

Event-Driven Architectures for Timely Information Availability in Public Administration

Nicolás Passerini^{1,2,3}, Valentina Lodigiani⁴ and Ivan Novotny^{3,4}

¹ Departamento de Ciencia y Tecnología, Universidad Nacional de Quilmes, Bernal, Provincia de Buenos Aires, Argentina

² Universidad Tecnológica Nacional, Facultad Regional Buenos Aires, Ciudad Autónoma de Buenos Aires, Argentina

³ Universidad de Buenos Aires, Facultad de Ciencias Sociales, Ciudad Autónoma de Buenos Aires, Argentina

⁴ Gerencia General de Tecnología e Innovación, Agencia de Recaudación de la Provincia de Buenos Aires (ARBA), La Plata, Buenos Aires, Argentina

Abstract

In organizations characterized by the management of large volumes of data and a long history of process computerization, batch processing constituted the predominant paradigm for managing critical information for decades. As demands for the availability and timeliness of information intensified, this approach began to reveal several limitations, including difficulties in providing real-time information. These limitations are not merely technological; over time, they have shaped organizations' internal processes and become embedded in their institutional culture.

This article presents an ongoing implementation that proposes the adoption of event-driven architectures as a strategy to reduce the delays inherent to batch-based approaches, with the goal of enabling organizational systems to react to changes as they occur and to maintain continuously updated information.

The theoretical framework underpinning this work invites an analysis of the relationship between organizational processes and the tools that enable them, as mediated by institutional structures and culture. Far from being solely a technological modernization effort, this transition is conceived as an opportunity to rethink value flows.

Finally, the paper presents preliminary results obtained from the initial implementations, along with a description of the project's next steps, aimed at consolidating these learnings into the definition of a reference architecture that can accelerate the modernization of existing software and improve the quality and timeliness of the services provided by the organization.

Keywords

Event-Driven Architecture; Incremental Processing; ETL; Legacy Systems;
Public Sector

Disponibilidad Continua de la Información mediante Arquitecturas Orientadas a Eventos en la Administración Pública

Resumen

En organizaciones caracterizadas por el manejo de grandes volúmenes de datos y con una extensa historia de informatización de sus datos, el procesamiento por lotes (*batch*) constituyó durante décadas el paradigma predominante para la gestión de información crítica. A medida que las exigencias sobre la disponibilidad y oportunidad de la información se intensificaron, este enfoque comenzó a evidenciar diversas limitaciones, entre ellas la dificultad para ofrecer información en tiempo real. Estas limitaciones no son únicamente tecnológicas; con el tiempo han moldeado los procesos internos de las organizaciones y se han arraigado en su cultura institucional.

Este artículo presenta una implementación en curso que propone la adopción de arquitecturas orientadas a eventos como estrategia para reducir las demoras propias del enfoque por lotes, con el objetivo de que los sistemas organizacionales puedan reaccionar a los cambios en el momento en que ocurren y mantener información actualizada de manera continua.

El marco teórico que fundamenta este trabajo invita a analizar la relación entre los procesos organizacionales y las herramientas que los posibilitan, mediada por la estructura y la cultura institucional. Lejos de tratarse únicamente de una modernización tecnológica, esta transición se concibe como una oportunidad para repensar flujos de valor.

Finalmente, el trabajo presenta resultados preliminares obtenidos a partir de las primeras implementaciones, así como una descripción de los próximos pasos del proyecto, orientados a consolidar estos aprendizajes en la definición de una arquitectura de referencia que permita acelerar la modernización del software existente y mejorar la calidad y oportunidad de los servicios que brinda la organización.

Palabras clave

Arquitecturas orientadas a eventos; Procesamiento incremental; ETL; Sistemas legacy; Sector público

Introducción

Desde los inicios de la computación, el procesamiento por lotes (*batch*) se consolidó como uno de los modelos predominantes para el tratamiento de grandes volúmenes de datos. En el ámbito estatal, el IRS¹ comenzó ya en 1959 la automatización del procesamiento impositivo mediante sistemas *batch*, siendo uno de los primeros organismos públicos en adoptarlos de forma sistemática (Norman, 2025).

Sin embargo, las organizaciones modernas operan en contextos crecientemente dinámicos, donde **la disponibilidad inmediata de información es un factor determinante**. En este contexto, esa lógica de ejecuciones periódicas y masivas, que introduce latencias estructurales (Dean & Ghemawat, 2008) al concentrar la actualización de los datos en procesos periódicos, se ha vuelto una limitación para poder brindar servicios interactivos en línea y/o información para la toma de decisiones (Duvvur, 2022).

Adicionalmente, muchas organizaciones públicas con alto volumen de procesamiento de datos, que iniciaron la digitalización de sus procesos en la década de 1980, enfrentan hoy fuertes condicionamientos derivados de su infraestructura de software *legacy* (Abu Bakar *et al.*, 2022). Con frecuencia, los sistemas responsables de procesos críticos concentran décadas de reglas de negocio, conocimiento organizacional y decisiones históricas; su costo de mantenimiento aumenta con el tiempo, pero su reemplazo integral resulta inviable tanto por su complejidad como por el riesgo de pérdida de información crítica (Fontela, 2018).

En este artículo se presenta una experiencia en curso, desarrollada en el marco de la Agencia de Recaudación de la Provincia de Buenos Aires (ARBA), que tiene por objetivo la adopción de arquitecturas orientadas a eventos (*event-driven architectures*, EDA) como estrategia para obtener un salto de calidad en la disponibilidad de la información. El foco del trabajo está puesto en **procesos de integración y transformación de datos**, comúnmente implementados mediante esquemas de extracción, transformación y carga (ETL), que en este ámbito exceden el uso analítico, dado que sus resultados alimentan de forma directa procesos operativos como la emisión de boletas y los procesos de cobranza o fiscalización.

EDA como estrategia para la evolución tecnológica

La adopción de arquitecturas orientadas a eventos en este contexto se fundamenta en dos capacidades complementarias. Por un lado, el **bajo acoplamiento y las interacciones asincrónicas** propias de este estilo arquitectónico facilitan la incorporación gradual de nuevos componentes, permitiendo definir estrategias de migración progresiva en coexistencia con sistemas existentes (*c.f.* Darif, 2025). Por otro, la orientación a eventos habilita un **cambio en la lógica de procesamiento de datos** respecto de los esquemas ETL tradicionales: en lugar de recalcular periódicamente conjuntos completos de información² —aun cuando los cambios afecten solo a una fracción de los datos—, los eventos permiten mantener el estado de

¹ Internal Revenue Service, agencia de recaudación de los Estados Unidos.

² *C.f.* Kimball, R., & Caserta, J. (2004)

forma continua, **recalculando únicamente aquellos casos impactados por cambios en el dominio** que afectan el resultado de los procesos.

En este marco, y como parte del proyecto orientado a mejorar la disponibilidad de la información en ARBA, se propone desarrollar una **experiencia concreta de migración** que funcione como caso piloto y espacio de aprendizaje institucional. Los resultados de esta experiencia buscan aportar insumos para la definición de una arquitectura de referencia, orientada a identificar patrones, criterios y decisiones reutilizables que faciliten una adopción gradual y coherente de EDA en otros procesos y sistemas de la organización.

Uno de los desafíos que se identificaron consiste en la tensión entre el enfoque orientado a eventos y la **cultura organizacional**. En organizaciones con una fuerte dependencia en sus sistemas para la ejecución de sus procesos operativos (como son las administraciones tributarias), dichos procesos fueron moldeados por las tecnologías disponibles en el momento en que fueron concebidos. Si bien las EDA habilitan la disponibilidad continua de la información y reducen significativamente la latencia entre la ocurrencia de un cambio y su reflejo en los sistemas, estas capacidades suponen un **cambio de enfoque en la forma en que la organización concibe y procesa la información**. Su aprovechamiento efectivo requiere que los procesos operativos y las prácticas institucionales evolucionen en consonancia; de lo contrario, la adopción de EDA tiende a reproducir las lógicas y restricciones de los esquemas *batch*, limitando su impacto transformador.

Desde esta perspectiva, el proyecto se estructura en torno a las siguientes preguntas de investigación: ¿cómo definir un marco de migración incremental desde esquemas *batch* hacia EDA, aplicable a procesos de transformación de datos en organismos públicos?, ¿qué criterios permiten organizar y secuenciar dicha migración en contextos operativos reales?, ¿de qué manera las restricciones organizacionales y culturales influyen en la adopción y el aprovechamiento efectivo de este enfoque?

Resultados preliminares: cálculo de deudas

Computar la deuda es requisito para muchas operaciones en una agencia tributaria, tanto en los servicios en línea ofrecidos a los contribuyentes, como en procesos *batch* ejecutados periódicamente para la generación de estadísticas, como insumo de acciones de cobranza y/o fiscalización. Dada la complejidad de estos cálculos y el volumen de *objetos imponibles*³ involucrados, ambos escenarios representan un desafío significativo para la infraestructura de procesamiento de ARBA.

En el caso de los servicios en línea, este desafío se manifiesta los días de vencimiento, caracterizados por altos niveles de tráfico. Como primera estrategia de mitigación, se incorporó una *caché* de las deudas calculadas por objeto, combinada con un proceso de *pre-warming*⁴ nocturno que permitió adelantar el cálculo de deuda para un

³ Se entiende por “objeto” u “objeto imponible” aquello sobre lo que se impone un tributo.

⁴ *Pre-warming* refiere a la precarga proactiva de una *caché* con datos o resultados computados antes de que ocurran las solicitudes reales (Zhang et al., 2020).

conjunto seleccionado de contribuyentes. La *caché* se mantiene válida mientras no se registren eventos que modifiquen el estado de deuda —en particular, el ingreso de un pago—, caso en el cual se invalida y el cálculo se actualiza en el siguiente acceso.

Esta estrategia permitió procesar 50% más de liquidaciones en la semana pico (*c.f.* Fig. 1), a la vez que mejoró significativamente los tiempos de respuesta. Los registros comparados muestran un incremento de las transacciones promedio por minuto de 221 a 333 en la semana del vencimiento, a la vez que el tiempo promedio de respuesta para el cálculo de deuda se reduce de 238ms en 2025 a 89ms en 2026 (reducción de 62.6%).

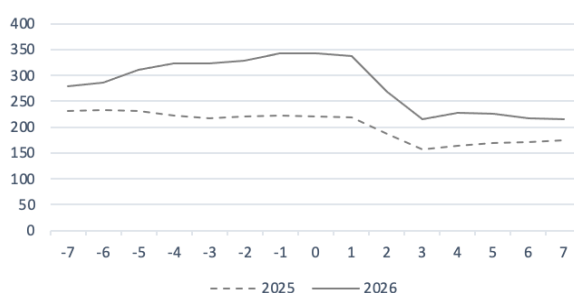


Fig. 1 Incremento en la capacidad de atención on line.

Comparación de capacidad de atención a usuarios, liquidaciones por minuto, 7 días antes y después del vencimiento de la primera cuota del impuesto inmobiliario urbano, media móvil de 7 días.

Para el caso de procesos *batch*, dado que el elevado costo computacional del cálculo de deuda limita su frecuencia de ejecución, la eficiencia se vuelve un requisito indispensable. No obstante, el análisis de los procesos vigentes reveló que (a) distintos procesos **realizan cálculos similares sin reutilizar resultados**, y (b) cada corrida implica **recorrer la totalidad del universo de objetos involucrados**, aun cuando los cambios afecten solo a una fracción de ellos.

Para mitigar estas limitaciones, se diseñó una estrategia orientada a centralizar el cálculo de la deuda en un repositorio operativo único, reutilizable por múltiples procesos de la organización. Bajo este enfoque, la deuda deja de calcularse ante cada requerimiento y pasa a mantenerse en una **capa operativa persistente** (ODS, *c.f.* Inmon, 1995), desde la cual distintos procesos —tanto analíticos como operativos, incluyendo el *on line*— pueden consumir, derivando las vistas específicas mediante filtrados o agregaciones, evitando accesos reiterados a los sistemas de origen.

Con este marco, se identificó que la arquitectura desarrollada para la *caché* utilizada por los servicios interactivos podía extenderse para configurar como un primer prototipo de repositorio operativo, susceptible de ser reutilizado por otros servicios de la organización. El desafío central de este enfoque radica en mantener actualizado el estado de la deuda **sin recurrir a reprocesamientos completos**, dado que la *caché* original se invalidaba diariamente, mientras que un ODS requiere mayor persistencia temporal. Esto llevó a la necesidad de **identificar un conjunto más amplio de eventos del dominio** que pudieran modificar la cuenta corriente de un objeto, que es la antesala de los cálculos de deuda.

Este enfoque ya fue aplicado al proceso de *mora temprana*, un proceso de notificación a los contribuyentes pocos días después un vencimiento de cuotas impagas, en este caso, patentes automotores. En la nueva arquitectura, cualquier acción sobre un automotor genera un evento, que es **procesado asincrónicamente por un sistema de colas**; uno de los consumidores de esa cola de eventos actualiza la deuda del automotor en cuestión.

Gracias a este cambio, la información necesaria para la mora temprana se encuentra actualizada casi en tiempo real, evitando las tradicionales bajadas de datos que solían tardar hasta 5 horas. Desde esta vista materializada, las estadísticas necesarias para continuar el proceso se pueden computar en menos de tres minutos, lo que permite a los usuarios ajustar parámetros, validar escenarios y realizar iteraciones previas al envío de las notificaciones, algo inviable bajo el esquema *batch* tradicional.

Aprendizajes, desafíos y próximos pasos

El principal desafío identificado no fue de carácter técnico, sino **organizacional y cultural**. La transición desde una lógica de datos consolidados en momentos discretos hacia un esquema de actualización continua generó tensiones con prácticas muy arraigadas. Por este motivo, se optó por pausar temporalmente la actualización vía eventos mientras el proceso de mora temprana se encontraba en ejecución y, una vez finalizado, reanudar el procesamiento de los eventos acumulados, conciliando la necesidad de consistencia operativa con un modelo de actualización incremental.

Asimismo, se detectaron resistencias vinculadas a la **confianza en los datos**, dado el hábito de validarlos directamente contra los sistemas transaccionales en el *mainframe*. Para facilitar la adopción del nuevo enfoque, se desarrollaron mecanismos de comparación y validación que permitieron fortalecer gradualmente la confianza en el ODS y su proceso de carga basado en EDA.

A partir de estos aprendizajes, el trabajo futuro se orienta a continuar **generalizando progresivamente la estrategia** implementada. En primer lugar, se planea extender este enfoque a otros tipos de cálculos de deuda, de manera que el ODS se consolide como el punto central desde el cual se calculen y mantengan actualizados todos los objetos imponibles. Esto implica incorporar gradualmente nuevos impuestos y nuevos tipos de operaciones, preservando una lógica incremental y controlada.

En paralelo, ya se han desarrollado tableros de consulta sobre la información consolidada y se espera avanzar en la construcción de **distintas vistas de los mismos conjuntos de datos**, orientadas a necesidades analíticas, operativas y de gestión. Finalmente, una línea de trabajo relevante será continuar integrando este esquema con otros sistemas existentes y extender el uso del modelo orientado a eventos a procesos organizacionales adicionales más allá del cálculo de la deuda, reforzando así el rol del ODS como componente estructurante de la arquitectura de información de la organización.

Referencias

- Abu Bakar, H., Razali, R., & Jambari, D. I. (2022). *A qualitative study of legacy systems modernisation for citizen-centric digital government*. **Sustainability**, **14**(17), 10951. <https://doi.org/10.3390/su141710951>
- Brey, G. A., Escobar, G., Passerini, N., & Arias, J. (2005). *Arquitectura de proyectos IT: Evaluación de arquitecturas*. Universidad Tecnológica Nacional, Facultad Regional Buenos Aires.
- Darif, I., Abdellatif, M., & El Boussaidi, G. (2025). *On the migration of legacy systems to an event-driven architecture: A survey*. In *Proceedings of the 49th Annual Computers, Software, and Applications Conference (COMPSAC 2025)* (pp. 1260–1269). IEEE. <https://doi.org/10.1109/COMPSAC65507.2025.00159>
- Dean, J., & Ghemawat, S. (2008). *MapReduce: Simplified data processing on large clusters*. **Communications of the ACM**, **51**(1), 107–113. <https://doi.org/10.1145/1327452.1327492>
- Duvvur, V. (2022). *Modernization of legacy batch processes: An upgrade to prolonged batch runs with advanced computing and parallel processing techniques* [Technical manuscript]. ResearchGate. <https://www.researchgate.net/publication/381087644>
- Fontela, C. (2018). *Migración de sistemas heredados hacia microservicios con el soporte de especificaciones mediante ejemplos*. En **XIX Simposio Argentino de Ingeniería de Software (ASSE), JAIIO 47**.
- Inmon, W. H., Imhoff, C., & Battas, G. (1995). *Building the operational data store*. John Wiley & Sons.
- Kimball, R., & Caserta, J. (2004). *The data warehouse ETL toolkit*. John Wiley & Sons.
- Norman, J. M. (2025, December 28). *The IRS automates income-tax processing (1959–1966)*. **History of Information**. <https://www.historyofinformation.com/detail.php?entryid=1064>
- Zhang, L., Yang, J., Blasiak, A., McCall, M., & Vigfusson, Y. (2020). *When is the cache warm? Manufacturing a rule of thumb*. In **Proceedings of the USENIX Workshop on Hot Topics in Cloud Computing (HotCloud 2020)**.