ESG Lab Review

InterSystems IRIS Data Platform: A Unified, Efficient Data Platform for Fast Business Insight

Date: April 2018 **Author:** Kerry Dolan, Senior IT Validation Analyst

Abstract

This ESG Lab Review documents our audit of InterSystems IRIS Data Platform performance testing. Tests were executed in a proof of concept environment, with comparisons to three common databases. Testing focused on performance of common queries run in the financial services industry.

The Challenges

Transactional databases are essential to many revenue-driving business processes, and data analytics can improve competitive advantage. But organizations often struggle to fit the traditional RDBMS model to the types of data they have and knowledge they seek. NoSQL and NewSQL databases offer valuable features, but using multiple databases means building out different infrastructures, each with performance, scalability, and flexibility needs. In addition, organizations are always looking to increase performance to speed business insight; in-memory databases provide speed, but become challenging with large amounts of data. All these demands drive up costs. It is no surprise, then, that when ESG asked IT decision makers about database challenges, 48% of respondent organizations cited managing database size and growth, 35% cited meeting database performance requirements, and 34% cited supporting cloud-based databases.[1]

Figure 1. Database Challenges

48% cited managing data growth and database size

35% cited meeting database performance requirements

34% cited supporting cloud-based database

Source: Enterprise Strategy Group

Most organizations end up using multiple databases, adding infrastructure and staffing costs. If organizations can reduce the footprint required to handle all their database needs—transactional and analytics workloads, including real-time, historical, natural language processing, and business intelligence—they can save money. But they must be able to deliver high performance, along with data durability, scalability, and management ease.

---

[1] Source: ESG Research Survey, *Enterprise Database Trends*, January 2017.

The Solution: InterSystems IRIS Data Platform

The InterSystems IRIS Data Platform was designed to help organizations build intelligent, data-rich applications that support multiple transaction processing and analytical use cases. InterSystems IRIS is a unified, real-time platform that combines powerful data management, analytics, and interoperability into a single solution. It was designed to handle and optimize mixed workloads and data-intensive applications at scale in the cloud, on-premises, or in hybrid environments, while providing built-in security, proven reliability, and support.

InterSystems IRIS handles the varied needs of businesses today, all in a single platform, with a single copy of data. This saves organizations from having to purchase and manage different point solutions, and from storing and mapping multiple data copies. Features include:

- *Multi-workload*. InterSystems IRIS processes both transactional and analytic workloads in the same engine at scale, eliminating the complexity associated with moving and transforming transactional data for analysis, and with multi-database deployments. InterSystems IRIS can process and index transactional data while simultaneously running real-time or historical analytics, delivering a faster response to the business and therefore better insight.
- *Multi-model*. InterSystems IRIS provides a single database that can represent multiple projections of the same data, rather than a traditional siloed database approach. It supports numerous data models, including relational, key-value, document, object, and a variety of schema-free NoSQL data.
- *Interoperability*. The comprehensive integration and open analytics platform provide DBAs with streamlined data coordination and business orchestration, making it simpler to connect data sources and business processes. InterSystems IRIS incorporates real-time and batch analytics, including distributed SQL processing, business intelligence, predictive modeling runtime engine, and natural language processing of unstructured data from sources such as emails and social media, and integrates with big data processing frameworks such as Apache Spark.
- *Analytics*. InterSystems IRIS incorporates real-time and batch analytics, including distributed SQL processing, business intelligence, predictive modeling runtime engine, and natural language processing of unstructured data from sources such as emails and social media, and integrates with big data processing frameworks such as Apache Spark.
- *Enterprise-level security*. Integration with Kerberos, LDAP, and KMIP, role-based access control, and encryption of data in transit and at rest provide rock solid security.

Figure 2. InterSystems IRIS Data Platform
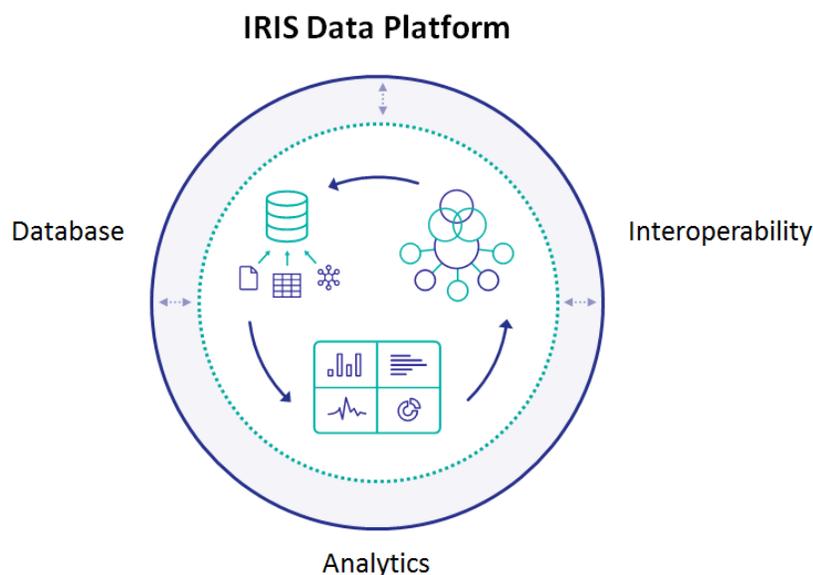

Source: InterSystems Corporation

InterSystems IRIS Architecture

InterSystems IRIS can ingest and store transactional data fast, while at the same time processing analytic workloads on batch and real-time, ACID-compliant data; this eliminates the delays from moving real-time data to an analytics platform for processing. It scales vertically with parallel SQL processing, and employs a differentiated approach to horizontal scaling

For large data sets and complex queries, InterSystems IRIS employs a specialized approach to horizontal scaling. Using commodity servers to distribute query processing, it leverages intelligent sharding and distributed caching. Workloads and tables can be distributed to multiple nodes by sharding the database; sharding is transparent to the application, and requires few if any code changes. InterSystems IRIS uses an intelligent Shard Master that pushes out application code for parallel execution, aggregates the results of each shard, and returns a final result (see Figure 3). Each shard is aware of the location of all of the other data; if data needed for a query resides in another shard, one shard can reach out directly to another to obtain *only* the data it needs. With other solutions, the entire table would have to be broadcast to the other nodes, increasing latency and timeouts. This enables InterSystems IRIS to achieve high performance even for queries that involve complex joins of partitioned data.

Each server's cache is independent, as the sharding creates disjoint partitions. Scaling out servers will therefore increase the cluster's memory in a linear fashion. As a result, when sized appropriately, InterSystems IRIS deployments can achieve levels of performance similar to in-memory databases, but without the requirement to fit all data into memory, as it uses the cache for frequently-used data, with "colder" data durably persisted on disk.

Figure 3. Horizontal Scalability
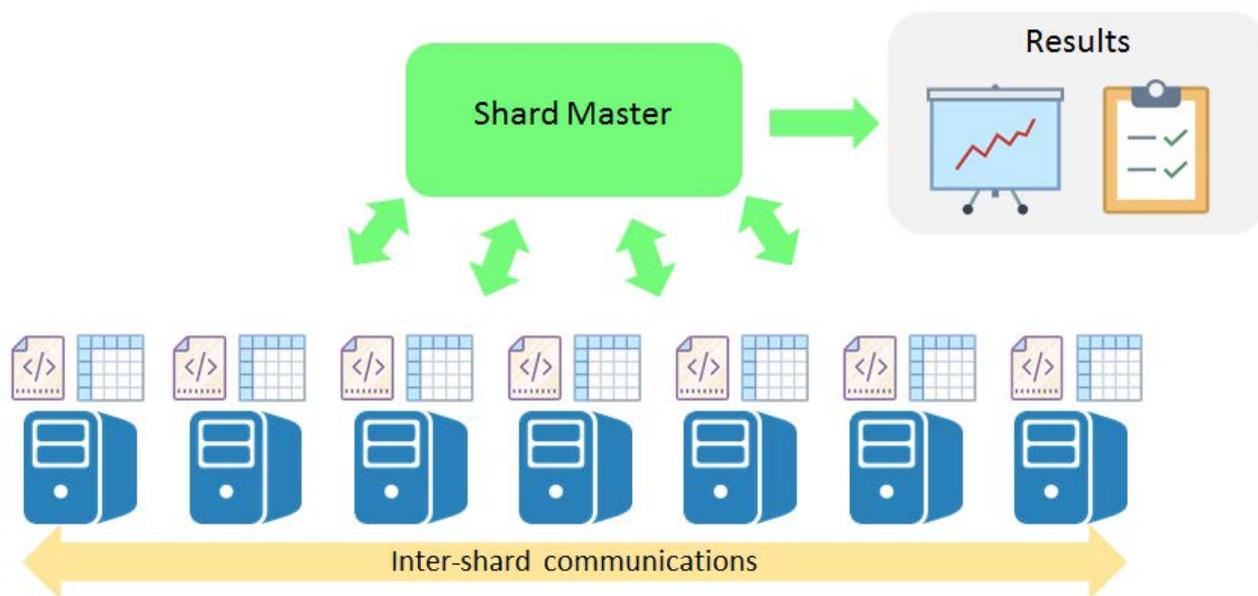
Source: Enterprise Strategy Group

ESG Lab Tested

ESG Lab audited results from tests performed on a database application proof of concept by a financial industry firm. Testing focused on analytical query performance of InterSystems IRIS and several other databases commonly used in the financial industry, leveraging thousands of financial instruments. The application supports multiple, firm-wide mission critical lines of businesses; algorithmic trading apps require the SQL queries to successfully execute within strict performance constraints, while hundreds of users across the firm need ad hoc access to the query results. Millisecond performance is critical; if the queries cannot be completed within the required performance SLAs, the algorithmic trades will not execute and the firm cannot successfully carry out its trading strategies.

The test bed included InterSystems IRIS and three other common database deployments. *Table 1 shows that the InterSystems IRIS configuration used less compute infrastructure (including fewer VMs, cores, and RAM), while running queries against more data than the competitors (measured in days, GBs, and rows in Table 1 and Figure 4).*

Table 1. Test Bed Configurations

| Database	Configuration	Days of Data	Dataset Size (GB)	Dataset Size (Rows)
InterSystems IRIS	**4 VMs: 8 cores, 96GB RAM, 400GB local storage**	**135**	**320**	**200,000,000**
Vendor A	8 VMs: 16 cores, 256GB RAM, 400GB local storage	14	33	20,740,000
Vendor B	8 VMs: 16 cores, 256GB RAM, 400GB local storage	14	33	20,740,000
Vendor C	3 dedicated physical servers, each with 24 cores and 256GB RAM, 2TB shared SAN storage	20	47	29,620,000

Figure 4. Dataset Size Comparison
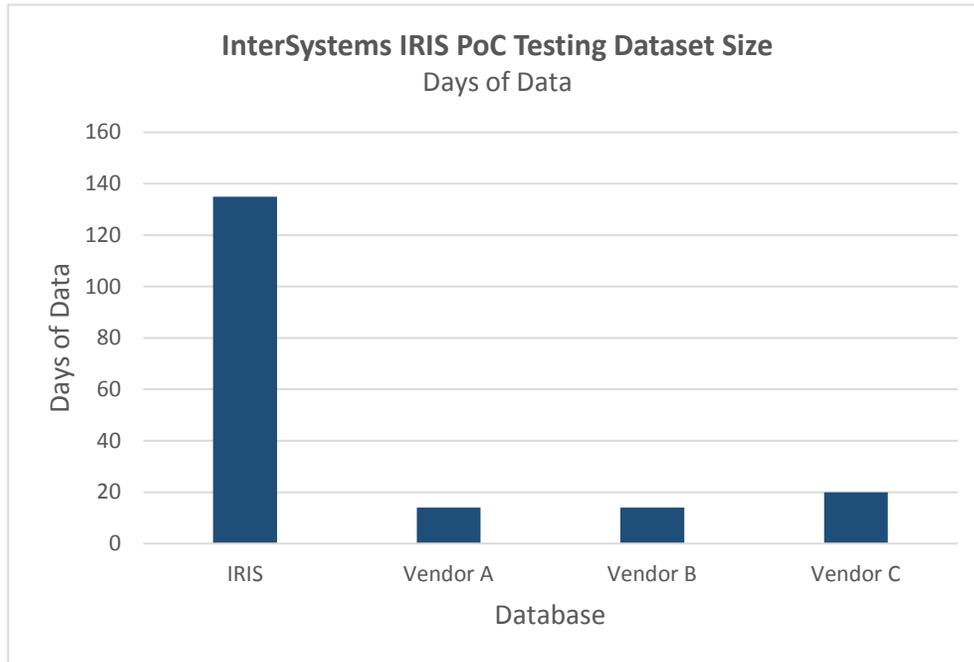
Source: Enterprise Strategy Group

Five queries were tested. All of them were complex joins among five to six tables with eight to ten predicates in the WHERE clause, with results sets containing between 19 and 6,504 rows, and each row containing hundreds of fields. Response times were measured as each query was executed against an increasing number of concurrent users: 1, 10, 25, 50, and 100.

Figure 5 presents the results of Query 1, returning 19 rows. Performance was measured by how many milliseconds each query required to complete as the number of concurrent users scaled.

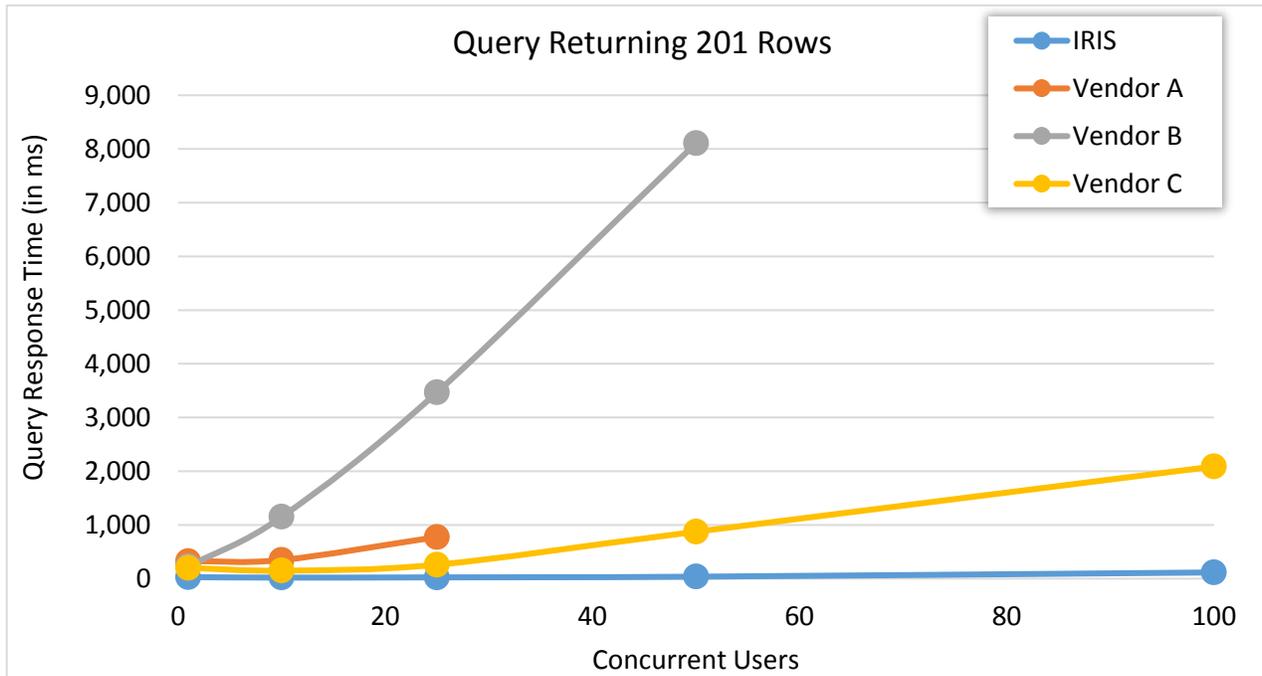**Figure 5. Query 1 Performance as Concurrent Users Scaled**
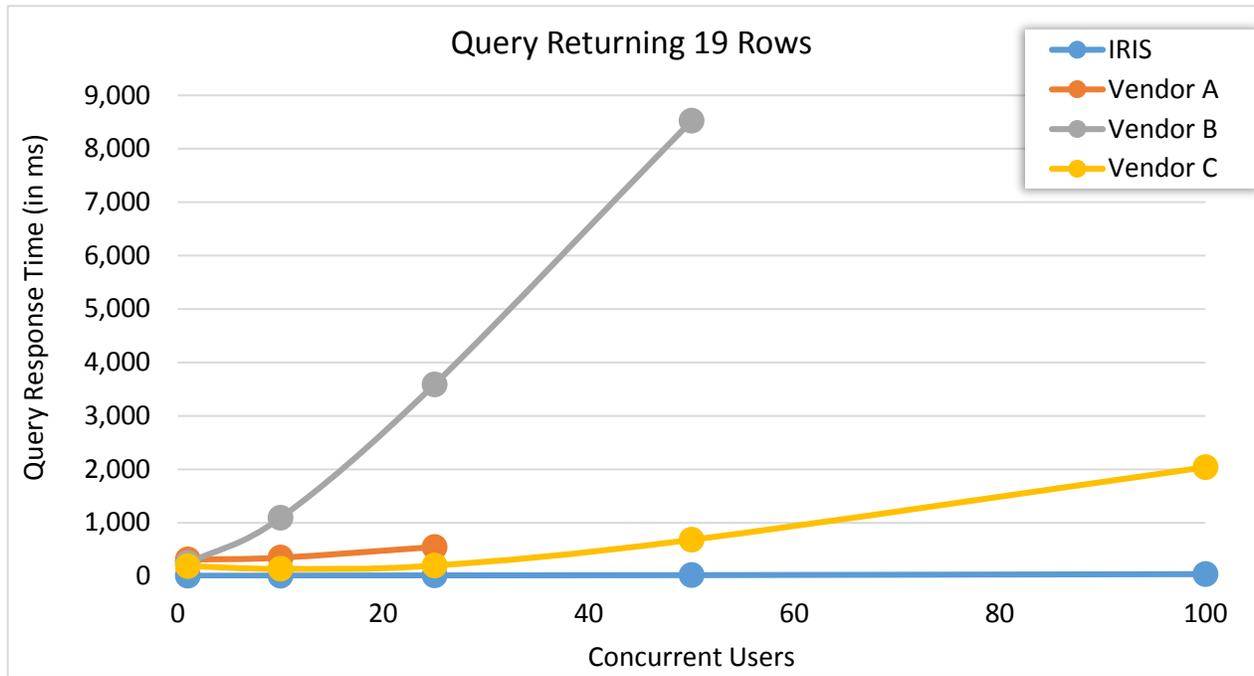
Source: Enterprise Strategy Group

What the Numbers Mean

- InterSystems IRIS was operating on less infrastructure, and accessing 5-8X more data than the other vendors.
- For all five queries, results were plotted only when no errors occurred. Vendor A was plotted through 25 users, since errors were encountered at 50 and 100 concurrent users; Vendor B was plotted through 50 users. These solutions were unable to deliver results at scale, limiting their usefulness.
- For Query 1, InterSystems IRIS performance was significantly faster than other databases for all concurrent user counts. InterSystems IRIS performance ranged from 8 ms to 38.97 ms, increasing by only 4.87X as concurrent users scaled to 100.
- Response time for a single user was 310 ms for Vendor A (37X slower than InterSystems IRIS), 245 ms for Vendor B (30X slower than InterSystems IRIS), and 188 ms for Vendor C (22X slower than InterSystems IRIS).
- Vendor C offered second best performance. At 50 users, Vendor C was 35X slower than InterSystems IRIS; at 100 users, Vendor C was 51X slower than InterSystems IRIS.

Queries 2-4 delivered similar results, with InterSystems IRIS consistently faster. Query 2 returned 201 rows, Query 3 returned 600 rows, and Query 4 returned 984 rows; Figures 6, 7, and 8 show these results.

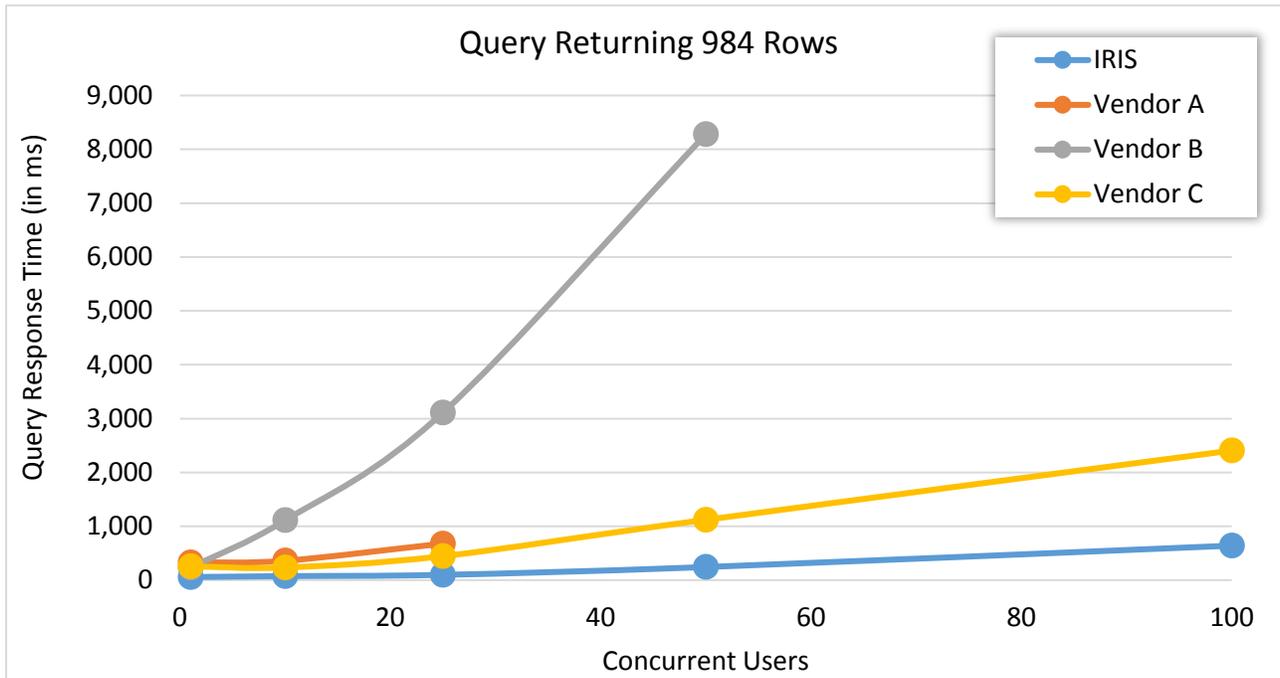**Figure 6. Query 2 Performance as Concurrent Users Scaled**

Query Returning 201 Rows

Source: Enterprise Strategy Group

Figure 7. Query 3 Performance as Concurrent Users Scaled

Query Returning 600 Rows
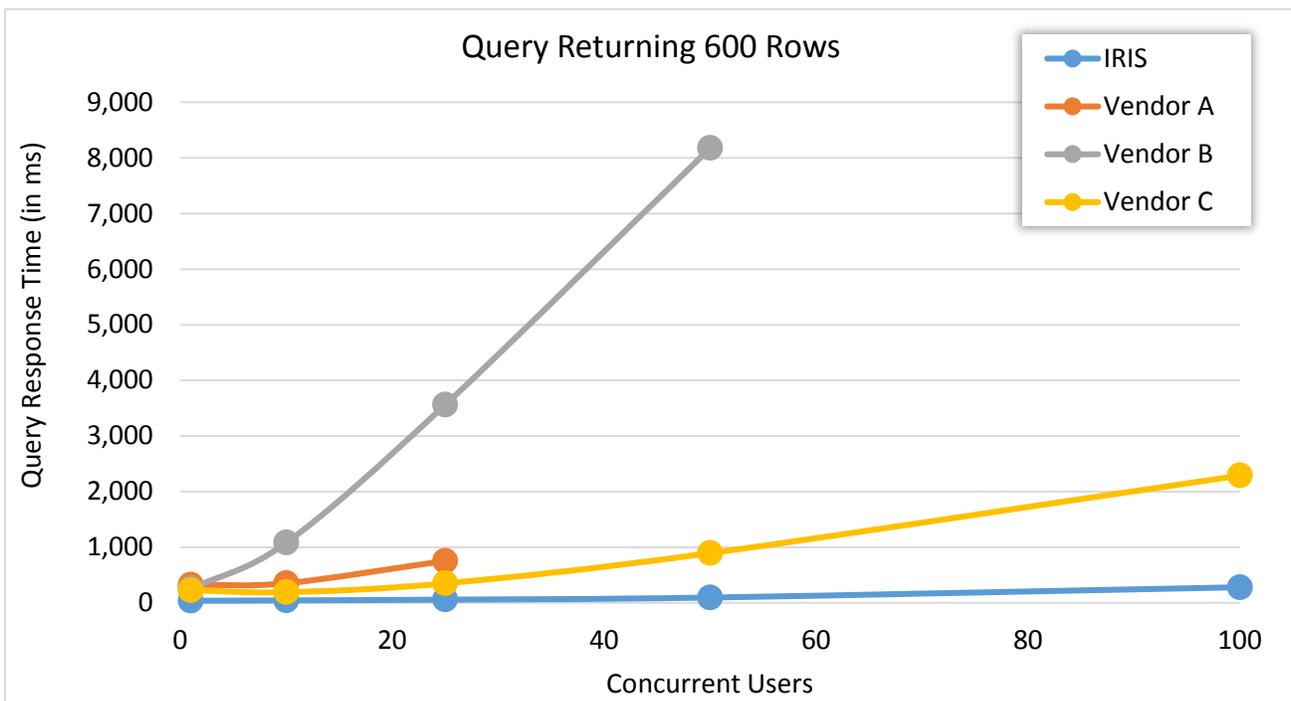
Source: Enterprise Strategy Group

Figure 8. Query 4 Performance as Concurrent Users Scaled

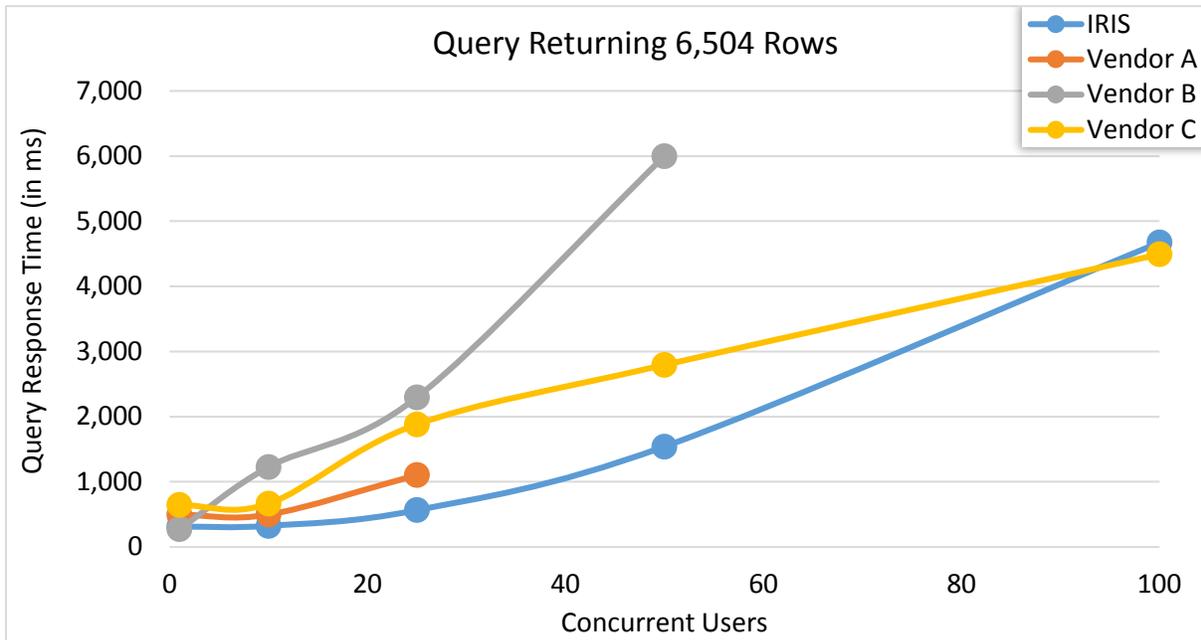*Source: Enterprise Strategy Group*

What the Numbers Mean

As the queries returned more rows, they took more time, as would be expected. In these three queries:

- For a single user, other vendors' performances were 3-10X slower than InterSystems IRIS.
- For 10 concurrent users, other vendors' performances were 2-59X slower than InterSystems IRIS.
- For 25 concurrent users, other vendors' performances were 3-147X slower than InterSystems IRIS.
- For 50 concurrent users, other vendors' performances were 3-218X slower than InterSystems IRIS.
- Only vendor C was able to support 100 users, and it ranged from 2-17X slower than InterSystems IRIS.

Figure 9 presents the results of Query 5, which returned 6,504 rows. For this more complex query, performance was more impacted as the number of concurrent users scaled.

Figure 9. Query 5 Performance as Concurrent Users Scaled

Source: Enterprise Strategy Group

What the Numbers Mean

- InterSystems IRIS was operating on less infrastructure, and accessing more days of data than the other vendors.
- InterSystems IRIS performance ranged from 312 ms for a single user to 4,673 ms for 100 users, an increase of 14X.
- Vendor C performed slightly faster than InterSystems IRIS at 100 users, achieving 4,492 ms to InterSystems IRIS's 4,673.

ⓘ Why This Matters

According to ESG research, two of the top three business initiatives driving technology spending over the next year are *cost reduction* and *improving data analytics for real-time business intelligence and customer insight*.[2] The only initiative cited more often was strengthening cybersecurity. It's no wonder analytics is a priority—gaining business insight using data analytics can improve decision making. The more data you can access, and the faster your queries return results, the better your decisions will be.

ESG Lab validated that InterSystems IRIS performed consistently faster than other tested databases, while using less infrastructure and accessing more data than the others. While one vendor performed slightly better for a single data point (on Query 5 at 100 users), that is unlikely to make it worth the cost and burden of adding another point solution. InterSystems IRIS handled all the query types with fast performance, scaled easily, cost less in infrastructure, and experienced no failures. By accessing more data and delivering faster results, InterSystems IRIS can deliver better business insight.

The Bigger Truth

Organizations gain competitive advantage by leveraging information. Companies in finance, healthcare, manufacturing, telecommunications, and just about every other industry know that analyzing trends leads to better decision making. But it can be frustrating when you finally get the results of analysis that reveals that customers prefer the red widget—while in the meantime, you had beefed up production of the blue one. The more—and faster—organizations can understand customers, suppliers, and market trends, the better they can make business-driving decisions. The challenge is getting insight from as much data as possible, as fast as possible, and turning that into action quickly.

InterSystems IRIS is a unified, cost-efficient platform that was designed to improve decision making by letting organizations combine data-driven intelligence with real-time business processes. InterSystems IRIS is an innovative knowledge platform that speeds analytical queries on real-time data by eliminating the need for different transactional and analytics databases. Organizations can run transactional and analytic workloads using the same engine, without having to move, map, or translate data; the data can be represented in multiple ways at the same time, eliminating the need to deploy multiple point solutions, and resulting in faster performance.

ESG Lab validated InterSystems IRIS performance against several common database solutions. Results demonstrated that InterSystems IRIS provided faster performance, leveraging more data, and using less infrastructure—enabling organizations to gain better insight, faster, for less money. With more in-depth insight, organizations can increase revenue, improve customer experiences, create tailored products and services, improve compliance, reduce risk, make operations more efficient, and reduce costs. If these are of interest in your organizations, we recommend evaluating InterSystems IRIS.

The goal of ESG Validation reports is to educate IT professionals about information technology solutions for companies of all types and sizes. ESG Validation reports are not meant to replace the evaluation process that should be conducted before making purchasing decisions, but rather to provide insight into these emerging technologies. Our objectives are to explore some of the more valuable features and functions of IT solutions, show how they can be used to solve real customer problems, and identify any areas needing improvement. The ESG Validation Team's expert third-party perspective is based on our own hands-on testing as well as on interviews with customers who use these products in production environments.