaría de Marina, como se muestra en la Figura 3.

Figura 3

Compuesto RGB resultado de la combinación de las imágenes nocturna y color verdadero, en disco completo, Estados Unidos continental y uno de los recortes en proyección geográfica que abarca México



Actualmente, cuatro nodos dentro del clúster se dedican a tareas exclusivas de procesamiento. Todos tienen una configuración similar, con el mismo hardware, por lo que son intercambiables. Si cualquiera de ellos falla o se desactiva para mantenimiento, sus tareas son asumidas por los nodos restantes. Conforme se adoptan programas con requerimientos más especializados, está en consideración como mejora futura el uso de contenedores tipo docker, con la misma filosofía de que pueda activarse en cualquiera de los nodos de procesamiento.

2.5 PUBLICACIÓN

Los productos procesados son puestos a disposición de los usuarios por medio de nodos especializados en publicación. Todos los nodos de publicación se conectan por un lado a la red local privada del LANOT y por otro a la red externa de 10Gb de la UNAM. El principal método de tolerancia a fallas es la redundancia pues las tareas las realizan dos nodos idénticos en cada caso.

Dos nodos idénticos se especializan en FTPS. Se implementó una configuración a modo del servidor PureFTP (PureFTP, 2024) por eficiencia y seguridad. Si falla uno de los dos, el que queda ofrece exactamente los mismos servicios. También provee balance de carga manual. Si un usuario encuentra uno de los servidores saturado, puede intentar entrar al otro. Cada nodo admite un máximo de 70 usuarios. El balance de carga automático se había implementado con Pacemaker y funcionaba asignando la IP pública al servidor menos activo, pero se suspendió su uso. Una mejora futura obvia será reimplementarlo junto con otras prestaciones que puede ofrecer un gestor de clúster moderno.

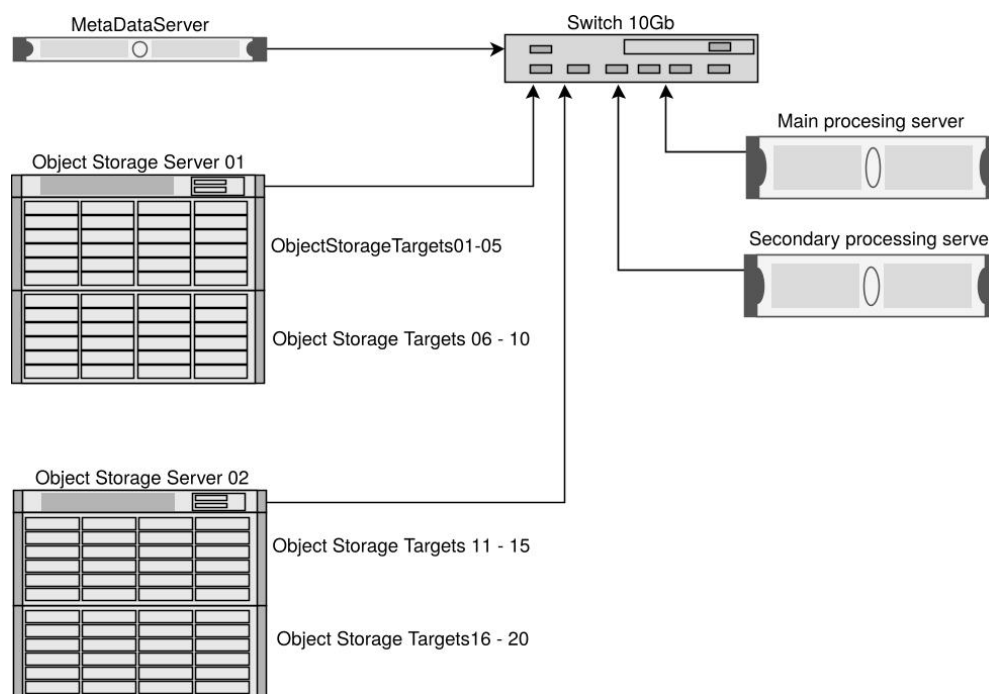
Otros dos nodos se especializan en publicación web con los protocolos HTTP y HTTPS. Usan Apache 2, Postgresql, Django para la página principal interactiva y GeoServer para visualización interactiva de mapas. Si cae uno de los servidores, el otro puede asumir sus funciones, pero al igual que en los servidores FTP, aún no está automatizado su sistema de balance de carga y queda para mejora futura.

2.6 ALMACENAMIENTO

La información guardada en el Sistema de Almacenamiento es persistente, no se borra hasta que sea necesario. El criterio para almacenar información en este sistema es que sean productos originales adquiridos por antena o internet, o que sea costoso volverlos a generar en procesamiento. Se usa un sistema de archivos distribuido con base en Lustre y sus componentes se ilustran en la Figura 4.

Figura 4

Componentes y diagrama de red del Sistema de Almacenamiento del LANOT



Se guardan 24 discos físicos de 12 TB en un gabinete o “cajón” y se organizan en arreglos RAID, cada cajón cuenta con una tarjeta controladora. El sistema Lustre se encarga de organizar estos arreglos para lograr una capacidad total aproximada de un PB con 4 cajones (una explicación más detallada de su armado y funcionamiento se encuentra en Aguilar, 2020).

2.7 SEGURIDAD

El *site* donde se encuentran las estaciones receptoras y los racks con los servidores, se encuentra en el primer piso del edificio anexo del Instituto de Geografía. Aparte de que el acceso al edificio está controlado con una estación de vigilancia, el *site* cuenta con acceso controlado mediante reconocimiento facial y de huella dactilar.

Se implementaron medidas de seguridad para los siguientes riesgos:

- Falla eléctrica. Se cuenta con generadores con base en gasolina, que se activan a los pocos segundos de que haya un corte eléctrico. Dentro del site, se cuenta con UPS redundantes (dos por rack) con batería de media hora que entran de inmediato ante el evento de una falla eléctrica, ya sea corte o descarga, y avisan por correo electrónico cualquier situación. Todos los nodos tienen fuentes redundantes, conectadas cada una a un UPS.
- Falla en discos. Los discos duros del sistema de almacenamiento se organizan con una tarjeta controladora en un arreglo RAID 6 que incluye dos discos de paridad y uno de repuesto por cajón, por lo que puede soportar un máximo de tres discos fallidos. Se monitorea automáticamente cada hora y cuando se daña un disco se envía un reporte por correo electrónico. Si se dañan más de dos discos en el mismo gabinete, se detiene el sistema para evitar daños mayores. Los servidores cuentan también con un sistema local que va de 4 a 10 discos organizados en arreglo RAID 5, lo que permite tener discos virtuales de 20 o 30 TB que soportan que falle un disco físico.
- Prevención de ataques. Los nodos con acceso al exterior tienen una estricta cortina de fuego en su sistema operativo, administrado con *nftables*, que permiten acceso solamente a los puertos autorizados para comunicación con SSH, FTP o HTTP y HTTPS. La mayoría de los nodos están en red local y no es posible entrar a ellos desde el exterior. El servicio *sshd* está estrictamente configurado para reducir riesgos, como por ejemplo, intentos de acceso directo al usuario *root*, lo cual no es posible. Para descarga de datos, sólo es posible entrar por FTP seguro, ya sea SFTP (basado en *ssh*) o FTPS (con base en TLS), y cuenta además con mecanismos para prevenir ataques de denegación de servicio y saturación; hay un límite preestablecido de accesos simultáneos, no se permite la entrada al usuario anónimo, es preciso contar con un usuario y una contraseña y, además, la IP desde la que se quiera entrar debe estar en una lista restringida de IPs.

3. RESULTADOS

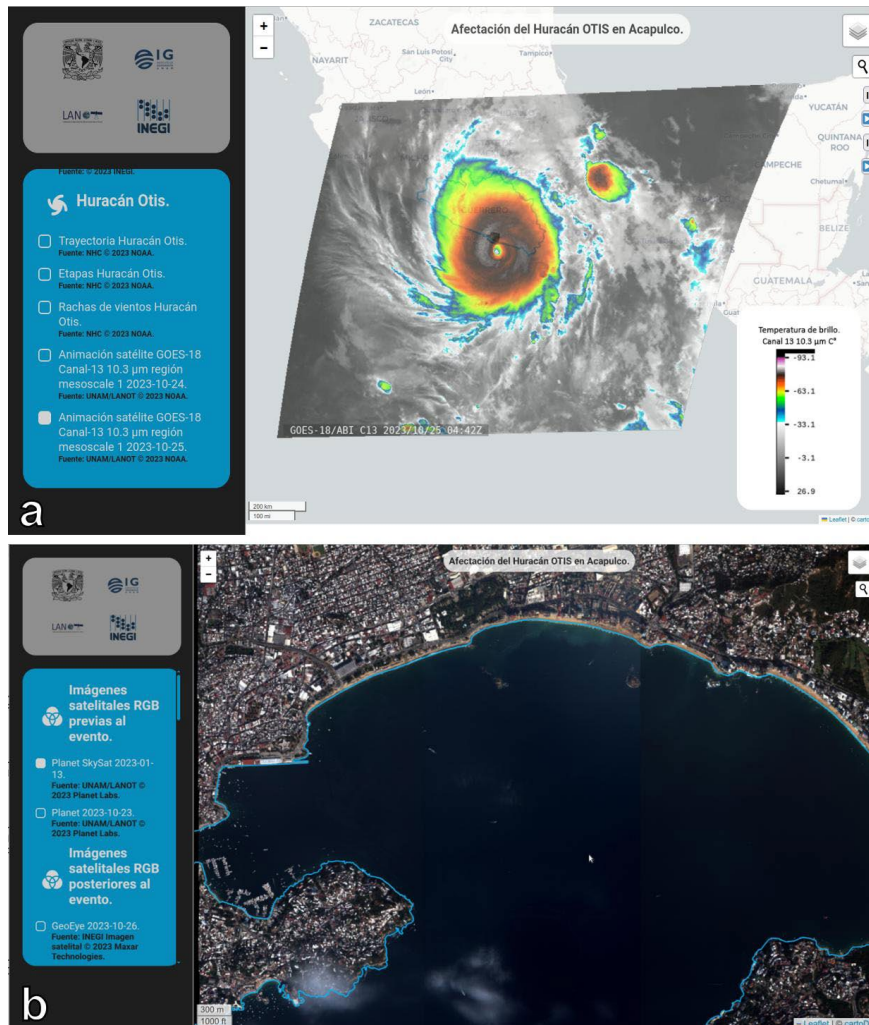
El funcionamiento del clúster permite la distribución de más de 100 productos satelitales de varios niveles, que pueden usarse en tareas de monitoreo ambiental, meteorología, prevención de desastres y estudios ambientales.

El LANOT es sede de los siguientes proyectos activos en los que participan distintas instituciones:

- Monitoreo de ceniza volcánica.
- Monitoreo del sargazo en el Caribe mexicano.
- Detección de puntos de calor para la prevención y el combate de incendios forestales.
- Monitoreo de tormentas tropicales. Un caso especial fue el huracán Otis que devastó Acapulco en octubre de 2023. En la Figura 5 se muestra el visualizador interactivo con el que se puede ver el mismo fenómeno con diferentes niveles de resolución, desde el satélite geoestacionario (5a, 1 km por pixel) hasta un detalle de 0.5 m con la constelación de satélites Planet (5b).
- Monitoreo de los mares territoriales de México.

Figura 5

Ejemplo de cobertura de eventos importantes, como el huracán Otis que devastó Acapulco en el otoño de 2023.



Nota. Se pueden ver vistas de diversas escalas, desde los 2 km por píxel del sector de mesoescala del GOES (a) hasta los mosaicos de 0.5 m por píxel de la constelación Planet (b), así como animaciones, vistas nocturnas y comparaciones de antes y después del evento.

Además de los datos recibidos directamente por las antenas, algunos de estos proyectos usan datos de otros satélites, como los mencionados de la constelación *Planet* o, en particular para el monitoreo del sargazo, imágenes de alta resolución de la constelación Sentinel, de la Unión Europea. En el caso de almacenamiento, antes de que se armara la infraestructura descrita en este reporte, simplemente no era posible almacenar la información recibida y procesada y debía desecharse a los pocos días.

4. CONCLUSIONES

Desde la introducción del primer servidor de software libre en el LANOT, en 2018, hasta la puesta en marcha del actual clúster, los mecanismos para procesar y distribuir datos de origen satelital han mejorado continuamente y atienden una demanda de información cada vez mayor. La construcción de un clúster heterogéneo con base en software libre ha permitido mantener el sistema actualizado al ritmo de las nuevas tecnologías a un costo bajo y satisfacer la demanda de productos satelitales para usuarios del LANOT.

En este trabajo se reportó el estado actual del clúster, cuya prioridad ha sido mantener el suministro de datos continuo y oportuno a los usuarios del consorcio. Entre las mejoras futuras se contempla utilizar contenedores que simplifiquen la administración de los nodos y adoptar un gestor de clúster apropiado a las necesidades del LANOT tipo *Pacemaker* o alguno más actual.

REFERENCIAS

- Aguirre, R. (2018). Laboratorio Nacional de Observación de la Tierra (LANOT). Investigaciones Geográficas, 96.
- Aguilar, A. (2020). Sistema de Almacenamiento Masivo para el Laboratorio Nacional de Observación de la Tierra (LANOT). Especialidad en Cómputo de Alto Rendimiento, IIMAS, UNAM.
- Aguilar, A. y Jiménez, V. (2025). Implementación de una infraestructura de procesamiento satelital con software libre en el Laboratorio Nacional de Observación de la Tierra. Sometido a Investigaciones Geográficas.
- CSPP Geo (2024). Community Satellite Processing Package for Geostationary Data. University of Wisconsin-Madison, Space Science and Engineering Center (SSEC). Consultado el 12 de noviembre de 2024. <https://cimss.ssec.wisc.edu/csppgeo/>
- CSPP Leo (2024). Community Satellite Processing Package for Polar orbiting satellite data processing. University of Wisconsin-Madison, Space Science and Engineering Center (SSEC). Consultado el 12 de noviembre de 2024. <https://cimss.ssec.wisc.edu/cspp/>.
- Frausto del Río, S. E. et al. (2024). Grid UNAM, la experiencia en su implementación. Cuadernos Técnicos Universitarios de la DGTIC, 2 (4).
- Gluster (2024). Un sistema de archivos de red escalable. Consultado el 12 de noviembre de 2024. <https://www.gluster.org/>
- NASA (2025). ARSET - Fundamentals of Remote Sensing. Consultado el 8 de mayo de 2025. <https://appliedsciences.nasa.gov/get-involved/training/english/arset-fundamentals-remote-sensing>
- NOAA (2017). GOES-R Mission Overview. Consultado el 10 de noviembre de 2024. <https://www.goes-r.gov/mission/mission.html>
- NOAA (2024). The Joint Polar Satellite System. Consultado el 10 de noviembre de 2024. <https://www.noaasis.noaa.gov/POLAR/JPSS/jpss.html>
- Pacemaker (2024). Pacemaker, un sistema de gestión de clústers altamente disponible. Consultado el 12 de noviembre de 2024. <https://clusterlabs.org/projects/pacemaker/>
- Warmerdam, F. et al. (2024). Geospatial Data Abstraction Library, GDAL. Consultado el 6 de noviembre de 2024. <https://gdal.org/>