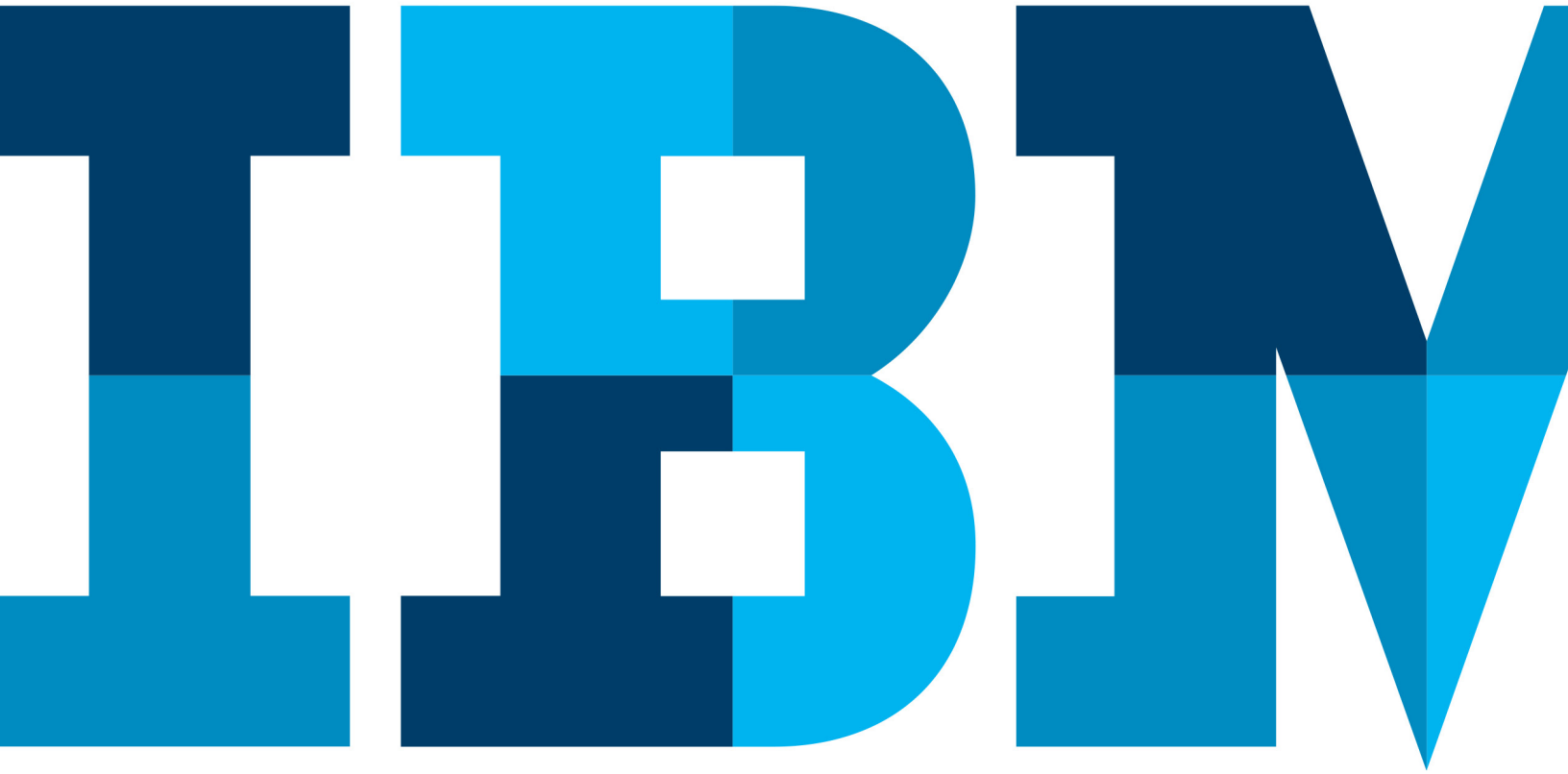
IBM Cloud
White paper

May 2018

Unleash digital intelligence with data and apps

Multicloud integration architecture provides secure gateway for data governance, analytics



IBM

Contents

→ See the interactive version of this paper on the IBM Cloud website.

Introduction

In a world of rapid market change, technological upheaval and massive restructuring, arguably, your most important business asset isn't your products, technologies, or processes — it's your data.

The ability to transform your data into digital intelligence is central to competitive advantage. Only then can it fuel and automate processes, accelerate decisions, bring unique client experiences to life, and ultimately, disrupt entire industries.

The problem is that data often remains trapped in siloed data stores or applications — both on premises, and increasingly, in the cloud. Different vendor APIs, architectures and data formats can make it hard to access and apply the advanced analytics needed to elevate data's value — and make it available to the right people.

What's needed is a more proactive approach that integrates data and applications in all its forms, on and off premises, and provides greater visibility and control. The right solution must also secure and govern all types of data, from any source, across a hybrid cloud infrastructure.

Is your data AI ready?

The key to AI is a strong data foundation. True innovation in this space cannot be achieved until an organization has a strong grasp on all of its incoming data and makes every decision with data as its backbone.

Your data is a valuable corporate asset. As such, it has the power to transform your organization, add monetary value, and enable your workforce to accomplish extraordinary things. If that data is not collected, organized, managed, controlled, enriched, governed, measured and analyzed, it's not just useless — it can become a liability.

In today's multicloud world, data often remains trapped in siloed data stores or applications — both on premises and in the cloud. The first step is to discover where your data lives. Next, ensure authorized users can access all data with confidence. This involves satisfying all matters of finding, integrating and cataloging data to create a fluid "single version of the truth," with built-in data lifecycle management, compliance and protection controls. It also includes understanding possible compliance issues posed by US and international regulations. Finally, data must enable the radical scaling of insights on demand using advanced analytics and data science to make better decisions and power smarter business processes.

→ Read the IBM and BizTechInsights survey findings that reveals the challenges organizations face in integrating data and their priorities for improving data integration.

In its August 10, 2017 "Magic Quadrant for Metadata Management Solutions" report, Gartner noted: "By 2020, 50% of information governance initiatives will be enacted with policies based on metadata alone."

→ Read the full report: "Magic Quadrant for Metadata Management Solutions."

Unified governance and integration

While a multicloud environment allows data to be created, used and shared anywhere – in cloud, on premises and across hybrid IT – it can also exacerbate siloed data challenges.

A trusted analytics foundation overcomes that challenge. By empowering your stakeholders to discover, understand, integrate, analyze, govern and self-serve structured and unstructured data at any scale, in multicloud data environments, that foundation accelerates the journey to AI.

Your Chief Marketing Officer needs current, trusted and complete customer data encompassing a 360-degree view. Your Chief Analytics Officer needs to crunch external and internal data to predict sales performance and revenue, and to identify new solutions and connections. Your Data Scientists require clean data so they can focus on their modeling versus spending time on data hunting, cleanup or integration. Your Chief Compliance Officer needs to protect brand equity and reduce risks.

Data-driven, innovative organizations are doing all of this while digitally transforming themselves. They all start by asking four fundamental questions in their data journey:

- **What data do we have?**
- **Where is that data?**
- **What systems are using that data, for what purpose?**
- **Does all data usage meet all regulatory and business requirements?**

These questions lead them to discover all relevant data; ensure agreement on what the information means and how it will be used; define rules for monitoring data quality levels; build data lineage to increase users' confidence; and most importantly, provide easy data user interfaces.

By doing so, organizations can trust and use their data to achieve their business goals: compete wisely, adapt to market changes faster and disrupt versus being disrupted.

Governed data lakes: Paving the path to insights

With multicloud integration, organizations are now moving to hybrid environments faster than ever before. As they do so, they need to decide where to keep which data, based on business requirements. A data lake architecture helps them to quickly offload structured, semi-structured and unstructured data to a lake for exploratory analytics and to derive meaningful insights from this data.

One of the primary challenges in data lake environments is finding and understanding the data within – where it came from, what it means and who owns it. When not governed, these unknowns turn the lakes into swamps.

With a governed data lake, using an intelligent metadata catalog coupled with industry-specific compliance enablers, organizations can address risk management, regulatory requirements and business performance within a singular framework. They have the flexibility to bring data from any source into a common platform for use and exploratory analytics projects.

Since data lakes store data in a raw, unrefined state, they're easy to build. At the same time, since the data is governed, data consumers can track the lineage of this data and use it with trust and confidence.

Governing a data lake can improve the trust your data consumers have on the data, help them use this data and, in turn, improve the quality of business insights derived from data.

"Cloud-first mandates make on premises application adoption an exception rather than the rule," notes IDC's July 7, 2017, white paper.

→ Read the full details in "The urgent need for hybrid integration."

Multicloud integration

When it comes to digital business success, what's needed is a more proactive approach that integrates data in all its forms, on and off premises, and provides greater visibility and control. The right architecture must also secure and govern all types of data, from any source, across a hybrid cloud infrastructure. It must be open and agile so it can connect new data sources quickly and deliver data securely and rapidly to the apps and people who transform the latest ideas into reality.

IT leaders must drive constant digital reinvention to bring cloud and cognitive together to make initiatives possible that focus on:

Accelerating the move to multicloud. A multicloud environment offers many of the benefits of cloud but it does raise some additional barriers such as what to move to cloud, where to move it and when.

Driving new revenue streams through APIs. Creating digital capabilities for end users and API experiences that set themselves apart in the minds of developers and the organizations they support.

Integrating blockchain into business. Distributed trust systems such as blockchain are redefining industries and go-to-market models.

Integrating AI into customer experiences. This challenge has manifested itself in AI-driven product recommendations, next-best action simulations and automated AI chatbots. IT leaders must look at ways to bridge the data divide and integrate not only all of their current data but also new data sources.

Organizations need to integrate applications across all the different cloud domains, connecting traditional on-premises enterprise applications with private cloud environments and new lightweight public cloud services. Since new services and connections are rarely exclusively governed by central IT, the result is a confusing network of point integrations. What's needed is a way to standardize integrations and data movement quickly and securely across multiple organizations and actors to help maintain control without slowing down innovation.

An agile, multicloud integration architecture is born when you break up your integration runtimes into smaller, more manageable and dedicated components so that new applications can perform the integration they need at the pace and scale of modern innovation.

When integration is broken down into separate pieces, you can opt to distribute those pieces differently from an ownership and administration point of view. There are many potential advantages to this decentralized integration approach:

- A common challenge for separate SOA teams is that they don't fully understand the applications they're exposing. Application teams know the data structures of their own applications better than anyone.

- Fewer teams need to be involved in the end-to-end implementation of a solution, significantly reducing the cross-team chatter, latency and inevitable waterfall development that typically occurs in these cases.

Ultimately, such an architecture should enable users to:

- Manage access to services in and out of your company with APIs

- Connect apps on premises or in the cloud to drive business transformation

- Protect APIs, the data they move, and the systems behind them

- Conduct reliable messaging communications across application boundaries

- Move huge amounts of data rapidly, securely and predictably

- And cleanse and prepare data for a consistent view of your business

→ Learn how to integrate across different cloud domains to bring apps and data together to power your latest initiatives.

Next steps

Together with a secure data foundation, an agile, multicloud integration architecture allows enterprise applications to become living, dynamic tools powered by machine learning models (and analytics) deployed for scalability.

As you think about next steps, consider these key questions:

- Is our data extremely complex?
- Is our data storage too costly?
- Is our data analytics straining our production?
- How will we apply data governance across a multicloud environment?
- How will we assure data quality and security in a multicloud environment?

Check out these resources for more details on IBM Cloud capabilities.

→ Learn how to integrate across different cloud domains to bring apps and data together to power your latest initiatives.

→ Read the IBM and BizTechInsights survey findings that reveal the challenges organizations face in integrating data and their priorities for improving data integration.

IBM Corporation
New Orchard Road
Armonk, NY 10504

Produced in the United States of America
May 2018

Please Recycle