ESG WHITE PAPER

Modernize the Data Stack to Transform the Data Experience

Next-generation Business Intelligence and Analytics with Looker

By Mike Leone, ESG Senior Analyst; and Josh Clark, Research Associate

September 2020

Contents

Introduction

As organizations look to revolutionize how they analyze and utilize data, modernizing the data-centric technology stack is critical to success. Today, the traditional stack poses several challenges—too many steps, too many tools, and too many integrations—all leading to operational complexity, time delays, and high cost. Simplifying the data pipeline, data lifecycle, and data stack offers organizations improved efficiency as well as cost savings, and provides them more value by freeing up resources to focus on deriving insight through the analysis of data. As organizations begin to transform their approach to analytics, modernizing the analytics stack is a top priority. To address both the data supply and data demand, data teams must look for ways to simplify and optimize the data pipeline. That means transitioning traditional solutions, like data warehouses and business intelligence tools, to modern architectures built with scalability, agility, and availability in mind. And what follows with a modernized stack is an ideal data experience rich in actionable insight, API-driven application integration, and the ability to address the real-time needs of a dynamic business.

The Traditional Data-centric Stack

A traditional approach to analytics consists of several tools and phases that must work in lockstep to deliver timely value. It starts with identifying, collecting, and capturing source data, whether from a traditional application, an event stream, a file share, or a collection of sensors. This data varies in type, size, rate of change, scale, and level of accessibility. With organizations looking to leverage more data to conduct more precise and accurate analysis, pulling data from many sources means unpredictability and complexity. And data is far from static and predictable. Existing data sources change, while new data sources are requested and added regularly, creating a never-rest, reactionary environment for data teams.

As data is collected, it must be ingested and processed. This is where ETLs/ELTs come into play, leveraging APIs to stream data from source to destination (e.g., a data lake), while working to ensure data quality and governance controls are enforced. The new integrated data is staged before entering its ultimate destination, typically a data warehouse, or data lake. Then the organization relies on data teams to extract and aggregate the data depending on the ultimate destination. Maybe that's a BI platform, a data visualization tool, or simply a spreadsheet. This is highly dependent on the line of business and the particular end-user requesting the data. And while a functioning data pipeline and analytics workflow may appear efficient and optimal, it can also be brittle, with several points of failure between the number of integrated components delivered by multiple vendors and differing interdependencies between components and services. One wrong update or change of one component can compromise the entire data flow. Complex data pipelines and other shortcomings in traditional approaches to data are a major reason why organizations are in the early phases of transforming their data stack with a focus on modernizing the key components: the data warehouse and the analytics platform.

A Modern Approach to Data Requires a New Way of Thinking

Traditional BI is heavily rooted in disparate data, cumbersome workflows, delayed responses from data/IT teams, and untimely insight. Put it together, and business analysts are often left waiting for an updated report or dashboard that doesn't accurately convey the latest insight. How can the business expect to quickly and confidently act on insight if the workflows are slow, the points of failure are everywhere, and the insights only tell part of the story? A new way of thinking is required to meet the analytical needs of a modern business.

The workforce is clamoring to use data in an open and transparent way. To get there, organizations must think more broadly and strategically about how to empower their employees to work with data. Making this more challenging is the emergence of new data: log and event data for modern applications, sensor data for edge devices, SaaS application data across different lines of business, the list can go on. The speed and scale of all this data is forcing organizations to take an approach that prioritizes real-time action as opposed to historical understanding. Data should be collected and analyzed
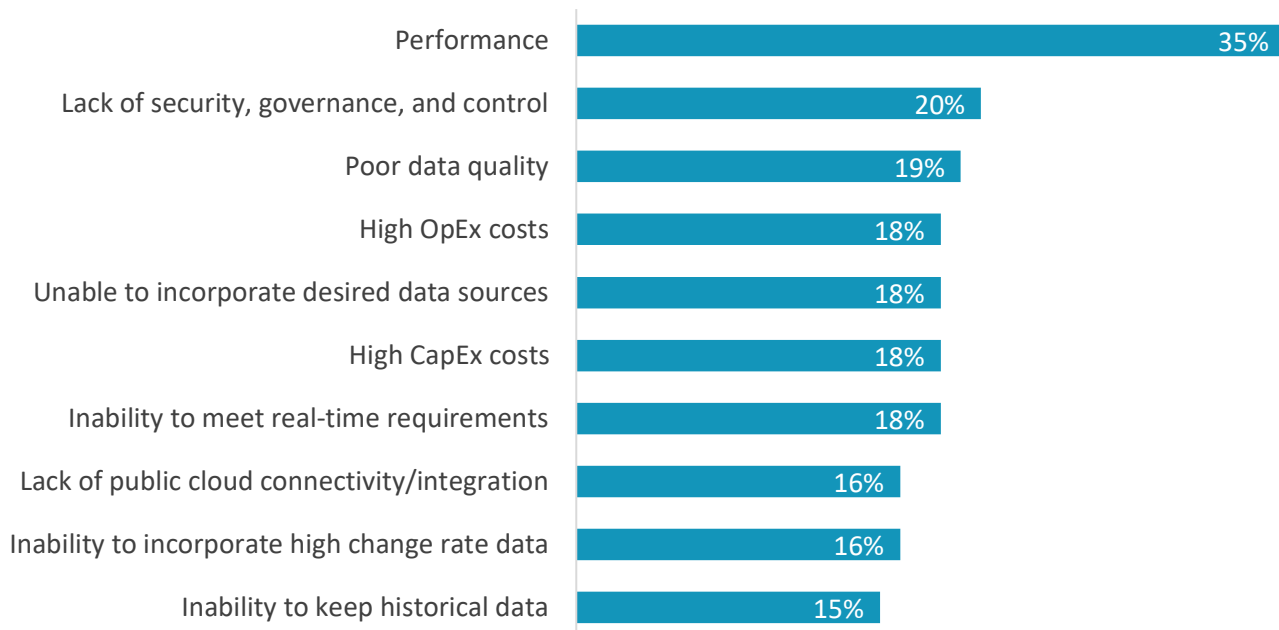
with a goal of informing change, driving action, and fueling different business behavior with the business outcome top of mind. As investments continue to be made in cloud-based architectures and the consolidation of data silos gets prioritized, organizations must rethink how to effectively tie structured business transactions with their ultimate analysis. How can businesses maximize the speed at which growing volumes of data can be collected and analyzed, insights can be reliably delivered, action can be effectively taken, and all of it can be done in a repeatable manner? Simply put, in today's fast-moving world, traditional BI can't keep up.

Data Warehouse Modernization

Today, organizations struggle with traditional EDWs that are not capable of doing much more than storing and batch querying structured data. EDWs are constrained by performance, with most handling only a subset of an organization's structured data. In fact, ESG research shows performance as the most often cited challenge experienced with data warehouses today (see Figure 1).[1] And to address the performance challenges, many organizations just look to add resources (compute and storage) in hopes of solving their performance woes. Not only is this approach expensive due to the proprietary nature of traditional EDWs and costly licensing structure, but resource inefficiency also becomes a major issue as organizations try to run workloads that were never meant to be run on an EDW. Data preparation is a great example of a workload that takes valuable resources away from analyzing data on an EDW as organizations look to integrate and conduct data preparation tasks on an EDW before analysis can even begin. Unless utilized properly, EDWs can quickly become disorganized storage systems with little attention to data quality, jeopardizing both trust in system reliability and trust in the data. As organizations look to prioritize incorporating more data of different types, the EDW is beginning to serve as the single storage location for structured and semi-structured data.

Figure 1. Top 10 Challenges with Enterprise Data Warehouses

What challenges have you experienced with your enterprise data warehouse solution? (Percent of respondents, N=310, multiple responses accepted)

Challenge	Percent
Performance	35%
Lack of security, governance, and control	20%
Poor data quality	19%
High OpEx costs	18%
Unable to incorporate desired data sources	18%
High CapEx costs	18%
Inability to meet real-time requirements	18%
Lack of public cloud connectivity/integration	16%
Inability to incorporate high change rate data	16%
Inability to keep historical data	15%

Source: Enterprise Strategy Group

---

[1] Source: ESG Master Survey Results, *The State of Data Analytics*, August 2019.

As an organization grows, the amount of data quickly grows with it. What was once a workable solution to storing and organizing data for reliable reporting and dashboarding quickly becomes an issue of maintenance and optimization. As organizations prioritize the need to bring data to every decision, more people need access to more data. Between constant changes to the underlying data, growth of that data, and ongoing updates to the technology stack, there is a never-ending cycle of more data, requiring more people, faster scaling requirements, and more overall agility. Organizations are practically forced to look for a new approach. In fact, according to ESG research, nearly 40% of organizations are likely to change their data warehouse vendor in the near future.[2] And they're looking to embrace a modernized data warehouse that can enable data-rich experiences that transform modern businesses into data-centric powerhouses.

Advantages to Modernizing the Data Warehouse

So, what advantages do organizations realize when modernizing their data warehouse?

- **Elastic scaling** – Serverless infrastructure can scale elastically to meet growing analytics needs. Because it is fully managed, users can stay focused on the analytics instead of worrying about database operations, like migration and staging of data before analysis can be conducted.

- **Integrating new data sources** – Organizations can use data at any scale, focusing on data volume as well as several data sources. As the organization grows and rapidly obtains more data, it needs the capacity to integrate new data sources and host the rising volume of data. A modern EDW can help meet the demands of near-real-time analysis in a data-centric organization.

- **Faster time to insight** – With less delays surrounding insights, organizations derive value from insight, so modern data warehouses emphasize speed to insight, providing end-users with the ability to quickly find value in data. They also ingest streaming data so users can analyze business events as they unfold.

- **Democratizing access** – Rather than siloing data and making it a burden to traverse the data landscape, modern EDWs focus on every business function by storing all data in one place. They empower even data analysts to run analytics without needing to acquire new skills.

- **Plan for the future** – With a modern data foundation, modern EDWs can serve as a launching pad for more advanced forms of analytics, including the incorporation of next-generation technology like artificial intelligence and machine learning.

And while these benefits are valuable, modernizing the data warehouse is just the start. Although the promise of modern data warehouses is appealing, the database itself still primarily serves as the processing engine. While many EDWs have integrations and hooks into scalable analytics platforms, a modern data warehouse requires an equally modern analytics platform to reach its full potential.

The Rise of Modern Analytics

Between the lack of accessibility to high quality and trusted data, the need to incorporate real-time data that is coming in too quickly for traditional systems to handle, and the pace at which business objectives change, organizations face several roadblocks that are forcing them to reevaluate their approach to analytics.

---

[2] Source: ESG Master Survey Results, *The State of Data Analytics*, August 2019.

Data is becoming a daily resource for many workers across organizations, and there is a continued emphasis on using data more openly and transparently to derive actionable insight. As workers are utilizing data more frequently, they expect it to always be readily available, from non-data-specialists and their demand for data to data experts like data scientists or generalists just getting started with data analysis. Non-data-specialists are the single fastest growing source of demand for data and insights. As with a traditional data warehouse, problems start to arise when there is a constant flow of new data without something to do with it yet. With the lack of ability to use or access that data effectively as well as scale on-demand, organizations are struggling to perform real-time analysis. And simply put, traditional tools do not have the capacity to help organizations derive real-time insight from their data. And gaining real-time insight is step one. Step two is applying that insight to the business. As business goals and objectives constantly change given new insights and strategies, organizations have not perfected how to take insights and accurately align them to ultimate business outcomes. In fact, ESG research shows that 44% of organizations take at least a month to act on new insight, and 21% take four months at a minimum. [3]

44% of organizations take at least one month to act on data insight

What Do Organizations Want?

Organizations want to empower their workers with fast tooling that allows them to perform near real-time analytics on all their data without worrying about managing the infrastructure. Businesses want to empower data-driven thinkers to access any data without limits, with the end goal being to bring data to all jobs and tasks regardless of where people sit in an organization. And to get there, organizations are investing in modern analytics tools and technology. In fact, nearly 1 in 3 organizations believe improving data analytics for real-time business intelligence and customer insight will be one of the business initiatives that drives the most technology spending at their organization in 2020. [4] And when it comes to capabilities of that modern technology, ESG research shows that security, reliability, performance, cost/TCO/ROI, and simplicity are the top 5 most important attributes/capabilities when it comes to analytics-centric technology (see Figure 2).[5]

Figure 2. Top 5 Capabilities/Attributes of New Technology to Support Analytics

When looking to adopt new technologies to support data analytics initiatives within your organization, what are the most important capabilities/attributes?(Percent of respondents, N=310, five responses accepted)

Security	47%
Reliability	38%
Performance	36%
Cost, ROI, and/or TCO	35%
Ease of use	31%

Source: Enterprise Strategy Group

---

[3] Source: ESG Master Survey Results, *The State of Data Analytics*, August 2019.
[4] Source: ESG Master Survey Results, *2020 Technology Spending Intentions Survey*, January 2020.
[5] Source: ESG Master Survey Results, *The State of Data Analytics*, August 2019.

Organizations want trust, speed, efficiency, and accuracy to support the need to be more action-oriented based on rapid insights derived from data. Organizations want a way to use their data beyond a top-to-bottom, traditional BI approach. They want integration of their data into advanced workflows. They want to leverage APIs within preferred tools. They want to integrate data into modern, event-driven applications in a user-friendly way.

So how do organizations get there? One way to tackle these problems and allow end-users to bring more data to the business decision making process is leveraging a modern approach to business intelligence (BI) and analytics. A modern BI approach greatly improves an organization's ability to trust, access, and operationalize data through automation and self-service. Data teams and data experts can focus on enablement and adding strategic value, as opposed to ongoing management, maintenance, and uptime. This is a big reason why ESG research shows approaches like serverless analytics are garnering interest. While just 15% of organizations currently use serverless analytics, 42% of organizations are planning to implement or evaluate serverless analytics in the next 12 months.[6] On top of simplifying administration, a modern BI approach can reduce the time to develop, deploy, and refresh data integration processes; improve data modeling; define key metrics automatically; and implement advanced dashboarding and reporting, all anchored in self-service. This arms the workforce with ways to quickly shift with the dynamic nature of the business, preventing them from falling behind and being inundated with the constant influx of new data. By modernizing the stack and consolidating tools, technologies, and services, organizations can add more business value, faster. And while addressing the needs of the present is important, organizations have not lost sight of what is to come soon: the integration of more advanced technologies like AI and ML.

The Modern Data Experience

Modernizing the data experience requires investing in a platform that delivers value across the entire data analytics lifecycle. That means unlocking access to the right data at the right time to conduct effective data storytelling with rich and interactive visualizations and dashboards. And when this is done right, a new level of trust can be achieved throughout the business on top of traditional trust pillars of downtime, security, and privacy. Business and people will gain trust in the platform, trust in the data, and trust in the outcomes.

A modern data experience incorporates:

1. **Multi-cloud analytics** – Incorporate data into analytic workflows regardless of where it is located across the public cloud landscape.

2. **Automation** – Focus less on administration and more on business outcomes via serverless analytics. And with self-service capabilities, organizations can lower the barrier to entry for more people to leverage the right data when needed. That level of effective self-service carries over to embedded analytics, where developers can seamlessly incorporate advanced analytics and BI directly in their modern applications.

3. **Flexibility and agility** – Scale to meet the needs of the business, whether addressing a burst of activity, a growing end-user base, or more complex interactions and integrations.

4. **Openness** – Ensure the vast ecosystem of open source and third-party tool integrations are available and able to be incorporated into the analytics lifecycle seamlessly.

5. **Democratization** – Provide the necessary guardrails to enable more generalized access across the business, without burdening the experts. That means having the right data available to the right teams while ensuring they can easily run analysis and share results. Self-service gives the data experts time back, whether by not having to hand-hold generalists or respond to inquiries/requests, or simply by making preferred tools more accessible and integrated into

---

[6] Source: ESG Master Survey Results, *The State of Data Analytics*, August 2019.

their custom workflows. That means the data experts get the granularity and customization they want within their working environments and custom applications.

6. Security – By prioritizing ubiquitous data access to everyone and the ability to integrate preferred tools and services, ensuring the proper levels of security are in place is critical to successfully delivering a modern data experience. Data governance should be prioritized to ensure the right users have access to the right data. Further, proper and effective data governance will enable organizations to meet various types of compliance, whether based on location (i.e., GDPR, or based on industry (i.e. HIPAA).

7. Advanced insights – Enable end-users to incorporate next-generation technology services infused with AI and ML to deliver more advanced insights based on historic and streaming data.

Looker Can Help

Looker is a modern data platform that helps companies deliver modern data experiences. Its services align with any organization's data-centric goals, providing users access to all their data and the ability to rapidly act on data-driven insights. Regardless of where an organization sees itself, whether needing to modernize its database, embracing multi-cloud, or advancing analytics initiatives, Looker can help. It is an organization aimed at arming companies with the ability to derive better insights and make data-driven decisions at the speed of the business. By empowering more end-users through an advanced self-service BI architecture, Looker enables reliable access to more trusted data quickly. And the platform does so in a way that enables rapid action to be taken depending on the dynamic business environment and objectives. Looker is helping shape the possibilities of what data can do for a company, and its value comes from helping organizations work with data in more ways, at whatever stage they are when it comes to data analytics maturity. And most importantly, it provides a powerful data analytics experience regardless of the type, size, speed, rate, or location of data.

> **With Looker, you can deliver actionable business insights at the point of decision, create new value streams, and infuse data into products and workflows to move the business forward.**

The Bigger Truth

Traditional data tools are unable to solve modern data problems. They were not built with the size, scale, and speed demands of today's data. And it's forcing organizations to look for help by embracing modern tools that can enable modern data experiences. It starts with a modern data warehouse, transitions to modern business intelligence and analytics, and ends when everyone is enabled to bring data to their jobs, tasks, and responsibilities on their terms. As Looker continues to prioritize the delivery of an advanced BI and analytics platform across cloud environments that enables a modern data experience, organizations have taken notice. And now being part of the Google Cloud team, a shared philosophy around openness, support, and democratization of analytics and business intelligence will enable customers to achieve their data modernization goals.