Progress®

2018 Data Connectivity Annual Report

Discover the latest trends in our 5th annual
database technology survey

REPORT

Table of Contents

Executive Summary

Progress, the trusted leader in data connectivity and integration, is excited to present the results of our 5th annual vendor-neutral Data Connectivity Survey. This year we had 1,400+ respondents – a number that has more than quadrupled since the inaugural survey in 2014. This growth in participants has given us a tremendous amount of insight into this rapidly-changing market.

Enterprise data has changed dramatically over the past decade in terms of volume, variety and velocity. This increase in the number and types of data sources and formats has fueled an increased demand for high-level data visualization and complex analytics. To overcome these challenges companies are adopting multiple BI tools and embedded analytics. The survey findings also highlight the role that cloud computing is expected to play as part of this ongoing digital disruption.

Along with those trends, there are also pressing business requirements for standards-based access to business-critical data and for real-time connectivity across the entire hybrid architecture, while complying with ever changing regulations and data security standards. For many organizations, these represent huge barriers to stay ahead of the complex data connectivity curve.

To help enterprises navigate these challenges, this year's report, which benchmarks how organizations are handling the impact of the changing landscape of disruptive data sources, takes an in-depth look at where organizations are on their data connectivity journey and uncovers the challenges they face.

Key Findings

- Relational Databases are still critical for many enterprises, making them very popular among ISVs as well.

- Standards-based access is becoming more important as the number of data sources continues to grow at a rapid pace

- REST API has become the standard framework for application integration , as industries continue to move away from SOAP

- OData, the industry Standard REST API, is growing in popularity

- Data integration has become the #1 challenge. The right solution should be based on industry standards and support hybrid access across a wide range of legacy and emerging data sources

- The increase in cloud adoption by enterprise organizations makes real-time hybrid connectivity critical

- Open analytics is on rise as organizations seek to query cloud applications with their favorite analytics tool or programming language

- Increased data security vulnerabilities, penalties and regulations are creating new challenges and opportunities for data integration

Progress®

Breaking Down the Numbers

1400+
PARTICIPANTS

43%
NA

33%
EMEA

17%
APJ

7%
SA

Role

Architects
BI Specialists
Business Analytics
CEOs/CIOs/CTOs
Consultants
Data Scientists
DBAs
Developers
Sys Admins

Industry

Education
Energy and Utilities
Financial Services
Government Military
Healthcare
IT Services
Manufacturing
Retail
Software
Telecommunications
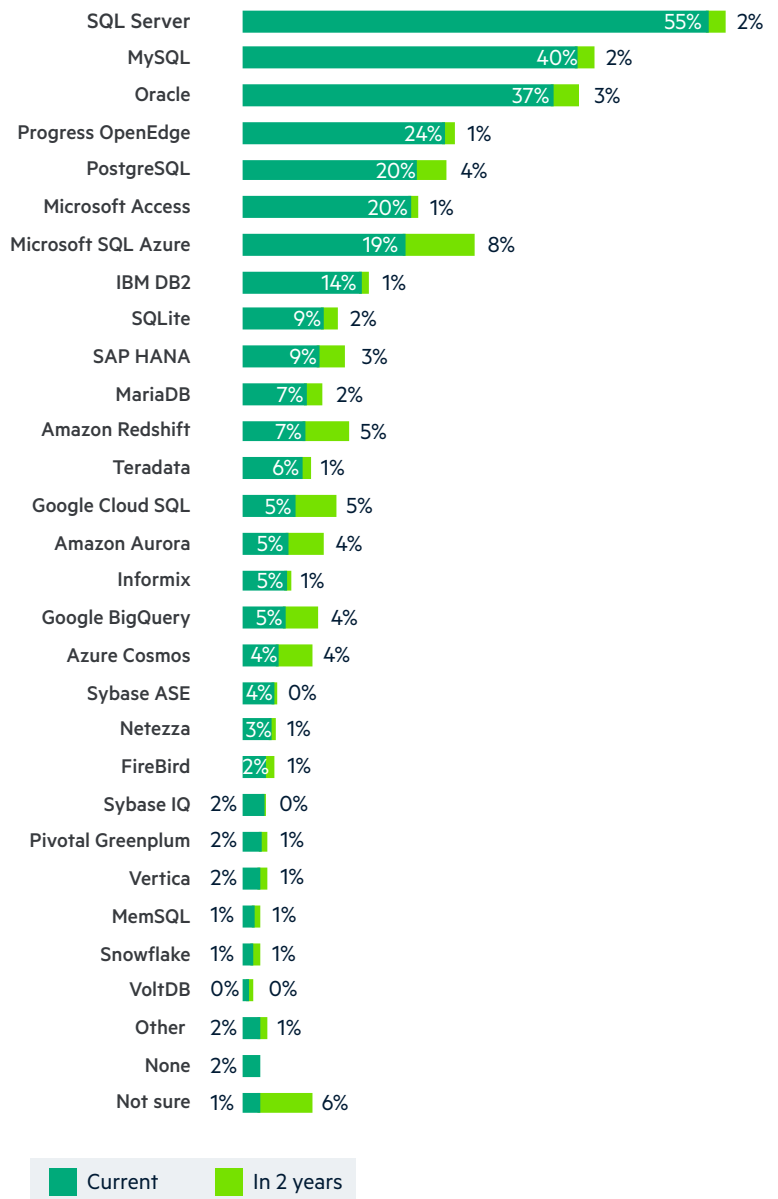Transportations

and more…

Progress®

Modern Data Platforms on the Rise

Data is the engine of modern business. Emerging technologies make data management more diverse, complex and challenging. The past year has seen the rise of powerful analytics and an embrace of new tools and platforms to more effectively tap into the power that data offers. Businesses modernizing their data platforms—platforms that allow fluidity of data movement and secure management from edge to cloud—will continue accelerate. But, for important data there is no substitute for Relational databases.

Relational Data Sources (RDBMS)

Many new additional relational databases for storing and processing data have emerged in the last 10 years, but SQL server, MySQL and Oracle remain the most popular databases. Although they've seen an on-premises decline of 13% (combined) from 2017 to 2018, that can be attributed to migration to cloud offerings. While cloud has its advantages such as cheap infrastructure and high availability, many enterprises remain on-premises due to potential security and compliance concerns. Also, modern options such as big data, NoSQL still aren't the right fit for all business needs, with analytical tooling for these modern databases still in its infancy. Thus, RDBMS databases are keeping pace and are here to stay in the enterprises for the foreseeable future.

Which relational/analytics database technologies do you or your customers currently use? Which do you or your customers plan to adopt in the next two years?

Database	Current	In 2 years
SQL Server	55%	2%
MySQL	40%	2%
Oracle	37%	3%
Progress OpenEdge	24%	1%
PostgreSQL	20%	4%
Microsoft Access	20%	1%
Microsoft SQL Azure	19%	8%
IBM DB2	14%	1%
SQLite	9%	2%
SAP HANA	9%	3%
MariaDB	7%	2%
Amazon Redshift	7%	5%
Teradata	6%	1%
Google Cloud SQL	5%	5%
Amazon Aurora	5%	4%
Informix	5%	1%
Google BigQuery	5%	4%
Azure Cosmos	4%	4%
Sybase ASE	4%	0%
Netezza	3%	1%
FireBird	2%	1%
Sybase IQ	2%	0%
Pivotal Greenplum	2%	1%
Vertica	2%	1%
MemSQL	1%	1%
Snowflake	1%	1%
VoltDB	0%	0%
Other	2%	1%
None	2%	
Not sure	1%	6%

Legend: ■ Current ■ In 2 years

>>> Relational Databases are still critical for many enterprises, making them very popular among ISVs as well. Cloud and NoSQL databases are not here to replace, but to solve different problems for businesses.
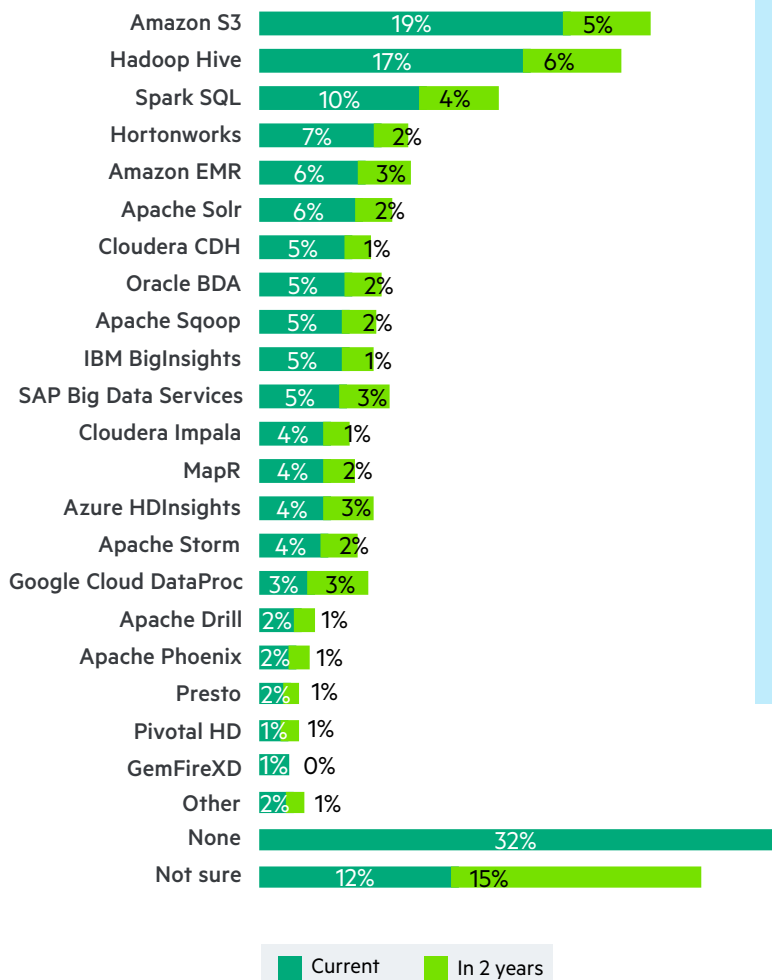
Progress®

Big Data Sources

Adoption of big data platforms continues apace, with the segment seeing growth of 7% since the 2017 survey. Amazon Simple Storage Service (S3) was the clear favorite among respondents. ETL tools such as AWS Glue which makes data loading chores simple and efficient enable adoption of S3 simple.

Hadoop Hive, the de facto standard for SQL on Hadoop, remains a popular big data interface. As demand for analytics continues to increase, Hadoop Hive adoption is expected to grow by 6% in the next two years.

Apache Spark is gaining more importance, with Spark SQL as its most popular interface. In addition to providing support for various data sources, it makes possible to weave SQL queries with code transformations which results in a very powerful tool. Its demand is evident from a 3% rise from 2017 and expected adoption growth of 4% over the next two years.

What Big Data platforms/interfaces do you or your customers currently use? And which do you or your customers plan to adopt in the next two years?

Platform	Current	In 2 years
Amazon S3	19%	5%
Hadoop Hive	17%	6%
Spark SQL	10%	4%
Hortonworks	7%	2%
Amazon EMR	6%	3%
Apache Solr	6%	2%
Cloudera CDH	5%	1%
Oracle BDA	5%	2%
Apache Sqoop	5%	2%
IBM BigInsights	5%	1%
SAP Big Data Services	5%	3%
Cloudera Impala	4%	1%
MapR	4%	2%
Azure HDInsights	4%	3%
Apache Storm	4%	2%
Google Cloud DataProc	3%	3%
Apache Drill	2%	1%
Apache Phoenix	2%	1%
Presto	2%	1%
Pivotal HD	1%	1%
GemFireXD	1%	0%
Other	2%	1%
None	32%	
Not sure	12%	15%

Legend: ■ Current ■ In 2 years

Big data will continue to grow and become more essential to every business. But with frequent release cycles, it's impossible to continuously certify and maintain connectivity for hundreds of big data components and versions. As a result, enterprises need to rely on technology partners for day one support for all these versions.
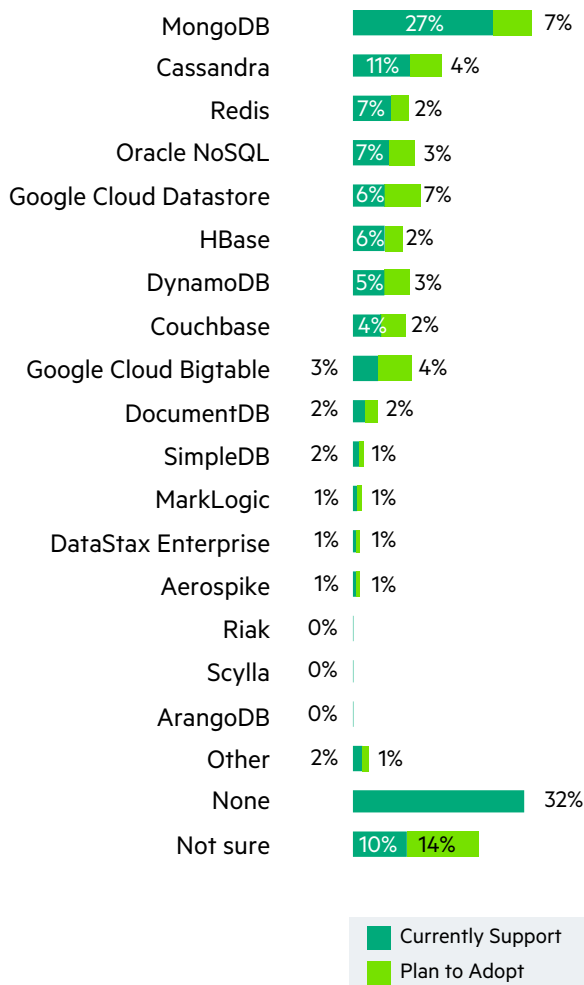
Did you know? ISVs resort to commercial Hive drivers for improved performance.

Progress®

NoSQL Data Sources

NoSQL databases ability to handle large volumes of structured, semi-structured and unstructured data while providing horizontal scalability make them the preferred choice for large or ever-changing data sets in real-time and batch data analytics. They continue to gain traction with nearly 70% of respondents using NoSQL databases, a 11% increase from 2017. Mongo DB remains the favorite, with an adoption rate of 27% among respondents.

The growing popularity of NoSQL is shaking up core business system integrations in organizations of all sizes. NoSQL databases have query languages and frameworks to better support SQL-like operations, but they are still incompatible with existing SQL-based infrastructures. As a result, third party ODBC and JDBC drivers for NoSQL databases are becoming popular.

Which NoSQL database technologies do you or your customers currently use or support?

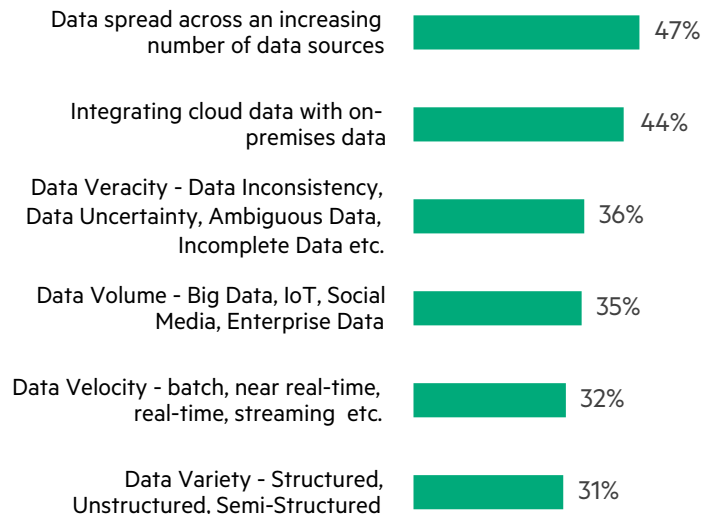| Technology	Currently Support	Plan to Adopt
MongoDB	27%	7%
Cassandra	11%	4%
Redis	7%	2%
Oracle NoSQL	7%	3%
Google Cloud Datastore	6%	7%
HBase	6%	2%
DynamoDB	5%	3%
Couchbase	4%	2%
Google Cloud Bigtable	3%	4%
DocumentDB	2%	2%
SimpleDB	2%	1%
MarkLogic	1%	1%
DataStax Enterprise	1%	1%
Aerospike	1%	1%
Riak	0%	
Scylla	0%	
ArangoDB	0%	
Other	2%	1%
None	32%	
Not sure	10%	14%

Legend: ■ Currently Support ■ Plan to Adopt

While adoption of NoSQL is on a steady rise, end users continue to demand SQL access to NoSQL data for analysis. Thus, organizations are looking for industry accepted SQL based ODBC and JDBC drivers.

Similarly, Apache Spark SQL is another recommended interface for Cassandra/DataStax to run interactive, real-time analytics.

Progress®

Increasing Data Sources is the Primary Data Integration Challenge

The scope of data integration has significantly changed over the past decade. This is due to the increase in sheer number of data sources, hybrid environments, constantly evolving APIs and new, disruptive data types. To provide a unified view of data for ad hoc analysis and business intelligence, most companies are combining data across several disparate sources. Each of these sources has a unique set of APIs, and 47% of survey respondents pointed to integrating all these sources as their most challenging task. 44% of respondents agree that the biggest challenge is incorporating all relevant data across an ever-increasing number of cloud, database with on-premises database. The rapid advancement of social media and IOT contribute greatly to soaring volume of data circulating in the networks, especially with the rise number of connected devices. And 35% are worried about the volume of data they're trying to handle.

What are your other biggest data integration challenges?

Challenge	%
Data spread across an increasing number of data sources	47%
Integrating cloud data with on-premises data	44%
Data Veracity - Data Inconsistency, Data Uncertainty, Ambiguous Data, Incomplete Data etc.	36%
Data Volume - Big Data, IoT, Social Media, Enterprise Data	35%
Data Velocity - batch, near real-time, real-time, streaming etc.	32%
Data Variety - Structured, Unstructured, Semi-Structured	31%

Data Integration can give you a true 360° view of your enterprise data, but has its own set of challenges, such as modern infrastructure (cloud and big data), hybrid connectivity and data variety. A traditional approach of ETL can be too complex and expensive when solving niche connectivity problems across the modern infrastructure. Modern ETL tools provide a holistic approach to this unique, multifaceted challenge.
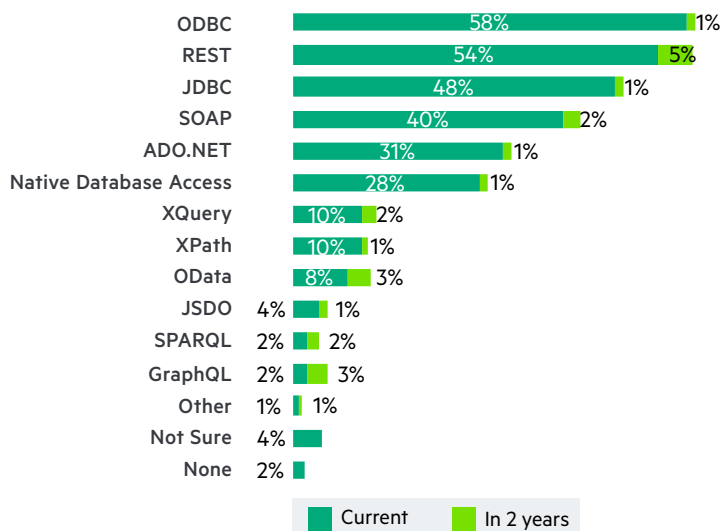
SaaS and microservices have also aggravated the data integration challenges by introducing hundreds of new data silos into the enterprise. Success of the ISVs depends on how well they can integrate with other enterprise apps, but they can't keep up with the exploding APIs. To address this challenge, Progress' upcoming "Project Sagamore"—a smart REST connector—helps you consume enterprise REST APIs with simple codeless configuration. Now, you can build new connectors in minutes. Contact us to learn more.

Progress®

Market Needs – Enterprise-Ready, Standards-Based Drivers

The number and variety of data sources continues to explode every year. Each new data source is unique and has its own API and requirements. As an enterprise/ISV, you could integrate all these apps by developing against the APIs yourself, but that requires expertise in each API and the cost of lifetime maintenance may outweigh the benefits of ease of communication between data. Enterprise-ready standard drivers will allow you to abstract the differences between all those APIs and simplify data integration through a unified interface.

Organizations are increasingly adopting standard SQL and REST data access standards for easier and faster connectivity to disparate data from popular BI, reporting and ETL tools, while ensuring interoperability and compatibility with existing systems. While ODBC remains at the top of the SQL heap and is expected to gain further share, the move from SOAP to REST is significant. REST—specifically OData—is expected to see the most growth in adoption over the next two years.

Which data access interfaces do you or your customers currently use? Which do you or your customers plan to adopt in the next two years?

|	Current	In 2 years
ODBC	58%	1%
REST	54%	5%
JDBC	48%	1%
SOAP	40%	2%
ADO.NET	31%	1%
Native Database Access	28%	1%
XQuery	10%	2%
XPath	10%	1%
OData	8%	3%
JSDO	4%	1%
SPARQL	2%	2%
GraphQL	2%	3%
Other	1%	1%
Not Sure	4%	
None	2%	

While standards are becoming more prominent as the number of data sources continues to grow exponentially, it's very important for the organization to make their data accessible to modern applications through standard APIs to stay ahead of the competition.

The OData REST API is gaining ground in hybrid and cloud environments because of its rich set of interoperable query capabilities to unify data access for disparate resources across the Internet.

Progress®

REST APIs have become increasingly popular for accessing enterprise data from web and mobile applications, with an impressive 65% of respondents opting for REST/web APIs for databases. A common challenge with the REST API approach is needing a different client codebase for every REST API in use, which can quickly become a technical burden. OData addresses this challenge.

Which interface bridges are the most interesting to you?

REST/Web APIs for Databases	65%
SQL for REST/Web APIs	46%
GraphQL for Databases	14%
Other	1%
None	6%
Not Sure	13%
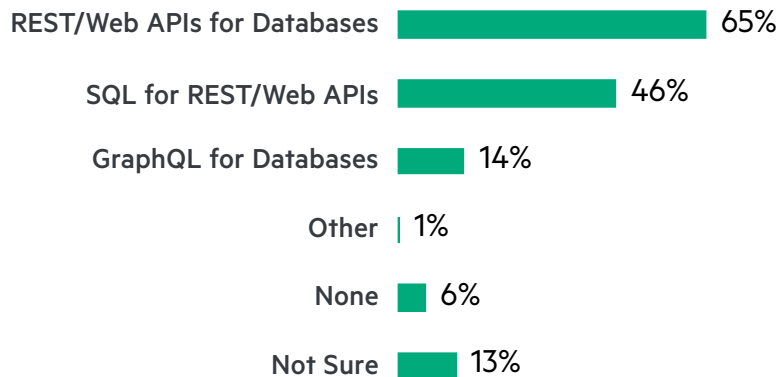
> There is a new query standard called OData that is considered "SQL for the Web"

The trend toward RESTifying databases can be seen across the industry. That's because companies have built business applications on top of these databases and now require modern apps to access this enterprise data. To explore this further, please see the Progress whitepaper OData: The Fastest Way to RESTify Your Databases.
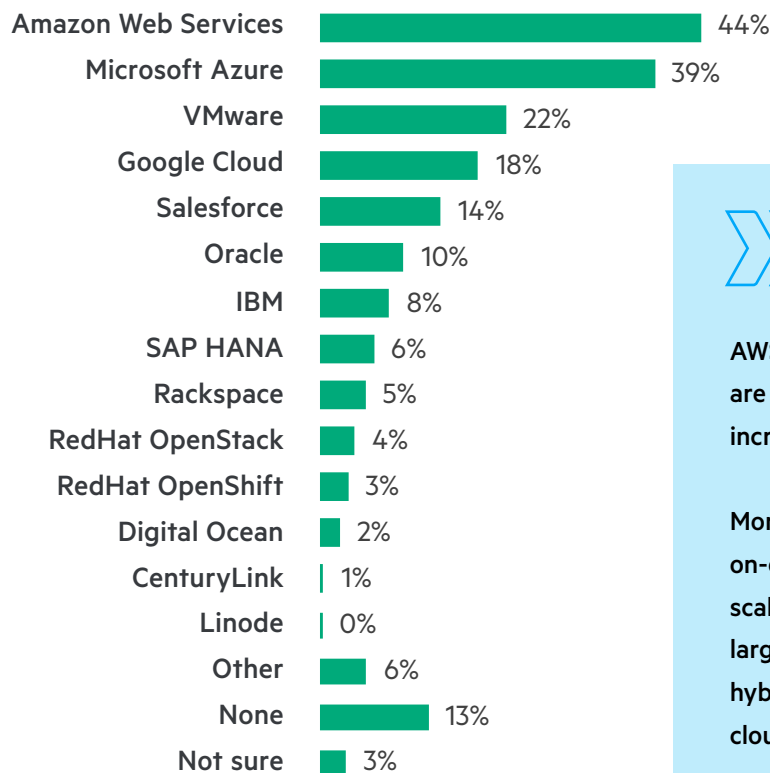
Other popular standard APIs to query data are emerging such as GraphQL and ORDS. Learn more about this industry debate in the blog: OData vs GraphQL vs ORDS.

Progress®

Need for Faster Time to Market is Driving Cloud Adoption

For staying relevant in a fast-paced world, more large enterprises are moving their workloads from traditional on-premises systems into the cloud—at a rate and pace that is expected to be far quicker than in the past.

The adoption of cloud infrastructures continues to escalate, according to 2018 respondents. The only three providers seeing gains in adoption over 2017 were Amazon Web Services (+12% among respondents) Microsoft Azure (+7%) and Google Cloud (+3%). Most of the participants in the survey cited faster time to market as a core reason to adopt cloud-based infrastructure.

Which cloud infrastructure do you or your customers currently use?

Provider	%
Amazon Web Services	44%
Microsoft Azure	39%
VMware	22%
Google Cloud	18%
Salesforce	14%
Oracle	10%
IBM	8%
SAP HANA	6%
Rackspace	5%
RedHat OpenStack	4%
RedHat OpenShift	3%
Digital Ocean	2%
CenturyLink	1%
Linode	0%
Other	6%
None	13%
Not sure	3%

AWS, Microsoft Azure and Google Cloud are the only three providers who have seen increase in share over the previous year.

More and more organizations are adopting on-demand clouds to help their businesses scale and grow. But for organizations with large investments in on-premises systems, hybrid connectivity is a necessary part of cloud adoption.

Progress offers a vendor-agnostic solution, to help you deploy hybrid connectivity on cloud computing platforms such as AWS, Azure, VMware and Google Cloud. See our cloud and hybrid tutorials.

Progress®

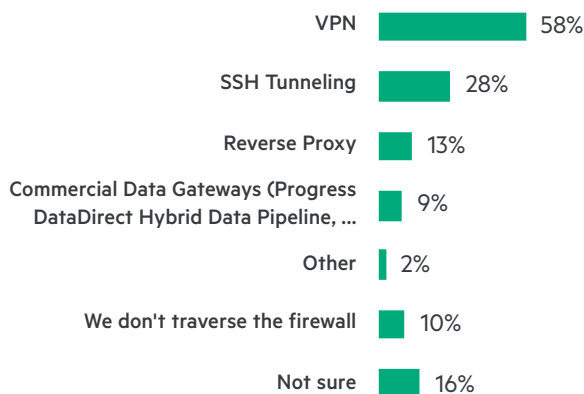# Hybrid Connectivity: A Foundation for Digital Transformation

With the increased adoption of cloud and the continued relevance of on-premises data (both relational and big data), there is greater interest among enterprises to integrate data sources both from within and outside of the organization's firewall. ISVs today seek real-time connectivity and a lightweight solution, which can provide a standard way to connect to on-premises data behind multiple firewalls—without having to deploy/maintain VPNs or SSHs for multiple customers.

Connecting cloud platforms to on-premises systems behind the firewall is a complex issue, with complex security policies, data privacy, and compliance as top challenges. To solve this challenge of connecting cloud and on-premises, respondents are primarily using VPN, SSH tunneling and reverse prox

However, these approaches are difficult to manage and not engineered for accessing business data from behind multiple firewalls. VPN is a secure option for traversing the firewall, but it becomes very complex when you have to scale it. It also involves plenty of manual effort because there are many points of failure to manage. SSH tunneling requires administration, poses security risks to manage SSH keys, and doesn't scale well.

That's why enterprises are turning to hybrid connectivity solution gateways that allow for remote connectivity from cloud services to data sources residing behind a corporate firewall. This approach ensures reliable, scalable and secure data transfer to support a company's digital transformation endeavors.

How do you currently connect your cloud apps to on-premises data that is behind the firewall?

VPN	58%
SSH Tunneling	28%
Reverse Proxy	13%
Commercial Data Gateways (Progress DataDirect Hybrid Data Pipeline, ...)	9%
Other	2%
We don't traverse the firewall	10%
Not sure	16%

Commercial gateways such as Hybrid Data Pipeline are gaining in popularity because they are engineered for the cloud to support standard data access for both on-premises and hybrid scenarios—without changing the firewall configuration.
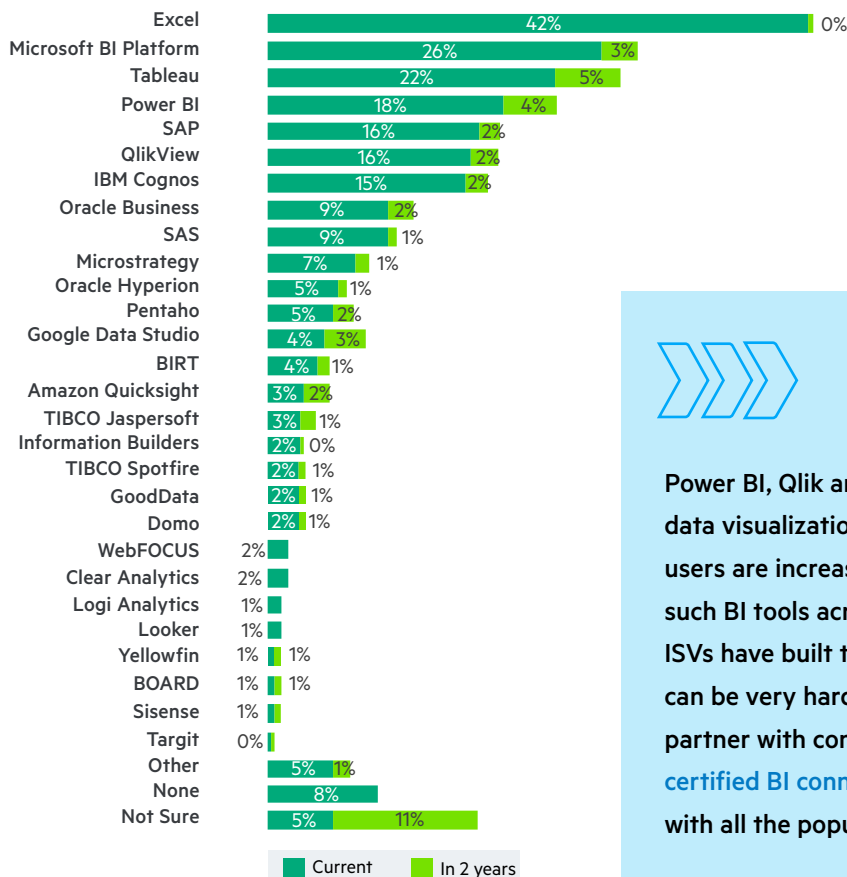
Did you know? A hybrid data gateway can be used to get connected to data and applications behind your customers' firewalls without opening any VPNs or SSH tunnels.

Progress®

Open Analytics is the emerging strategy for Delivering Data Through an Open Data Access Layer in the Cloud

In the next wave of digital transformation, organizations are adopting business analytics to drive decisions and compete in a highly unpredictable marketplace. Organizations are using on average 2.5 different BI reporting tools, underscoring the need for enterprises to have Universal BI connectors to support their diverse BI needs.

Among respondents, Tableau usage has increased from 14% in 2016 to 22% in 2018. Power BI adoption is expected to grow around 4% in the next two years.

What BI and reporting tools do you or your customers currently use? And which do you or your customers plan to adopt in the next two years?

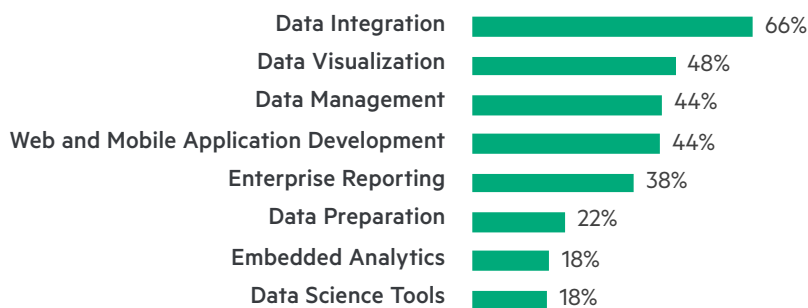| Tool	Current	In 2 years
Excel	42%	0%
Microsoft BI Platform	26%	3%
Tableau	22%	5%
Power BI	18%	4%
SAP	16%	2%
QlikView	16%	2%
IBM Cognos	15%	2%
Oracle Business	9%	2%
SAS	9%	1%
Microstrategy	7%	1%
Oracle Hyperion	5%	1%
Pentaho	5%	2%
Google Data Studio	4%	3%
BIRT	4%	1%
Amazon Quicksight	3%	2%
TIBCO Jaspersoft	3%	1%
Information Builders	2%	0%
TIBCO Spotfire	2%	1%
GoodData	2%	1%
Domo	2%	1%
WebFOCUS	2%	
Clear Analytics	2%	
Logi Analytics	1%	
Looker	1%	
Yellowfin	1%	1%
BOARD	1%	1%
Sisense	1%	
Targit	0%	
Other	5%	1%
None	8%	
Not Sure	5%	11%

Power BI, Qlik and Tableau are among the top data visualization tools in the market. Business users are increasingly demanding support for such BI tools across every enterprise app. Some ISVs have built their own BI connectors, but this can be very hard to maintain and scale. So, ISVs partner with connectivity vendors to deliver certified BI connectors that work seamlessly with all the popular BI tools.

Meanwhile, data visualization has grown in importance, increasing by 11% from 2017's survey. Organizations have seen the impact of visualizing complex data sets to help them make smarter, data-driven decisions to manage growth and performance.

The major forces driving the embedded analytics market (18%) are the rise in data-driven organizations, higher adoption of self-service analytics, and the increasing demand to integrate analytics with business applications. The embedded analytics market is growing rapidly, a 6% increase from 2017, because of the transformation from traditional business data analytical techniques to advanced techniques and the massive surge of structured and unstructured data.

Which types of applications do you connect to from standard SQL/REST interfaces?

Application	Percentage
Data Integration	66%
Data Visualization	48%
Data Management	44%
Web and Mobile Application Development	44%
Enterprise Reporting	38%
Data Preparation	22%
Embedded Analytics	18%
Data Science Tools	18%

However, embedded analytics alone isn't enough to meet the unique enterprise data needs. This exposes the importance of a "bring-your-own-analytics" (BYOA) capability to supplement embedded analytics in business applications. In response, companies are adopting embedded analytics for quick insights and supporting external analytics tools such as Tableau, PowerBI, Qlik or MicroStrategy for advanced analytics.

Building tool-specific BI data connectors to integrate data with analytics tools is neither economical nor scalable for ISVs. That's given rise to Open Analytics where ISVs can open their customer data to a wide range of new data stores, using a standard coding interface (ODBC/JDBC/ADO.NET/OData/REST). Open Analytics has enabled many ISVs to increase their reach to all the popular BI tools in the market.
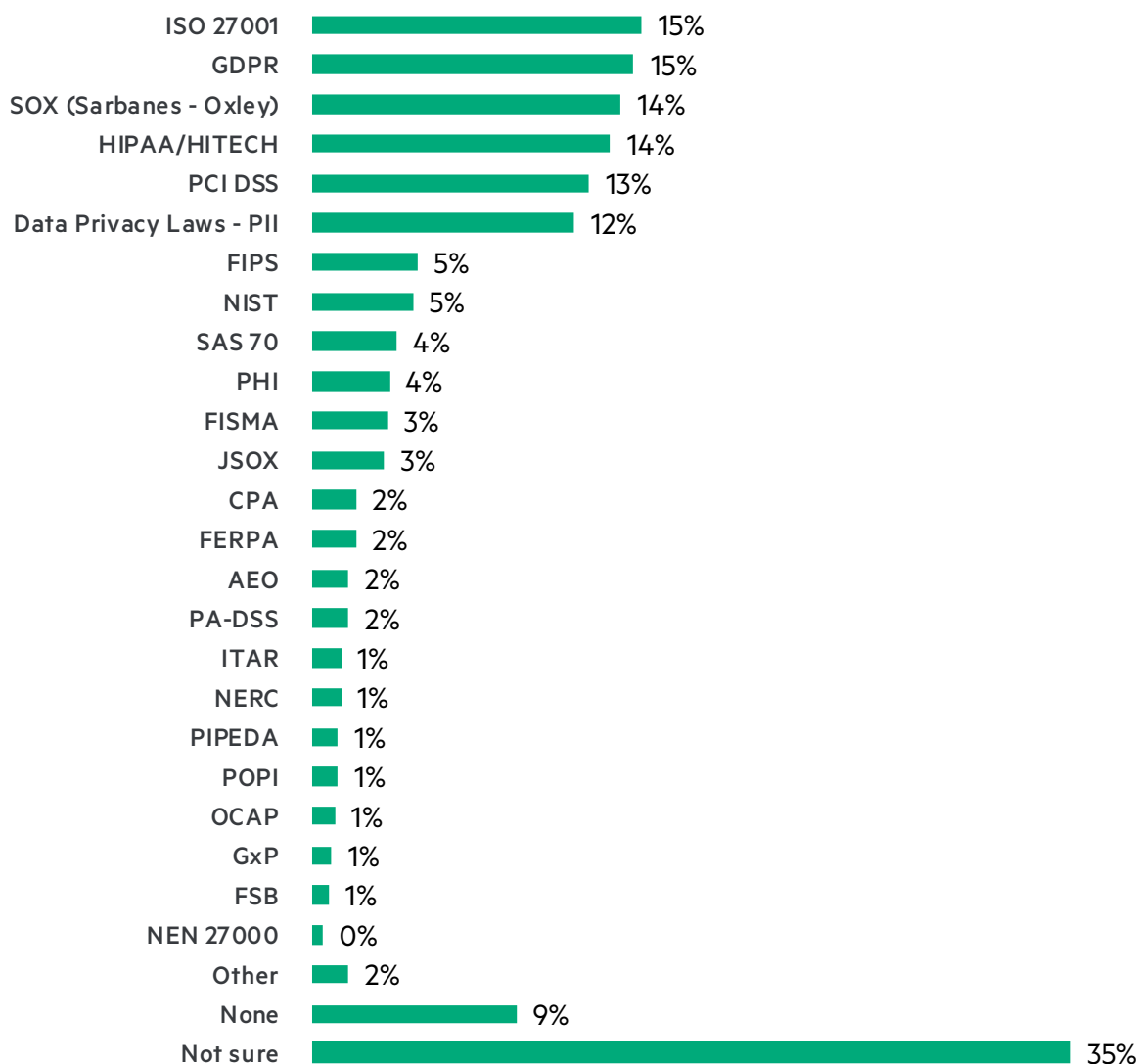
One of the latest trends among ISVs is embedding analytics functionality in their solutions to give users data, dashboards and reporting. But users also want to perform complex analysis on aggregated data using BI tools and programming languages. With Open Analytics, customers can use the tool or language of their choice to query cloud applications. These strategies work together to meet the diverse needs of business users, data analysts and data scientists.

Progress®

Enterprise Data Access Security is Becoming More Complex

As organizations continue to invest in business intelligence, big data, IoT and cloud, IT teams are introducing increasingly complex security mechanisms to authorize and encrypt access to data. Also, organizations have security threats, users misusing their credentials, compliance standards and laws to understand and satisfy. More than 65% of respondents said they must comply with one or more standards.

Industry-specific regulations such as Sarbanes-Oxley, PCI DSS and HIPAA, and government regulations such as GDPR, top the list of data protection/privacy laws. Respondents wrote in many more, included in the Other category, such as PIPEDA in Canada, FERPA in the U.S. and standards required by the E.U. The landscape of privacy and protection laws around the world continues to evolve and grow more complex.
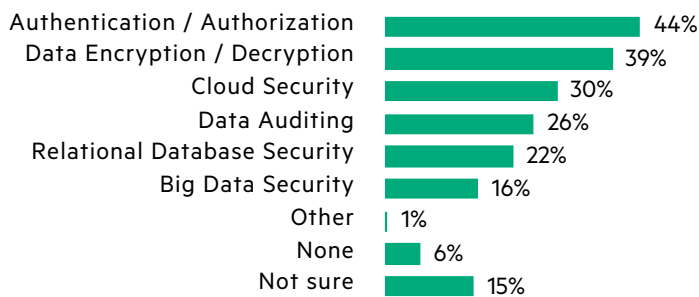
Which security/data standards are you required to comply with?

Standard	Percentage
ISO 27001	15%
GDPR	15%
SOX (Sarbanes - Oxley)	14%
HIPAA/HITECH	14%
PCI DSS	13%
Data Privacy Laws - PII	12%
FIPS	5%
NIST	5%
SAS 70	4%
PHI	4%
FISMA	3%
JSOX	3%
CPA	2%
FERPA	2%
AEO	2%
PA-DSS	2%
ITAR	1%
NERC	1%
PIPEDA	1%
POPI	1%
OCAP	1%
GxP	1%
FSB	1%
NEN 27000	0%
Other	2%
None	9%
Not sure	35%

Progress®

Data security is complicated for a variety of reasons. For applications to access data, there are a substantial number of distributed technologies that are required to securely access data. There's also the issue of ensuring what data actions are appropriate for each client at each stage of data transfer. Thus, respondents are seeing increased security complexity across the board, whether using traditional relational sources, big data or cloud, —from authentication and authorization, to the cloud, data auditing, big data and relational database security.

Authentication and Authorization are bigger problem for ISVs as they need to support diverse security policies of all their customers.

Where are you seeing increasing complexity with security when accessing data?

Authentication / Authorization	44%
Data Encryption / Decryption	39%
Cloud Security	30%
Data Auditing	26%
Relational Database Security	22%
Big Data Security	16%
Other	1%
None	6%
Not sure	15%

Security is a complex topic, and that complexity continues to grow. ISVs in particular need to stay on top of their game because every customer has a different security ask. Learn more about the challenges and best practices for implementing enterprise data security across big data, cloud and relational databases, including network security, firewalls, Amazon AWS IAM, Hadoop Cluster Security, Apache Knox, and more our the blogs:

- Data Security Basics for Cloud and Big Data Landscapes
- Enterprise Data Security Guide for Big Data, Cloud and Relational Databases
- Latest security enhancements in Data Access & Connectivity

Progress®

Conclusion

As enterprises grow, a wide variety of data is produced, consumed and stored at different parts of the organization. The challenge is to effectively manage and leverage the volume, variety and velocity of data. At a time when budgets and resources remain tight, the old approach of trying constantly to increase expenditures on infrastructure and hardware assets can't keep pace. The success of a business lies in its seamless data connectivity and integration technology in an increasingly hybrid cloud/on-premises world.

The 2018 Data Connectivity Report underscores the need for standard data connectivity to all your data sources, whether they reside in the cloud, in legacy systems, or both, and the importance of investing in standards-based drivers for better API integration. This year's survey shows that relational databases are going strong, a trend toward RESTifying databases can be seen across the industry with OData as the new industry standard REST API, and the need for hybrid connectivity is on rise as ISVs move towards cloud applications. Also, the industry is marching towards deploying Open Analytics as a strategy to deliver open interfaces for cloud apps, with multiple analytics strategies becoming available.

For more about how Progress® DataDirect® is transforming how ISVs function in the modern enterprise, and case studies on how a number of companies have overcome data access obstacles to deliver seamless data connectivity to their end customers, please visit Progress DataDirect.

Progress®

About Progress DataDirect

Progress® DataDirect® delivers powerful data connectivity for applications running on-premises or in the cloud. Get connected to the vast data landscape using a single standard interface with SQL or REST. DataDirect connectivity solutions support the full range of relational, cloud, NoSQL and Big Data sources across trusted industry standards, including ODBC, JDBC, ADO.NET and OData (REST).

Learn more at www.progress.com/datadirect-connectors

About Progress

Progress (NASDAQ: PRGS) offers the leading platform for developing and deploying mission-critical business applications. Progress empowers enterprises and ISVs to build and deliver cognitive-first applications, that harness big data to derive business insights and competitive advantage. Progress offers leading technologies for easily building powerful user interfaces across any type of device, a reliable, scalable and secure backend platform to deploy modern applications, leading data connectivity to all sources, and award-winning predictive analytics that brings the power of machine learning to any organization. Over 1,700 independent software vendors, 100,000 enterprise customers, and two million developers rely on Progress to power their applications.

Learn about Progress at www.progress.com or +1-800-477-6473

Progress®