

Built a **Robust Foundational Data & Reporting Platform** for a Leading **OTT Solutions Provider**

Designed and built a foundational data model (FDM) that is scalable to many tenants and can provide analytics common to an online video system with advertising, subscriptions and eCommerce innovations.

— The Client

The client provides a service suite that encompasses a revolutionary end-to-end digital media management and distribution solution. Global brands rely on their media-as-a-service platform to distribute content to thousands of destinations around the world, enabling their clients to keep pace with evolving consumer demand and emerging multi-channel distribution models.

— Executive Summary

The transformation of entertainment mediums from analog such as cable and televisions to digital and OTT platforms has given rise to the use of powerful tools such as big data & data analytics. By tracking the aggregated user activity and behaviour, Veltris was able to draw very detailed inferences about content consumption with several dimensions such as time, geography, user categories, hardware usage, user satisfaction etc. This in turn, gave our client a holistic understanding of what drives content consumption on their platform.



Analytics Dashboard

While privacy of user data is of utmost importance, aggregated data can tell various stories about content consumption and behaviours to give extremely useful insights. With the advent of cloud architecture, big data and digital OTT platforms, we were able to aggregate user activity and stream it into cloud-based data pipelines and crunch them for such insights.

— Challenge

The client was looking for insights with the kind of data granularity that required drilled down reporting. They needed a big data pipeline for ingesting real-time data collected from their viewers globally. We build a robust platform that generates quantitative analytics reports based on customer usage patterns to provide the client with insights they need to make better, more informed decisions.

— Approach

The player and ad events are processed using spark structured streaming in Databricks. Structured Streaming is a scalable and fault-tolerant stream processing engine. One can express the streaming computation the same way one would express a batch computation on static data. The Spark SQL engine takes care of running it incrementally and continuously and updating the result as streaming data arrives continuously.

Technical Deliverables



1

DASHBOARDS AND REPORTS

Configure, reskin with V3nue L&F, integrate with SSO and build initial dashboard and video, user and commerce reports. Estimated 30-40 reports in total, 10-12 dashboard reports.



2

MULTI-TENANT DATA STORE

Configure and deploy a multi-tenant instance of Azure SQL using Elastic Pools. Tenants share the same V3nue core schema but have separate schemas each.



3

MULTI-TENANT DATA LAKE

Configure and build out a partitioned instance of Azure Data Lake to support raw data ingested for different tenants across the platform.



4

DATA INGEST ORCHESTRATION

Configure, script and deploy a multi-tenant-aware instance of Azure Data Factory to handle ingestion of data from Tealium, Heartbeat, User and Commerce databases and GA.



5

DATA TRANSFORMATION

Configure and build Azure Data Bricks for cleaning and processing data in the lake and transferring new views of that data into the warehouse.



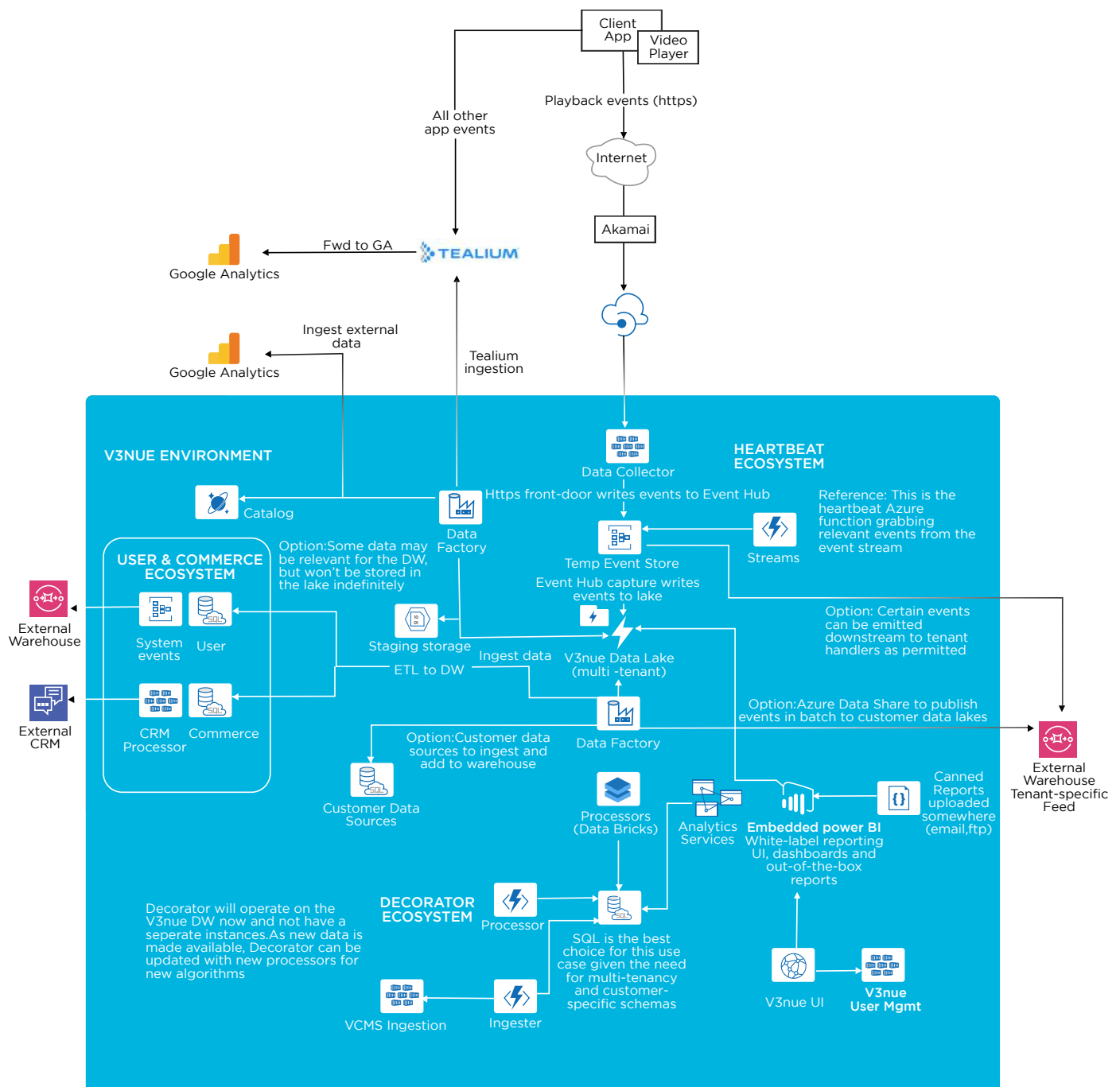
6

EXTERNAL DATA SYNC/ACCESS

Configure and build external data access mechanisms, including the ability for a customer to access their warehouse securely from an external tool like Tableau.

— Solution

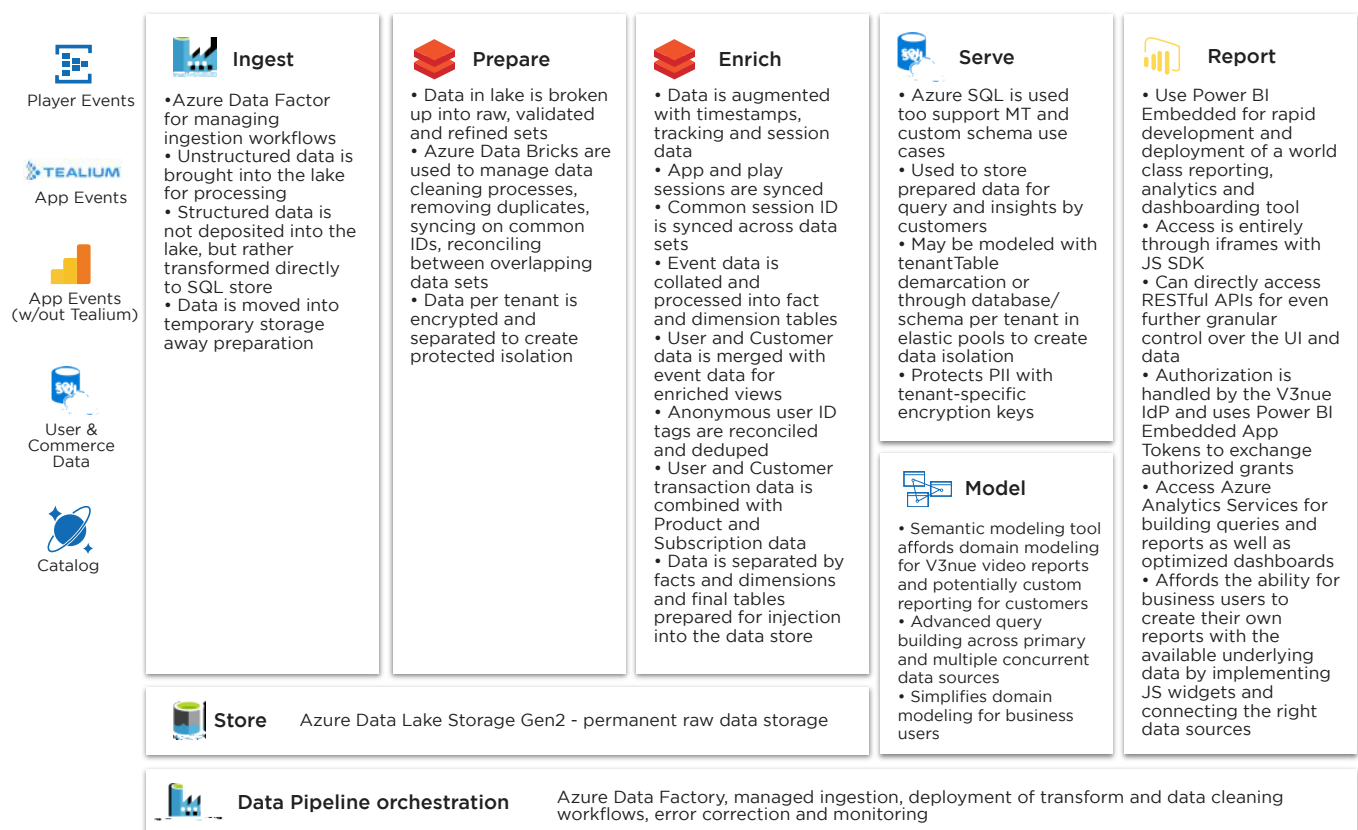
Veltris built a multi-tenant solution on Azure full stack with state-of-the-art, end-to-end pipelines that can analyze more than 250 million events each day to provide analytics to help the client with insights to increase audience engagement and generate new revenue streams.



Multi-Tenant Architecture for a Robust Foundational Data & Reporting Platform

When users are active on the OTT platform, we use analytics that stream data such as browsing, content selection, play/pause, stop, forward, rewind, mute, volume/subtitles/ audio language change, ads played etc. This data is captured using pipelines and stored in raw formats to be processed and reported at later stages. This data when aggregated and filtered based on reporting requirements can give us insights such as;

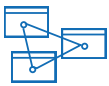
- Which was the highest consumed content in any given time period or location?
- What search terms were most used in any given location?
- What kinds of devices are mostly used to watch content? What is the percentage distribution of wired vs mobile devices?
- What is the average bandwidth of users consuming content?
- What is the consumption and monetization rates per tile or per content banner?
- What kind of content or genre is being consumed with respect to specified age and income groups?
- Where is the content most prominently being skipped or paused?
- What are the different kinds of errors and how are they spread over time and location?
- How many ads are being displayed per session at any given time and location?



Components



Azure Data Factory - used as a managed data orchestration and integration platform, Data Factory provides scripted and code-free ETL and ELT pipelines for transferring data into or out of a data lake and warehouse



Azure Analytics Services - used for domain modeling and query generation for data in the data warehouse. Aides in the creation of dedicated view tables for rapid and repeated dashboard data lookup



Azure Data Lake - used for the long-term storage of both structured and unstructured data, the data lake is a central collection point for a tenant's data, can be accessed directly for configured queries or machine learning and is used to feed the data warehouse



Azure Data Bricks- used for cleaning, transforming and preparing data for injection into the database/data warehouse, Data Bricks is a managed and extended instance of Apache Spark, a big data processing and query engine.



Azure SQL - used as the central clearing house for report-ready data, the data store will be where a fully realized and prepared schema will exist for V3nue tenants and could be extended on a per-tenant basis with a tenant's own data. The solution entails a single Azure SQL resource but with multiple different schemas and separate security per which affords maximum flexibility and data isolation.



Azure Power BI Embedded - used as a managed reporting and analytics solution that is hosted and white-label, allowing V3nue to acquire enterprise grade analytics and data visualization while retaining brand identity through a skinned application.

We are delighted to have helped the client create a centralized, analytics-ready repository for their big data and look forward to helping them meet their strategic goals using our cloud capabilities.

Better understanding of the users

to help customer retention and reduce churn

Improved the lifetime value of each user

by nudging them up the revenue chain with ads and subscriptions

Enhanced user experience

by providing deeper insights for recommendation cues

— How the Product Helped

The reporting platform helps the client discover and better understand consumer behavior and their content consumption patterns. Our data analytics and insights help the client deliver a richer, deeper, and a more customized user experience.

— Tech Stack



Ingest , ELT, ETL,
Workflow & Orchestration
Azure Data Factory



Serverless Functions
Azure Functions



Azure Storage
Azure Data Lake



ETL/ELT
Azure Data Bricks



Workflow & Integration
Logic Apps



Integration Message
Broker
Service Bus



Reporting
**Azure Power BI
Embedded**



Database
Azure SQL