



El análisis de datos en tiempo real más rápido con InterSystems IRIS

Este white paper describe cómo se utiliza hoy en día la tecnología de InterSystems para replicar datos de un sistema de producción operativo a un "nodo de informe", donde pueden consultarse y analizarse sin afectar al sistema operativo. Los clientes de InterSystems han estado utilizando esta arquitectura con éxito durante muchos años, y este documento describe el caso de uso general con más detalle. Presenta una perspectiva sobre cómo las empresas usuarias de las soluciones de InterSystems pueden utilizar este enfoque dentro de su implementación existente.

De la aplicación al análisis

Los datos de una aplicación son valiosos, obviamente, para los propios fines de la aplicación. Los sistemas de gestión de pedidos gestionan éstos. Los sistemas de planificación de recursos empresariales ayudan a gestionar las empresas. Los sistemas de historias clínicas electrónicas organizan la información del paciente para que pueda recibir el tratamiento adecuado. Sin embargo, los datos de las aplicaciones pueden ser aún más valiosos si se analizan. En la práctica, eso puede significar analizar las tendencias de muchas órdenes para detectar las tendencias del mercado. Puede significar mejorar la planificación de la producción utilizando previsiones automatizadas. Podría consistir en aprovechar las métricas de calidad clínica para mejorar la atención al paciente.

Los widgets analíticos a pequeña escala que encajan con el propósito de la aplicación, como un panel de control en vivo, que muestre las camas disponibles en una sala de hospital, tienen sentido funcionando en la propia aplicación. Sin embargo, esto tiene sus límites, y los sistemas que ejecutan una aplicación generalmente no tienen el tamaño adecuado para soportar cargas de trabajo de análisis completas. Los patrones de uso del análisis difieren significativamente de los de una aplicación. Obtener el archivo actual de un paciente supone utilizar unas características de CPU, memoria y E/S diferentes a las de realizar un análisis de salud de la población para cientos o miles de pacientes.

El enfoque tradicional para ofrecer estas capacidades analíticas, a mediana y gran escala, ha sido extraer los datos de estos sistemas fuente, transformarlos en un esquema apto para el análisis y, a continuación, cargarlos en una base de datos analítica dedicada, normalmente, un almacén de datos. Las herramientas ETL especializadas suelen ser potentes y ricas en funciones, pero su instalación es costosa, requieren personal especializado con conocimientos distintos a los de los operadores de las aplicaciones y, por lo general, las ventajas sólo pueden obtenerse tras un proyecto de implantación completo y no en pequeños incrementos.

Una plataforma unificada

Con la plataforma de datos InterSystems IRIS, las cosas son distintas. InterSystems IRIS es una plataforma unificada, diseñada para soportar cargas de trabajo operativas y de análisis en una única tecnología. No solo ofrece una base de datos de alto rendimiento para ejecutar aplicaciones de misión crítica, sino que también incluye un conjunto completo de capacidades analíticas, incluidos informes, inteligencia empresarial y machine learning.

Por ello, muchas aplicaciones desarrolladas en InterSystems IRIS incluye complejos componentes analíticos integrados, como cuadros de mando y gráficos totalmente integrados sobre datos en directo a gran escala. InterSystems IRIS también se despliega en escenarios de almacenes de datos dedicados, soportando análisis y machine learning en grandes volúmenes de datos.

Esta capacidad de soportar tanto una escalabilidad extrema para el uso de aplicaciones, como el máximo rendimiento para consultas analíticas en una única tecnología hace que InterSystems IRIS esté en una posición única para un modelo que no depende de un costoso ETL. Con la tecnología de duplicación de InterSystems, los datos se pueden replicar de un sistema a otro, sin incurrir en una carga adicional perceptible en el sistema de origen. Se trata de un elemento clave para ejecutar aplicaciones y análisis a gran escala con los mismos datos en tiempo real. El resto de este white paper describirá cómo InterSystems mirroring facilita esto a nivel arquitectónico y presentará casos de uso de dos clientes que han implementado este enfoque.

Tecnología probada

InterSystems Mirroring es utilizado habitualmente en modo síncrono, por los clientes de InterSystems, para implementar estrategias de alta disponibilidad (HA) en las que un miembro espejo de backup permanece en espera en caliente si el nodo primario se cae. El mirroring también se implementa a menudo en modo asíncrono en topologías de recuperación ante desastres (DR) que cruzan regiones de nube o centros de datos. El mirroring funciona a bajo nivel y, por lo tanto, alcanza tasas de rendimiento extremadamente altas utilizando técnicas como la compresión diaria y el IO asíncrono. Utilizando el mirroring asíncrono, los clientes también pueden configurar un nodo de informes que se dimensiona y utiliza independientemente del sistema primario, con un coste de configuración casi nulo. Esta copia de los datos puede ser consultada tal cual por los analistas familiarizados con el esquema de la aplicación, o utilizarse como fuente para soluciones de business intelligence, nativas de InterSystems o de terceros, que proyecten un modelo analítico sobre este esquema. Los datos replicados a través del mirroring asíncrono son prácticamente en tiempo real, con una latencia que suele estar en el rango de los milisegundos de un solo dígito, y siguiendo un modelo de consistencia eventual adecuado para la mayoría de los casos de uso de análisis e informes.

Los clientes que necesiten una reestructuración más sustancial del esquema de la aplicación en una estructura de almacén de datos pueden seguir haciéndolo, a partir de los datos del nodo de informes, utilizando las capacidades SQL nativas de

InterSystems IRIS, su soporte dbt u otras herramientas aplicables. Este modelo es similar al enfoque ELT anunciado por algunas herramientas -en el que los datos se cargan primero en la plataforma de destino y luego se transforman allí-, pero tiene la ventaja clave de que los datos también pueden consultarse inmediatamente.

Dado que el uso de la duplicación de InterSystems no requiere ningún desarrollo adicional ni tecnología de terceros, los clientes obtienen una rentabilidad mucho más rápida. En caso de que alguna transformación siguiera siendo beneficiosa, se puede desarrollar en el nodo de informes en pequeños incrementos, utilizando un proceso ágil basado en necesidades analíticas específicas, en lugar de requerir largos proyectos de diseño e implementación de almacenes.

El mirroring asíncrono se utiliza ampliamente hoy en día para replicar datos en diferentes regiones de la nube o desde una instalación local a la nube y, como tal, la tecnología ofrece una flexibilidad total en el despliegue de nodos de informes en la nube más adecuada, local o topología híbrida.

Mirroring también soporta la replicación a nodos de informes en una versión diferente y más reciente de InterSystems IRIS, permitiendo a los clientes aprovechar las nuevas características en el nodo de informes, independientemente del calendario de actualización del servidor operativo principal.

Esto abre la puerta a capacidades que se introdujeron recientemente en InterSystems IRIS, como InterSystems Adaptive Analytics, Columnar Storage y Foreign Tables. Utilizadas por separado o combinadas con las herramientas existentes, estas capacidades permiten crear soluciones analíticas extremadamente potentes que aprovechan al máximo los datos operativos de los sistemas de origen.

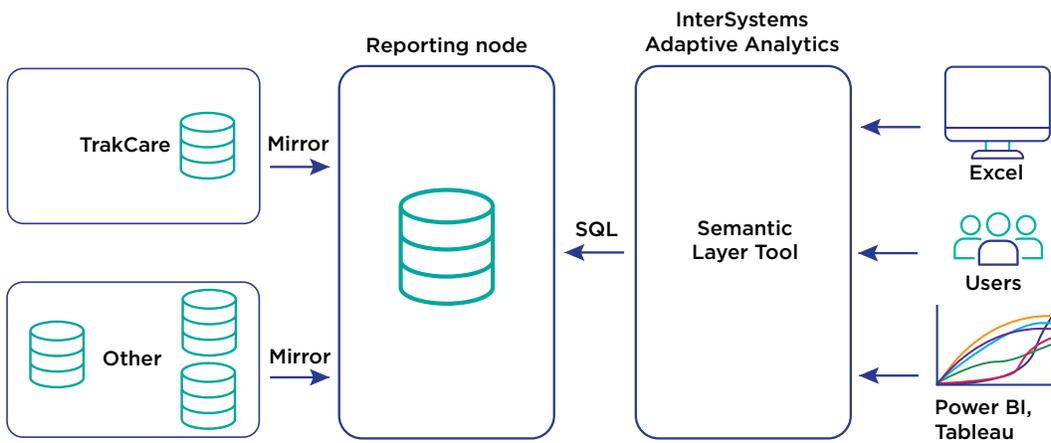
De las tecnologías a las soluciones

InterSystems Mirroring es una función de plataforma disponible en todas las aplicaciones y soluciones basadas en InterSystems IRIS data platform. La mayoría de las soluciones sanitarias actuales de InterSystems se centran en casos de uso de aplicaciones e integración. Por ejemplo, InterSystems TrakCare® es un sistema de información sanitaria que captura datos tales como historias clínicas electrónicas para operaciones hospitalarias generales. La funcionalidad Advanced Analytics de TrakCare mejora estos casos de uso añadiendo un conjunto de cuadros de mando integrados para escenarios específicos, como la monitorización de eventos adversos y los tiempos de espera en urgencias. Sin embargo, estos datos también pueden tener una finalidad ajena a esos casos de uso operativo. Añadir un nodo de generación de informes IRIS de InterSystems que contenga una copia de los datos del sistema de producción, pero que pueda consultarse libremente sin afectar al sistema operativo, desbloquea un enorme potencial para los clientes que se enfrentan a este tipo de casos de uso analítico. Todo el esquema TrakCare puede consultarse directamente, o utilizarse para poblar mercados de datos específicos para una necesidad analítica concreta.

Ejemplo de cliente #1

Muchos clientes de TrakCare de todo el mundo aprovechan la tecnología InterSystems Mirroring para replicar los datos de su instancia TrakCare de producción en uno o varios nodos de generación de informes dedicados al análisis. Por ejemplo, varios clientes en Tailandia están consultando estos nodos de informes utilizando tecnologías de terceros, como Crystal Reports, Tableau y soluciones a medida para complementar los cuadros de mando integrados de TrakCare.

Un hospital líder en Francia está utilizando InterSystems Adaptive Analytics en su nodo de informes para construir cubos OLAP virtuales sobre el esquema operativo de TrakCare, para análisis como el seguimiento de KPI sobre ocupación de salas y pacientes con alertas sobre condiciones específicas como COVID. Los casos de uso futuros combinarán datos de TrakCare con datos de fuentes externas para servir a análisis no operativos y explorar oportunidades para la ciencia de datos y el machine learning.



Ejemplo de cliente #2

Shift es el principal proveedor brasileño de sistemas de información de laboratorio (SIL) y procesa más de 370 millones de pruebas clínicas al año en toda su base de clientes. Uno de sus mayores clientes, el Laboratorio Sabin, presta servicio a casi 6 millones de ciudadanos y recientemente ha revisado la arquitectura de su sistema para ajustar el tamaño de la infraestructura. La volátil carga de trabajo de unos 3.000 usuarios simultáneos se distribuye en un nivel informático de 10 servidores de aplicaciones. La carga de trabajo de generación de informes se descarga a un nodo de generación de informes, aprovechando la réplica asíncrona de InterSystems para replicar el conjunto de datos completo de 20 TB desde el SIL. Esto permite elaborar informes e inteligencia empresarial sobre todos los datos en tiempo real, sin interrumpir los procesos operativos, en una arquitectura asequible y fácil de ampliar a medida que evolucionan las necesidades de procesamiento y elaboración de informes.

Limitaciones

InterSystems Mirroring no impone ningún límite sobre qué versiones de productos InterSystems IRIS pueden ser el origen o el destino cuando se replican datos de aplicaciones. Para el esquema de aplicación, una versión de sólo lectura utilizada para consultas SQL puede usarse en diferentes versiones sin problemas, siempre que el número de versión principal sea el mismo. Para replicar un esquema de aplicación en una versión con un número de versión principal diferente, consulte con su equipo de cuentas de InterSystems para determinar qué combinaciones son compatibles. En muchos casos en los que el código fuente del esquema está disponible, basta con recompilar el esquema en el nodo de informes para permitir el acceso a los datos de sólo lectura.

Conclusión

Empresas de todos los sectores han implantado sus aplicaciones operativas de misión crítica en InterSystems IRIS. Muchas están enriqueciendo estas aplicaciones con análisis integrados, pero cada vez identifican más casos de uso de análisis totalmente nuevos y encuentran oportunidades para monetizar los datos de las aplicaciones. Para estos casos de uso, la tecnología InterSystems Mirroring les permite implantar rápidamente nodos de generación de informes, que replican los datos de su sistema operativo, en un entorno dedicado a la generación de informes y análisis. Un número cada vez mayor de clientes de soluciones de InterSystems están implantando este patrón con éxito y experimentando una rápida rentabilización de sus proyectos analíticos, con amplias oportunidades de expansión gracias al amplio conjunto de tecnologías analíticas de InterSystems IRIS. Toda la tecnología necesaria está disponible en la actualidad y se incluye con todos los productos de InterSystems, a sólo un paso de la configuración.

Si desea ejecutar una carga de trabajo analítica en sus datos operativos con un impacto mínimo en las aplicaciones, póngase en contacto con su equipo de cuentas para analizar su arquitectura.

