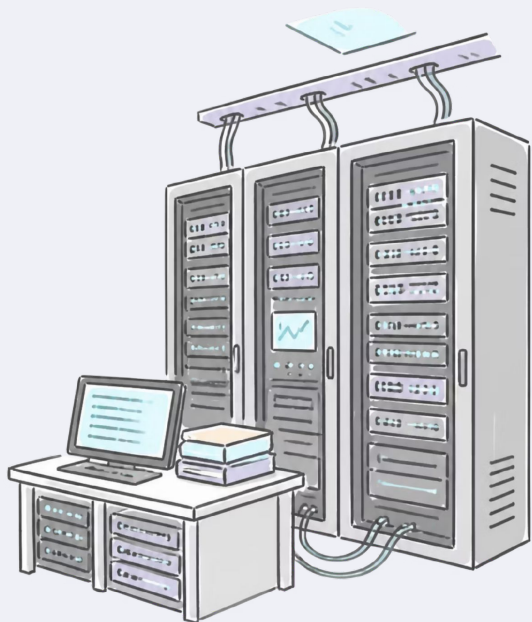




Experiencia en la Universidad Nacional de Rosario, Argentina

Cómo la actualización del DSpace mejoró la saturación de RAM causada por bots de IA

 CASO DE ESTUDIO

DSPACE · INFRAESTRUCTURA

Saturación por Bots de IA y SSR en DSpace 7.6.5

Tráfico de Bots

- Ráfagas masivas de scrapers de IA e indexadores
- Peticiones no cacheadas forzaban renderizado dinámico continuo

Fuga de Memoria

- Motor SSR de Angular/Node.js no liberaba memoria correctamente
- Saturación de RAM al **100%** causando caídas (*Out of Memory*)

El motor de renderizado del lado del servidor (SSR) de Angular/Node.js, bajo carga de scrapers de IA, no liberaba memoria entre peticiones, acumulando consumo hasta el colapso del sistema.

Mitigación Temporal

Geobloqueo

Bloqueo preventivo de todo el tráfico originado fuera de Argentina para reducir la carga de scrapers externos.

Handles y Google

Habilitación selectiva de tráfico hacia servidores de Handle y Google Scholar para mantener la resolución de identificadores persistentes.

El Dilema

Un repositorio abierto cerrado es un contrasentido. La restricción geográfica contradice el mandato de acceso abierto institucional.

⚠ La mitigación resolvió la estabilidad a corto plazo, pero comprometió el principio fundamental de acceso abierto del repositorio.

La Solución: Migración a DSpace 9.3

1

DSpace 7.6.5

SSR con fuga de memoria · Geobloqueos activos · Sistema inestable

2

Migración

Actualización de Angular SSR · Optimización del ciclo de vida de peticiones · Gestión de memoria corregida

3

DSpace 9.3

Sistema estable · Geobloqueos eliminados · Mismo volumen de bots sin impacto

La versión 9.3 incorpora mejoras sustanciales en el motor Angular SSR: gestión correcta del ciclo de vida de peticiones y liberación de memoria, permitiendo absorber el mismo volumen de tráfico de bots sin degradación del servicio.

Nuestros aprendizajes

1 La importancia de actualizar

Mantener versiones desactualizadas no es una elección sin costos: implica asumir riesgos de estabilidad, seguridad y rendimiento que pueden comprometer el servicio institucional.

2 Los scrapers de IA redefinen el tráfico web

Los bots de IA generan patrones de carga que los mecanismos tradicionales de protección —como bans por IP— no pueden contener. La resiliencia debe tener en cuenta la arquitectura del sistema.

¡Muchas gracias!

¿Preguntas?