



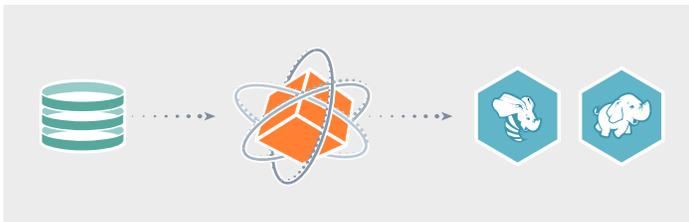
HVR 5.2: Data Lake Enhancements

HVR 5.2 provides even faster, more secure, and accurate data movement in a modern environment. Latest enhancements include features that help customers optimize their adoption of data lake technologies.

Native Hive Support

The data lake is typically the initial landing zone for the data. Data scientists commonly use access to the data lake to identify data sets of interest in order to slice-and-dice the data in a different, more suitable, high performance analytical engine.

Native Hive Support facilitates data retrieval on file systems S3 and HDFS. Tables are created automatically with the correct data types and changes to the table definitions on the source are automatically propagated. Users can now use SQL access to perform data discovery with “off-the-shelf” BI tools.



S3 Optimizations

Taking large volume scenarios to the next level. The data lake release introduces a number of performance enhancements to deliver data into files. With S3’s unique file handling specific enhancements let to up to 5x faster data loads into S3.

Amazon Key Management Service Support

Most data stores provide data security at-rest. However in-flight data exposure is a common concern with cloud computing. With HVR 5.2’s support for client-side encryption, integrated with Amazon’s Key Management Service (KMS), data is truly encrypted end-to-end.

Evolution of a Data Lake

The first data lakes were built for specific use cases around IoT data like log files and other machine-generated or sensor-generated data. As organizations are becoming analytically smarter and analytic outcomes become more prescriptive in nature – influencing the primary business – there is a need to consolidate data from more different types of sources in the data lake. Traditional transactional data must now be added to the data lake.

With more data from different types of sources organizations are looking for the most cost-effective deployment platforms for their data lakes. The Hadoop Distributed File System (HDFS) has been a popular platform choice for data lakes, and more recently native cloud storage like Amazon’s S3 is commonly selected.

Metadata Manifest

Traditional relational databases follow commonly accepted transactional rules to enforce data consistency. With most data lakes built on file systems and changes commonly stored per table, a lot of the consistency between data objects is lost, and many use cases cannot handle this challenge well. Metadata Manifests enable data publication at a transactionally consistent representation of the source system, even when multiple integrators deliver data in parallel.

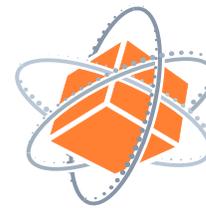
Big Data Compare

Wide data lake adoption requires users' trust in the data. HVR 5.2 enables programmatic data validation against the data in your data lake, even on HDFS and S3.

HVR's distributed architecture, log-based change data capture, and efficient use of network resources generally play well with continuous, large volume data integration into a consolidated data store. The Data Lake Release takes this support to the next level with exciting new capabilities to help organizations get more value from their data.

Other 5.2 Enhancements

- SAP HANA capture enables log-based change data capture from an SAP HANA database, allowing organizations to unlock data in SAP HANA, e.g. for delivery into a data lake.
- SQL Server log-based CDC from backup transaction files only enables efficient and fast, direct log-based capture from SQL Server's transaction logs on a separate system, with no elevated operating system or database privileges.
- Notifications through Slack allows HVR's automatic notifications to be delivered into Slack as an alternative to email or SNMP-based notifications.



ABOUT

We Accelerate Data Movement so that you can Revolutionize Your Business. HVR is designed to move large volumes of data FAST and efficiently in complex environments for real-time updates.

Our goal is to keep your data moving and in sync as you adopt new technologies for storing, streaming, and analyzing data. Our scalable solution gives you everything you need for efficient data replication from beginning to end so that you can readily revolutionize your business for the modern world.

Interested in trying HVR for your data lake project?
CONTACT US FOR A FREE TRIAL: info@hvr-software.com