SAP Vora

SAP® Vora™ and Big Data
Table of Contents

4  Big Data Analysis for Today’s Business

7  Diverse Data Types Pose a Challenge to Enterprise-Wide Analytics

8  A Unified Big Data Environment Across the Enterprise

9  Data Management Solutions from SAP for Big Data

10  Summary
EXECUTIVE SUMMARY
Huge amounts of raw data are being gathered because so much more of it is available and the volumes are growing each year. The data is being generated by diverse sources – the Web, enterprise network infrastructure, cameras, telemetry devices, software logs, enterprise applications, mobile apps, and the Internet of Things (IoT). Critical Big Data use cases are maturing, including IoT applications and the ability to develop a 360-degree customer view.

Making all the data usable for business analysis is a challenge, however. Some data is structured, some is unstructured, and some is semistructured. There is graphical data, Structured Query Language (SQL), time series, relational, and JavaScript Object Notation (JSON), to name just a few. While this data can now be gathered in massive amounts, unless it is shaped and defined based on context and semantics, it is of little use in helping businesses gain strategic insights. At first, the Big Data journey was about storage. Now it’s about processing giant, diverse data sets to make them useful for business transformation.

SAP Vora helps businesses tackle this complexity by unifying processing engines for different data types into one unified data pipeline to feed analytics solutions. This paper provides an overview of SAP Vora and includes examples of its use in leading analytics use cases. Now with SAP Vora and other products included in the data management solutions from SAP for Big Data, businesses have powerful features that simplify, expand, accelerate, and lower costs for computation of all data managed by the enterprise. You can:

- **Get actionable insights and create business value from Big Data** with an in-memory distributed computing solution that unifies data from enterprise systems and data stored in Apache Hadoop in near-real time without the need to move data between systems for computation.
- **Reduce complexity and drive results** with user-friendly access for everyone from business analysts and data scientists to engineers and developers. They can use familiar tools, programming languages, and templates to analyze huge amounts of data quickly and efficiently.
- **Deploy SAP Vora** in the cloud on a full-service Hadoop and Apache Spark as-a-service platform. This option is available to enterprises that don’t have the staff, budget, or desire to deploy and manage their own Hadoop environment.

Once an exclusive domain of scientific and engineering applications, Big Data analysis is now available to every business. It’s dramatically changing how businesses operate, strategize, and innovate, influencing business strategy and bottom-line results and serving as inputs to enhance products and services. The SAP® Vora™ solution tames the complexity of the data glut that businesses are grappling with. The result is **better insights from a broader and more diverse data pool.**
Big Data Analysis for Today’s Business

Today’s businesses can extract huge amounts of data from a vast variety of sources. And they can derive many different benefits from analyzing that data. The following are some examples of the data sources and benefits.

THE INTERNET OF THINGS
SAP believes that IoT applications will soon be among the biggest use cases for Big Data analytics among businesses. Network-connected products and processes are already generating huge amounts of data, but only a tiny fraction of this data is being captured and stored for analytics. With cloud and edge computing, however, and unified products such as data management solutions from SAP for Big Data, that’s expected to change. In particular, the time-series engine in SAP Vora provides efficient compression and processing for IoT data.

According to a forecast from ABI Research, the volume of data captured by the IoT will exceed 1.6 zettabytes by 2020.

How will businesses use analytics to benefit from the IoT? Here are three examples:
• Companies with fleets of trucks can use IoT data from vehicles to monitor driver behavior, track vehicle location, assess engine health, and do many other things. Benefits may include lower insurance premiums for good drivers, lower fuel consumption, and more-accurate forecasting of delivery times.
• Remote patient monitoring through body sensors allows healthcare professionals to keep track of a patient’s health remotely. Benefits include shorter hospital stays, preventive care that is less expensive, and lower health insurance premiums.
• Farmers can use sensors to measure soil moisture, air temperature, wind speed, humidity, and other factors. Benefits include better decisions about irrigation and crop placement, lower costs for water due to less usage, and higher-quality crops.

1. Edge Analytics in IoT, ABI Research, April 2015.
THE 360-DEGREE CUSTOMER VIEW

With so many different touch points that customers use today to interact with companies (for example, e-mail, phone, in-store, Web, text, mail, IoT, social media, demographics), the ability to gather this data could be the basis for a more complete, 360-degree view of the customer. But achieving this view has until now been problematic.

Data from different channels exists in different silos and data lakes and is hard to integrate and process together. The ability to combine and correlate the massive amounts of data from different touch points and sources for a 360-degree customer view is a major challenge. The graph engine in SAP Vora is particularly useful in this case for revealing hidden relationships between data collected at multiple touch points. Integration between SAP Vora and the SAP HANA® platform also provides the capability to utilize insights gained in near-real time at the point of sale.

Benefits from deep insight into customer behavior and preferences include more-accurate predictions of customer buying habits; more effective, targeted marketing campaigns; more competitive pricing, packaging, and branding; reduced support resolution times; and many other improvements in real time.
CONTEXT-BASED LOG ANALYSIS

Every transaction and many interactions between customers and businesses leave digital footprints that are captured in logs. By capturing the data in user logs, businesses can gain a clearer understanding of online user behavior. Benefits include better Web design based on what pages were viewed and for how long or what e-mails were opened and clicked to open Web content. Such insights help marketers create a more compelling and effective buyer’s journey.

By extracting the data in machine-generated event logs and making it available to analytics applications, businesses can gain important insights into company operations. Benefits include uncovering cyber threats, seeing and acting on opportunities for greater efficiencies, predicting system failures before they occur, optimizing resource utilization by anticipating server loads, and other enhancements based on data from real-time monitoring.

For both types of context-based log analysis, a Big Data environment such as Hadoop is essential for handling the large data volumes associated with this use case. The ability to unify log data in a Hadoop repository with application data from different lines of business, or IT operational data from within the enterprise, for analytical systems would result in valuable insights. The ability of SAP Vora to efficiently process large volumes of data regardless of data source can help provide benefits, including cost savings, better resource utilization, and new monetization opportunities.
Diverse Data Types Pose a Challenge to Enterprise-Wide Analytics

Hadoop is a powerful and robust processing environment for Big Data that incorporates massive, low-cost storage. The Hadoop infrastructure can scale to handle colossal volumes of data. The challenge of integrating Big Data into enterprise business analytics, however, comes from the lack of uniform data formats between the two environments.

Much of the data within the enterprise is hierarchical and structured. Big Data storage lakes in Hadoop contain a wide variety of unstructured data. Making that data useful for processing has required the use of different engines based on distinctive technologies that do not interoperate. These different data silos remain separate and proprietary, connected to relational databases, data warehouses, and Big Data infrastructures.

Currently, some organizations are stitching together multiple systems to build data pipelines. This approach requires a significant investment in time and resources to execute batch computations. The expertise of highly skilled engineers is needed. Even when available, these solutions are often inflexible and have limited scalability.

This is not an efficient or effective strategy for delivering the timely analytical insights and decision-support capabilities that users have become accustomed to in enterprises.

Because of these challenges, the full range of enterprise data today is seriously underutilized. On average, 60% to 73% of all data collected within a typical enterprise is left unused for business intelligence (BI) and analytics. So, while tools may be available to extract insights and business value out of data in each separate format, what’s needed is a single tool to provide processing across all data types.

As all engines in SAP Vora are connected with a single access layer, you will be able to combine results from different processing engines, revealing a comprehensive view of your business using all types of data from multiple sources.

A Unified Big Data Environment Across the Enterprise

With the in-memory engines in SAP Vora, including SQL, time series, graph, and JSON, integrated into a single tool, you can more easily process different types of data without stitching together multiple systems. As all engines in SAP Vora are connected with a single access layer, you will be able to combine results from different processing engines, revealing a comprehensive view of your business using all types of data from multiple sources.

SAP made a huge impact in the business analytics market with the introduction of SAP HANA, providing tremendous speed of in-memory computing for enterprise applications. With SAP Vora, you can unite the enterprise and Hadoop and Spark domains to derive greater value from your burgeoning information assets.

SAP Vora supports all major Hadoop distributions. Recently, we’ve added cloud deployment options to offer more flexibility and speed your time to value. This includes deployment on a public cloud, such as SAP’s own Big-Data-as-a-Service offering or Amazon Web Services, which offers 10 times faster performance at a lower total cost than generic Hadoop infrastructure in public clouds. In addition, data management solutions from SAP for Big Data comprise multiple strategic offerings to help you aggregate, federate, secure, manage, and optimize your data for use in analytics across the enterprise.
Data Management Solutions from SAP for Big Data

Four complementary products, shown in the table below, comprise the data management solutions from SAP for Big Data.

Products in the data management solutions from SAP for Big Data

<table>
<thead>
<tr>
<th>Products in the data management solutions from SAP for Big Data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>SAP® Vora™</strong></td>
</tr>
<tr>
<td><strong>SAP Cloud Platform Big Data Services</strong></td>
</tr>
<tr>
<td><strong>SAP HANA®</strong></td>
</tr>
<tr>
<td><strong>SAP solutions for enterprise information management</strong></td>
</tr>
</tbody>
</table>
SAP Vora is a product within the data management solutions from SAP for Big Data. It is designed to unify diverse data types for more inclusive and impactful analytics, and to unify data processing – from the enterprise to Hadoop environments – to support a new class of massively scalable Big Data analytics for businesses like yours.

By deploying SAP Vora in an existing Hadoop installation, you can use Spark – and optionally SAP HANA – to concurrently work with data from Hadoop and enterprise systems. By doing so, you unlock the full potential of your organization’s entire data inventory. You can apply familiar tools and techniques such as SQL and online analytical processing (OLAP) to correlate, interpret, and process the masses of raw, unstructured, or semistructured data in Hadoop.

Unification and usability of different data types from diverse sources, simplification of your processing and analytics environment, reduced costs, broader utilization within your organization – these are some of the main benefits that SAP Vora and other SAP products that work with Big Data bring. These leading-edge offerings from SAP enable you to:

- **Unify data access** to enterprise and Hadoop data, regardless of their sources. This means your organization can make decisions based on the entire, up-to-the-minute collection of information instead of relying on disconnected data sets. Linking archived and live operational data improves compliance and yields better insight by including the complete data set.

- **Simplify your technology landscape**, allowing your organization to reduce the number of technologies it uses by using a single tool for processing different types of Big Data. In addition, SAP Vora bridges the gap between SAP HANA and Hadoop, merging enterprise systems and Hadoop into a single landscape. SAP Vora also opens up access to the growing Spark ecosystem for enterprise analytics users.

- **Extend Big Data analytics apps to a wider audience**. SAP Vora allows a wider set of users to work with Big Data. By supporting widely adopted information access techniques such as SQL and OLAP, SAP Vora allows anyone familiar with these tools to start querying in minutes.

Data management solutions from SAP for Big Data comprise **multiple strategic offerings** to help you process, federate, secure, manage, and optimize your data for use in analytics across the enterprise.

**FIND OUT MORE**

For more information about SAP Vora, visit [www.sap.com/vora](http://www.sap.com/vora).