

# **Administración de IBM DataStage en IBM Cloud**

Instalación, seguridad, gestión de proyectos y operación

Curso de Capacitación Técnica

Consultoría y Gestión Operativa

# ¿Qué incluye la administración?

 Instalación y configuración.

 Administración de proyectos.

 Respaldo y recuperación.

 Gestión de usuarios e IAM.

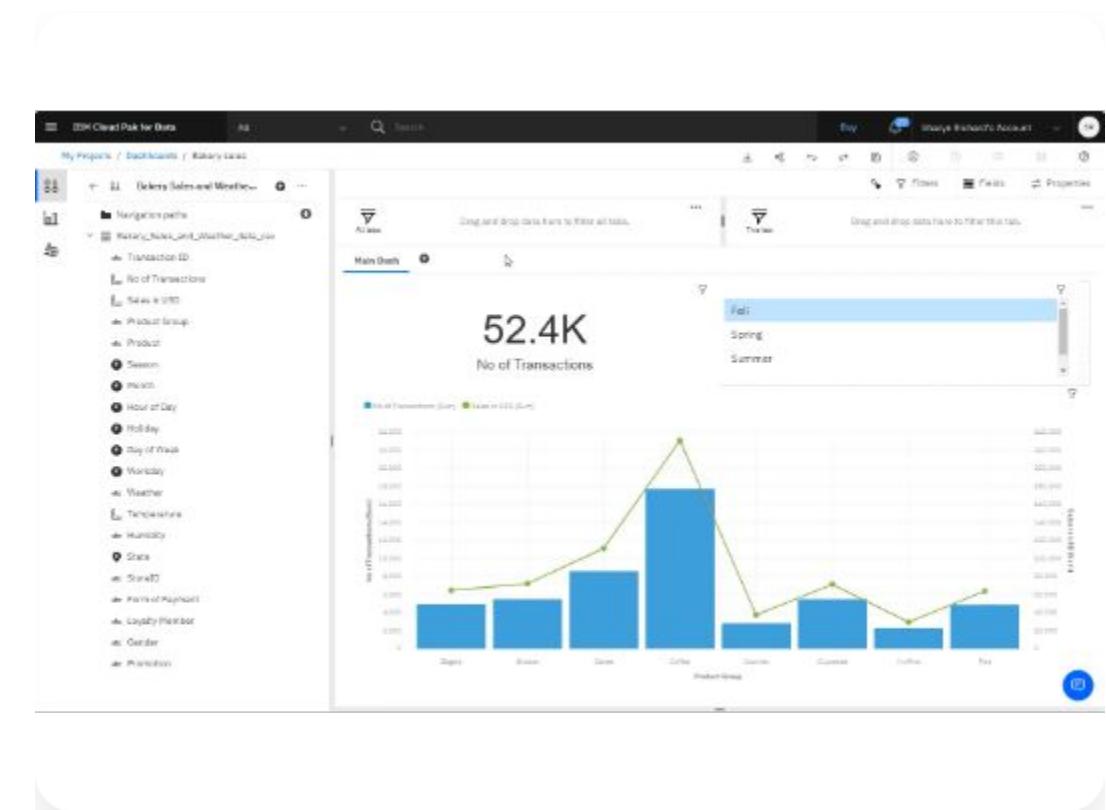
 Programación y monitoreo.

 Optimización de recursos Cloud.

# IBM DataStage en IBM Cloud

Parte integral de IBM Cloud Pak for Data:

-  Basado en contenedores (Kubernetes).
-  Escalabilidad elástica e inmediata.
-  Alta disponibilidad gestionada.
-  Acceso universal vía navegador.

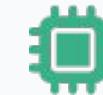


# Componentes de la Arquitectura



## Red Hat OpenShift

Orquestación de contenedores y base de la plataforma.



## DataStage Engine

Motor de ejecución optimizado para la nube.



## Cloud Storage

Servicios de almacenamiento persistente y seguro.

# Instalación y Requisitos

## Opciones de Despliegue

- IBM Cloud Pak for Data (PaaS)
- Despliegue sobre OpenShift (IaaS)
- Entornos DEV / TEST / PROD

## Requisitos Clave

- Cuenta IBM Cloud Corporativa
- Cluster OpenShift Activo
- Recursos CPU/RAM Escalables
- Almacenamiento Persistente

---

# Seguridad y Usuarios

Gestión de IAM y Roles en Cloud

# Roles de Usuario



## Admin

Control total del entorno, usuarios y configuración global.



## Developer

Diseño y modificación de Jobs en proyectos asignados.



## Operator

Ejecución y monitoreo. Sin permisos de modificación.

# Buenas Prácticas

**Mínimo Privilegio:** Solo los roles estrictamente necesarios.

-  **Integración LDAP:** Uso de directorios activos corporativos.
-  **Vaults:** Credenciales cifradas en bóvedas de seguridad.
-  **Auditoría:** Monitoreo constante de accesos.



# Gestión de Proyectos

## Organización

Un proyecto es un entorno lógico que separa equipos, activos y conexiones. Facilita el gobierno de datos.

## Segregación

- Proyecto\_DEV (Desarrollo)
- Proyecto\_QA (Validación)
- Proyecto\_PROD (Producción)

# Variables y Parámetros

## 0 Hard-coding

### Flexibilidad Dinámica

El uso de **Parameter Sets** permite que los Jobs sean reutilizables entre entornos sin modificar el diseño original.

- Rutas de archivos dinámicas.
- Cadenas de conexión por ambiente.
- Fechas de proceso parametrizadas.

# Monitoreo y Operación

Herramientas de control en tiempo real:



**DataStage Director:**  
Consola web de  
gestión.



**Logs:**  
Análisis detallado de  
eventos.



**Alertas:**  
Notificación  
automática de fallos.



**Reinicios:**  
Recuperación de  
procesos fallidos.

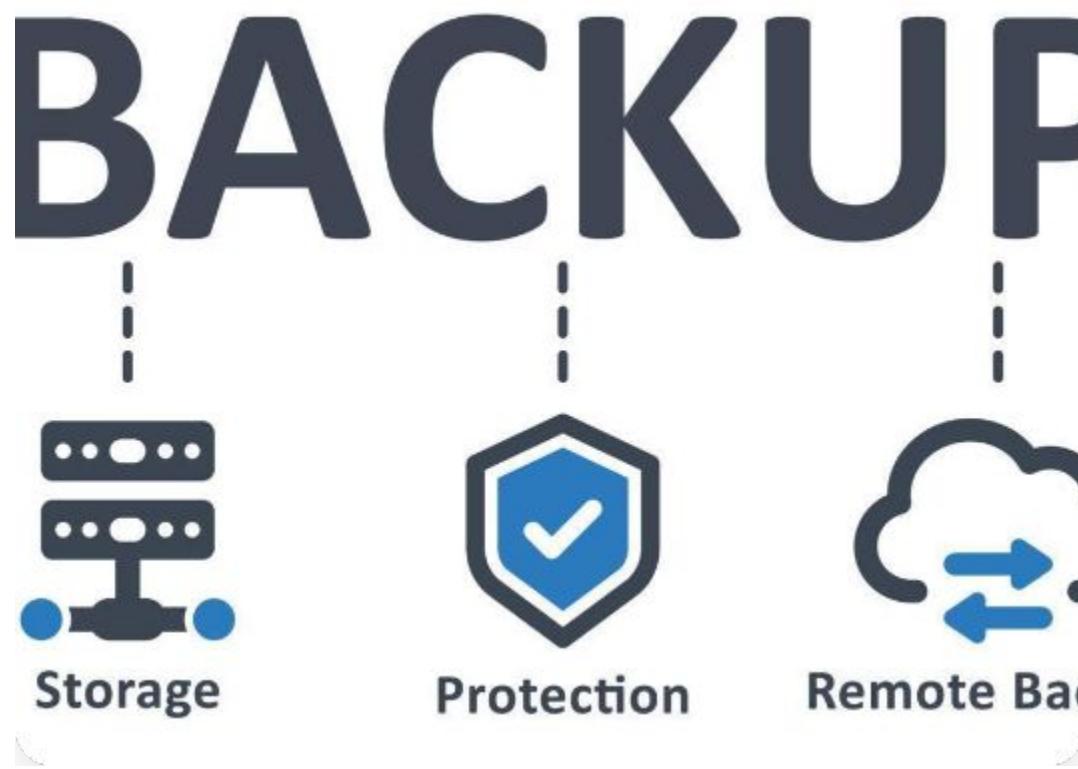
# Respaldo y Recuperación

Garantizando la Continuidad del  
Negocio

# ¿Qué se debe respaldar?

Activo de Información	Método en Cloud	Frecuencia
Proyectos y Jobs	Exportación ISX / Snapshots	Diario / Por Cambio
Metadatos y Repositorio	Backup de Plataforma CP4D	Semanal
Configuraciones e IAM	IBM Cloud IAM Backup	Mensual
Logs Críticos	Servicio de Logging Cloud	Tiempo Real

# Estrategia de Recuperación



Plan de contingencia ante fallos:

- ⌚ Restauración: Reinstalación selectiva de proyectos.
- 📷 Snapshots: Regreso rápido a estados anteriores del almacenamiento.
- 📝 Validación: Pruebas de integridad post-recuperación.

# Beneficios del Modelo Cloud



Menos administración de infraestructura = Más foco en el valor del dato.

# Resumen Ejecutivo

## **Cloud Native**

Administración simplificada mediante Cloud Pak for Data.

## **Seguro**

Control de acceso centralizado con IAM y roles RBAC.

## **Resiliente**

Backups integrados y estrategias de recuperación ágiles.

# ¿Consultas?

Muchas gracias por su participación.