Data Unification:
The Bridge to Better Data Management

THE CURRENT LANDSCAPE

The critical imperative of businesses today--across all sizes and markets--is to be data-driven. This has led companies to purchase an ever-growing number of software applications, both on-premise and cloud-based, in order to be more responsive to customer needs, provide deeper insight into their own business and manage their operations more smoothly and efficiently. However, this is a double-edged sword because while each new application may make sense in its own right, whether it's a marketing automation system, an HR benefits software package, or a cloud-based team collaboration tool, each is solving a specific need—but in isolation from other parts of the company's business systems.

The problem with this growing number of applications is that as a company's arsenal of systems (new and legacy) proliferates, its data becomes more distributed and disparate, making it more difficult (and often impossible) to access and manage.

While there are many reasons why managing data in these sorts of typical environments is difficult there are these three core challenges, any one of which can be daunting:

1. Each application inevitably has its own underlying database structure with a unique set of schemas, tables, joins, queries, etc. It is very difficult to get applications to integrate and share their data at the level needed for effective data management without enormous ongoing expenditures of people, time and money.

2. End-users are constantly raising the bar in terms of the accuracy, speed, scalability and security of their applications. They have limited tolerance for system faults or downtime, which in turn raises the criteria of what is needed for effective data management.

3. Data management, itself, is very complex and multidimensional. To be effective a business must have expertise in data modeling, data architecture, master data management, security, data governance, data cleansing, data synchronization, database administration, metadata management, data integration, and much more. For large companies with sophisticated IT organizations this is hard enough. For small to medium-sized businesses it can be impossible to keep up.

As a response to this fragmented picture there are hundreds of data management tools to address specific pieces of this puzzle. These include data governance, master data management, data synchronization, data integration and business intelligence software too numerous to mention by name. These are often highly capable tools that each serves a specific purpose, and therein lies a fundamental problem: to be effective, a company often needs many tools from many different vendors to tackle data management. Larger companies can afford to do this but often get frustrated with the expense and inefficiencies of using many tools, most of which don't inter-operate very well. For small to mid-sized companies they often don't have the resources to purchase all of these tools or sufficient IT resources to use them.

A DIFFERENT APPROACH: DATA VIRTUALIZATION

There is another approach that is gaining headway in the market and that is to bring data virtualization technology to data management. At core, it allows the unification of a company's siloed data. The following is a definition:
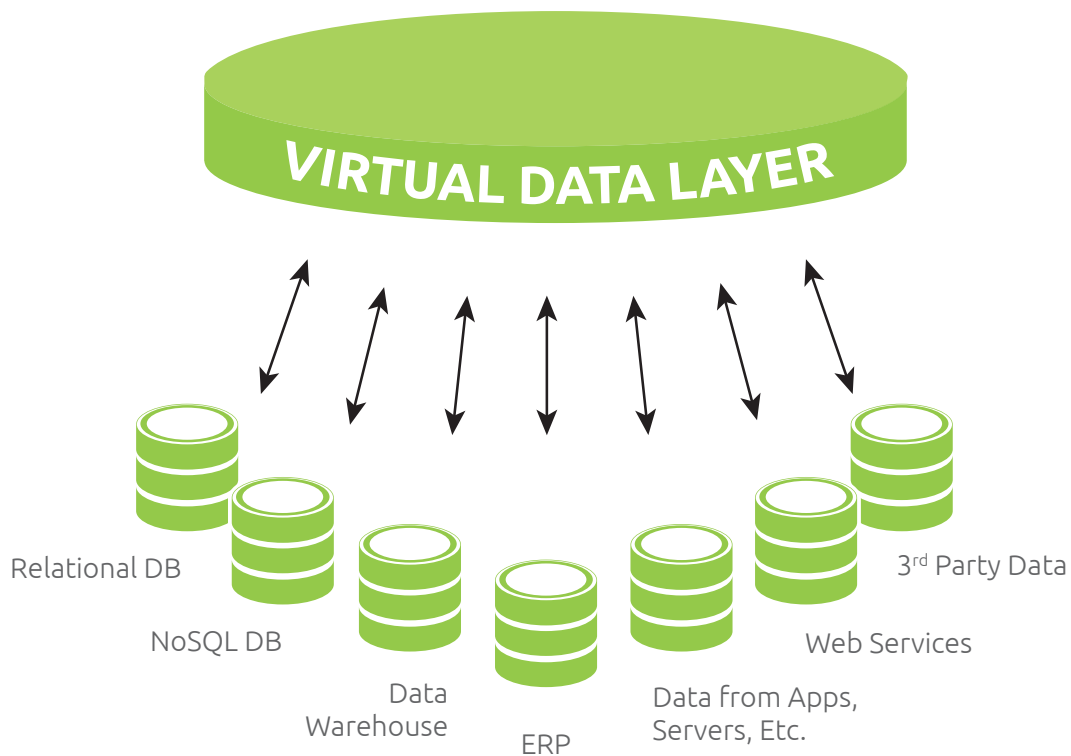
"Data virtualization abstracts data from multiple disparate sources creating a unified virtual data layer that provides users easy access to the underlying source data."

In other words, Data Virtualization enables a business to unify and use all of its data, regardless of data source or format, without building out complex data warehouses, or managing expensive infrastructure.

Figure 1

Graphic depiction of how data virtualization works.

VIRTUAL DATA LAYER

Relational DB

NoSQL DB

Data Warehouse

ERP

Data from Apps, Servers, Etc.

Web Services

3rd Party Data

The diagram below describes Accur8's technological approach to data virtualization. The first step is to model each of the underlying data sources to a metadata repository. This includes the schema, location of tables, type of joins, security and other key bits of information for each data source. The resultant metadata repository contains a detailed map on how to locate all the underlying data in each database. To access the data a user (which can be a person or a program) directly queries the virtual data layer.

03

Accur8
SOFTWARE

From here, the request is automatically processed by our query engine that augments it with pertinent information pulled from the metadata repository in real time. This, in turn, creates a transformed query that can retrieve and stitch together data from multiple data sources back into a single result set, thereby providing a unified view of, and access to, the underlying data for users to consume. The Accur8 query engine handles all of the typical back-end issues such as security, business logic, auditing, monitoring, etc. that typically make integrating disparate data overwhelmingly complex and expensive.

Importantly, this loosely coupled architecture means there is no need to copy or replicate data from each constituent data source to a central repository in order to gain a unified view. It also offers the ability to create, update or delete information in the underlying data sources in real time, a cornerstone of effective data management.

Figure 2

Detailed view of the mechanics of data virtualization
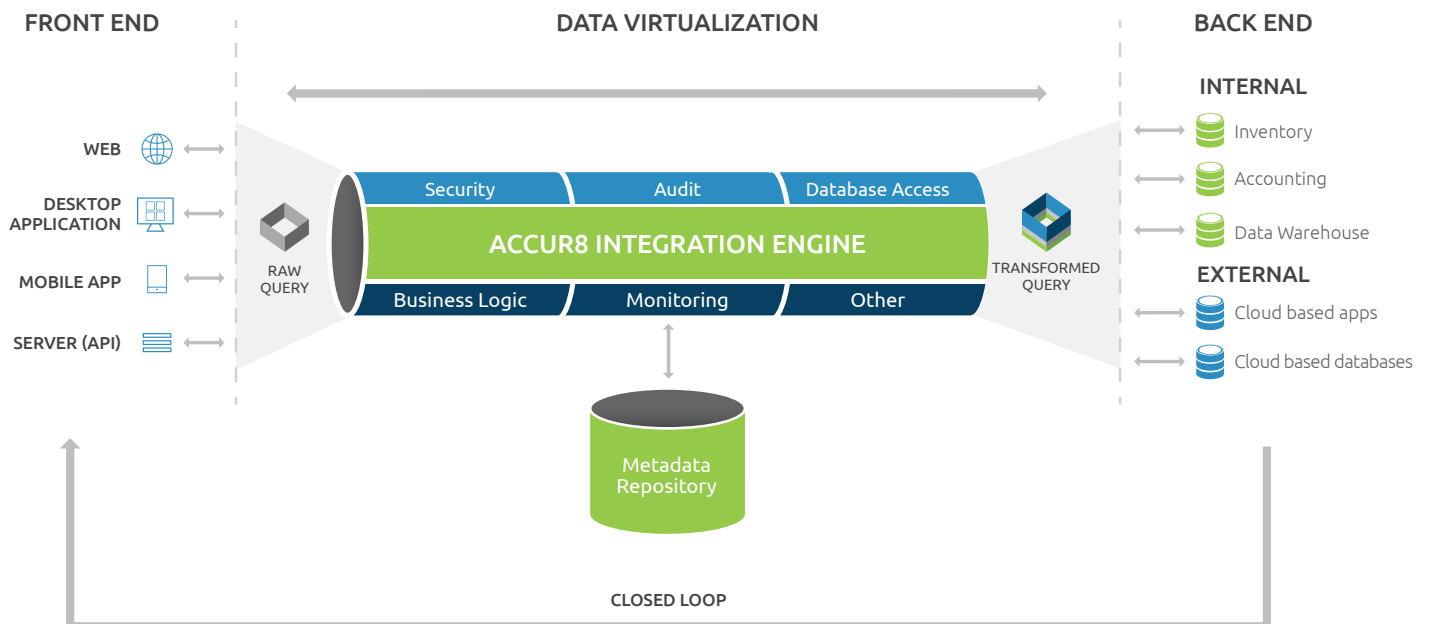
Critically, with data virtualization a company has both an abstracted, unified view and access to its discrete applications and data to complement its existing physical view and access. The combination of the two changes the game for data management because it enables functions such as data management and reporting and analytics to be performed with access to a company's underlying data from a unified place. At the same time it doesn't preclude direct access to applications if required.

Figure 3

Depiction of Unified Data
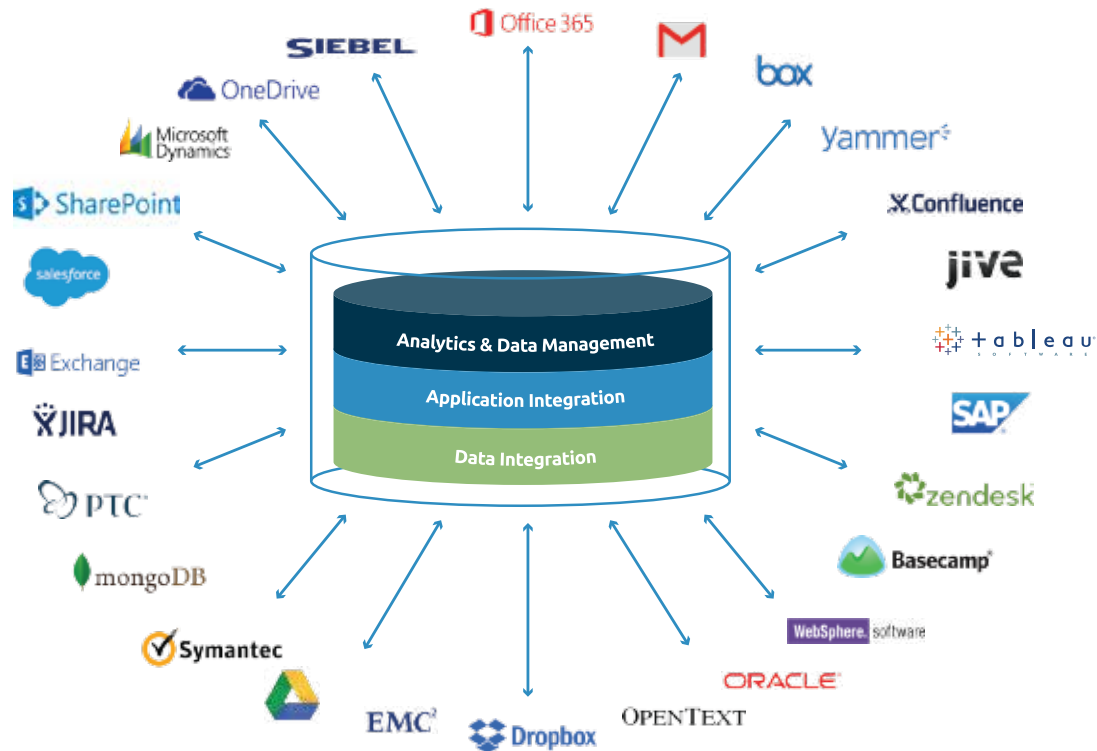
This diagram illustrates the hub and spoke architecture of data virtualization. Each application is connected to the hub by having its underlying database mapped to the virtual data layer. From this vantage point data management and reporting and analytics can be performed much more easily than with traditional integration approaches as there is easy access to read from and write to all of the underlying data sources without the typical complexity of back-end data source integration.

HOW DATA VIRTUALIZATION IMPROVES DATA MANAGEMENT

A key benefit of the data virtualization approach is that the metadata repository provides **a single point of reference** to access the entirety of the mapped data, regardless of source. By approaching many of the data management functions with unified access to all of a company's data it can greatly improve many of the traditional data management functions. The following are a few examples:

1. **Data Synchronization**

   An important component of effective data management is to be able to synchronize data between applications so that data is consistent across all data locations. Data virtualization provides an overall intelligence about the "DNA' of each data source, which provides very powerful synchronizing of a company's data. This allows the following:

   - Master to slave migration
   - Unidirectional synchronization
   - Bidirectional synchronization
   - One to many synchronization (from master source to multiple slave databases)

   Because data virtualization effectively turns all of a company's data into one big, virtual database there is the ability to monitor and track all of the syncs performed with self-serve reporting and analytics. This enables developers and database administrators to quickly build the specific views they need to see and manage the performance of their syncs. The workflow can be geared around each user's needs rather than the business needing to change their workflow to match the tool.

2. **Master Data Management**

   This is another cornerstone of data management and is based upon having a central repository of master data such as customer, product, employee and vendor data that acts as the gold standard for the underlying databases. This enables good governance and control of a company's data and builds upon the data synchronization capabilities mentioned above. With data virtualization the virtual data layer becomes the single source of truth for a company's data. There is no need to copy or replicate the underlying data to a central repository or data warehouse to do so. This in itself represents great savings. Company specific views and automated programs can be easily created to leverage the virtual data layer to ensure underlying data sources are consistent with the company's masters. All historical syncs are stores in a stand-alone data source integrated with the virtual data layer giving painless "time machine" capabilities to more easily restore data to a past point in time which in turn makes it easier to diagnose errors and track data lineage. Accur8 tools leverage the virtual data layer which easily enables custom views to be built of various fields across multiple data sources so they can be compared and evaluated for housekeeping, cleansing and de-duplication purposes.

3. **Application Management**

   Application management is the process of managing the operation, maintenance, versioning and upgrading of an application through its lifecycle (1). Application management clearly gets more complicated the more applications that a company needs to manage. However, with data virtualization the underlying data schema, data model, security, queries, version of each application are mapped to the virtual data layer. For example, if a new security, diagnosis, or configuration table is needed in several applications, it can be a time-consuming process to implement them

using traditional methods. This is so because the specified table needs to be copied into the identified applications while paying heed to each application's unique configuration and rules. With data virtualization the specific table can be lodged in the virtual data layer and each application can draw this data by being joined to the virtual data layer. Alternatively, the specific table can be pushed directly into the required underlying databases using the one to many synchronization tool described above. Views of all of the attributes of each application can also be created.

4. Security

When security is performed on top of a virtual data layer it leverages the virtual data layer's rich security model. Coarse grain or fine grain constraints can be set up to limit the underlying data that can be accessed by various data management tools; data syncing can be set to only access and update the data it is authorized to see. This provides much more accuracy and control while reducing the chance of unnecessary errors. Importantly, when a company develops the security model it needs for a particular function, say data syncing, it can be re-used in other data management tools such as master data management.

5. Unified Tool Set

When approaching data management through data virtualization most of the tools required by a developer or DBA are consolidated into one tool set. This means a single toolset for data syncing, master data management, generating various views of the underlying data, etc. A single toolset provides a much more efficient environment for users, rather than having them struggle with multiple tools where each tool has its own login, menus, workflows, etc. This saves significant time and expense and greatly improves productivity.

SUMMARY

The paradox companies face is that the more applications they purchase to become data-driven, the more difficult being data-driven becomes because their data gets ever more distributed and unwieldy to manage. However, by unifying their data with data virtualization technology, businesses are empowered with a new, more efficient and effective data management capability to tackle this challenge. Core features such as data synchronization, master data management, and application management can be approached from a unified view of, and access to, the underlying data. Data virtualization is a leading edge technology that can transform data management and should be considered by any business struggling to get better control and value from its disparate data.

Accur8 Software

Accur8 Software is a leading data unification company, focused on high performance, scalable and accessibly priced integration technology and tools. We recognize that companies' application environments are growing in complexity as they deploy more and more software applications to drive their businesses forward. This complexity hampers business performance because valuable data from across the company is not readily available to business users or systems. It forces IT staff to waste significant time and money dealing with the never-ending cycle of trying to integrate and share needed data across the organization.

The Accur8 Integration Engine is a data unification tool designed to help companies address the issue of complex application environments. It provides a flexible, agile way to unify data across processes and applications without coding. This means being able to integrate data and applications together whether they are in-cloud, on-premise or separated by geographical distance. It allows companies to access data and have it flow across the organization to users and systems as needed. Its capabilities include data integration, application integration, master data management, and reporting and analytics. It can be deployed as a point solution to integrate data between two applications or as a tool to unify all of a company's data and applications.

We have customers ranging from growth stage to Fortune 150.

Accur8 Software is one of CIO Review's Top 20 Most Promising Data Integration Providers for 2017.

For more information please visit our website at www.accur8software.com
or send an email to info@accur8software.com

(1) Techopedia, https://www.techopedia.com/definition/28008/application-management-am

08