GROWTH ACCELERATION PARTNERS

snowflake®

Automating Cloud Infrastructure to Leverage Snowflake

The Snowflake Data Cloud is taking the big data world by storm, as more and more organizations discover the huge advantages Snowflake brings over its more traditional competitors like Teradata and Oracle.

Building a modern data analytics platform based on Snowflake can be time consuming and challenging, whether the deployment is on Amazon Web Services (AWS), Microsoft Azure or other cloud platforms. But it doesn't have to be that way.

Growth Acceleration Partners (GAP) brings speed and simplicity to an otherwise thorny IT problem: automating infrastructure as code (IaC) for Snowflake cloud build-outs.

A multi-dimensional problem

One of the most common infrastructure projects is building a modern data analytics platform with ETL (extract, transform and load), staging and dashboards. Let's look specifically at creating a cloud-based modern data analytics platform using Snowflake.

What does that require?

Data ingestion, storage and transformation are key. Both structured and unstructured data require pipelines, possibly including ETL scripts before storing it on AWS or Azure in a scalable and flexible manner. The method of storing different data lakes will dictate how easy and effective using the data will be. Rarely is data ready for use directly from its raw source, so tools and automated processes to transform, normalize and enrich the data will be required as well.

Analytics tools like PowerBI or Tableau offer the ability to easily create and share reports, dashboards and insights across teams. Additionally, data analytics should include capabilities for machine learning, predictive analytics, data visualization and more.

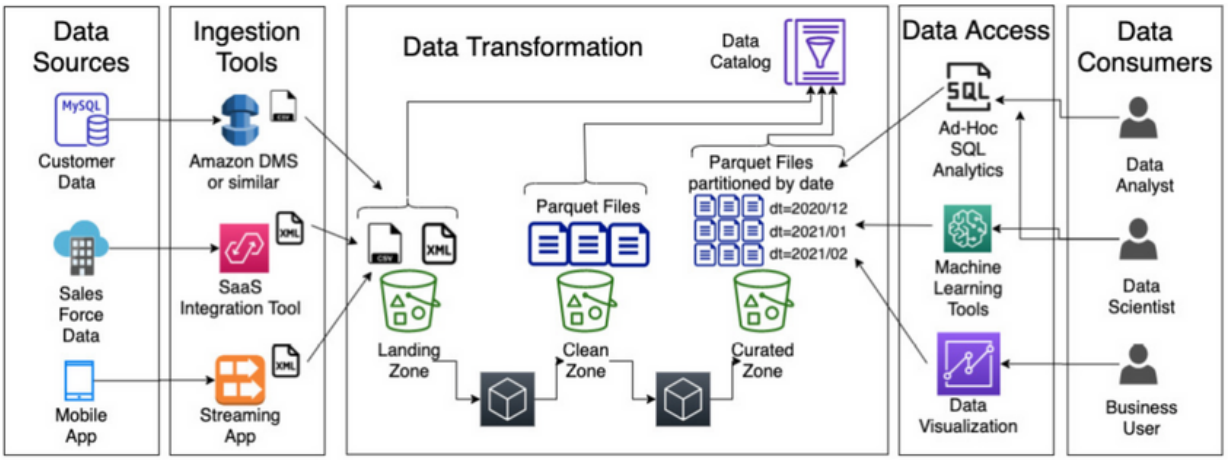
Scalable, secure and responsive architecture allows you to easily and quickly respond to changes in demand and volume without incurring unnecessary costs and burdens.

Security and privacy of data — including conformance with regulatory statutes such as HIPAA, GDPR and CCPA — is paramount for encryption of all data, both in transit and at rest.

Infrastructure management includes monitoring and control of all cloud resources to guarantee high availability with failure/fall back plans.

Getting it right is critical; getting it wrong can be catastrophic

Here's an example of a modern data analytics cloud deployment. As you can see. it's a complex architecture:



Chris Fregly and Antje Barth. *Data Science on AWS*. O'Reilly Media, 2021.

It goes without saying how critical it is to put a correct version of this into production, especially since the resultant data analytics, visualization and dashboards will drive organizational decisions and even investments. The consequences of even a small misconfiguration could be both invisible and also significant. Thus, the design, implementation and QA audit of a data analytics platform is perhaps the most important aspect of the entire system.

A better way

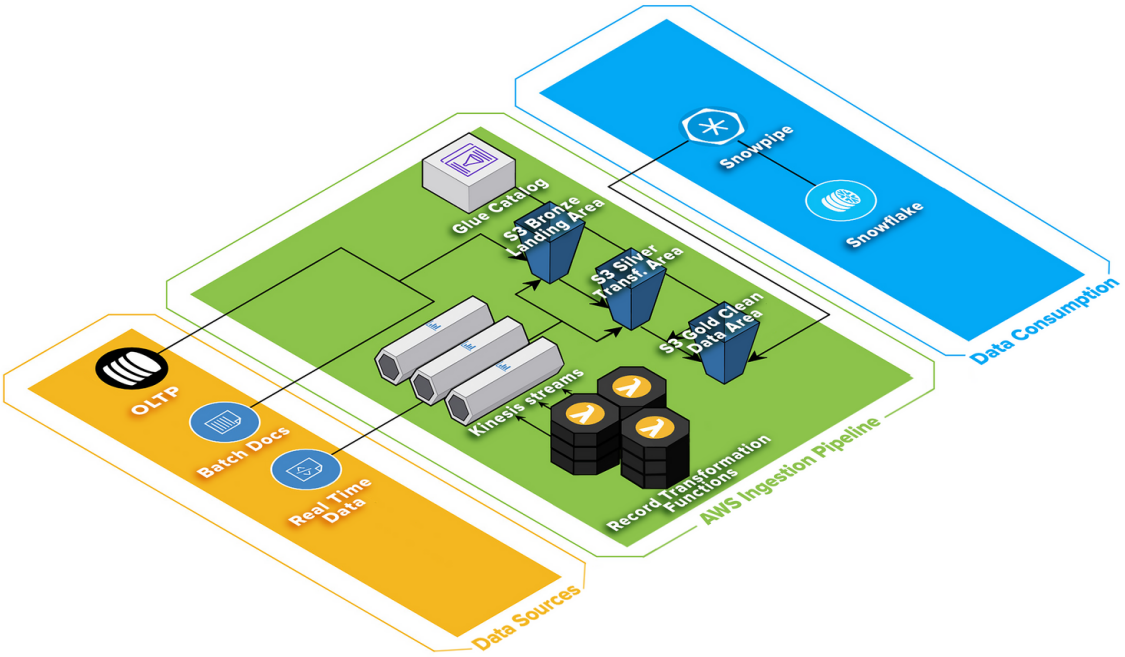
GAP has years of experience building out these kinds of cloud-based platforms for customers. Taking into consideration both customer requirements and current industry best practices, GAP built multiple cloud infrastructure accelerators to cut the time needed to build out a complex Snowflake deployment and also to ensure correctness and simplify modifications.

Based on Terraform, an IaC toolset, GAPBuilt Accelerators allow the customer to easily tailor the templates to their specific requirements. These templates — using the Terraform IaC engine — can then create an entire cloud-based serverless web application or modern data analytics platform as pictured above in just minutes.

The power of IaC using Terraform and carefully-developed templates is two-fold:

1. Build out all the cloud infrastructure with a single button push.
2. Gain the ability to modify, delete or update any piece without expensive and painful duplication or risk of error.

As an example, here is the structure of one platform build-out from the accelerators, created with a single “go” button push and finished within a few minutes:



This platform is now ready for production, saving weeks or months of painful build-out using the AWS or Azure management consoles, while avoiding mistakes, mis-sizing of components or wrong-sizing of the architecture.

To learn more about GAP’s tools for automating cloud infrastructure build out and creation, please visit **WeAreGAP.com**

> These are non-trivial implementations; they take on the order of months or years, not weeks or days. These are tricky and tough to build, but they're very, very powerful when built. GAP has done the hard work, and has some of the magic happening around this data platform.
> — Steven Borg, Big Data & AI/ML Expert and former Senior Director at Microsoft