

Incidencia: Servicio de Scoring y Flujos de Crédito

ID de Incidente: INC-20251528-01
Fecha: 28 de diciembre de 2024
Severidad: Crítica (SEV1)
Estado: Resuelto





Objetivo:

Presentar el informe de los eventos ocurridos a partir de las incidencias presentadas en la funcionalidad de solicitudes de crédito y el servicio de scoring.



Índice

1. Resumen y Contexto
2. Impacto en el Negocio y Usuario
3. Cronología del Incidente (Timeline)
4. Análisis de Causa Raíz (RCA)
5. Resolución y Acciones Inmediatas
6. Plan de Mejora
7. Lecciones Aprendidas



Resumen y Contexto

El 28 de diciembre, entre las 09:00 y las 10:00, se detectó una degradación crítica en el servicio de solicitudes de crédito. El proceso de finalización de solicitudes cayó un 40%, impactando directamente en la conversión de nuevos clientes. El problema fue originado por una actualización en el sistema de scoring realizada la noche anterior, la cual, ante una inestabilidad del proveedor externo, generó un efecto de saturación en nuestra base de datos. El servicio se normalizó tras realizar un rollback de la versión y ajustar las configuraciones de red.



Impacto en el Negocio y Usuario

- **Duración:** 1 hora (09:00 - 10:00).
- **Alcance:** 100% de los usuarios que intentaron completar una solicitud de crédito en el periodo afectado.
- **Métricas de Negocio:**
 - Reducción del 40% en flujos completados.
 - Incremento de latencia en el endpoint /loan/apply de 300ms a >15s.
 - Pérdida estimada de 450 solicitudes potenciales.



Cronología de Eventos

Hora	Evento
09:00	Negocio reporta caída del 40% en finalización de solicitudes.
09:10	Dashboard muestra reducción en créditos procesados.
09:15	Observabilidad detecta latencia alta en /loan/apply y errores hacia scoring.
09:25	Logs revelan timeouts y errores de conexión hacia Aurora.
09:40	Se identifica despliegue nocturno del servicio de scoring con nueva política de reintentos.
10:00	Mitigación: rollback + ajuste de reintentos → normalización del servicio.



Análisis de Causa Raíz (RCA)

El incidente fue causado por una Tormenta de Reintentos (Retry Storm) exacerbada por una mala configuración en la política de reintentos del servicio de scoring.

1. **El Problema Externo:** El proveedor de scoring tuvo intermitencias.
2. **La Falla Interna:** La nueva versión desplegada no tenía un Exponential Backoff (retroceso exponencial). Ante cada falla del proveedor, el sistema reintentaba inmediatamente múltiples veces.
3. **Agotamiento de Recursos:** Estos reintentos mantuvieron abiertas las conexiones a la base de datos Aurora y los hilos de ejecución de la aplicación.
4. **Cascada:** Al agotarse el pool de conexiones (Connection Pool Exhaustion), el servicio dejó de responder incluso para operaciones que no dependían del proveedor externo.



Acciones Correctivas Inmediatas

- **Rollback:** Se revirtió el servicio de scoring a la versión estable anterior.
- **Ajuste de Configuración:** Se modificaron los parámetros de reintento en caliente para incluir tiempos de espera más largos entre intentos.
- **Reinicio de Instancias:** Se reciclaron los nodos de la aplicación para liberar las conexiones "colgadas" a Aurora.



Lecciones Aprendidas

- ¿Qué se hizo bien?
 - **Respuesta Rápida:** Una vez identificada la causa, el rollback se realizó en menos de 20 minutos.
 - **Visibilidad Técnica:** Los dashboards de observabilidad permitieron ver rápidamente que el problema escaló de un proveedor externo a nuestra base de datos.
 - **Coordinación:** Comunicación efectiva entre las áreas de Producto, Negocio y SRE para dimensionar el impacto.
- ¿Qué debemos mejorar?
 - **Monitoreo Proactivo:** El incidente fue detectado por Negocio antes que por las alertas técnicas automáticas.
 - **Aislamiento de Fallas:** El sistema permitió que una falla externa "contagiara" a nuestra base de datos interna. Debemos implementar mecanismos de corte (Circuit Breakers).
 - **Validación de Cambios:** Los cambios en políticas de red deben probarse bajo escenarios de estrés (simulación de caída de proveedores).



Plan de Mejora

Acción	Objetivo	Plazo
Implementar Circuit Breaker	Evitar que fallas externas saturen recursos internos.	Corto
Retry con Backoff Exponencial	Espaciar los reintentos para no auto-atacar el sistema.	Corto
Alertas de Conversión	Notificar automáticamente si el flujo de créditos baja del 10%.	Medio
Pruebas de Caos (Chaos Engineering)	Simular fallas de terceros trimestralmente.	Largo



TECHNICAL CHALLENGE

SITE RELIABILITY ENGINEER