

# Extend the Value of SAP HANA with Big Data

Practical Use Cases for a Scalable, Simplified Converged Analytics Platform

Enterprises recognize the tremendous value of using big data to generate new insights, and many are turning these insights into a competitive advantage. A challenge persists, however, in determining how to layer business context onto the steady influx of new and non-traditional data sources that are coming into the business.

For many companies that have started on a big data journey, SAP HANA has provided a firm foundation in the shift toward becoming more analytics focused; the in-memory relational database system provides the ability to perform predictive analytics, spatial data processing, and text analysis, among other analytic functions. To extend the business value of SAP HANA, Lenovo developed the Lenovo Converged Analytics Platform for SAP Vora and SAP HANA, a scalable, end-to-end big data and analytics platform that allows companies to improve decision making by harnessing unstructured and structured data from a variety of environments.

## A Simplified Approach

Recognizing the value of big data solutions, Lenovo established the Big Data Center of Competency to build on its solid foundation of engineered solutions, bringing together teams from development, marketing, sales, and professional services for a cohesive approach to fulfill customer needs.

The Lenovo Converged Analytics Platform for SAP Vora and SAP HANA represents a significant output

---

*For many companies that have started on a big data journey, SAP HANA has provided a firm foundation in the shift toward becoming more analytics focused.*

from this project. A fast and resilient big data platform that incorporates SAP HANA and SAP Vora, the Lenovo Converged Analytics Platform is powered by the MapR Converged Data Platform, SUSE Enterprise Linux, and Lenovo systems, and is scalable to hundreds of nodes.

For SAP customers, the platform solves the challenge of how the business can use enterprise data in an SAP HANA environment to enable decision making in parallel with unstructured data from a big data environment. Simplifications to the platform ensure that SAP customers do not have to worry about external equipment to fulfill storage requirements.

These simplifications are possible because there is no storage area network (SAN) requirement to run the Lenovo Converged Analytics Platform. Eliminating the external SAN requirement greatly simplifies configuration and implementation of the platform, allowing for easy scalability from five nodes to thousands of nodes, with no downtime required.

The MapR file system (MapR-FS), accessed by SAP HANA nodes via a highly available NFS mount, immediately stores and replicates SAP HANA data among the SAP Vora nodes. This means that the moment data is written from SAP HANA, it is replicated for resiliency, offering protection in the event of failure of any node while also allowing for the data to be extended to additional SAP HANA or SAP Vora nodes.

## Driving Value

With this as the framework for extending SAP HANA into a big data environment, the next question from many companies concerns the practical benefits. The value propositions are essentially endless, but let's explore three key use cases that demonstrate value.

### 1. Run SAP Vora Independently from SAP HANA

The first use case is that the platform enables SAP Vora to be used for analytics independently of SAP HANA. Any big data environment will contain a mix of



**Brian Finley**  
Principal Architect  
**Lenovo Big Data Solutions**



**Al Kalafian**  
SAP Solutions Architect  
SAP Center of Competence  
**Lenovo**

structured and unstructured data, the latter including Internet of Things (IoT) data or image data that doesn't necessarily fit neatly into rows and columns. With the Lenovo Converged Analytics Platform, SAP Vora can take that unstructured data and apply a structure on top of it, which can then be queried against like a database. Visualizations and other analytics with SAP Vora can thus be run independently of SAP HANA, even while leveraging some of the included SAP analytics engines such as graphing and time series.

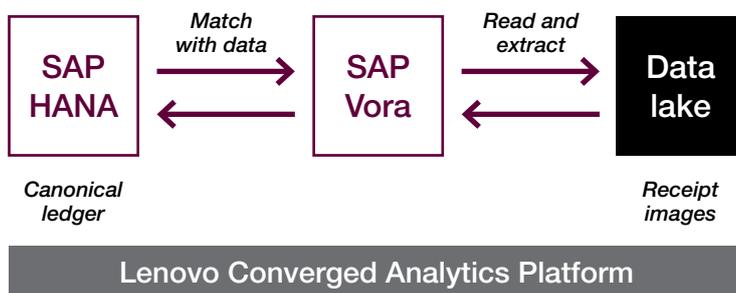
### 2. Incorporate SAP Vora with SAP HANA

The second use case involves incorporating SAP Vora with SAP HANA. The aforementioned table structure created by SAP Vora allows for a user to perform SQL queries against SAP HANA and SAP Vora at the same time; a single SQL query can incorporate data from the SAP HANA and SAP Vora tablespaces, including, for example, the use of a JOIN clause.

### 3. Use SAP Vora for Additional Storage

Third, SAP Vora can act as a second, lower tier of storage for cold data for SAP HANA, keeping it in-memory on the SAP Vora nodes, ensuring higher performance than if the data were to be moved directly to disk.

It's easy to see how a company can benefit from a platform that acknowledges the difference between data that's appropriate for SAP Vora and data that's appropriate for SAP HANA — and how that data can be used together. Take a retailer, for example, that wants to drive actionable insights from a trove of receipt images from hundreds of retail outlets (see **Figure 1**).



**FIGURE 1** Retailers can leverage SAP Vora and SAP HANA in tandem to compare vast image data with the canonical ledger

*A big data platform that can extend its value to an existing analytics environment provides customers with the contextual awareness needed to turn insight into action — and move the business forward.*

This retailer's canonical ledger resides in SAP HANA, but it wouldn't make much sense to also store image receipts there. To extract value from the images using the Lenovo Converged Analytics Platform, the retailer can instead keep the image data in a data lake and have SAP Vora read the receipts into the system, extract the text on the receipt images, and perform optical character recognition. By creating a tablespace in SAP Vora that represents that data and running queries against SAP HANA and SAP Vora at the same time, the retailer can corroborate that the canonical ledger matches up with the data from the physical receipts or otherwise detect anomalies.

### Begin a Big Data Journey

So much of the business value that organizations are looking for as they extend their digital footprint is tied to big data infrastructure deployments, and for SAP customers that value extends to leveraging their existing SAP analytics environment. A big data platform that delivers this capability without a SAN requirement, regardless of the sizing of a multi-node SAP HANA and SAP Vora environment, provides customers with the contextual awareness needed to turn insight into action — and move the business forward.

A Lenovo Converged Analytics Platform demo environment, including one SAP HANA node and five SAP Vora nodes, is available if you'd like to learn more. All stakeholders — SAP, MapR, SUSE, and Lenovo — have remote and on-site access to the environment. On-site demonstrations can be hosted at the Lenovo Executive Briefing Center in Morrisville, NC. For more information, go to [www.lenovo.com/sap/hana](http://www.lenovo.com/sap/hana). ■