PROOF

Future States

Faster Custodial Data Processing

Challenge: The T+1 Problem

One of the biggest problems with trying to convert raw custodial data into actionable intelligence for advisors is time: specifically, the T+1 Problem. At the close of every market day, custodians must undertake the process of data reconciliation which takes no less than 2 hours but can often take up to 8.

The result of this process is a data archive reaching hundreds of thousands of records that must be transferred over public networks to reach data consumers. Only then can the data be processed and converted into useful information such as analysis and alerts—a variable, time-consuming process in its own right.

The result is, advisors often don't have what they need to make trading decisions until well into the market day.

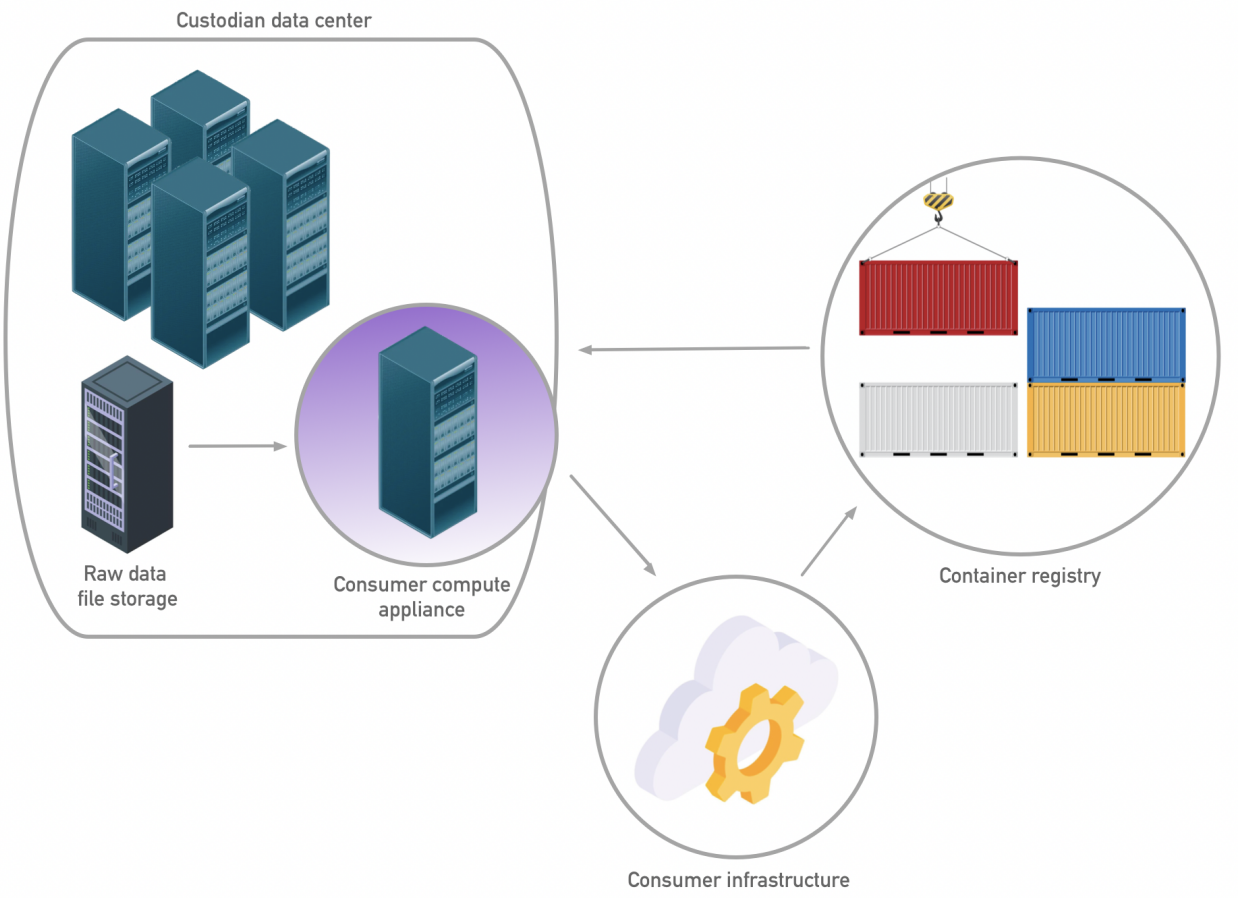
Opportunity: Custodian-As-Cloud-Provider

The first step to a solution is to reframe the problem: instead of forcing consumers to move the data to themselves, custodians can help consumers & provide a strategic value-add by allowing consumers to move their processing workloads to the data.

Using modern container orchestration technology, and our hardware, data engineering, and MLOps expertise, Proof Partners is designing a compute appliance that custodians can plug directly into their data centers to securely and efficiently run consumer-controlled workloads right next to the data they need to process.

Design

Once installed and configured, the appliance begins monitoring consumer container registries for new images. Once found, new images are deployed into a secure network environment, fully isolated from the rest of the custodian's data center and other consumers.



The appliance operates in a locked-down configuration, with access limited to data storage. Rather than simply expose raw files, the appliance packages up the data into industry-standard DataFrame formats, ready to be consumed by data analytics & machine learning jobs. Each secure network environment exposes the set of DataFrames for their respective consumer.

These consumer processing jobs can then stream alerts, insights, and other high-priority data to their own applications in real-time and in any format over any protocol they choose—while deferring bulk data for lower-priority processing as necessary (eg. ledger or compliance).

Details

With namespacing, configurable resource limits, Kubernetes offers an ideal enterprise-ready foundation for companies looking to offer infrastructure-as-a-service to their customers.

Proof aims to offer a turnkey solution, with a streamlined process for onboarding new customers, which is a 3-part process:

Some images from vecteezy.com.

- *Namespace & data volume mount*: Provision the new customer namespace, and map the network drive containing the customer's data into it. This is a small lift for operations teams to ensure network access to consumer data from the appliance, but the rest is able to be streamlined by management tooling built into the product.
- *Container registry*: Consumers deploy software as containers delivered to a registry. The appliance monitors the registry for new container versions and deploys them in 'quiet periods' between data processing jobs.
- *Consumer user identity*: Consumer-side user identities are registered with an isolated management dashboard built into the appliance for monitoring and telemetry, and to perform administrative tasks.

The Next Step: Solving for Reconciliation with Machine Learning

Moving consumer workloads local to custodial data is only the first step. Version 2 of the appliance would accept raw trading data and leverage specially-trained machine learning models to automatically validate, match, and enrich transactions—flagging only true exceptions for human-in-the-loop review.