

Un data center para big data astronómico



Con una capacidad de 1 petabyte, permite almacenar cinco años de información captada por el Observatorio ALMA en Chile



El Campus San Joaquín de la Universidad Técnica Federico Santa María (USM) en Chile acoge desde 2015 un centro de proceso de datos para la astroinformática. Este data center nace fruto del acuerdo entre la USM y la Academia China de Ciencias, junto con la empresa Huawei Chile.

Según indican desde la propia Universidad, es el CPD de mayor capacidad del país y uno de los más grandes de Latinoamérica, y albergará el primer Observatorio Virtual Chileno (ChiVO), una plataforma de astroinformática para administrar y analizar la ingente cantidad de datos proveniente de los observatorios astronómicos repartidos por el país andino.

"ChiVO nos permitió ingresar como miembros a la Alianza Internacional de Observatorios Virtuales (IVOA por sus siglas en inglés), y en ese contexto es importante contar con un centro de datos que esté operativo 24x7 todos los días del año", explica Mauricio Solar, subdirector de vinculación con el medio del departamento de informática de la Universidad.

Este centro de datos es uno de los más grandes en Latinoamérica, con una capacidad de 1 petabyte que permite almacenar cinco años de información captada por el Observatorio



Rittal – The System.

Faster – better – everywhere.



¡Enhorabuena por estos 10 años!

ENCLOSURES

POWER DISTRIBUTION

CLIMATE CONTROL

FRIEDHELM LOH GROUP



ALMA (Atacama Large Millimeter/submillimeter Array), que equivalen a más de 30 años de grabación de videos en alta definición.

Según Solar, es importante comentar como antecedente que Chile tiene acceso al 10% del tiempo de observación en los telescopios instalados en el país. En 2020, Chile tendrá instalado casi el 70% de la capacidad de observación astronómica mundial, consolidándose como potencia y transformándose en la capital en términos de capacidad de observación astronómica.

Estos observatorios producirán grandes volúmenes de datos. "Por este motivo nos propusimos alojar estos datos astronómicos localmente en el país, y evitar pérdidas de tiempo innecesarias para los astrónomos que los requieren en Chile. No tiene sentido que todos los datos astronómicos obtenidos en el país viajen hacia el hemisferio norte para después traerlo de vuelta a Chile, siendo que nuestro data center ofrece una capacidad inigualable de almacenamiento", asegura Mauricio Solar.

Falta de espacio. Para el centro de proceso de datos se eligió un data center modular, que se ubicó en el Campus de San Joaquín. Había diversas razones para la elección, pero la fundamental era la falta de espacio físico para instalar un data center indoor. "En las universidades a veces no es fácil conseguir un espacio para instalar un laboratorio, y nuestro caso no es la excepción. La solución de un data center modular calzaba con nuestra disponibilidad de espacio. No es necesario construir un edificio con las capacidades exigidas por los estándares tipo Tier, sólo se

requiere una conexión de fibra óptica, una conexión trifásica, una cañería simple de agua, y eso es todo", afirma Solar.

Según el investigador, además de la rapidez de instalación o el sistema de monitoreo remoto, destacan sus estándares de seguridad: la resistencia al agua (lluvia e inundaciones), al fuego, puertas de seguridad, etc.

En cuanto al apartado de TI, dispone de 1 Petabyte de capacidad de almacenamiento, lo que constituye la característica más importante del data center, dado que fue dimensionado como un data center de storage.

Deep learning astronómico. Todas las aplicaciones del data center son de búsqueda y procesamiento inteligente de datos astronómicos. El proyecto cuenta con varias publicaciones científicas reportando los resultados obtenidos en la implementación de machine learning, especialmente deep learning, y diversas técnicas de inteligencia artificial para procesar datos astronómicos. "Uno de los problemas que hemos estado enfrentando en el último tiempo es la detección y la clasificación de líneas moleculares en los datos de ALMA. No es un problema simple de resolver, y tienen diversas dificultades para realizarlo en forma automática", describe Solar.

Como se trata del procesamiento de grandes volúmenes de datos (big data), deben ofrecer la capacidad de procesamiento en forma local. Esto requiere desarrollar aplicaciones con técnicas de procesamiento paralelo y distribuido para optimizar el uso de los recursos computacionales disponibles en el data center. ●

Proyectos de futuro

- Segunda etapa del proyecto, en la que se está ampliando la disponibilidad de datos astronómicos de otros observatorios en Chile, tales como Gemini, Paranal, y los otros ciclos de datos públicos de ALMA.
- Desarrollo de nuevos servicios de procesamiento inteligente de datos, tales como detección y clasificación de líneas espectrales y moleculares en forma rápida y confiable.
- Mejora de los servicios de búsqueda con la elaboración de índices apropiados en astronomía.
- Aceleración de los procesamientos con técnicas de paralelización de código para ofrecer los servicios en la nube (cloud computing) con buenos tiempos de respuestas, es decir, que sea imperceptible para el usuario.



Edge Data Center

- Escalable de 2, 4 y 6 racks
- Infraestructura completa del Data Center (rack, energía, climatización, monitorización, detección y extinción de incendios)
- Ahorro de tiempo (data center plug & play)
- Diferentes opciones de instalación: edificio existente, sala de seguridad certificada, container

IT INFRASTRUCTURE

SOFTWARE & SERVICES

www.rittal.es

