There's No Gas Cap on a Tesla

Requirements for Cloud Data Warehousing

An ENTERPRISE MANAGEMENT ASSOCIATES® (EMA™) White Paper

Prepared for Snowflake
August 2017

Table of Contents

Setting the Stage for Cloud Data Warehousing

The world is changing at a faster pace than ever before, with constant technological advancements. This is true whether at work or at play. Driverless and electric cars, such as Tesla, provide great examples of how interaction and expectations have changed regarding technology use. Tesla revolutionized the driving experience with technology that changed the way cars are designed and driven. On a smaller note, the same can be said for Apple and the creation of the iPhone and iPod in terms of how people consume music, or the transition from mobile phones to smartphones, where actual phone use is simply one of hundreds of applications on a singular device. All of these shifts in technology and the way people interact with the world lead to more efficiency, quicker access, and increasing demand for self-service. Both IT and business users demand this level of ease in their interactions with technology. Within IT, this leads to agile development cycles and self-service developer interfaces to ensure quicker time to delivery. Overall, organizations expect to have the Tesla and iPod experience with their information systems on a daily basis.

This is also true for the way in which organizations adapt their business intelligence and data warehousing platforms. As organizations expand their data warehousing infrastructure and gain more value out of their data, many companies are researching how to adopt technology that supports their increasingly complex business and technical requirements. With cloud adoption increasing due to overall expanding capabilities and platform expansions, the analytics market is starting to shift. More organizations are taking advantage of cloud platforms to address scalability, big data, and advanced analytics needs. In many ways, cloud shifts the focus from complicated development to easier implementations and from internal maintenance to a services model. Organizations want their data warehousing infrastructure to fit their business requirements and not have to develop a whole new skill set in order to get a solution up and running. Basically, companies have the same expectations for technology adoption that we as consumers have in our everyday lives, but they need to make sure they adopt cloud platforms that meet their complex requirements.

At the same time, the value of cloud and increasing adoption of analytics and data warehousing in the cloud support a change in the overall expectations from companies evaluating how they manage their data and the requirements that go along with it. For instance, companies expect to support agile development and require support for flexible schemas and types of data being stored. However, not all cloud data warehousing options are created equal, and it becomes important for organizations to understand the capabilities they should be evaluating to make sense of the market and consider the best-fit solutions for strategic and analytical needs.

Evaluating Cloud Data Warehouse Options to Take Advantage of Increasing Data Complexities

Data diversity is a reality for most organizations, with data existing in both internal and external locations and becoming increasingly more complex. Organizations now require a way to integrate semi-structured and unstructured data into their existing information ecosystem to gain the proper insights required to remain competitive. This creates a challenge for businesses with data warehousing environments that fit the traditional, relational models and do not support the flexibility required to integrate the diverse and complex data. With big data projects, IoT, sentiment analysis, machine learning, etc., the effort required to standardize data within a traditional data warehouse is astronomical.

New data complexities give organizations the challenge to achieve broader data management and better contextual integration. For instance, visibility into sales requires more than just collecting sales data.

To get a better understanding of customer behavior, buying patterns, and the competitive landscape, organizations need to collect both external and internal data to the organization. Social media data, historical customer buying habits, and data from sensors can give organizations more insight into sales performance and supply chain needs. Third-party and IoT data support greater sales performance and insight into sales targets and strategic goals.

Businesses should make sure the platform they select supports the level of data complexity they require. Some organizations seek to leverage cloud data warehousing to expand their information access and add new sources that are not supported in their traditional environments. However, not all cloud data warehousing and analytics providers can offer sufficient integration with new data sources such as big data or IoT in a timely fashion. Other organizations may find their cloud-based providers can support the level of complexity for a short term, but they are pressed to continue to scale over time. Organizations need to understand their choices, how their cloud data warehousing providers currently support their data requirements, and how they will support future data source additions and overall scalability.

The Ability to Scale and the Allocation of Cloud Resources

With the changing technological and business landscape, the ability to scale becomes important. Traditional data warehousing infrastructures no longer meet those needs. Organizations need to understand the range of cloud data warehouse options before they make a decision. Some cloud-based options either cannot scale or are costly to implement. Companies need to make sure they select the most cost-effective options that will meet their objectives. Each platform offering provides different options and some are more limiting than others, meaning companies may think they are getting the flexibility they want but will actually be limiting themselves in the future.

Organizations require the flexibility to change scale according to their data and analytics needs as changes occur, without the underlying technology limiting those changes. This is what cloud-based data warehousing provides with a flexible management environment. This overall scalability includes the following:

- Need for expanding storage
- Ability to address changes in diverse analytics requirements
- Support growing data integration complexities
- Support concurrency of users and processing execution

For most organizations, these requirements are in constant flux depending on what the business requires. Within many cloud offerings, organizations can scale up and scale down as needed. They are basically able to allocate and de-allocate resources so they only pay for what they need when they need it, but are still provided with consistent service irrespective of the scale of their demands. Organizations should ensure that their cloud DW providers can support their business agility related to their initial implementation requirements and scalability needs over time.

Scalability cannot be overlooked as data requirements shift over time. As market needs evolve, organizations should make sure their cloud DW platforms support their scalability requirements. Some organizations have industry-specific considerations that platform providers may or may not meet. In addition, the ability to add new data sources, support an acquisition or new data types, and ensure real-time data access are requirements that all cloud providers may not support. However, the reality is that these are becoming more common capabilities and organizations should understand their current and future requirements and select cloud DW providers that support overall scalability needs.

Leveraging a Services-Based Approach to Data Management With SaaS

One of the benefits of cloud data warehouse adoption is the ability to take advantage of a services-based approach to strategic analytics solutions. In general, this approach encompasses organizations leveraging software providers that host solutions and platforms in the cloud. This means the business subscribes or stores and pays for access based on use, while relying on a third party for storage and support. When cloud first became popular, its appeal remained largely with organizations lacking internal IT infrastructure, a data center, and/or required skillsets. In some cases, organizations chose to take advantage of cloud environments because of their limited ability to manage day-to-day systems or provide internal DBA administration. Now, this services-based approach is gaining momentum and appeals to a broader spectrum of organizations. At the same time, some cloud offerings (specifically Infrastructure as a Service) provide the platform and storage capabilities, with overall maintenance taking place within the organization.

When evaluating platforms, organizations should consider solutions based on the levels of service they require. Some companies have internal skill sets required to manage their IT infrastructures and they adopt cloud data warehousing platforms as a way to expand their analytics strategies. This means the amount of services related to their cloud platform adoption will be less than organizations leveraging cloud-based DW platforms as a way to enable analytics within the organization due to a lack of internal expertise. These businesses need to adopt a managed services approach and make sure the cloud solution provider they select can provide this level of continued maintenance and support.

The spectrum of a services-based approach allows organizations to select the level of services based on skill sets, budget, current infrastructure, and scalability requirements over time. How an organization wants its IT infrastructure managed across the organization will also affect the level of services adopted. Depending on the level of services provided, organizations can take their technical resources and make sure they are more strategic in nature by focusing on enhancing performance, providing overall value, and creating a more structured approach to data management. Companies newer to analytics will want to take advantage of managed services, while organizations with more experience will base the amount of services on their overall analytics goals and direction. However, they will want less interference when managing their day-to-day technology requirements.

Enabling Easier Data Access

With cloud-based data warehousing, organizations are no longer limited to traditional development models that can lead to data mart redundancies and information silos. For effective strategic planning and analytical insight, organizations need a more holistic approach to data access. Leveraging cloud-based DW platforms provides companies with a better way to store and manage their data. A good example includes Customer 360 initiatives and the ability to leverage data more effectively to target customers better. The concept of Customer 360 is limited unless those who deal with customers actually understand all the components of customer behaviour to provide better service, keep retention levels high, and enhance competitive positioning. In essence, data without context limits business value. This applies to relationships with suppliers and partners, developing a strategic approach to sales and marketing, and virtually any business goal, since silos limit the contextual visibility into data. Many traditional data warehouse design models require the development of data marts, limiting contextual insights to a specific business view.

Organizations need to ensure their cloud-based data warehouse platform providers support this level of holistic data access and do not limit the types of data implemented within their analytical environment. Organizations should take into account that cloud-based data warehouse providers may handle and integrate data in different ways. These differences can affect integration times and overall delivery. The number of APIs, how or if data is cleansed, how it is stored, and how organizations leverage the data for analytics will differ based on the solution provider. These differences can affect delivery and data consumption.

Overall, companies should develop a strategy that involves:

- A tightly-coupled approach to data access and design that coincides with their business decisions
- A cloud-based data warehouse implementation to leverage elastic infrastructures that support growth and changing requirements.
- An understanding of how the selected cloud implementation providers handle changes to data domains and formats to eliminate technical or implementation silos.

All of these attributes create a more efficient data management approach that meets the needs of organizations.

Taking Advantage of Flexible Pricing Models

Whether cloud data warehousing uses pay as you go, general subscriptions, or other pricing options, it requires no upfront capital investment so initial cost is the subscription fee. Alternatively, organizations pay what they use, making initial startup costs zero. This includes identifying how usage is defined and whether storage and compute are metered together, and the importance of elasticity contributing to aligning pricing models with benefits associated with overall costs. Additionally, organizations can leverage cloud platforms to avoid overprovisioning while still ensuring performance and addressing the increasing unpredictability of demand.

At the same time, organizations need to understand that cloud providers can offer different payment models and some will be more beneficial to strategic goals than others. For instance, some payment models are based on data access or use, or on the amount of data stored. Companies need to understand that some of these models might seem beneficial up front but depending on scalability requirements, future costs may be very different than initial OPEX outputs. It becomes important to understand these differences to make sure the payment methods match organizational goals. The more mature the cloud market, the more diverse the pricing models available. For an organization with continued data access and query requirements, subscription fees might be the most practical. This may include scheduled reports and consistent use. Agile companies with self-service and data discovery needs leveraging big data stores might require a payment strategy that supports pay per use. How often and how much data is being queried will affect the payment choice. Some businesses only query at specific times of the month. These use cases might warrant a pay-as-you-go approach for computing/processing usage, whereby the company only pays for what is used as opposed to what is stored.

EMA Perspective

Corporate expectations are different than they were ten or even five years ago. The promise of business intelligence and data warehousing has always been to provide business value through data consolidation and analysis. The reality of the market, however, is that traditional implementations of BI, analytics, and data warehousing platforms fall short.

This is where cloud-based environments come into play. With more flexibility, centralized data access, and broader scalability, organizations can adopt a more agile approach to their data-driven analytics implementations through cloud-based adoption strategies. Because of this level of flexibility, organizations can leverage technology in a way that best meets their needs and not be held to the limits of technology. Mid-sized organizations[1] such as small- to medium-sized businesses without traditional, on-premises, internal IT infrastructure are better able to leverage analytic environments with cloud-based resources. Without having to construct and manage an on-premises environment and build up a warehouse on their own, mid-sized organizations can empower their business models and employees with the same advantages as larger competitors. Large enterprises can establish a flexible infrastructure that supports a global footprint without the supporting level of effort. In either scenario, the benefits described throughout this paper identify that most, if not all organizations, can benefit from cloud-based infrastructures for their analytical environments.

At the same time, all cloud-based environments are not created equal. Organizations need to evaluate their cloud choices by understanding scalability and the ability to have a single access point to data. For most organizations, the question no longer resolves around whether to remain on-premises or move to a cloud-based infrastructure. Hybrid data ecosystems are a reality and organizations are continuing their shift to cloud storage and computing for their analytical environments. The question now becomes, which cloud offering best meets short-term technical, business, and strategic needs as well as long-term?

---

[1] EMA defines a mid-sized organization as one with less than 500 headcount. These include the class of companies often referred to as small and medium sized (SMB) businesses.

About Enterprise Management Associates, Inc.

Founded in 1996, Enterprise Management Associates (EMA) is a leading industry analyst firm that provides deep insight across the full spectrum of IT and data management technologies. EMA analysts leverage a unique combination of practical experience, insight into industry best practices, and in-depth knowledge of current and planned vendor solutions to help EMA's clients achieve their goals. Learn more about EMA research, analysis, and consulting services for enterprise line of business users, IT professionals, and IT vendors at www.enterprisemanagement.com or blogs.enterprisemanagement.com. You can also follow EMA on Twitter, Facebook, or LinkedIn.

Corporate Headquarters:
1995 North 57th Court, Suite 120
Boulder, CO 80301
Phone: +1 303.543.9500
Fax: +1 303.543.7687
www.enterprisemanagement.com

3606.080117