

Técnicas de Recuperación y Procesamiento de Grandes Volúmenes de Datos Sísmicos. Un Repositorio Público Basado en Serverless.

Marcelo Moreno, María Murazzo, Nelson Rodríguez, Ricardo Sifón, Valeria Nicolía, Ignacio Benemerito, Leonardo Celador

mmoreno@inpres.gob.ar, marite@unsj-cuim.edu.ar, nelson@iinfo.unsj.edu.ar, rsifon@inpres.gob.ar, vnicolia@inpres.gob.ar, ignacio.benemerito@gmail.com, leonardomiguelcelador@gmail.com

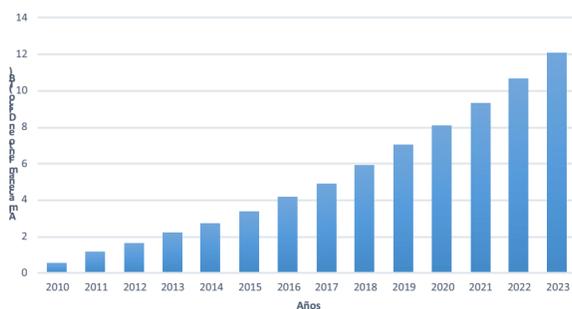
Contexto

El presente trabajo se encuadra dentro del área de I/D Sistemas Distribuidos y Paralelos y es una de las líneas de investigación internas del proyecto: Soporte Serverless para aplicaciones móviles de nueva generación, cuya propuesta está en etapa de evaluación para el período 2023-2024. El grupo de investigadores viene trabajando en proyectos relacionados con la computación móvil y distribuida desde hace más de 22 años.

Línea de Investigación

La presente línea de investigación surge como resultado de la cooperación obtenida con el INPRES gracias a que un docente investigador perteneciente al grupo de trabajo desempeña tareas en dicha institución. Se contará con la colaboración de profesionales del INPRES, gracias a lo que el grupo de trabajo tendrá acceso a una importante fuente de datos para poder realizar las tareas necesarios para trabajar con un back end serverless.

Tendencias de formas de onda continuas y de eventos. INPRES 2010-2022

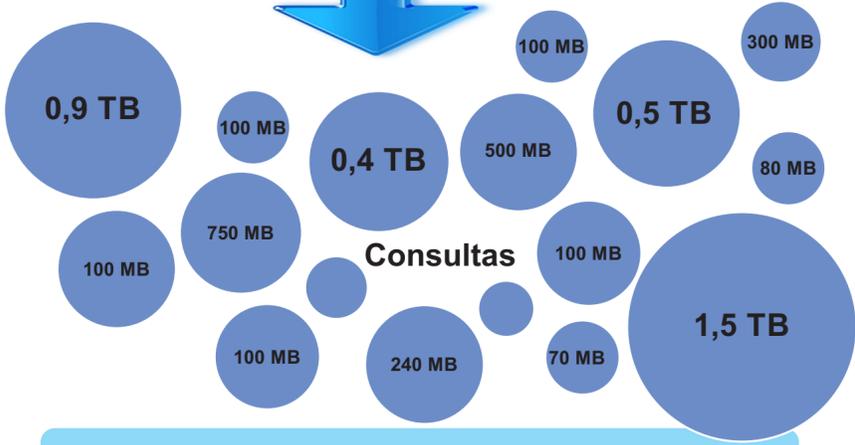


Repositorio Público

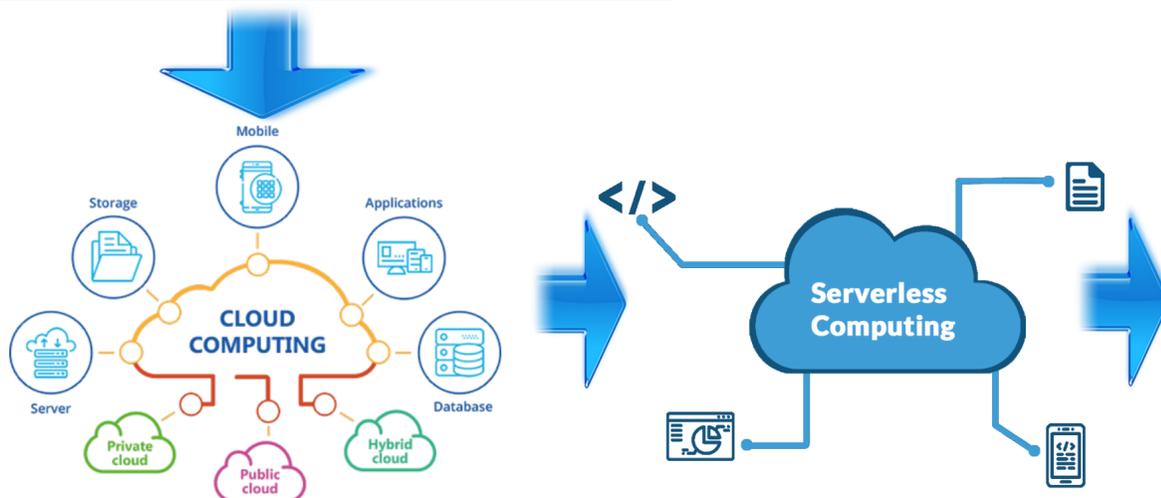
Catálogo

Forma de Onda

Archivos de Respuesta Instrumental



Los pedidos de datos, sobre todo de formas de onda, suponen un importante esfuerzo de cómputo y generalmente se refieren a grandes volúmenes de datos



Se hace necesario contar con una solución que provea el acceso a los datos y al análisis de los mismos en forma eficiente.

Debe haber un mínimo esfuerzo por parte del usuario y con escasa administración y gestión de la plataforma de soporte.

Se plantea usar una infraestructura cloud pública para alojar los datos y FaaS para implementar las técnicas de manipulación de los datos.

Objetivos

Desde el punto de vista institucional:

1. Proveer un entorno en el cloud que permita dar cumplimiento con la Ley 26.899 de Creación de Repositorios Digitales Institucionales de Acceso Abierto, para los datos de forma de onda sísmica disponibles en INPRES.
2. Proveer alta disponibilidad y escalabilidad de datos de forma de ondas sísmicas en el cloud, que facilite investigaciones que utilizan procesamiento computacionalmente intensivo y grandes volúmenes de datos sísmicos.

Desde el punto de vista científico:

1. Analizar la factibilidad de manipular grandes volúmenes de datos con FaaS.
2. Estudiar las técnicas de analítica de datos sobre FaaS
3. Evaluar la performance de serverless cuando procesa grandes volúmenes de datos.

Formación de Recursos Humanos

El equipo de trabajo de esta línea de investigación está compuesto de seis investigadores que figuran en este trabajo de las universidades Nacional de San Juan y un alumno de grado.

El proyecto marco ha establecido vínculos con investigadores de la Nacional de San Luis, de la Universidad Champagnat y de la Universidad Nacional de Salta y dos alumnos de grado.

Se está desarrollando una tesis doctoral sobre paralelismo híbrido y Big Data, una tesis de maestría en áreas afines y dos tesinas de grado en el área de Serverless computing, Concurrencia y Computación distribuida.

Se espera aumentar el número de publicaciones y se prevé la divulgación por medio de cursos de postgrado y actualización o publicaciones de divulgación y asesoramiento a empresas y otras instituciones públicas y privadas.