Actian Avalanche Federated Query for Hybrid Data Analytics

For the past decade, enterprise data warehouse strategists have espoused the virtues of a centralized location for all decision support data. With the advent of mobile apps, IoT devices, landing data into multiple clouds and data centers, the single warehouse notion has become increasingly tenuous. Data lakes are springing up in multiple locations, on-premise and in the cloud. This development has reinforced the importance of forward-thinking organizations incorporating a cloud data strategy that fully embraces hybrid data through a seamless federated data warehouse approach to harness data across the enterprise. Actian Avalanche's, a fully managed cloud data warehouse service, advanced architecture is purpose-built to address the demands of today's distributed hybrid data environments.

Hybrid Data is here to stay
Actian recognizes the inevitability of hybrid computing, giving data-driven organizations the ability to store and harness data wherever they chose to and employ all forms of data regardless of source, or delivery vehicle.

Cloud-based Analytics
Actian Avalanche, a third-generation cloud data warehouse, delivers a fully managed service that has been architected to be cloud agnostic. In concert with its on-premise cousins, Actian Vector analytics database and Actian-X hybrid transactional and analytical data management offering, these three compatible solutions, all based on Actian's industry leading X100 columnar database technology can deliver seamless hybrid data analytics across any enterprise data architecture.

Feeding Real-time Analytic Dashboards
Whether it's a national retail merchandising team trying to determine cross-sell and up-sell recommendations while a customer is browsing inventory or a healthcare benefits team trying to determine if millions of reimbursement claims processed per day are valid or form a pattern of fraud, the ability to query large sets of distributed hybrid data quickly and join them in a fashion that delivers 360° insights can be make or break for those organizations. Increasingly the data they use may reside not just across their organization but can span customers, partners, even government sources. With large distributed teams analyzing the data at the same time, the combination of

data size, variety of sources, and the parallel queries is beyond the capabilities of traditional data warehouse systems.

Federating disparate data sources to a single whole

Though most organizations profess to have a cloud strategy, few have completed their journey for production environments outside of leveraging SaaS apps such as Salesforce or using cloud storage for backup, document storage and DR. The vast majority of enterprises have their business-critical data spread across a combination of on-premise data centers and in cloud silos created by these native cloud apps often spanning multiple clouds. The need to support a cohesive corporate data strategy continues to exist – even more so with mobile, IoT, and outsourced elements of their business. This need can be best satisfied through a federated set of data warehouses that is abstracted to a single source.   Afterall, your analytics and the decisions made based on them are only as good as the completeness of the underlying data management feeding one's decision systems.

Introducing the distributed query

Imagine being able to transparently access data wherever it resides, regardless of physical location or technology platform. Actian provides the capability to register local and remote tables that can participate in distributed join queries. Once defined, remote table references are indistinguishable from local table references and become a part of the natural query landscape.

Why federated query?

In many cases, distributed queries are critical because data cannot be moved outside of specific domains or jurisdictions. Yet operations on the data within those remote locations is prohibited either by associated policy or technical resource restrictions. In other cases, the joins require combining data across multiple external repositories and then it becomes a matter of choosing the right virtual yet centralized operational data warehouse. In many of these cases, iterations of different federated queries may lead to the need for additional remote repositories and applications as well as additional parties leveraging the platform based on a series of prior successful results. Examples might include everything from consortiums studying communicable diseases, gene therapies, or other scientific collaboration to high street retail brands that must look across datasets in their own branded stores and department store chains as well as their suppliers.

A Game Changer – Actian Avalanche Intelligent Federated Query for Hybrid Data

Actian Avalanche cloud data warehouse features advanced federated query capability to incorporate virtually any source of hybrid data regardless of location, both in the cloud and on-premise. Our federated data operates via external table support and utilizes integrated Spark technology to access external data sources, register the data from those sources as external tables within Avalanche to be able to treat that data like its local Avalanche data.   Alternative vendor solutions often require extensive customer coding and other integration products to access and ingest external data, with Actian Avalanche everything you need to conduct federated queries of hybrid

data sources is built in to the service offering. Advanced, intelligent built-in capabilities including the ability restrict portions of a dataset from being accessed e.g. limiting the dataset to a particular zip code, or account id can be extended to be seamlessly applied to the remote source and only bring into Avalanche the data required to solve the query. Since all queries utilize Actian's industry leading X100 technology any query developed on an Actian analytics systems such as Actian Avalanche can be targeted to any other Actian analytics system without changes.

Actian Avalanche Supporting Multi-Cloud Hybrid Data in the new Age of Silos

While the examples above may not fit your particular use case, virtually every organization has organically grown in multiple directions when it comes to cloud. For example, with small to medium business (SMB) or large enterprises that support or sell to SMB, there has been a general move to Microsoft Azure as the cloud platform of choice. Also, for internal employee use, Office360 and other productivity and collaboration has migrated to Microsoft and or Google cloud environments. In both these Microsoft cloud trends, there is typically a mixture of Microsoft public cloud and Microsoft-based private cloud. Except for the truly legacy systems in the data center, most internal data sources can't easily be pigeon-holed into on-premise as legacy. Data center consolidation, virtualization, and internal cloud journeys have led to more efficient private clouds in many organizations, dedicated to critical business functions tied to sensitive data that cannot be taken out of the confines of the enterprise firewalls. And finally, there is AWS which has rapidly become the platform of choice for SaaS vendors and IT to provide support across multiple and very separate virtual systems residing on AWS.

Actian Avalanche, next generation cloud data warehouse is a managed service that runs on AWS, Microsoft Azure and soon Google Cloud. It incorporates Actian's proven X100 columnar engine technology to deliver a single big data analytics capability across all manifestations of private cloud and on-premise environments, including Linux, Windows, Hadoop, and all popular VMs and containers as well.

Avalanche Cloud on AWS to Actian Vector On-premise Federated Database Example

As an empirical real-world example, a large retail chain manages its vast historical sales data on Actian Avalanche Cloud Data Warehouse running on AWS. An Actian Vector instance running on a Hadoop cluster on-premise in a corporate data center might contain current supply chain data being continuously updated by micro-batch every 10 seconds and customer sentiment could be delivered in real time via a Kafka stream. The procurement manager's BI dashboard could draw from all three data sources completely transparently, employing Actian's distributed query manager to optimize the most cost-effective query processing location based on the cardinality and location of the base tables in both databases. Given that both Actian solutions utilize Actian's X100 columnar engine queries can be seamlessly run on either system with no changes required.

Avalanche Cloud to Legacy Enterprise Datawarehouse Example

An on-line retailer was reaching the data storage limits of their legacy enterprise data

warehouse appliance. They decided to off-load their customer review data, the most voluminous element of any product description, to Actian Avalanche on AWS to free up space on the existing legacy appliance. As customers browse items on the ecommerce site all the data is delivered at the click of a button and the customer has no idea that it's being joined from two very separate data sources. The performance delivered by Actian Avalanche far exceeded the retailer's expectations and CDO plans to migrate, over time, all their decision support data to Avalanche to provide increase analytical throughput, flexibility and cost savings.

Actian Hybrid Data Approach Delivers Superior Solution for both Today & Tomorrow

Hybrid cloud means more than just selecting best-of-breed native cloud apps and moving your on-premise systems into consolidated private clouds. To make the most of these assets and expenditures, you will need to leverage the data associated with these systems to take better decisions and execute more quickly and accurately at the point of action. Your analytics to support this will be only as good as what your team can leverage in terms of their algorithms, queries, and their visualization of their results. A best-in-class team needs best-in-class data management and integration to support them. Actian Avalanche Cloud Data Warehouse provides unbeatable price/performance and supports access and manipulation of data spanning multiple enterprise and remote systems across virtually any cloud app or data repository, bringing together a federated view and support for complete analysis of today's challenging business problems.

About Actian – Activate your Data™

Actian, the hybrid data management, analytics and integration company, delivers data as a competitive advantage to thousands of customers worldwide. Through the deployment of innovative hybrid data technologies and solutions Actian ensures that business critical systems can transact and integrate at their very best – on premise, in the cloud or both. Thousands of forward-thinking organizations around the globe trust Actian to help them solve the toughest data challenges to transform how they run their businesses, today and in the future. For more, visit http://www.actian.com

2300 Geng Rd, Suite 150, Palo Alto, CA 94303
+1 888 446 4737 [Toll Free] | +1 650 587 5500