Choosing a DBMS to Address the Challenges of the Third Platform

An IDC InfoBrief, sponsored by InterSystems | May 2017
Introduction

In October of 2016, IDC conducted research sponsored by InterSystems that focused on the requirements for data management solutions that support both transactional and analytic workloads for global organizations, and benefits such solutions might bring.

A total of 502 organizations were interviewed across Australia, Brazil, China, Germany, Japan, UK and the United States.

Companies participating in the study represented diverse industries and company sizes.

Respondents were chosen based on your familiarity with both transactional and analytic databases in use at their company.

This report represents the key findings, analysis, and take aways from this study.
Companies Want to Speed Innovation and Streamline Operations

What Is Your #1 Business Objective This Year?

- Speed innovation: 33.9%
- Streamline operations: 31.9%
- Reduce cost: 18.4%
- Simplify architecture: 15.8%

Q. Please rank your organization’s top 3 business objectives for the upcoming year.

Source: 3rd Platform Information Management Requirements Survey, IDC, October, 2016, n=502
The migration of IT systems to what IDC calls The 3rd Platform brings with it new IT requirements that impact information management systems.

New workloads are requiring support for more data types, larger data sets, and faster turnaround.

The variety of problems to be solved calls for a variety of data management tools to solve them.

Users are recognizing the need for better performance, greater agility, and support for more data types.

Q. Please rank your organization’s top 3 business objectives for the upcoming year.

The Third Platform is collective of trends and technologies that represents the third major wave of IT. The result is an enterprise that is nimble, adaptable, and responsive to changing business conditions.
### All Data Types Are Important

#### How Important Are the New Data Types?

(rating scale, 1 = not very important, 5 = very important)

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relational</td>
<td>4.31%</td>
</tr>
<tr>
<td>Data from the Internet of Things (IoT)</td>
<td>4.30%</td>
</tr>
<tr>
<td>Streaming data from external sources</td>
<td>4.27%</td>
</tr>
<tr>
<td>Sensor data</td>
<td>4.22%</td>
</tr>
<tr>
<td>Graphs</td>
<td>4.22%</td>
</tr>
<tr>
<td>Key Value</td>
<td>4.17%</td>
</tr>
<tr>
<td>Video/audio/image</td>
<td>4.17%</td>
</tr>
<tr>
<td>Object</td>
<td>4.16%</td>
</tr>
<tr>
<td>JSON documents</td>
<td>4.13%</td>
</tr>
<tr>
<td>Geospatial data</td>
<td>4.10%</td>
</tr>
</tbody>
</table>

Source: 3rd Platform Information Management Requirements Survey, IDC, October, 2016, n=502
Nearly 2 out of 5 are evaluating new database platforms while few are moving to open source.

What Is Your #1 Technical Objective This Year?

- Evaluate new database technologies: 37.0%
- Retire older versions of databases: 24.9%
- Move applications to SaaS: 20.8%
- Move to open source databases: 17.3%

Q. Please rank your organization’s top 3 technical objectives related to databases for the upcoming year.

Source: 3rd Platform Information Management Requirements Survey, IDC, October, 2016, n=502
How Important Is Open Source?

It is commonly assumed by many that the new third platform workloads require open source DBMS as a foundation. Our survey determined that this is not at all the case.

Q. Which of the following statements is most true regarding your policy in relation to open source?

- 45.2%: All new DBMS technology must be proprietary (e.g. not open source)
- 21.5%: We prefer open source for new DBMS technology, but are open to other technology as well.
- 17.1%: All new DBMS technology must be open source.
- 16.2%: We just want the best DBMS fit with respect to technical characteristics, overall cost, and fit with the problem at hand.

Source: 3rd Platform Information Management Requirements Survey, IDC, October, 2016, n=502
The Great Divide: Transactions vs. Analytics

Transactions

- Record-oriented processing.
- Drives business operations.
- Designed for write, not query, speed.
- Minimal operational schema.

Analytics

- Involves data from many transactional databases.
- Optimized for query speed.
- Schema designed to help answer key questions.
Real-Time Decisioning

**Strategic**
- Deep analytic models and data warehouse reporting inform executive decision making
- Data analysts comb tons of data, executive management makes the decisions

**Operational**
- Project or product-focused data is collected and analyzed
- End-users run queries against data marts and adjust their project or product plans

**Tactical**
- Immediately available current data and streaming data are presented
- Line and field staff, or automated computer algorithms, make decisions

---

Tactical decisions drive the business day-to-day, and depend on real-time operations.
Steamlining the Business Requires Real-time Analytics

Slow Data Integration

Data transformed and merged from many sources takes time, and removes its relevance from the decision window. Any change to any source requires changes to the transformation, and possibly to the analytic database schema, resulting in inflexibility.

Two Separate Database Types

Databases optimized for transactions are not usually able to perform complex analytic queries in a timely manner, if at all. Databases optimized for analytics are usually too slow at transaction processing to meet application throughput requirements.
Choosing a DBMS to Address the Challenges of the Third Platform

Large Amounts of Transactional Data is Moved At a Slow Rate Via Extract / Transform / Load (ETL) technologies

- **1%-24%**: 65.4%
- **25%-49%**: 13.5%
- **50%-74%**: 18.6%
- **75%-99%**: 1.0%
- **100%**: 1.5%

Q. What percent of data is moved between transactional and analytical systems using Extract / Transform / Load (ETL)?

Source: 3rd Platform Information Management Requirements Survey, IDC, October, 2016, n=502
Data moved via ETL is Not Moved Quickly

Q. On average, for each ETL task, how old is the data by the time it reaches the analytic database?

Source: 3rd Platform Information Management Requirements Survey, IDC, October, 2016, n=502
Can You Use Live Data to Make Decisions “In the Moment”?

- If you want to track business during the day, some latency may be acceptable
- Making a decision based on live data requires the ability to blend analytical queries into transactions
- Such a capability demands zero latency
- Even CDC (changed data capture) cannot deliver zero latency

**Q. On average, for each CDC connection, what is the latency of data moving into the analytic database?**

![Age of CDC Data (%)](chart)

- **Less than 1 minute**: 4.0%
- **Between 1 minute and 10 minutes**: 30.6%
- **Between 10 minutes and 1 hour**: 49.6%
- **Between 1 hour and 2 hours**: 13.5%
- **More than 2 hours**: 2.3%

Source: 3rd Platform Information Management Requirements Survey, IDC, October, 2016, n=502
More than 3/4 of Respondents Say That Untimely Data Has Inhibited Business Opportunities

Q. Has the inability to analyze current live data inhibited your organization’s ability to take advantage of business opportunities?

Source: 3rd Platform Information Management Requirements Survey, IDC, October, 2016, n=110
Untimely Data is Inhibiting in Many Ways

<table>
<thead>
<tr>
<th>Category</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Productivity/Agility</td>
<td>27%</td>
</tr>
<tr>
<td>Analysis</td>
<td>25%</td>
</tr>
<tr>
<td>“Customer Understanding”</td>
<td>15%</td>
</tr>
<tr>
<td>Sales/Marketing</td>
<td>14%</td>
</tr>
<tr>
<td>Other</td>
<td>7%</td>
</tr>
<tr>
<td>Security</td>
<td>6%</td>
</tr>
<tr>
<td>Revenue/Margin</td>
<td>5%</td>
</tr>
</tbody>
</table>

Q. Briefly describe in what way the inability to analyze current live data has inhibited your organization’s ability to take advantage of business opportunities?

Timely Data Benefits - Voice of the Customer

“Fraud can be detected the moment it happens and proper measures can be taken to limit the damage”

“To make robust Research and development we need this”

“Live data analysis speeds up decision-making processes on various fronts”

“Errors within the organization are known instantly”

“We can understand our customer better”
Untimely Data Also Limits Operational Efficiency

- 17% Yes, somewhat
- 32% Yes, moderately
- 37% No
- 9% Don’t know
- 5% Yes, significantly

Q. Has the inability to analyze current live data inhibited your organization’s ability to improve operational efficiency?

Source: 3rd Platform Information Management Requirements Survey, IDC, October, 2016, n=110
The Cost of Many Databases

Most respondents had more than 5 production transactional DBs

25% HAVE MORE THAN 10

OVER 60% HAVE MORE THAN 5 ANALYTICAL DATABASES

AND OVER 30% HAVE MORE THAN 10
In all cases, they allocate 1-2 DBAs per database.

Growing Complexity
- Fixed resources (servers and storage)
- Every allocation change is a major project
- Application isolation leads to system glut

Mounting Cost of Database Management
- Conventional Database Deployment = Fixed Costs
- Deployment is for the High Water Mark
- Result: Persistent Hardware Underutilization
- Staffing costs escalate as # of servers increases, and systems are allocated for the high water mark
The Need for Consolidated Database Operations

We have seen that most users need a **broad variety of data type support** that goes well beyond what native RDBMSs provide.

We have seen that tactical decisions cannot be supported as long as data is segregated into transactional and analytical databases.

We have also seen that the maintenance of many databases leads to excessive cost and complexity in the data center.

The answer is to **concentrate key, closely related data** together for combined analytic and transaction processing, ideally including a range of data types in support of digital transformation.

How important did our respondents regard the need to blend data management capabilities in one DBMS? See next.
More than 90% of Respondents Think Blended Features are Very Valuable

- Moderately valuable: 46.0%
- Significantly valuable: 45.4%
- A little valuable: 7.9%
- Don’t know: 0.6%

- Flexible scalability, AND
- Blending relational and nonrelational data, AND
- Performing complex analytical queries against live operational data without impacting transaction or operational performance

Q. How valuable do you believe a database system capable of delivering all of the following would be to your organization?

Source: 3rd Platform Information Management Requirements Survey, IDC, October, 2016, n=502
Conclusion

The era of digital transformation enabled by third platform technologies brings with it the need to analyze an array of datatypes and to analyze transactional data at near real-time speed.

Many organizations today are still using ETL data transfer approaches which cannot meet requirements for real-time data analysis.

Untimely data significantly inhibits businesses ability to compete and remain agile.

A solution to these challenges is to consolidate database operations combining transactional and analytic capabilities in a single DBMS solution with blended features.