4/15/2017

Unlock Data's Full Potential with Segment: A Cloud Data Integration Approach

Karu Lal

Venkata Koteswara Rao Ballamudi

技术与管理回顾

HTTPS://UPRIGHT.PUB/INDEX.PHP/TMR/

Technology and Management Review 技术与管理回顾

Vol 2, No 1 (2017)

Unlock Data's Full Potential with Segment: A Cloud Data Integration Approach

[1]**Karu Lal,** *Integration Engineer, Ohio National Financial Services, USA*
[2]**Venkata Koteswara Rao Ballamudi,** *Sr. Software Engineer, High Quartile LLC, USA*

Abstract:
For several years, every one of your products, services, and even the internal processes that required the use of technology but were housed in data centers that you owned and managed constituted your whole IT infrastructure. Since virtualization became widespread in 2000, VMWare has filled these data centers with storage devices and servers over time. IT personnel realized they could write code that would allow everyone to use fundamental models of servers and storage devices. Operations and developers could create code to compose apps and the related infrastructure code (IaaS and PaaS). The availability of cloud services that are both cost-effective and safe has facilitated the delivery of software code-based services. This has made it easier for organizations to swiftly change how they work to meet the demands of their customers or to resolve problems with the productivity of their internal workforce. The research has demonstrated the step-by-step procedure for cloud-based data integration, which differs from business to business and is driven by the requirements of each company. This article presents valuable resources for anyone considering implementing a cloud, hybrid cloud, or multi-cloud integration strategy. In addition, we have discussed the most critical aspects of cloud integration, including its many advantages.

Keywords: Cloud Data, Data Cloud Integration, Cloud Integration, Hybrid Data Integration

INTRODUCTION

The process of merging data from various sources into a cloud-based storage system (such as a data lake, relational or non-relational database, data warehouse, etc.) is called cloud data integration. This data may have originated from other cloud-based databases or applications, an on-premises system, or a combination. Alternatively, it may have arrived from a variety of both.

Think about how it will affect the data! What you handled may now be stored on cloud infrastructure and in several other locations owned and administered by third parties. The cloud makes it more difficult to keep data since it can be instantly relocated to a different part of the world in response to changes in demand (Dekkati & Thaduri, 2017).

The process of creating, managing, governing, and maintaining your technology across numerous locations, regardless of whether or not those sites are yours, is called cloud integration. Integration with the cloud is essential if you want to reap the benefits of using a hybrid combination of technologies across all facets of your company. For instance, a platform for sales and marketing that is entirely distinct from one used for finance could be utilized. To accomplish this purpose, the flow of data and control over these services will be streamlined to the greatest extent.

Integration with the cloud has become more popular among companies and organizations of all kinds to convert data into business intelligence. The explanation for this is relatively straightforward: more and more corporate processes are taking place in hybrid cloud environments, or even cloud-to-cloud environments, and if the appropriate tools to manage data in the cloud are not utilized, then the data might become compartmentalized, ignored, or even completely lost (Desamsetti, 2016).

The process of linking data stored on local servers' located on-premises with data stored in remote software as a service (SaaS) applications and cloud services is referred to as cloud integration. It acts as a conduit to highly effective data analytics platforms, customer relationship management (CRM) systems like Salesforce, and other applications hosted by third-party providers. These applications include data warehouses like Google BigQuery, Snowflake, Amazon Web Services, and Microsoft Azure.

HYBRID CLOUD SOLUTION TO TECHNOLOGY REQUIREMENTS

One of the reasons may be that very few companies have the necessary technological expertise and healthy financial standing to construct and manage their data services. SaaS and cloud vendors can absorb and adapt the newest technology securely, ensuring that their clients can attract and keep consumers while also simplifying their work processes (Gutlapalli, 2016a).

It is essential to remember that cloud integration combines the computational capacity you require with the ability to deliver data to any application or client when desired. We want to be off the top page of the newspaper because your service is reliable and fast. We also want information to be available to decision-makers promptly and updated when you have to make an important choice.

The benefits that have resulted from switching to a hybrid data integration cloud model are as follows:

- Establishing a single point of reference for all of your data
- A decrease in the price of technology (mainly in terms of computing power and data storage)
- Elimination of data that is either redundant or repeated
- Increasing adherence to regulations and improving governance
- The computerization of workflows through the elimination of data entry performed manually across teams
- Cost reductions in operating and capital expenditures

- An improvement in operational efficacy and efficiency results from elasticity, which is the capacity to scale infrastructure, particularly computation, in response to shifting demands posed by a business.
- Increased flexibility and agility to satisfy the needs of customers better while reducing the amount of time needed to bring products to market
- Improvements in the management of service incidents and the continuity of company operations
- Makes it easier to implement the ideas of agile development, DevOps, and enterprise service and data management throughout your organization
- Increased levels of satisfaction among both employees and clients

ADVANTAGES OF CLOUD INTEGRATION PLATFORMS

The traditional method of developing a firm relied on the specialized subject expertise of professionals working in departments such as warehousing, sales, marketing, and finance, amongst others (Gutlapalli, 2016b). This strategy took full advantage of the depth of knowledge available, but in the end, crucial information was kept isolated in data silos. Because the data generated by each team was kept separate within their departments, there were few opportunities for couples to share data or insights. As a direct consequence of this, the entire company experienced losses.

However, with the emergence of business intelligence, companies have acquired a taste for data mining, discovery, and sharing across departments to improve overall performance. Cloud integration enables business intelligence projects and contributes to optimizing a wide variety of essential business processes (Gutlapalli, 2016b). This is accomplished by configuring application programming interfaces (APIs) and connectors, freeing data flow between silos. An organization can enrich the local data held in on-premises systems by adding significant information pulled from cloud apps when they use cloud integration platforms. This comprises a variety of information that may or may not be provided to each department, including (but not limited to) the following particulars:

- The patterns of network traffic
- Behavior of the User
- Incidents of security, both internal and external to your environment
- Information on compliance requirements
- Errors and abnormalities in performance data, which can have an impact
- Application of resources

Thanks to cloud integration, teams from throughout the enterprise may obtain a comprehensive perspective of all the crucial — and frequently extremely complicated — interactions within their business environment. They now have access to the much-needed data, and they can use the insights they receive from going through this process to develop and keep a competitive edge in their industry.

BUSINESSES AND THE CLOUD

There is not a solitary argument that can be made in favor of adopting a cloud integration platform. Today's leading industries use cloud integration to accomplish breakthroughs in various domains, going beyond the primary objectives of enhancing business operations and maintaining compliance with increasingly complicated compliance rules. These advantages cover a broad spectrum and encompass a variety of benefits, including the following:

- Decreased costs — By delving deeply into data from individual sources and applications, you get a better view of areas in which resources may be over- (or under-) committed, enabling you to design a cloud integration strategy tailored to your particular company's requirements.
- Quicker delivery times – The pace of modern business is faster than ever. Your applications will be able to communicate in real-time with one another when integration processes and other essential procedures are moved to the cloud. The distribution of information via digital means enables you to make significant decisions more swiftly.
- Human resources have been optimized. Even a crack team of information technology professionals could not possibly ingest and comprehend the amount of data pouring across a whole firm if they relied on human power. Your current resources can have a far more significant effect on the larger picture if you have the correct integration tools and design, proper task automation, and comprehensive reporting options.
- The autoscaling feature allows you to cut costs and enhance scalability by automatically increasing available resources at peak times and decreasing them during reduced-demand periods. Automation of resource management processes is becoming increasingly vital as many businesses migrate their operations away from on-premises applications and toward virtual platforms hosted in the cloud.

This final one is essential: leaders in the sector are aggressively adopting cutting-edge methodologies such as continuous delivery and DevOps. Both schools of thought advocate for delivery cycles that are as close to real-time as possible, impossible to achieve without substantial automation (Mandapuram, 2016). Inevitably, this will include a convoluted network of software-as-a-service (SaaS) programs and cloud computing solutions. Finding solutions that are fast requires rethinking the idea of what modern data integration even is.

CHALLENGES OF CLOUD INTEGRATION

Turning on seamless cloud integration is more complex than switching to enable all its perks and advantages. Even organizations with a lot of experience and many people working for them can need help developing a model that works for them. A combination of storage space and computing power easily accessible through the cloud is required to create a procedure to access data and ensure that your apps do not fail (Thaduri et al., 2016). However, with such versatility comes several obstacles, including the following:

- Ensuring that data stored in warehouses or lakes are protected from unauthorized access and can only be accessed by authorized individuals or services
- Reducing the amount of bespoke data formats you use, as these may, in the future, hinder your flexibility.
- Providing access to non-technical staff members to use templates (data) that assist them in carrying out their day-to-day responsibilities
- Capability to demonstrate to clients, employees, and authorities that information is appropriately governed
- Planning the flow of data across different apps while ensuring that no mistakes are made in translation
- Adjusting the amount of data stored and the computer power available

CLOUD COMPUTING INTEGRATES SECURITY

Other than people, data matters most. Data, like humans, must be protected, and there are many regulations around data governance before you examine all the software options. A cloud provider or SaaS app lets someone else into your technology environment. Consider leaving a door to your house open so your neighbor can come in. Feel secure? Home safety and cloud data safety are still your responsibility.

- All-access to the database needs to be tracked and recorded.
- The Open Web Application Security Project, also known as OWASP, is a service that does not seek to make a profit from assisting businesses in maintaining the safety and security of cloud-based and internet-based applications.
- In situations where it is feasible, ensure client and data security between your services and the SaaS provider is maintained using two-factor or multi-factor authentication.
- What kinds of safeguards are in place to prevent intrusion by hackers?
- What precautions have been taken to ensure the SaaS provider would not mine your data and sell it to a rival company?• You need to ensure no open doors or data leaks by doing penetration tests.
- Although cloud redundancy is easy to set up, it can be challenging to manage. Conduct regular checks to ensure that your data and services are not being hosted mistakenly from a geographical location, jeopardizing you legally or in some other way.
- Decrease the number of high-level authorization roles or eliminate them using API access or microservice credentials.

4 WAYS TO INTEGRATE DATA IN THE CLOUD

Although the specifics of each data integration strategy will vary, a few best practices can be followed when integrating data into a cloud repository.

A central location for cloud integration: Software-as-a-Service (SaaS) applications, cloud ecosystems, and on-premises apps can all be connected and share data through a cloud

integration hub. It eliminates unnecessary and expensive cloud synchronizations while providing higher agility and efficiency than traditional point-to-point data integration techniques. An integration hub can decouple applications at both the source and the destination level, orchestrate complex data processing, and enable self-service publication and consumption of data.

Data integration without a server: The operational expenses of elastic compute clusters are reduced, and deployment is made more accessible. When data integration jobs that need to be processed are sent to a collection, the collection will scale up or down depending on the workload, and it will stop processing when all of the jobs have been finished. This removes the need for server management, makes it possible to charge customers depending on their use, and makes monitoring integration jobs easier.

The intake of data: You can collect and manage an increasing number of data sources, formats, and protocols by utilizing cloud-based centralized mass ingestion in conjunction with more conventional batch-oriented data-collecting techniques. It can filter and manage data drift resulting from high-performance streaming and edge data processing, in addition to supporting multi-latency data management. The ingestion of data can take place from a variety of sources, including databases, streams, and files.

Partner integration for business-to-business (B2B) transactions: Using a B2B gateway that is hosted in the cloud enables you to quickly set up business partners, configure communication protocols, monitor and manage EDI and other standard message exchange, and process trade partner messages in your backend systems. Accelerating onboarding for customers and partners, mitigating the burden of managing non-standard data originating in systems outside of your control, cutting operational expenses, and reducing the need to dedicate developers to B2B data integration projects are all benefits you may reap from utilizing this technology.

UTILIZE SEGMENT TO MAXIMIZE DATA

The customer data platform Segment provides contributes to integrating cloud data in several essential ways. Businesses can easily integrate new tools and technology into their tech stack in minutes when they use Connections (Thaduri et al., 2016). This saves those hours that would have been spent manually building the integration. These pre-built connections include many cloud-based data warehouses, enabling enterprises to set up one of these storage systems as a destination for their data quickly and seamlessly.

Second, Protocols make it possible for companies to validate data on a massive scale and implement a tracking plan to protect the integrity of their data. When quality assurance tests are automated, firms can proactively block the integration of inaccurate data and prevent it from being utilized in decision-making. In addition, you can restrict access to data based on roles. Personally identifiable information can be blocked automatically, all of which contribute to increased security and compliance. Other key characteristics include the following:

- Replays: This feature enables companies to transfer a subset of their existing data to a new tool or tools to test them and confirm that the result is accurate.
- The Debugger allows you to verify, in real-time, that API calls (made from servers, mobile devices, websites, and so on) have been received by your Segment Source in the required format.
- Regional Segment in the EU: Infrastructure hosted in the EU for data input, processing, storage, and audience generation to comply with the requirements of Schrems II.

CONCLUSION

At this point, there is no more extended room for debate: data management on the cloud is unquestionably the way of the future. Regardless of whether or not they are aware of the fact, even businesses with a sizable legacy on-premises infrastructure are inevitably dealing with such a large number of software-as-a-service (SaaS) apps that they spend most of their time in a cloud environment. A data integration platform can help you achieve a smooth, quick, and effective transfer to the cloud if you are prepared to explore the possibilities cloud computing presents. Without the laborious process of hand coding, Several Data Integration Platform gives you access to a comprehensive and straightforward set of tools you can utilize to get your data where it needs to be.

REFERENCES

Dekkati, S., & Thaduri, U. R. (2017). Innovative Method for the Prediction of Software Defects Based on Class Imbalance Datasets. *Technology & Management Review*, *2*, 1–5. https://upright.pub/index.php/tmr/article/view/78

Desamsetti, H. (2016). Issues with the Cloud Computing Technology. *International Research Journal of Engineering and Technology (IRJET)*, *3*(5), 321-323.

Gutlapalli, S. S. (2016a). An Examination of Nanotechnology's Role as an Integral Part of Electronics. *ABC Research Alert*, *4*(3), 21–27. https://doi.org/10.18034/ra.v4i3.651

Gutlapalli, S. S. (2016b). Commercial Applications of Blockchain and Distributed Ledger Technology. *Engineering International*, *4*(2), 89–94. https://doi.org/10.18034/ei.v4i2.653

Mandapuram, M. (2016). Applications of Blockchain and Distributed Ledger Technology (DLT) in Commercial Settings. *Asian Accounting and Auditing Advancement*, *7*(1), 50–57. Retrieved from https://4ajournal.com/article/view/76

Thaduri, U. R., Ballamudi, V. K. R., Dekkati, S., & Mandapuram, M. (2016). Making the Cloud Adoption Decisions: Gaining Advantages from Taking an Integrated Approach. *International Journal of Reciprocal Symmetry and Theoretical Physics*, *3*, 11–16. https://upright.pub/index.php/ijrstp/article/view/77