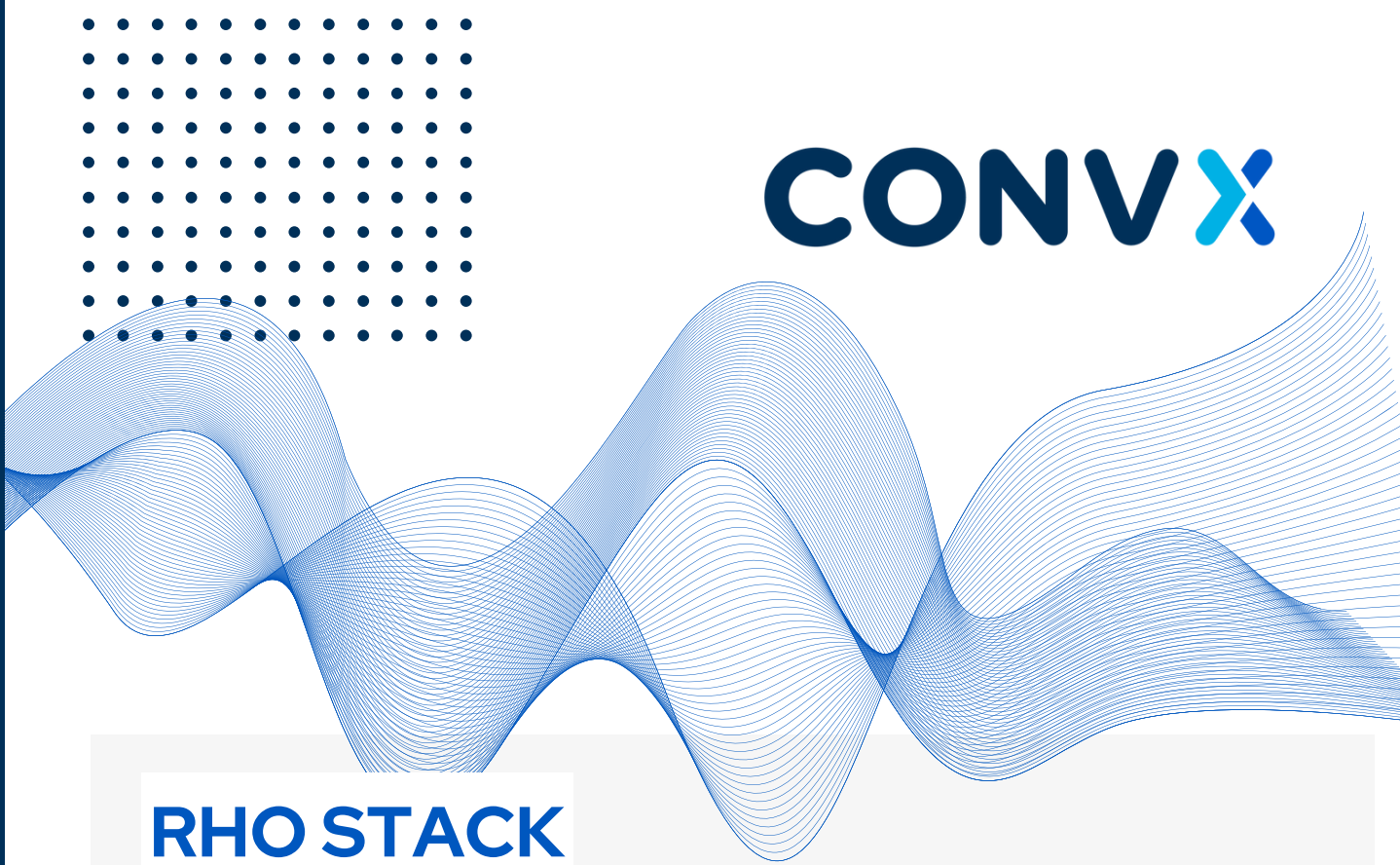


The background of the slide features a dark blue vertical bar on the left. In the top left, there is a grid of small dark blue dots. A large, flowing, light blue wave graphic spans across the middle of the slide, partially overlapping the text area.

RHO STACK

Managed Cloud Data Platform

DATA ARCHITECTURE ON A RESTAURANT'S BUDGET

Technical Considerations and Best Practices
for Restaurant Data Architecture

BUILD OPERATIONAL EFFICIENCY

INCREASE PROFIT MARGIN

convx.com

©2024 CONVX LLC 2301 Blake Street Denver, CO 80205

TABLE OF CONTENTS



INTRODUCTION	1
ABOUT CONVX	2
CONVX RHO STACK SUCCESS	3
EVALUATING SOLUTIONS: WHAT TO ASK AND CONSIDER	4
CONVX SOLUTIONS FOR RESTAURANTS	10
RHO Data Pipeline Solutions	10
RHO Data Warehouse Solutions	11
RHO Data Apps	11
APPENDIX: UNDER CONVX’S HOOD	12
I. Data Warehouse Design	12
II. Ingest Pipeline Design	13
III. Data Transformation Design	18
IV. Data Enrichment and External Data Sources	19
V. Analytics and AI	19

[Contact CONVX](#)

Introduction

Restaurants run on as many systems as menu items: POS, BI and analytics, loyalty, customer experience, franchise, HRS, LMS, and supply chain, just to name a few. However, they're left with only a few options when it comes to managing, pooling, and integrating their data:

1. The Spaghetti Tangle: Connect everything, everywhere. This yields messy, inconsistent data that's either siloed or at risk with each new system change or update.
2. Invent an architecture: Hire a team of data engineers, cloud infrastructure engineers, DevOps engineers, and analysts to try to build something great from nothing. This is a multi-year process with a high risk of failure, especially with turnover and continuous changes to the restaurant technology ecosystem.
3. The **CONVX RHO Stack**: The data hub that unifies everything in a single platform. Connect once, automate workflows, and gain crystal-clear insights.

Efficient data clarity starts with the right data architecture: **CONVX RHO Stack**. This paper discusses the key technical considerations underpinning our architecture and the best-in-class practices that every restaurant should incorporate into its own data platform.

But first, here is why you should consider CONVX:

CONVX RHO Stack's managed solution puts you in control of your data. No more vendor lock-in for fear of losing your valuable guest intel and sales history. No more weekends spent trying to recover file backups. No more time or money wasted on hiring consultants or staff. Instead, you will have:

- **Safe, seamless data management with pre-built integrations, custom data apps, secure architecture, and our expert team oversight.**
- **An affordable cost effective all-in-one solution delivering near-real-time data.**
- **Actionable insights with AI-powered analytics, predictive models, and custom data apps.**
- **Streamlined operations and automated workflows that let you focus on growing your business.**

Our results are indisputably impressive: 50% gains in integration efficiency and up to 70% reduction in data analysis tools and expenses.

CONVX RHO Stack is more than the restaurant industry's single source of data truth. It's how you'll increase your profit margin, and unite your multiple data streams into one seamless platform to automate workflows, scale operations, and accurately forecast guest traffic.

About CONVX

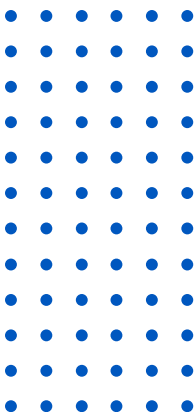
Founded in 2017, our expertise in data and cloud infrastructure runs deep. Serving over 110 customers and completing 36 projects just last year, we accelerate time-to-value (TTV) by 250% with 60% to 75% cost savings and a 100% success rate compared to in-house development or consulting firms.

Our collective decades of experience in IT and data services, DevOps, software development, and cloud infrastructure have earned us the role of most-trusted solution provider to an impressive portfolio of clients spanning government, mid-size to enterprise companies, and even small businesses.

Our U.S.-based team has built the leading data and cloud infrastructure for restaurants based on our decades of core expertise in guiding our clients through digital transformation and data modernization with security and compliance at the heart of it all.

For more information on our success in implementing our RHO Stack, see below.

70%
Cloud Integration Savings



CONVX RHO STACK Success

Customer Metrics

Operational Improvements

Cloud Integration Server Savings

Slashed 70% off cloud integration server costs and eliminated excess CPU and memory.

70%

Data Services Cost Reduction

Reduced data services costs by 40% with cost-optimized data architecture.

40%

Payroll Savings

Saved more than \$80,000 per year by eliminating the need for a data and DevOps engineer.

\$80K

Shortened Transaction Time

Data insights from A/B testing multiple configurations led to a 6% reduction in POS transaction time.

6%

- Solved integration issues between loyalty and customer experience platforms with custom CONVX connectors
- Combined POS, loyalty, and customer experience into a unified view of customer behavior
- Eliminated the need for expensive analytics tools for many users
- Eliminated irrelevant questions from customer surveys by customizing surveys to the customer's menu selections and dining choices (e.g. omitting bathroom cleanliness questions for drive-thru customers)
- Identified menu anomalies at different restaurants by easily comparing menus across all locations

Evaluating Solutions: What to Ask and Consider

When so much of your success is riding on the integrity and frictionless processing of your data, you need to ask probing questions like these below to identify what to focus your limited time and budget on.

CONVX's proven data architecture, the RHO Stack, was developed and refined for moderately-sized restaurants to overcome the challenges of cost-prohibitive data solutions. It unlocks efficiencies, enables scale, and increases profit margins. It provides a recipe for any restaurant business to build its data platform, either on its own, or as a CONVX-managed service.

We believe:

- Restaurants with data insight access have an outsized advantage for profitability and growth.
- Small or midsize restaurants that lack a dedicated data team can still win with the right data architecture.
- The best data architecture for restaurants uses common design patterns based on established data engineering principles.

Are you in the cloud or on the ground?

Restaurants are inherently a distributed IT architecture. SaaS now reigns supreme for restaurant apps (POS, ordering, loyalty, etc.). Cloud-based tools for a shared tech stack deliver distributed efficiency, which makes the cloud the natural fit for a restaurant data warehouse. It's also CONVX's standard platform. Although we support many traditional databases, we see new deployments starting in, and legacy deployments migrating, to the cloud. Our customers, from restaurants to highly-regulated gaming companies, are finding efficiencies in shifting analytics and data integration workflows to the cloud, implementing our best-in-class and state-of-the-art solutions.

Are you able to build your architecture from scratch?

Building a complex data platform from scratch yourself can take years. It requires significant payroll to develop, maintain, and update. The capital investment and operational costs can be huge. Even with open-source tools and technology, the infrastructure and effort are still significant (and technical support inconsistent or non-existent). You will need to integrate data from the systems your business has adopted. Each integration must be built and maintained.

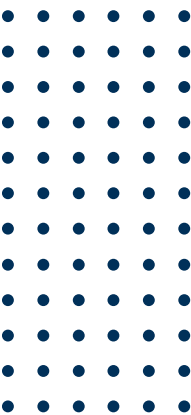
The solution? Minimize complexity and maximize ROI: leverage existing tech, proven patterns, and efficient processes for a smoother, more cost-effective operation. Base your data warehouse structure on a known design, like a star schema, that data professionals and tools readily understand. CONVX offers a pre-built architecture and integrations with common restaurant solutions that significantly reduce development effort.

Is data sprawl derailing you?

Cloud object storage is inexpensive and easily available, and you pay for only what is needed. Yet cloud object storage's ease of data replication and instant availability can lull many businesses into preserving duplicates of large amounts of data that will never be needed and make it difficult to hone in on the data that matters. It can be a large effort to identify and reclaim this wasted storage space.

The CONVX architecture eliminates data sprawl with a standardized architecture and multiple tiers that match each use case and budget.

40%
Data Cost Reduction



Are your insights really hindsights?

Traditional data architecture uses nightly batch transfer, so your insights are stale by at least a day. Switching from daily batch ingestion to modern streaming technology reduces this to minutes. It opens up new possibilities: near real-time performance dashboards, immediate response to issues, and more. What could you do with that readily available analysis?

Is data transfer eating your margin?

Data transfer is fast in the cloud, but transferring data to on-premises servers or other clouds is expensive. With the major cloud providers, data is free to transfer within the cloud, but fees apply for data leaving the cloud or moving to other cloud regions. These fees are significant for large, frequent data transfers and can eat at your profit margins. Optimize your pipelines to minimize the number of times data is moved and carefully evaluate the need for transfers between clouds.

Transferring over a slow link also impacts your agility. What takes seconds or minutes over a hyperscale cloud's links may take hours over a slower Internet connection. Consider where your applications are going and establish your data footprint there.

To minimize cost and delays, many restaurants opt to keep data in the cloud where it originates, eliminating the need for time-consuming transfers. Keeping data in the cloud closest to your most important applications maximizes application performance while minimizing network costs.

Are you sending cloud compute costs into the ether?

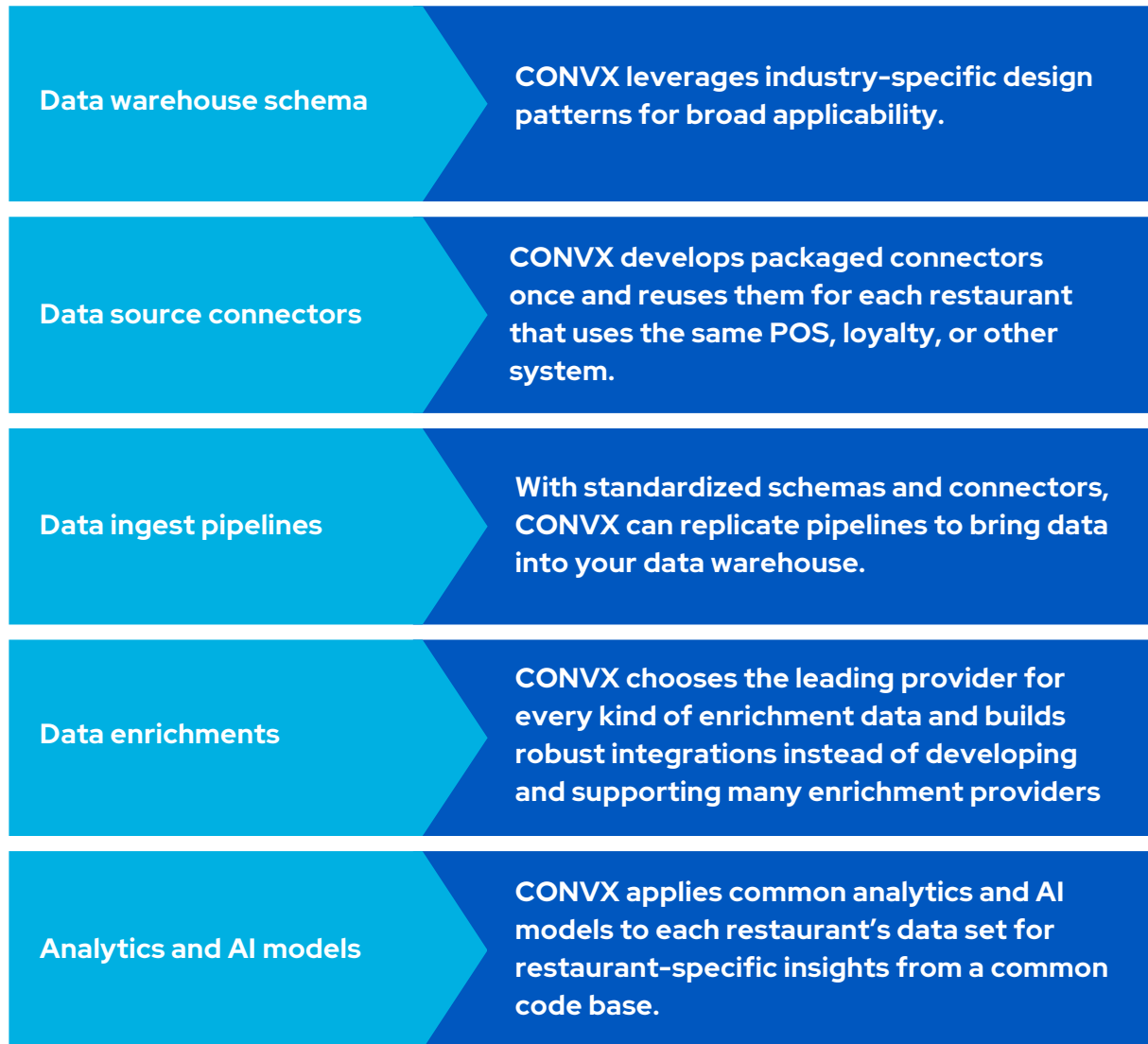
Cloud computing costs can be exorbitant. Unlike object storage, fixed compute allocations in traditional cloud models leave resources idle during restaurant off-peak hours, incurring unnecessary expenses. Why pay for a data integration server 24 hours a day when it runs jobs only one hour of that time?

Consider ephemeral computing for infrequent jobs. This simply is server capacity that can be destroyed and recreated anytime. Container-based applications can be deployed in seconds or minutes, then destroyed when their job is done. Major cloud vendors support serverless "functions as a service" (FaaS) to automate this whole process. CONVX's modern serverless options provide on-demand compute allocation and dynamic scaling, allowing you to optimize costs for infrequent jobs.

Are you stuck in a loop, reinventing the wheel?

Don't Repeat Yourself (DRY) is a well-known principle in software development. Restaurants with a standardized data architecture, integrations and pipelines will find components and design patterns that can be reused for the next use case.

This DRY principle applies to each component of CONVX's architecture:



Is your platform fragile or agile?

Restaurants need components that can be maintained independently or swapped out. CONVX components have minimal interdependency, limited to defined interface specifications. This loose coupling enables independent maintenance, technology swaps, and reduced testing overhead.

The simplified workflow below demonstrates CONVX's modular architecture and its ability to accommodate component changes without disrupting system-wide functionality. The data loader step, responsible for populating a table with raw unstructured data, is decoupled from the data transformer step, which solely reads the raw data and transforms it to our standard schema. This loose coupling allows for independent modification or automation of the data loader without affecting the downstream transformer.



Inter-step communication relies on a simple interface: an S3 bucket or Snowflake table with the raw data. Independent maintenance is frictionless, thanks to this loose coupling and well-defined interfaces, requiring collaboration only when interface updates are needed.

CONVX's pipeline design makes major pipeline components autonomous. They have well-defined input/output interfaces. Internal workings may evolve rapidly (schema changes, new systems). Loose coupling minimizes ripple effects, keeping most changes isolated.

\$80K
Payroll Savings

Do you need a house by the lake?

CONVX recommends maintaining both a data warehouse and data lake. The data warehouse provides the structured subset of data for current analytics and integrations. The data lake preserves all the raw data for any future need.

By preserving the raw data, we don't have to plan for all use cases upfront. We quickly develop for current priorities, then go back to the raw data and extract more information when needed. This agile approach is the key to dramatically reducing the time to value (TTV) on a data warehouse project from years to months.

The most efficient solution is a combined "lakehouse," where the data warehouse and data lake are stored alongside each other in a common repository. Our lakehouse solution allows us to directly query unstructured data in place. We can reprocess data without the cost and delay of large network transfers from an external solution.

For the most critical data, our architecture supports an additional tier of data archive in a second cloud. For the most vital historic data, like POS records, this may be warranted. It assures no single vendor owns your data, and no single incident can destroy all your data. It also offers broader tool options, better resilience (fault tolerant backup) and governance, but costs more to store and maintain. Consider data origin, value, and recovery options before duplicating. Less critical data may be re-ingested from the source or lost altogether.

Do data reruns cause meltdowns?

Idempotency protects data integrity during multiple loads, ensuring consistent results for downstream analytics and reports. It enables reliable reloads of past data in case of failures, handling duplicates through upsert operations that update existing records or add new ones.

While essential for data integrity, upserts can be resource-intensive, requiring more processing time than simple inserts. CONVX offers several strategies to mitigate this while ensuring idempotency, such as optimizing loaders for insert-only operations wherever possible.

Is your MDM a symphony or cacophony?

Large companies have the staff and resources for expensive master data management (MDM) tools that can generate master records for each restaurant, customer, or inventory item. Smaller businesses lack the budget and employees.

Excel spreadsheets are a common solution for small companies; but they're siloed from the data warehouse, dimension tables, POS, etc. They often become inconsistent with the production systems.

For small companies, CONVX eliminates the need to replicate master data changes into the data warehouse by managing this data directly in the data warehouse. This would traditionally require a DBA or other technical administrator to safely update database tables directly. With granular role-based access controls and simple tools, even non-technical users can safely administer master data without the risk of corrupting the database.

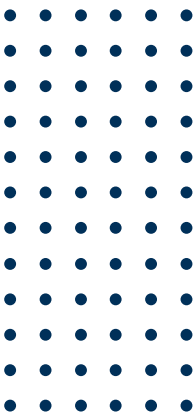
CONVX Solutions for Restaurants

CONVX RHO Data Pipeline Solutions

Forget endless integration development. Traditional ETL/ELT solutions leave you coding integrations for countless data sources and target systems. Even with extensive pre-built connectors, the gaps are endless.

CONVX breaks you free from the integration burden. Our managed, cloud-based data pipeline transforms restaurant data without the open-source complexity or custom coding nightmares. We leverage enterprise-grade, open-source tools, but eliminate the skills gap and support headaches that come with it.

Don't waste resources building custom integrations for every restaurant tool. CONVX delivers pre-developed, industry-specific pipelines for popular POS, loyalty, customer experience, CRM, databases, files, finance, and other applications. We eliminate your friction with expert custom development for your unique needs, all seamlessly integrated into your data ecosystem.



CONVX RHO Data Warehouse Solutions

Unleash restaurant data insights with CONVX. We manage and optimize:

- Cloud-native Snowflake data warehouses: Leverage industry-leading scalability, elasticity, and performance for lightning-fast analysis and maximized ROI.
- Other cloud data warehouse solutions: We support Microsoft, AWS, and Google managed data services.
- On-premise stalwarts: We get it; not everything is in the cloud. We help you get the most from traditional Microsoft SQL, Postgres, and MySQL databases, integrating them seamlessly with your data ecosystem.

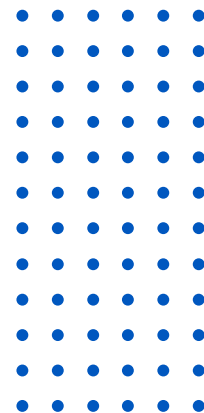
Snowflake is the database technology built into the CONVX RHO Stack. Snowflake's cloud architecture and industry-leading capabilities have proven to be the best solution for the restaurant industry, producing significantly better ROIs for our customers and much higher development velocity than alternatives.

CONVX RHO Data Apps

Publish those KPI dashboards, so your operators and stores know exactly how they are performing right now. Replace tedious daily or monthly data workflows with a self-service console. Keep everyone in the loop with team messaging for critical alerts, reminders, and team messages. These are all common restaurant use cases of a CONVX Data App. Let us build your own app, quickly and cost effectively with the CONVX Data App framework.

6%

Shorter POS Order Time



Appendix: Under CONVX's Hood

This deeper dive into the workings of CONVX's restaurant solutions illustrates the architecture and the common software development principles and standards it is based on. For specific questions not addressed here– or better yet, to get a demo– contact us: info@convx.com.

I. Data Warehouse Design

Our data warehouse design accommodates structured and unstructured data together in a lakehouse to support current and future use cases.

Star Schema Data Marts

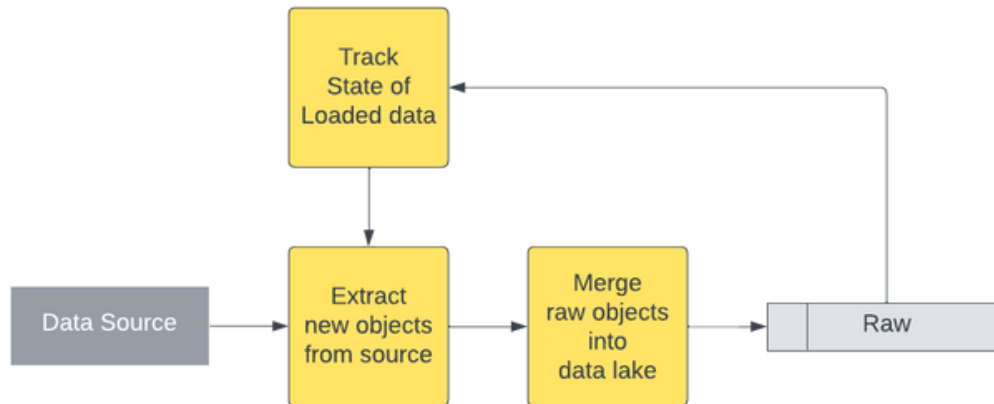
Data marts present data to client-facing applications, such as business intelligence (BI) and analytics tools, in an optimized structure suitable for these use cases.

Decades of successful use, analyst-friendly design, BI tool compatibility, and simplified insight generation make the star schema CONVX's data mart standard. This well-established pattern leverages dimensional modeling principles to structure data related to specific analytics areas, such as sales, loyalty engagement, or inventory management.



II. Ingest Pipeline Design

The ingest pipeline is the first step to fetch data into the data warehouse. CONVX uses an ELT (Extract, Load, Transform) design pattern for our data warehouse architecture.

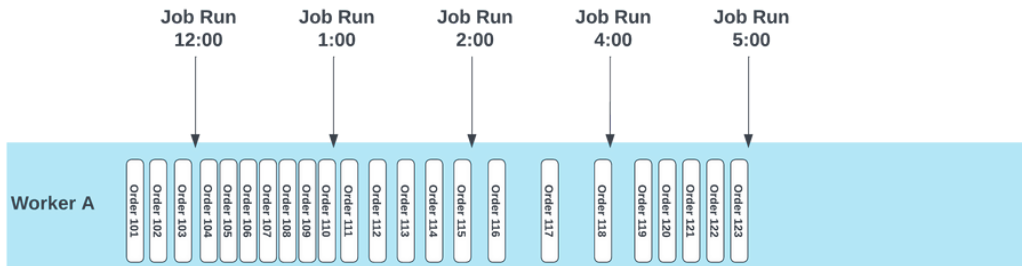


This is not to be confused with an alternate pipeline design called ETL (Extract, Transform, Load) pipeline design. CONVX recommends ELT over ETL in data warehouse projects for several reasons:

1. Data preservation: ELT ingests all the original data, to easily return to it later if needed.
2. Transformation efficiency: SQL transformations performed within the data warehouse are much more efficient than external transformations outside the data warehouse.
3. Enrichment: Since the data warehouse is the hub for all data, ELT allows access to other data within the data warehouse in a way that's harder and less efficient in an ETL pipeline.
4. Simplicity: ELT pipelines are easier to build and require less maintenance than ETL pipelines. They are not as sensitive to schema changes.

Sequential Data Loader

The simplest data loaders start with the earliest records from the source data and load them in sequence. Once the first data records are successfully loaded, the loader fetches the next records in sequence.



Many solutions use a simplistic sequential loading method that doesn't handle failures well. If it encounters an error in a particular source data object or timeframe, it can't move on. It will continually polling for the same records until it succeeds.

Interval Data Loader

CONVX improves the sequential loader by polling for specific intervals instead of just the latest records.

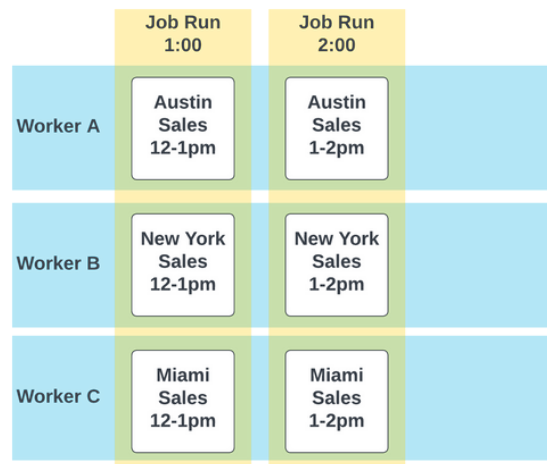


This has several advantages:

- If the poll in a particular interval fails, it can be skipped and polled in the next period. A single bad record can cause a specific interval to fail but will not affect others.
- This provides the ability to poll separately for individual restaurants. A failure in pulling data for one restaurant will not impede getting data from others.
- If data is required from a previous period, the interval loader identifies the specific intervals that have a failure. It is easier to re-run jobs for specific periods than with a sequential loader.

Parallel Data Loaders

Sequentially loading data for many restaurant locations is time-consuming: 100 restaurants takes 100 times as long as one restaurant. CONVX uses parallel processing to load multiple restaurants simultaneously. Compute usage is roughly the same, but total load time is reduced by a factor of the number of workers.



Scalable Compute

CONVX's on-demand scaling adapts compute costs to workload: max workers during peaks, minimal at night, serverless for short tasks.

Change Data Capture (CDC)

Many databases use Change Data Capture (CDC) for incremental loading, instead of APIs. CDC technologies use a few mechanisms to record changes to a database, so those changes can be replayed in another environment to replicate the changes. CONVX selects the appropriate mechanism for the database and use case to minimize impact on database performance while supporting timely and reliable replication.

Bulk Data Loading

To create a data warehouse, a large volume of historical records often need to be loaded. API extraction is not efficient here. It's preferable to bulk export the entire history to files and then securely transfer those to the data warehouse. The initial bulk load of raw data is often manual because automating a one-time process isn't necessary. This approach does not offer real-time updates like an API, but is more efficient for resource usage.

CONVX supports bulk loads, both initial and subsequent, from cloud object storage or a customer SFTP server. Object storage is preferred for simplicity. It is easy to set up and use without compromising security.

Pipeline Orchestration

The hardest part of developing ingest pipelines is the orchestration required to automate loading new or changed records while gracefully handling failures, ensuring consistency in any circumstance, and providing administrators with the necessary visibility, auditing, and manual controls when needed.

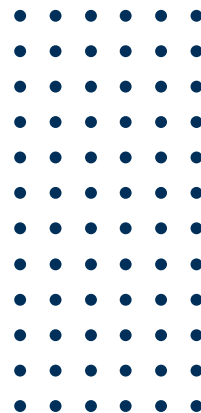
State Tracking

CONVX ingest pipelines safeguard your data integrity by meticulously tracking the state of loaded data. This ensures no missing or skipped records, maintaining data consistency and accuracy. For straightforward sequential loading, simple timestamp-based tracking efficiently manages incremental updates, even handling duplicate timestamps gracefully through unique ID tracking.

For more complex interval-based loading patterns, CONVX employs high-throughput state management techniques that go beyond just timestamps, ensuring precision for each time interval. These capabilities translate into key benefits: zero data loss, smooth incremental updates, and the ability to accommodate a wide range of loading requirements with adaptable strategies.

State Recovery

CONVX pipelines are configured to assume failures can happen in the middle of a job. Data pipelines are especially susceptible due to their myriad of integration sources and targets that we don't control. Failures occur in the source system, target system, and the pipeline itself. If an ingest job fails for a particular interval, CONVX can re-run the load in question with state recovery.



Streaming Pipelines

Many pipelines bringing new records into the data warehouse only need to run hourly or daily, but can run as frequently as every few minutes, or be triggered by events. Incremental pipelines are optimized for frequent transfers of small batches of records. These pipelines may run in seconds or minutes. CONVX can decrease the window size and set the frequency to match the business needs for prompt data.

Incremental Load Pipelines

Many pipelines bringing new records into the data warehouse only need to run hourly or daily but can run as frequently as every few minutes, or be triggered by events. Incremental pipelines are optimized for frequent transfers of small batches of records. These pipelines may run in seconds or minutes. CONVX can decrease the window size and set the frequency to match the business needs for prompt data.

Efficient Updates

Many data sources generate record updates. A customer may change their address. A POS transaction may be refunded. These result in updates to previously-loaded records. Updates are expensive database operations. Inserts are much more frequent than updates in sales transactions, so we optimize our DML (data modification language) to perform the most efficient DML technique for the pipeline by one of the following techniques:

1. Creating a new record instead of changing/updating an existing record. It is more efficient to add an offsetting transaction, such as a refund, instead of editing a record.
2. Differentiating updates from new records at the initial extract, then performing the optimal transaction without requiring a table scan for a matching record. Tagging each object with the right DML transaction can efficiently propagate those changes downstream.
3. Optimizing updates (or upsert) transactions as efficiently as possible, via table keys, partitions, and segmenting as delete+insert transactions.

Reporting and Auditing

CONVX pipelines include data lineage tracking from the data source to the data warehouse, along with significant controls and transformations. This provides prompt resolution of issues in case of a failure, and satisfies an organization's specific needs for documentation of data lineage and actions performed.

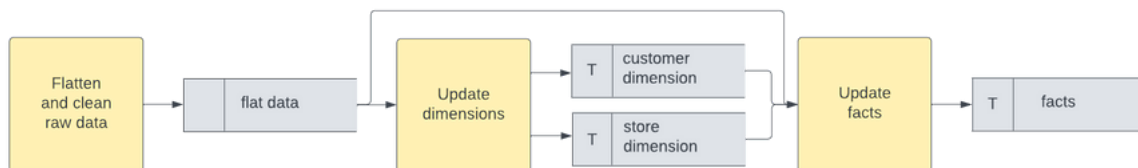
Ephemeral and Elastic Compute

The CONVX architecture uses both ephemeral FaaS and dedicated always-on compute resources to fit the job at hand. We also support elastic resources that scale out with additional compute instances to meet demand. They scale back down automatically when demand subsides.

A key to the ephemeral strategy is separating state from the server. Ingest jobs store no artifacts on the server. All data and state are stored externally, other than temporary files required for running the job. This architecture allows us to create and destroy resources as needed. Since most jobs only run for a minute or two, they are not required 95% of the time.

III. Data Transformation Design

In CONVX's ELT pipeline design, cleaning and transformation happen after the raw data lands in the data warehouse. Our staging process is dramatically streamlined to transform raw, unstructured objects into facts and conformed dimensions, depicted below in a simplified process:



Change Data Capture (CDC)

The handoff from data loader to data transformation leverages proprietary Snowflake table streams and pipelines to perform CDC efficiently and automatically. CDC captures any new or updated records in a raw table. CONVX monitors the stream for changes and then reads the new or updated records from the stream.

Data Flattening

CONVX's transformation job flattens unstructured data (usually XML or JSON) into a temporary table. This table includes all columns required for downstream operations. The table is visible only within the job context. After completing the job, the system deletes the temporary table to free up resources.

IV. Data Enrichment and External Data Sources

CONVX's architecture supports a growing list of external data sources to enrich and complement a customer's data:

- Name enrichment helps customers complete their guests' profiles. CONVX architecture can query external sources to complete the customer profile and correctly segment the customer.
- IP enrichment identifies an internet device by analyzing its IP address, which can geolocate internet visitors or block or ignore suspicious activity from bots or untrusted parties.
- Event data, such as holiday schedules and local sporting events, contributes to forecasting, shift staffing, and assessing sales performance.
- Weather data also provides important context to predict sales performance and customer activity for a specific location.

V. Analytics and AI

CONVX's data mart architecture features standardized dimensions for efficient rollup and slicing across restaurant, product, and customer dimensions. Its integrated AI engine powers training and predictive models for sales forecasting, shift planning, and inventory management, optimizing restaurant operations and profitability.

86 your data headaches. Boost your restaurants' profitability.

Let CONVX take the wheel. [Schedule a demo today](#) and see how our unified, optimized data solutions can help you serve up greater revenue and less waste. Call 720-605-1341 or visit convx.com.

Contact CONVX

