Overcoming Cloud Limitations with a Hybrid Data Hub

Leveraging Data Virtualization to Supercharge Cloud Deployments

Organizations are moving to cloud-based infrastructure, to enable SaaS products and other service-oriented initiatives, and forgo many of the costs associated with maintaining on-premises infrastructure. But transitions to the cloud are not easy: As a first step, many organizations move data from multiple sources into cloud based data lakes, only to find that though the data is now stored in a single location, the data is still organized into silos. Organizations also find that security presents additional challenges, since the data is no longer under their direct control. And finally, they struggle with latency, since the data must now travel across the network, competing with other traffic.

The Denodo Platform provides a ready solution. Using data virtualization, which provides seamless, real-time access to data from multiple sources, the platform enables the creation of a hybrid data hub fed by both on-premises and cloud sources, simultaneously. Such hubs also provide a single access point for managing security, and hybrid data hubs enabled by the Denodo Platform offer specific features for overcoming network latency. Several companies have successfully leveraged the Denodo Platform to overcome the typical challenges associated with a migration to a cloud-based infrastructure, including Autodesk and Asurion.

SOLUTION
Hybrid Data Hubs for Cloud Infrastructures

PRODUCT OVERVIEW
The Denodo Platform offers the broadest access to structured and unstructured data residing in enterprise, big data, and cloud sources, in both batch and real-time, exceeding the performance needs of data-intensive organizations for both analytical and operational use cases, delivered in a much shorter timeframe than traditional data integration tools.

What Is a Logical Data Fabric?

Organizations struggle with cloud deployments for three primary reasons:

The silos remain. Cloud storage is inherently limitless, so many organizations begin their migrations to the cloud by moving data, department by department, into single, monolithic, cloud-based data lakes for analytics purposes. However, they cannot query across the entire data assets within a data lake, because the data is stored in different formats and schemata. Also, data lakes cannot exist in a vacuum. Enterprises also rely on SaaS applications, operational applications and other data repositories for their information needs and thus relegating a data lake to be yet another data silo.

Cloud security can be a deal-breaker. Because users can access a cloud source from any browser, and on most devices, securing a cloud deployment can be prohibitively daunting. If organizations opt to restrict access, they will be unable to gain the benefits of a cloud deployment.

Data is delayed by network congestion. Finally, because data is dispersed among many different servers, connected by the network, it is, by its very nature, subject to normal latency, based on the amount of traffic taking up bandwidth at any particular juncture.

Hybrid Data Hubs for Cloud Initiatives

Hybrid data hubs, enabled by data virtualization, overcome common cloud challenges. Such hubs provide seamless, real-time access to on-premises, cloud, and SaaS sources, all without data replication, and its associated costs and risks. Because of data virtualization, the hybrid data hub doesn't actually store any replicated source data; instead, it contains the metadata required to access each of the applicable sources, creating unified, virtual views across all of the various heterogeneous sources, including the silos within data lakes. For SaaS sources, it abstracts each of their associated APIs, so that organizations can point any of their standard reporting tools directly at the hub, which in turn gets the data directly from the SaaS application.

Hybrid data hubs seamlessly unify data stored in both on-premises and cloud sources. And since users are abstracted from the complexities of accessing each source separately, they suffer no impact whatsoever when organizations switch sources, or move the data from one to another. For this reason, hybrid data hubs enable organizations to migrate to cloud infrastructure on their own time, with zero downtime, and no impact on day-to-day operations.

Because consumers access all data through the hybrid data hub, it provides a single point from which to manage all security protocols, greatly simplifying the security challenge, and enabling security to be more effectively monitored and controlled.

Additionally, advanced hybrid data hub solutions can scale only as needed, to meet the changing needs of any organization, and prevent IT from purchasing unnecessary server time.

Figure 1: **Hybrid Data Fabric Architecture**

Finally, advanced hybrid data hub solutions offer specific features for overcoming network latency. The Denodo Platform, for example, enables hybrid data hubs that support caching with incremental queries. With this feature, a user's query across the hybrid data hub delivers three results:

1. Cached data, taken at regular intervals.
2. Changed data, based on user defined thresholds.
3. A merge of both of the above.

This results in a significant reduction in latency, simply by caching the static data, and only reporting on the changed data.
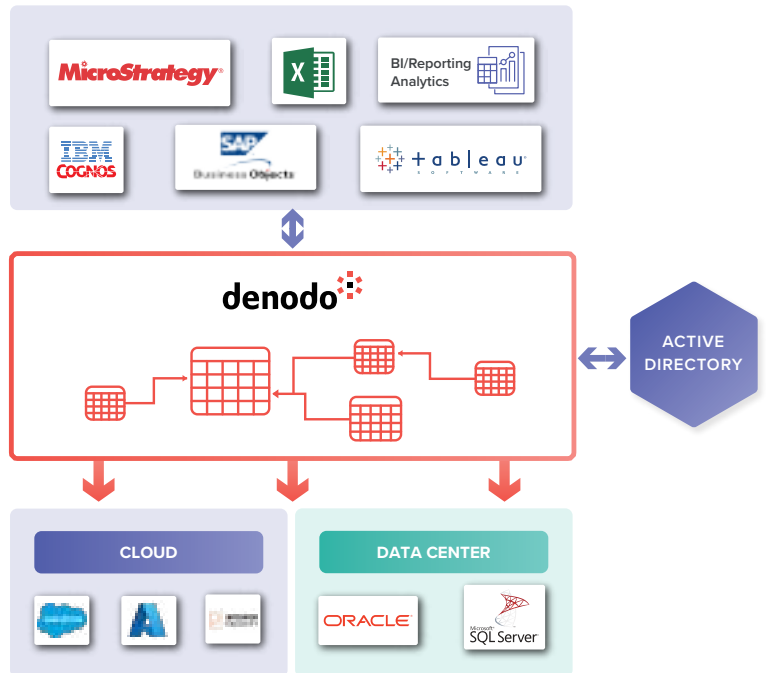
Benefits of a Denodo Hybrid Data Hub

Integrate data from silos within data lakes, from cloud and on-premises sources simultaneously, from SaaS applications, and other sources, without replication, in real time.

Reduce the inherent latency of accessing data from network sources.

Enable migrations from on-premises to cloud sources, with no impact on day-to-day operations.

Track cloud usage by department or individual.

Manage cloud and on-premises security from a single point.

Scale the solution only as needed, to hold down costs.

Hybrid Data Hub Deployment Models

Cloud-only Deployment Model: The Denodo Platform is deployed in a cloud environment such as AWS or Microsoft Azure, enabling unification of both cloud and on-premises data sources. This is suitable when most of the data resides in the cloud environment.

Hybrid Deployment Model: One instance of the Denodo Platform is deployed in the cloud while another instance is deployed on-premises. The on-premises instance functions as a gateway for exposing on-premises data for cloud access. This is suitable when the data is distributed across the two environments and an enterprise chooses to initially unify the on-premises data locally before integrating with Cloud data. The added benefit is in establishing a single-point-of-security within on-premises Denodo instance for data access.

Case Study: **Autodesk**

Autodesk is a leader in 3D design, engineering, and entertainment software. For years, the company sold its products in shrink-wrapped boxes via brick-and-mortar points of sale, using a conventional perpetual licensing model, and then as downloads, using the same model. But when Autodesk decided to switch to a more modern subscription-based licensing model, the company found that its infrastructure would not be able to support the kinds of BI and agility that the new licensing model required. Autodesk wanted to implement an agile BI 2.0 architecture with a logical data warehouse at its core to track subscriptions, renewals, and payments. But first, the company faced the challenge of how to manage the transition from its current, traditional data warehouse and data marts.

Solution

To integrate the new, subscription-based transactional systems with the legacy systems that manage perpetual licensing, Autodesk leveraged the Denodo Platform in AWS Cloud, which made it possible for Autodesk to continue reporting to stockholders and the SEC as though the company had never changed its revenue model. The new logical data warehouse, enabled by the Denodo Platform, acted as a hybrid data hub composed of both legacy, on-premises sources as well as modern, cloud-based sources, and it served as a unified enterprise access point for companywide data.
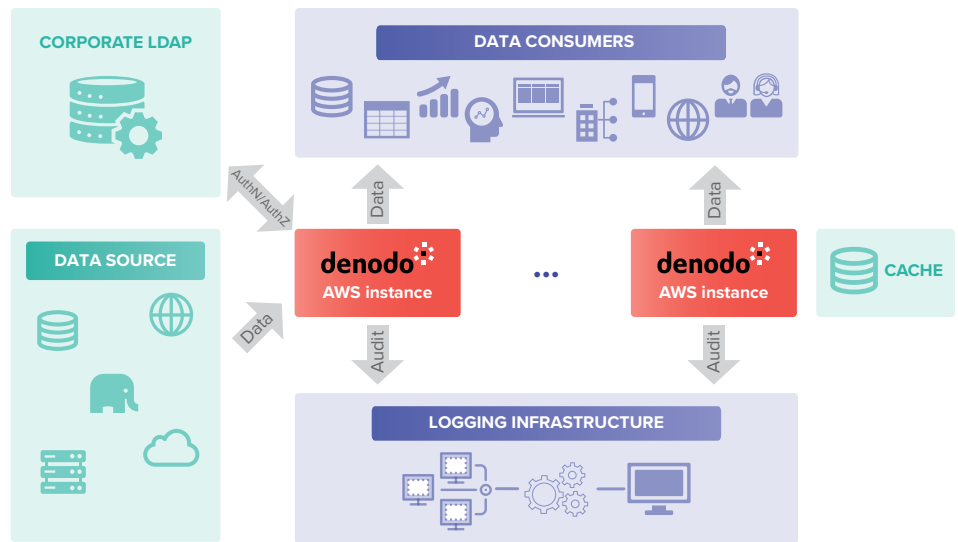

Figure 2: **Autodesk DV Architecture**

Results

Autodesk successfully transformed its revenue model from a perpetual, license-based model to a more modern subscription-based model, with no impact on business users or financial reporting. This change not only improved business performance and collaboration, but also profitability. Autodesk saved at least four months of effort redesigning the integration to systems supporting the new licensing model. Autodesk is also able to leverage a single point of security enforcement across its cloud infrastructure, along with consistent controls for data access.
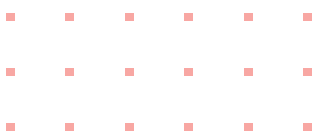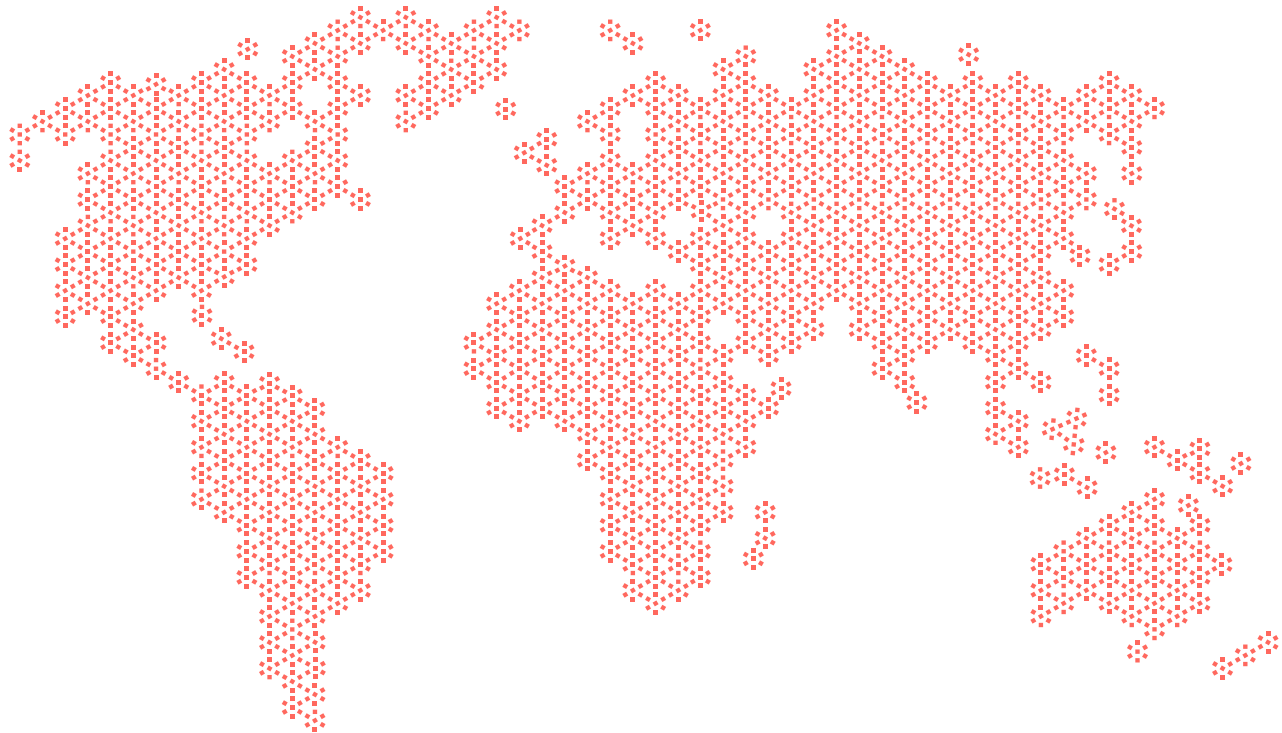
denodo

Denodo Technologies is the leader in data virtualization providing agile, high performance data integration, data abstraction, and real-time data services across the broadest range of enterprise, cloud, big data, and unstructured data sources at half the cost of traditional approaches. Denodo's customers across every major industry have gained significant business agility and ROI.

Visit **www.denodo.com** | Email **info@denodo.com** | Discover **community.denodo.com**