

PANOPLY WHITEPAPER

Eradicating Efficiency Roadblocks with Panoply

Standing up a data platform usually takes months. A managed data warehouse + ELT capability can take your business from zero to insights in minutes.



Becoming data-driven

Startups and SMBs are investing heavily in data initiatives.

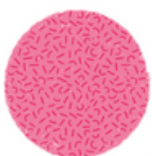
Is your company trying to become data-driven? A robust data platform is key to getting insights from data, but there are a number of inefficiencies that can slow down the process for businesses embarking on their data initiative. For many organizations, the lead time required to procure storage hardware for an on-prem solution can delay – or even derail – the start of a data and analytics initiative.

Standing up a data platform can take months, not to mention time and effort for ongoing, constant maintenance, and business leaders are impatient when it comes to realizing the value: insights and monetization, agility, and customer satisfaction to name a few. Data insights and monetization are critical to growth and gaining a competitive advantage. In this paper, we'll discuss the roadblocks¹ in many data initiatives and how to solve them.



Data insights and monetization are critical to growth and gaining a competitive advantage.

In this paper, we'll discuss the roadblocks in many data initiatives and how to solve them.



(1) <https://bespokeunit.com/articles/style/icons-of-mens-style-book-review/>

(2) <https://www.marketingprofs.com/2/buyingnotrational.asp>

(3) <https://www.inc.com/logan-chierotti/harvard-professor-says-95-of-purchasing-decisions-are-subconscious.html>



Efficiency blocker 1: Siloed Resources

The reality is most business units are competing for budget and various IT resources to support their strategic initiatives and are executing their projects in a siloed manner. The friction caused by these IT resource silos can complicate the progress and launch of your project unnecessarily. Tension between the multiple teams involved often lengthens the project lifecycle. For example, consider the following teams that are typically involved in a data warehouse initiative, and how siloed requirements for the disparate groups could impact the project:

- **IT Infrastructure**
- **IT Storage**
- **Database administration**
- **Development**
- **IT business intelligence**
- **IT project management**
- **Business team requesting projects**

Managing and aligning all of these resources to reach a common goal is quite challenging. So it's no surprise that an undertaking like building a data platform can take months to years to launch—not to mention complete. Worse yet, sometimes the projects aren't even successful. There is often competition for these team resources, especially shared ones like the infrastructure and database teams, between the project stakeholders and elsewhere in the business. This means the data warehouse may not be the highest priority for the IT organization.

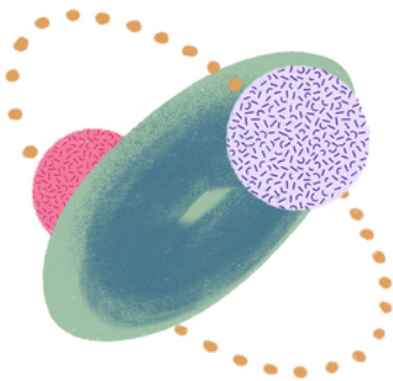


Deploying data platforms is tough

Besides the political and organizational challenges of deploying and maintaining a data platform, there are also significant technical challenges. The following are some of the most intimidating hurdles for any organization attempting a deployment or refresh of these types of systems today.

4

Efficiency blocker 2: Development and data logic



When trying to build a data and analytics platform in-house, many companies reach out to third-party consulting firms to support development. While this approach can be beneficial since the hired resources are fully dedicated to the project, it often causes delays and missteps due to a lack of in-depth knowledge of your business. Third-party teams aren't always properly brought up to speed, including the transfer of historical knowledge and undocumented intricacies of the business.

Even when working with internal development staff, it can be challenging to convey the necessary requirements. Data warehousing projects tend to follow more of a waterfall model of development than an Agile one, which means rework tends to be more comprehensive and costly in terms of time and resources. The other challenge is the cumbersome and fragile nature of the ETL process. Every change to one downstream process also potentially impacts other processes all the way back up to the top. This cascade means reconfiguring and redeploying code in addition to going through new testing cycles.

With legacy tools, ETL (extract, transform, load) was nearly always custom-built. When an organization wanted to develop a data platform, they typically had to start from scratch with ETL. Many attempted internally to better automate their ETL processes, but often still relied on manual techniques or primitive scripts that can be greatly impacted by changes. Since ETL development is customized for each customer, there is limited use of frameworks and repeatability between industries. Now we have modern, managed ELT (extract, load, transform) tools to handle these processes and eliminate the time spent on manual tasks.

Efficiency blocker 2: Development and data logic

For smaller companies, one of the benefits of cloud computing is that they are effectively outsourcing the management of their infrastructure. As mentioned earlier, large organizations have a significant percentage of their IT staff dedicated to keeping the lights on. These IT operations teams usually performs some of the following duties:

- **Managing storage and free space**
- **Patching servers and databases**
- **Replacing failed hard drives**
- **Managing backups and restores**



In a smaller business or startup, all of these tasks might be handled by just one or two people with additional responsibilities within the IT team. They can quickly become over-burdened, making it tricky to proactively implement new initiatives. Just getting the hardware for a project can become challenging in a smaller company, much less planning and launching a large project like a data platform.

Efficiency blocker 3: Performance

In organizations of all sizes, there are often performance issues in large data platforms. When you start working with terabytes and petabytes of data, your ELT process, data warehouse, and visualizations must be optimized for ideal performance. In the traditional relational database world, this meant optimizing indexes across the data warehouse. That task used to be a time-consuming, tedious art, but it has evolved over time, driven by the need to handle larger data sets, varied analytical workloads, and the evolution of new data storage and processing technologies.

The focus has shifted from manual index creation to automated or data-driven approaches that account for specific characteristics of the data warehouse. In larger companies, the solution is sometimes as primitive as purchasing more robust hardware to meet the needs of the system. While this can be an adequate fix, it is expensive, sometimes even wasteful, and does not solve the underlying inefficiency.

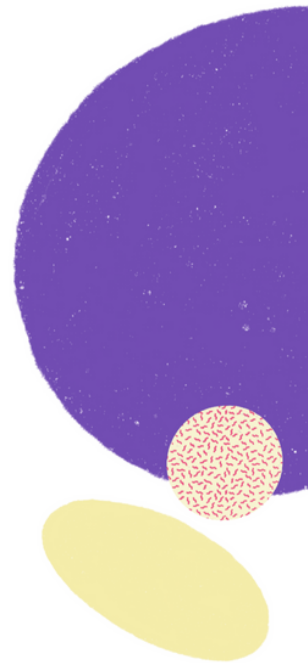
For businesses that have moved to big data systems like Hadoop, performance tuning is even trickier. Of course there are techniques for optimizing performance of these systems, but the expertise in that space is rare and expensive⁷. In some cases, adding more nodes to the cluster can solve the problem, but “throwing hardware at the problem” is not a resolution method that can scale forever.

In recent years, many relational database vendors have made inroads in large system performance by using columnar storage techniques. This structure, used in Google BigQuery (upon which Panoply runs), involves taking tables of data and turning them on their side; the rows turn 90 degrees and become columns.

Structuring the data in this way has the benefit of introducing a great deal of compression potential into the data.

Since the data in columns tends to have a higher rate of duplication, compression of columns can be up to five times more effective than other reduction techniques. The other major benefit of this technique is that the columns not involved in a given query never get scanned; this greatly reduces the storage operations needed to return the query results.

While this technique is good, many organizations are stuck on older versions of database software that do not support columnar data storage and other modern features. Friction concerning upgrades can happen for many reasons: licensing, organizational standards, or dependence on third-party software that does not offer support. Being stuck on legacy database software means that a great deal of time is spent troubleshooting performance rather than writing code and getting insights from your data that deliver business value. Further, smaller organizations may lack the expertise necessary to perform appreciable performance tuning on these systems.



Tuning data warehouses involves a tedious process of capturing queries, evaluating execution plans, and gradually implementing improvements through a testing cycle. It can take several days, and often several weeks, to resolve a particularly problematic performance issue. It is important not to underestimate the resources necessary to meet these needs, and automation tools should be considered to speed things along.



Automation is an inevitable part of the process for companies that want to deliver successful IT projects in the post-cloud era.

Another challenge for many organizations is a lack of knowledge around newer solutions. Many leaders spend time and money researching data solutions when they could be better spent looking into industry-related tools (SEO and ads for marketers, better payment platforms for eCommerce, etc.) and leveraging a managed solution for their data platform.

Their business may be in a position to take advantage of emerging techniques such as machine learning and predictive analytics, but they lack the organizational knowledge to deploy these types of systems. Or in a large company with a large IT team, these systems may be “off-menu” items that would require extra time to configure and deploy.

How does Panoply address these problems?

By delivering a [fully cloud-based solution](#), Panoply eliminates the costly delays associated with deploying hardware and lets you quickly get to the true value of your data. Instead of months of lead time followed by months of development data, you can have your platform ready in minutes.

Panoply’s cloud-based, low-code model allows us to eliminate some of the problems we discussed earlier, while creating some advantages for our users. Below, we’ll explore a few ways Panoply drives value over legacy data warehouse solutions.

Ease of use

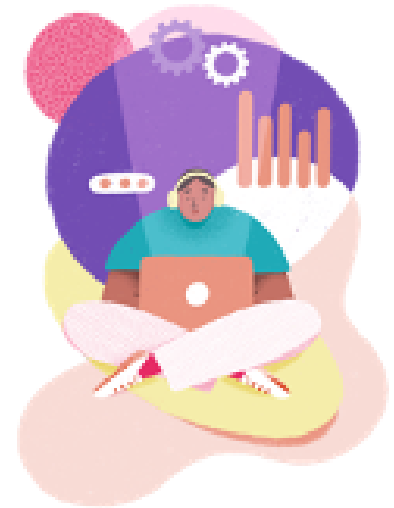
Banishing resource silos

Panoply's ease of use eliminates resource silos and long project timelines by standing up your ELT pipeline and managing your data warehouse for you.

Your business units won't be competing for already strained IT resources trying to get the project off the ground, because Panoply is designed to be connected in minutes, even for non-technical team members.

Our no-code ELT connectors allow users who need the data to connect to any data source and begin ingesting right away.

With data from all your APIs, CRMs, databases, and other sources in the data warehouse, your company has a single source of truth for better, actionable insights. The platform allows you to manage permissions and access so that each business unit and user can get the data they need as they need it, without chasing down the IT team. The concept of "self-service BI" has been around for quite some time, but Panoply takes it a step further with a self-service data platform.



Eliminating the build timeline

Additionally, Panoply's ease of use means that your company doesn't have to wait for a third-party data platform vendor to get up to speed on your business and product. Decision-makers and other users without a technical background can spin up the data warehouse, connect sources, and start aggregating the data that matters most for their objectives. Users can also automate data collection, significantly reducing time spent on manual tasks.

This isn't to say that Panoply is only for the less technical users - developers can explore the data more deeply in our SQL-based workbench.

(11) <https://page.clever-touch.com/rs/868-BYZ-669/images/Clevertouch%20State%20of%20Martech%20Full%20Report%202022.pdf>

(12) <https://www.gartner.com/en/newsroom/press-releases/2022-09-15-gartner-survey-reveals-marketing-analytics-are-only-influencing-53-percent-of-decisions>

(13) This Panoply customer prefers to remain anonymous

Aggregation

Another part of legacy data warehouse projects is building an aggregation model. Typically, aggregation is accomplished using an OLAP cube, which allows users to query the data warehouse in a more ad-hoc fashion. Users can slice and dice data based on key values and filters.

Building this OLAP model requires additional development time, and OLAP queries are batch-processed daily, meaning that the business may be looking at day-old data at times. You also have the ability to extend this functionality by building transformation views, which are instantiated views with programmable, user-defined functions. By building these, you can customize your warehouse to meet all your business needs, and eliminate the tedious process of getting to where you can finally use your data to make decisions.

Looking ahead

For today's businesses, a data platform isn't simply a "nice to have" - it's a must. While every company embarking on this project should be aware of the potential roadblocks, understanding the ways to solve them will make your initiative more streamlined and efficient. Leveraging a solution like Panoply means dramatically reducing your time from "data everywhere" to clear business insights.





If you're inspired and looking to take your marketing analytics to the next level, book a call to see how Panoply can work in your data stack today.

[Book a demo](#)

Get in touch, visit us at panoply.io



panoply
by SQream