

SOLVING DATA CONSUMPTION CHALLENGES WITH ENZO UNIFIED

Abstract: This white paper explains how companies can compress development and data integration project timelines, improve data quality, and increase operational efficiency by solving some of the most complex real-time data consumption challenges with the Enzo Unified platform.

Audience: Developers, Managers and Technology Leaders

Published on August 3rd 2015

Copyright © 2015

Enzo™ is a registered trademark of Blue Syntax Consulting, LLC

CONTENTS

03 INTRODUCTION

03 DATA CHALLENGES

03 REAL-TIME DATA AND SERVICE ACCESS

04 ENTERPRISE DATA VIEWS

04 EDGE CACHING

05 AUDIT HETEROGENEOUS DATA ACCESS

05 DATA ACCESS FROM DATABASES AND MOBILE DEVICES

05 THE ENZO UNIFIED SOLUTION

07 WHERE ENZO UNIFIED FITS

09 SOLUTIONS

09 REAL-TIME DATA AND SERVICE ACCESS

09 ENTERPRISE DATA VIEWS

10 EDGE CACHING

10 AUDIT HETEROGENEOUS DATA ACCESS

11 DATA ACCESS FROM DATABASES AND MOBILE DEVICES

11 CONCLUSION

11 FOR MORE INFORMATION

INTRODUCTION

Today's enterprises depend on the ability to manage many sources of data, both from complex internal systems, and internet data sources. The need to combine information from multiple source systems is vital for many scenarios, including data acquisition for customer sentiment monitoring, data integration, enterprise data virtualization and IoT device management, just to name a few. In addition, the need to acquire data is even more acute when the information needs to be as real-time as possible.

This white paper introduces you to the major real-time data consumption challenges organizations face today, and how Enzo Unified is uniquely positioned to solve these challenges.

DATA CHALLENGES

Many data integration projects are still being solved by implementing batch operations, and by writing complex jobs that run on a schedule. However most businesses need to achieve greater flexibility by accessing, changing, and interacting with data and services in real-time. And with the increasingly popularity of devices, the Internet of Things (IoT) wave places more pressure on the need to interact with systems in real-time. The very nature of real-time data integration creates certain key challenges that most organizations are not fully equipped to absorb; let's review some of them.

REAL-TIME DATA AND SERVICE ACCESS

Many companies rely on the ability to access Internet data and services, such as Twitter feeds, current weather conditions, geocoding services, address validation services, Internet security services, or even government data sources. Some organizations depend on even more complex sources of data, such as accounting packages, legacy systems, Enterprise Resource Planning (ERP) products and Human Resources packages just to name a few. While many of these data sources and services are available with Application Programming Interfaces (APIs) and/or Software Development Kits (SDKs) for developers to consume, some companies still can't consume them easily for the following reasons:

- Limited development staff with the experience needed to consume complex APIs
- Limited time to build the necessary software logic to consume the APIs

While APIs are usually viewed as the answer to solving data integration problems, they usually tend to increase the complexity of a company's data landscape. More APIs means more software development, which usually translates into slower projects, more complex solutions, and a more skilled/expensive workforce. This creates an interesting paradox: APIs enable data exchange, but typically significantly increase complexity. The increase in complexity can be traced to many factors, including a lack of common authentication

across APIs, disparate representations of data, disparate and sometimes sparse documentation, and as a result a heavy learning curve that varies depending on the vendor providing the API.

“APIs enable data exchange, but can significantly increase complexity.”

As a result, many companies either wait until their key development resources become available, hire consultants to perform the work, or in some cases place these projects on hold.

ENTERPRISE DATA VIEWS

For some organizations, providing a single view of their systems can be critical for customer support, sales, or even IT operations. For example, some organizations use best-of-bread Human Resources platforms, usually in a mixed environment of internal systems and hosted platforms. The ability to create enterprise views that represent a single customer, or a single employee, can be central to an organization’s ability to integrate with ancillary systems and business partners. In addition to creating a uniform view of corporate data, some companies also need simpler access to information from multiple systems that provide disparate data formats, such as web services, databases and legacy systems.

Organizations facing this challenge typically include those that have:

- Multiple systems storing related information
- Heterogeneous systems and need to normalize their information
- Multiple departments working mostly in isolation
- Two or more operating system environments (Windows, Linux...)

The need to build enterprise data views is driven by the desire to simplify the data landscape, so that enterprise data can be more easily consumed by other systems such as big data environments, data warehouses and/or data analysts, and applications. Enterprise data views also increase the quality of data in the organization because everyone system accesses the same information. .

EDGE CACHING

In some cases, the information that organizations need to access can be stored across geographic regions, or is slow to query during peak season, or is temporarily unavailable due to network disruptions or other reasons. At times, organizations need to reduce access to their critical systems because the underlying data stores are under heavy usage during peak times, and processing resources should be reserved to the systems that need it most. Generally speaking, edge caching is needed when:

- Systems are under stress and cannot sustain additional load
- Data sources are located in various geographical locations
- Data sources are not always available, but data access is needed regardless

For some organizations, edge caching can provide relief to underlying processing systems, and make data available very quickly even when the source systems are far or unavailable.

AUDIT HETEROGENEOUS DATA ACCESS

While most database systems provide good data access auditing, it can be very difficult to provide data governance for heterogeneous data sources, such as flat files, web services, or even Internet Services. Organizations that need to monitor data access for heterogeneous systems include:

- Highly regulated industries such as Healthcare and Insurance
- Hosted environments that need to bill data access based on consumption

DATA ACCESS FROM DATABASES AND MOBILE DEVICES

Some organizations need to quickly expose their internal data in a way that mobile devices can consume easily, usually using a Representational State Transfer (REST) interface and JSON data formats. However, many organizations also need the ability to present data in rows and columns, like a real database server, to enable data analysis, reporting, data transfer and more. This creates yet another challenge for most companies: the ability to store, and query data in the most efficient way, regardless of the client technology (mobile data consumption or data-intensive applications).

Companies likely to face this challenge include:

- Organizations with a distributed field workforce
- Organizations with many remote locations
- Organizations exposing data with business partners and clients

This duality can create a challenge for organizations trying to decide where the source data should be stored. Should the data be stored in a raw database, because free-form querying is more important, or as JSON documents because mobile devices will be interacting with this data? Considering that most database servers cannot service JSON documents natively, the choice is rarely simple. In many cases, companies need to address both needs.

THE ENZO UNIFIED SOLUTION

Realizing the challenges corporations face related to consuming unstructured data and accessing data services in real-time, we created Enzo Unified to address them. Enzo

Unified is designed to simplify data access unlike any other solution available today.

Enzo Unified is a server platform that transparently makes data access a breeze to any developer, business analyst, application or service. With Enzo Unified, your organization can:

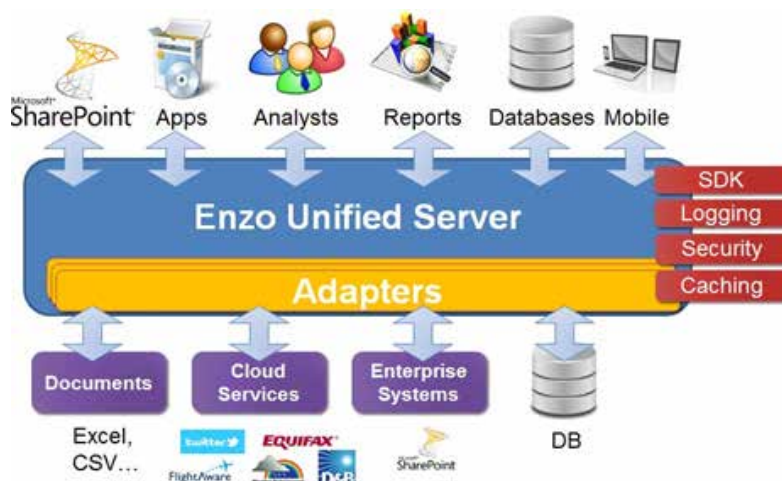
- Remove the learning curve associated with APIs and SDKs
- Create Enterprise Views of your most important systems
- Build edge caches of your data
- Audit access to heterogeneous data sets
- Expose any system as a REST and SQL service for modern consumption

At its core, Enzo Unified removes the challenges introduced by APIs while keeping the benefits; indeed, the ability to query any data, regardless of its origin, using a simple query, eliminates most of the symptoms of API development, such as a lower learning curve, simplified authentication, common format across data sources and built-in documentation. The ability to remove the challenges of APIs translates into significant benefits for many organizations, including:

- Faster development and implementation times due to reduced learning curve
- Greater data quality due to pre-tested components and common enterprise views
- Increased operational efficiency as more junior staff can accomplish more complex tasks

“Enzo Unified removes the challenges introduced by APIs while keeping the benefits.”

As shown in the picture below, Enzo Unified is a platform that abstracts access to any other data source or service, so that they can be accessed easily from any client (users, applications, reports, mobile devices...). Because Enzo Unified is a server platform, it can offer the key capabilities discussed previously, such as auditing, edge caching, enterprise views and more.



In addition, Enzo Unified also offers a Software Development Kit (SDK) that allows organizations with development staff to build their own extensions to expose desired data sets, or services, and automatically benefit from all the capabilities of Enzo Unified.

WHERE ENZO UNIFIED FITS

When it comes to data integration and data movement, companies have multiple options. Understanding how these options compare is important to better understand where Enzo Unified fits in this landscape. Let's review the main data integration choices available: ODBC Drivers, Extract Transform Load (ETL) tools, and Integration Platforms.

In order to understand how each integration approach deals with the major data challenges outlined previous, let's look at the following aspects:

REAL-TIME AND UNIFORM DATA ABSTRACTION

Represents ability to abstract the complexities of underlying data sources with access in real-time, without first storing elsewhere.

DATABASE AND MOBILE ACCESS

Represents the ability to service data to databases and mobile devices equally well, merging the capabilities of database servers and web servers.

EDGE CACHING

Represents the ability to selectively cache certain data sets so that access to data is faster and more resilient to network or service failures.

CENTRAL DATA GOVERNANCE

Represents the ability to centrally audit data access so that compliance groups can determine who accessed what, and when.

Generally speaking, ODBC drivers can deliver good abstraction to underlying data sources, but fail to provide additional capabilities, such as mobile data access, uniform enterprise views, caching and governance. That's due to the fact that ODBC drivers do not offer a centralized solution; ODBC drivers must be deployed on the machines where the client software runs. As a result, ODBC drivers have limited capabilities above and beyond data abstraction.

ETL tools are good at moving data in batch, through jobs, but are usually not meant to be used as real-time data access technologies. Some companies use ETL to move large amounts of data, with very high frequency (such as every 5 minutes or less), in order to increase the freshness of the data. However, high frequency jobs still do not provide real-time access to data, and are usually purpose built: changing ETL jobs over time to meet evolving business needs can be difficult. As a result, ETL tools solve different problems and

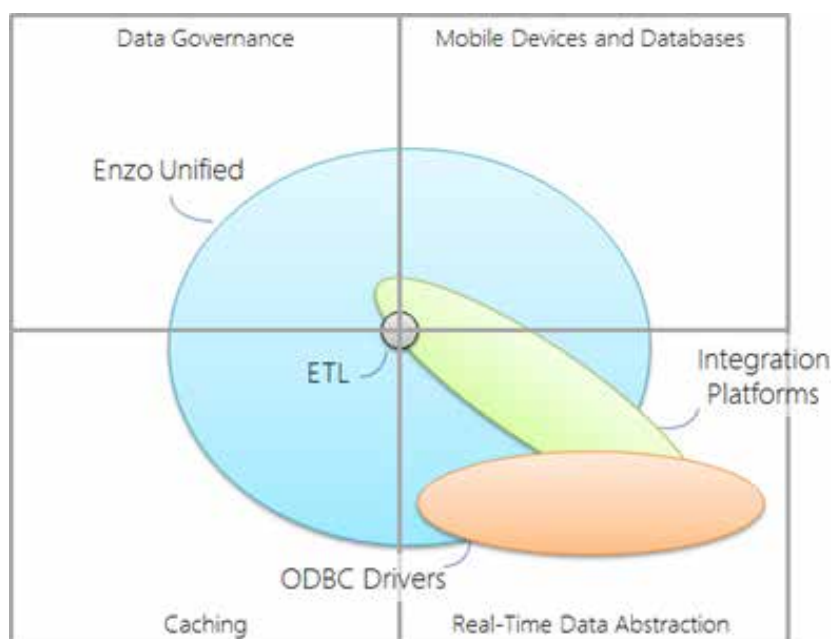
cannot answer the challenges defined above effectively.

Integration tools are usually heavy platforms and provide good abstractions for developers, including the ability to access and abstract data in real-time. However, Integration platforms can be hard to configure and lack certain capabilities, such as data governance and data caching, and are also single purpose-built. In addition, they usually work as background services to connect systems together, but are not available as pure data services that database servers understand, which means that Integration platforms cannot be easily queried. As a result, Integration platforms answer the need for data abstraction, but require heavy plumbing to achieve this result, and do not offer the necessary data servicing capabilities that lead to data governance, caching and the necessary data access flexibility of a data store.

Enzo Unified is designed to bridge the gap left by the currently technology landscape. Enzo Unified offers data abstraction, configurable enterprise views to hide the complexities and origin of underlying data sources, edge caching for remote data sources, and data governance, in real-time.

Because of its unique design, Enzo Unified can easily be used to enhance existing data integration strategies. Enzo Unified can be used, for example, to encapsulate ODBC drivers and immediately gain simpler authentication, caching, and enterprise views. Enzo Unified can also be used with existing integration platforms to expose key integration processes at the database level.

“Because of its unique design, Enzo Unified can easily be used to enhance existing data integration strategies”



SOLUTIONS

Let’s review how Enzo Unified can provide a simple solution for organizations that need to address the data challenges outlined previously.

REAL-TIME DATA AND SERVICE ACCESS

Accessing data in real-time, or interacting with services (internally hosted, or from a third party), usually requires a learning curve because of the complexity and the effort needed to consume APIs. Accessing data from Twitter requires learning the Twitter API, and posting a new entry in a SharePoint list requires finding the right API and sifting through the vast SharePoint documentation. As the number of systems grows, so does the complexity and the expertise necessary to build and maintain such systems.

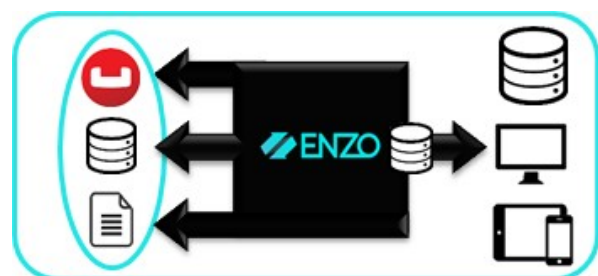
Enzo Unified removes the learning curve associated with APIs, and standardizes the data sources by presenting information in a uniform way regardless of the source of the data. In addition, clients use the same authentication mechanism regardless of the data source, further simplifying data integration. The ability to communicate with virtually any data source in real-time without having to learn complex APIS provides a significant boost in productivity and overall quality.



ENTERPRISE DATA VIEWS

The ability to simplify data access to underlying data sources, regardless of their origin or format, is a pre-requisite to building enterprise data views. Because data sources are exposed in a common format through Enzo Unified, they can easily be mixed as needed to present a single view of the truth. An enterprise view allows organizations to create a virtual data source that merges information from multiple systems, such as an ERP platform, a NoSql database and a text file for example.

With Enzo Unified, organizations can create enterprise data views on top of existing data sources, and make them available to internal



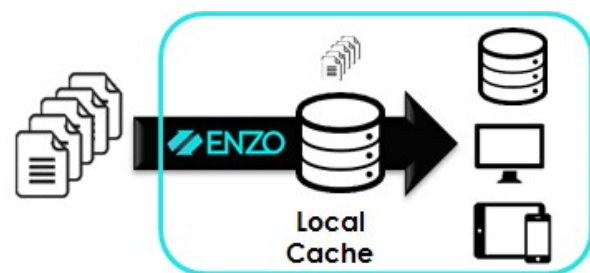
applications, business users, and even business partners if desired. Some of the major advantages of building Enterprise Data

Views include:

- Abstract underlying data source complexities for simpler data access
- Increase security by building a proxy authentication layer
- Create a single view of the truth for records stored in multiple source systems
- Migrate, merge, or retire underlying data sources transparently

EDGE CACHING

In some cases, data sources are slow to access, or temporarily unavailable. For larger organizations, some key systems are under heavy load, and administrators try to prevent frequent access to minimize performance degradation. Edge caching provides a buffer mechanism for systems that may



not be best suited for real-time access, or when a form of local indexing makes sense. In an edge cache configuration, Enzo Unified can be configured to refresh its cache on a specific frequency, and ensure that data is always available even if the actual data source is not. Some of the key benefits of edge caching include:

- Shield certain systems from potential performance degradation
- Improve performance of data access
- Eliminate temporary connectivity issues from remote data sources

AUDIT HETEROGENEOUS DATA ACCESS

Most organizations face challenges with controlling access to databases and heterogeneous data sources, and as a result have difficulties auditing and governing data access in general. The ability to centralize data access provides a unique opportunity to redefine access policies and auditing actual data usage, regardless of the type of data source being accessed (such as servers, files, services, and devices to name a few).

Enzo Unified solves this challenge by providing a proxy authentication to the underlying data sources, both enabling, and securing actual system



credentials. This allows administrators to also add granular access control to underlying data sources, and audit user and application access.

DATA ACCESS FROM DATABASES AND MOBILE DEVICES

The ability to access data both from database-friendly systems and mobile systems is becoming more important. While organizations typically favor one kind of consumer over the other, most companies still need them both: access data from database-friendly systems and mobile devices is necessary.

The challenge stems from the fact that mobile devices consume data using a very different format (JSON through REST) than traditional database clients (raw data through SQL). Enzo Unified closes this gap by automatically exposing the underlying resources, regardless of their native storage format, as both REST and SQL data sources.

This capability enables mobile developers, business analysts, report writers, integration systems and business partners to choose the language of their choice and seamlessly access information.



CONCLUSION

The need to move data and integrate with services in real-time, both internal and external, has never been greater. While APIs have helped organizations integrate their data ecosystem by building single-purpose integration solutions, the sheer number and complexity of internet systems, API specifications and SDKs is creating a learning curve that most organizations have difficulties solving. The data challenges that result include real-time data access, a common view of enterprise data, edge caching, auditing and governance of heterogeneous data, and mobile vs. database access.

Enzo Unified solves these challenges by offering a unique server platform that transforms the way organizations build systems to achieve greater agility, speed to market, and stronger security. Enzo Unified promotes a unified, yet flexible data consumption model that applications, business users and business partners can leverage easily without the learning curve of APIs, leading to faster development and implementation times, greater data quality, and increased operational efficiencies.

FOR MORE INFORMATION

Enzo Unified is a data-centric platform that focuses on answering today's data integration challenges by simplifying the data landscape of organizations, and by servicing data in a common format regardless of its origin. To learn more on how Enzo Unified can transform your business, contact us at info@enzounified.com.