



## Fastest Time to Value for Analytics on Live Data with InterSystems IRIS

- This white paper describes how InterSystems proven technology is used today to replicate data from an operational production system to a “reporting node,” where it can be queried and analyzed without impacting the operational system. Customers have been using this architecture successfully for many years, and this document describes the overarching use case in more detail, and presents an outlook on how current InterSystems IRIS® and InterSystems® solution customers may use this approach within their existing deployment.

### From Application to Analytics

Application data is obviously valuable for the application’s own purposes. Order management systems manage orders. Enterprise resource planning systems help run businesses. Electronic health record systems organize patient information so they can get the right treatment. However, application data can become all the more valuable if it is analyzed. In practice, that can mean analyzing trends across many orders to spot market trends. It can mean enhancing production planning using automated forecasting. It could take the form of leveraging clinical quality metrics to improve patient care.

Small-scale analytical widgets that fit with the application’s purpose, such as a live dashboard showing the available beds in a hospital ward, make sense running in the application itself. This has limits though, and systems running an application are usually not sized for serving full-on analytics workloads on top. The usage patterns for analytics differ significantly from the application’s. Fetching one patient’s current file has a different CPU, memory, and IO characteristic from performing population health analysis for hundreds or thousands of patients.

The traditional approach to deliver such medium to large scale analytics capabilities has been to extract the data from these source systems, transform it into an analytics-friendly schema, and then load it into a dedicated analytical database, usually a data warehouse. Dedicated ETL tools are often powerful and feature-rich, but they are inherently costly to set up, they require dedicated staff with a different expertise from the application operators, and usually benefits can only be achieved after a full implementation project rather than in smaller increments.

## A Unified Platform

With the **InterSystems IRIS data platform**, things are different. InterSystems IRIS is a *unified* platform, architected to support both operational and analytical workloads in a single technology. Not only does it offer a high-performance database for running mission-critical applications, but it also includes a comprehensive set of analytics capabilities, including reporting, business intelligence and machine learning. As such, many applications developed on InterSystems IRIS include complex embedded analytics components, including fully integrated dashboards and charts on live, large-scale data. InterSystems IRIS is also deployed in dedicated data warehouse scenarios, supporting analytics and machine learning on large data volumes.

This ability to support both extreme scalability for application use cases and top performance for analytical queries in a single technology makes InterSystems IRIS uniquely positioned for a model that does not rely on a costly ETL step. With InterSystems proven mirroring technology, data can be replicated from one system to another, without incurring noticeable additional load on the source system. This is a key enabler to run both your applications and analytics at scale, on the same live data. The rest of this white paper will describe how InterSystems mirroring facilitates this at an architectural level, and present use cases from two customers who have implemented this approach.

## Proven Technology

**InterSystems Mirroring** is routinely used in synchronous mode by InterSystems customers to implement high availability (HA) strategies, in which a backup mirror member remains on hot standby should the primary node go down. Mirroring is often also deployed in asynchronous mode in disaster recovery (DR) topologies that cross cloud regions or data centers. Mirroring operates on a low level and thereby achieves extremely high throughput rates using techniques such as journal compression and asynchronous IO.

Using asynchronous mirroring, customers can also set up a reporting node that is sized and used independently of the primary system, at almost zero configuration cost. This copy of the data can be queried as-is by analysts that are familiar with the application schema, or be used as the source for InterSystems native or third-party business intelligence solutions that project an analytical model on top of this schema. Data replicated through asynchronous mirroring is practically real-time, with latency usually in the single-digit millisecond range, and following an eventual consistency model fit for the majority of analytics and reporting use cases.

Customers with a need for a more substantial restructuring of the application schema into a data warehouse structure can still do so starting from the data in the reporting node using InterSystems IRIS native SQL capabilities, its dbt support, or other applicable tools. This model is similar to the ELT approach advertised by some tools – in which data is loaded into the target platform first and then transformed there – but still has the key advantage that the data can also be queried immediately. Because the use of InterSystems mirroring does not require any additional development or third-party technology, customers experience a much faster time-to-value. In case any transformations would still be beneficial, they can be developed on the reporting node in small increments, using an agile process based on specific analytical needs, rather than requiring long warehouse design and implementation projects.

Asynchronous mirroring is widely used today to replicate data across different cloud regions or from an on-prem installation to the cloud, and as such the technology offers total flexibility in deploying reporting nodes in the most appropriate cloud, on-prem, or hybrid topology.

Mirroring also supports replication to reporting nodes on a different, more recent version of InterSystems IRIS, enabling customers to take advantage of new features on

the reporting node, independently of the upgrade schedule of the main operational server. This opens the door to capabilities that were recently introduced to InterSystems IRIS, such as InterSystems Adaptive Analytics, Columnar Storage, and Foreign Tables. Used separately or combined with existing tools, these capabilities enable customers to build extremely powerful analytics solutions that take full advantage of the operational data in the source systems.

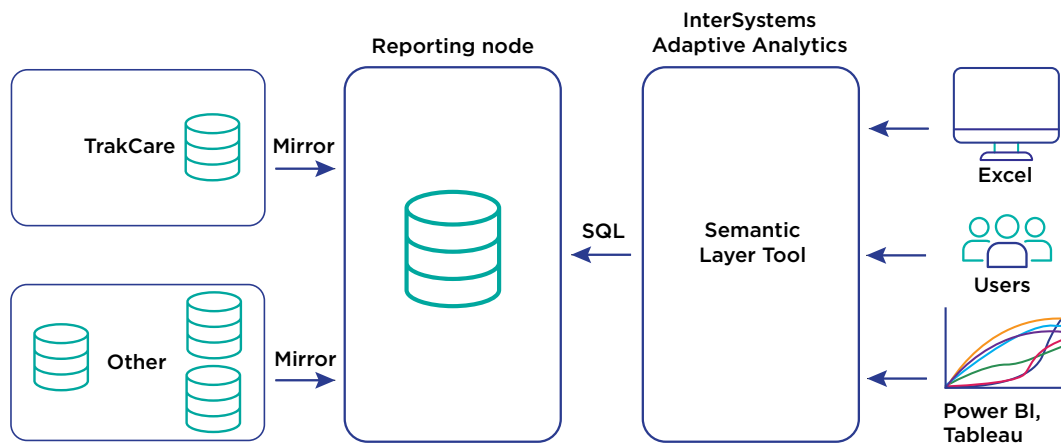
## From Technologies to Solutions

InterSystems Mirroring is a platform feature available on all applications and solutions based on the InterSystems IRIS data platform. Most of InterSystems current healthcare solutions are focused on application and integration use cases. For example, InterSystems TrakCare® is a healthcare information system that captures data such as electronic medical records for general hospital operations. The Advanced Analytics functionality within TrakCare enhances those use cases by adding a set of embedded dashboards for specific scenarios including adverse event monitoring and emergency department wait times. However, this data can also serve a purpose outside of those operational use cases. Adding an InterSystems IRIS reporting node that holds a copy of the production system's data, but can be queried freely without impacting the operational system, unlocks a vast potential for customers facing such analytical use cases. The entire TrakCare schema can be queried directly, or be used to populate data marts specific to a particular analytical need.

### Customer Example #1

Many TrakCare customers around the world are leveraging InterSystems Mirroring technology to replicate data from their production TrakCare instance to one or more reporting nodes that are dedicated to analytics. For example, several customers in Thailand are querying these reporting nodes using third-party technologies including Crystal Reports, Tableau, and bespoke solutions to complement TrakCare's embedded dashboards.

A leading hospital in France is using InterSystems Adaptive Analytics on their reporting node to build virtual OLAP cubes on top of TrakCare's operational schema, for analyses such as tracking KPIs on ward occupancy and patients with alerts on specific conditions such as COVID. Future use cases will combine TrakCare data with data from external sources to serve non-operational analyses and explore opportunities for data science and machine learning.



## Customer Example #2

Shift is Brazil's leading provider of laboratory information systems (LIS), processing more than 370 million clinical tests annually across their customer base. One of their largest customers, Sabin Laboratory, serves almost 6 million citizens and recently revised its system architecture to right-size the infrastructure. The volatile user workload stemming from about 3000 concurrent users is distributed across a compute tier of 10 application servers. The reporting workload is offloaded to a reporting node, leveraging InterSystems asynchronous mirroring to replicate the full 20TB dataset from the LIS. This enables reporting and business intelligence on all the data in real time without interrupting the operational processes, in an architecture that is both affordable and easy to extend as the processing and reporting needs evolve.

## Limitations

InterSystems Mirroring does not impose any limits on which InterSystems IRIS product versions can be the source or target when replicating application data. For the application schema, a read-only version used for SQL querying can be used across different releases without issue, as long as the major version number is the same. For replicating an application schema to a release with a different major version number, please consult with your InterSystems account team to determine which combinations are supported. In many scenarios where the schema's source code is available, recompiling the schema on the reporting node is all that's needed to support read-only data access.

## Conclusion

Customers across industries have deployed their mission-critical operational applications on InterSystems IRIS. Many customers are enriching these applications with embedded analytics, but increasingly identifying all-new analytics use cases and finding opportunities to monetize the application data. For those use cases, InterSystems Mirroring technology enables customers to quickly deploy reporting nodes that replicate the data from their operational system in an environment that is dedicated to reporting and analytics. A growing number of InterSystems solution customers are deploying this pattern successfully and experiencing fast time-to-value for their analytics projects, with ample opportunities for expansion through InterSystems IRIS broad suite of analytics technologies. All the required technology is available today, and included with every InterSystems product, only a simple configuration step away.

If you'd like to run an analytics workload on your operational data with minimal impact to applications, please reach out to your account team to discuss your architecture.

