Modern Data Analytics

Solve big data challenges with a converged appliance, advanced analytics and services

The big data dilemma

Big data and powerful predictive analytics tools are now the key to competitive advantage. By putting data-driven decision making at the heart of the business, your organization can harness a wealth of information to gain immediate insights, drive innovation and create an unparalleled competitive advantage.

That's the opportunity, anyway. The reality on the ground for many enterprises is quite different. The push to capitalize on big data is often hindered by slow data loading times, slow database query performance and the need to constantly tune complex data warehouse systems. At Dell, we have seen customers monopolize their data warehouse resources with ETL processes that run for 8 to 24 hours—just to prepare and load data—and then spend hours to complete complex reports and queries. Is that the new normal? We say not.

The big data opportunity is also hindered by fragmented data sources and incompatible data types. In the typical enterprise, data is spread over many disparate systems that might be on-premises or in the cloud. To complicate the challenge, many organizations don't have easy ways to join structured data from enterprise relational databases with unstructured and semi-structured data from social media, the Internet of Things and end-user systems.

These challenges make it difficult to meet today's demanding service level agreements for data query times. In an age when the business needs immediate insights from big data, the backend systems can't deliver—and the business pays the price in the form of costly lost opportunities that stem from a lack of real-time analytics.

While these problems are widely recognized, the solution to them is more elusive. Today's data warehouses already face bottlenecks in their shared RAM and compute resources, and many weren't designed to bring together structured,

Massively Parallel Processing: Real-Time Decision Support of Relational and Non-Relational Data

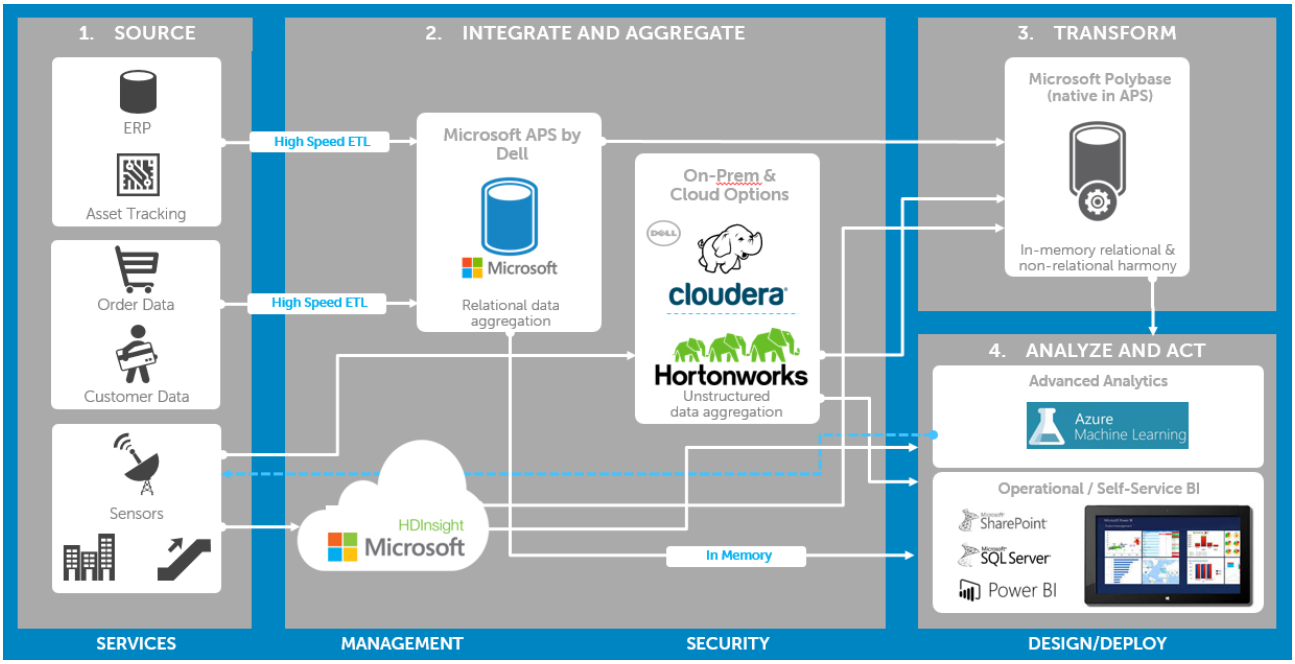


Figure 1. Bringing the analytics to the data. Relational and non-relational data, both on-premises and in the cloud, come together in harmony with the Microsoft Analytics Platform System by Dell.

unstructured and semi-structured data. These are all arguments for change, but the path forward is hindered by the costs of forklift upgrades and the complexity and staff resource commitments that come with custom-built solutions. Other barriers include the need to maintain investments in skills—new technology usually requires new skill sets, so many organizations have an incentive to stick with what they know best.

This is where the Dell Analytics Platform System enters the picture. This turnkey solution was designed from the ground up to address some of today's biggest data challenges.

One appliance for analytics on all your data

The system at a glance

The Dell Analytics Platform System brings together technologies for analytics on structured, unstructured and semi-structured data from enterprise systems, social media, the Internet of Things and countless other sources.

The system incorporates all the hardware, software and services components you need for a complete, end-to-end analytics solution in a Microsoft-certified system. The appliance is co-engineered and jointly supported by Dell and Microsoft for a "white glove" experience from delivery to support.

This is not a reference architecture to guide your deployment. This is a fully packaged solution that ships to your door as a prebuilt appliance that is ready to tackle your data challenges. You simply power up the system and start integrating your data into the solution, with the assistance of the included deployment and training services.

Once your system is up and running, the management layer enables you to query both structured data from your enterprise systems and unstructured and semi-structured data held in Apache Hadoop systems, using native software technology in Microsoft APS called PolyBase. The Hadoop environment can be based in the Dell Analytics Platform System rack, co-located in a rack that is connected to the appliance, or located in the cloud. And you have choice: PolyBase allows you to connect to data stored in Hadoop (Cloudera, Hortonworks or Azure HDInsight) for harmony with your relational (SQL) data.

The blueprint in Figure 1 depicts the big data harmony that can be achieved with the Dell Analytics Platform System in a common use case, such as improving customer profitability with advanced analytics. Disparate data sources that hold customer and order data, as well as location and device data, can be transmitted both on-premises and in the cloud, and deposited to a relational SQL Parallel Data Warehouse Store on APS and a Hadoop distribution for streaming data.

Microsoft PolyBase is used to align both data sets, while familiar tools from the Microsoft Office and Business Intelligence suite can be used to run operational reports. Advanced reporting and analytics solutions like Power BI, SQL Reporting Services and Azure Machine Learning work in harmony to assist in forecasting customer profitability and reducing fraud and churn.

Business-driven benefits

PERFORMANCE

To deliver performance that is up to 100 times faster than that of SMP systems, the Dell Analytics Platform System's massively parallel processing (MPP) architecture allows data processing to be distributed across nodes, so multiple processors can work on a problem or query. The system further accelerates time to insights with in-memory analytics, which moves the data closer to the processors.

With its MPP architecture and in-memory, updateable columnstore indexing capability, the Dell Analytics Platform System can help you get answers to queries in seconds or minutes instead of hours or days. Better still, with its ability to load and query databases simultaneously, the system offers exponentially faster data loading compared to conventional alternatives. It can load up 1 TB – 2 TB per hour.

SCALABILITY

With its highly scalable architecture, the Dell Analytics Platform System allows you to follow a pay-as-you-grow model. You can buy the capacity you need for today and add nodes to scale out in a seamless manner as your needs grow, without the costs or disruption of forklift upgrades that comes with traditional scale-up systems. The Dell Analytics Platform System scales out without downtime.

To help you be prepared for unpredictable growth, the system scales from 3 to 54 compute notes and from 45TB to 6PB of data. Each scale unit increases both the size and the compute capability of the system for linear scale.

TOTAL COST OF OWNERSHIP

With the included deployment and Jumpstart training services and optional customer workload development, Dell gives you everything you need to achieve fast time to value to solve your big data challenges. It enables you to create harmony between your SQL/relational data and unstructured Hadoop data to reduce the cost of aggregating, storing, analyzing and innovating.

The components of a complete big data solution

The Dell Analytics Platform System brings together the hardware, software and services components you need for a complete solution for data analytics. The appliance is shipped with all core software components installed, configured and tuned.

HARDWARE

Hardware components include 13th generation Dell™ PowerEdge™ servers, storage arrays, network switches, power, racks and more. The system architecture is fully redundant within the cabinet and designed for high availability.

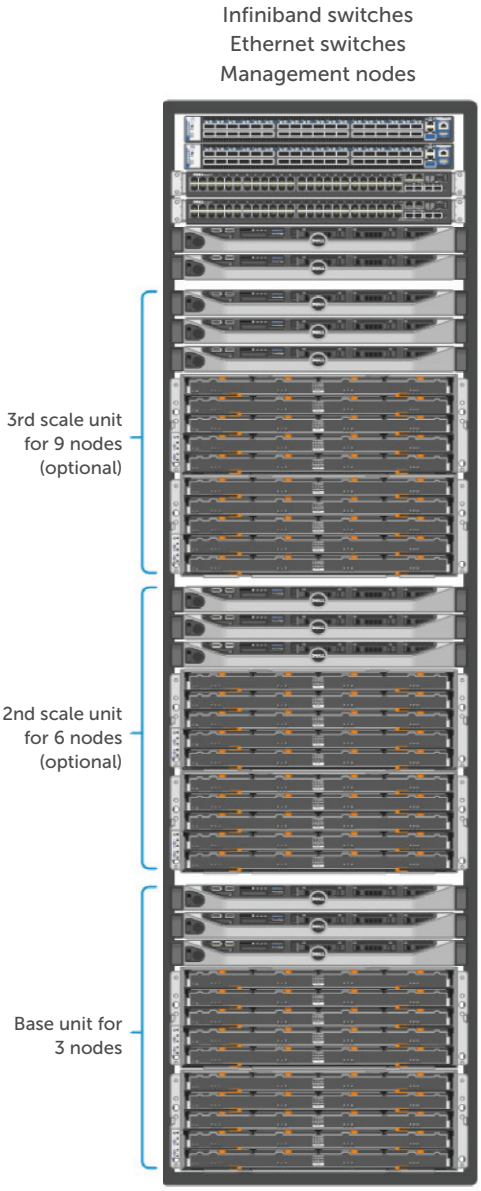
SOFTWARE

Core software components include the Microsoft Analytics Platform System (APS) with built-in Microsoft PolyBase software for the integration of data in SQL and Hadoop environments. PolyBase is a next-generation execution engine for querying, transforming and indexing relational data and non-relational Hadoop data using familiar SQL syntax.

Various tools are available to enable flexible configurations for the Hadoop component, whether you place it in the appliance, next to the appliance or in the cloud. Dell offers a range of options for harmony with Hadoop, including Microsoft Azure HDInsight, the Dell QuickStart for Cloudera Hadoop bundle, and reference architectures for Hortonworks, which work with Microsoft PolyBase.

To enhance your solution, you can drive more value with unique technology from Dell that drives insight and innovation. Dell Statistica can be used to glean insights and predictive analytics from the relational and non-relational data harnessed within the Microsoft APS ecosystem. Insights can also be gathered from other Microsoft tools, such as Excel, Power BI, SQL Reporting Services and PowerView.

Microsoft Analytics Platform System by Dell



An end-to-end analytics solution

- Realize up to a 100X performance increase over SMP database architecture.
- Combine SQL Massively Parallel Processing (MPP) plus Hadoop (optional) in a single appliance.
- Scale from 45TB to 6PB.
- Gain the benefits of a highly available and fault-tolerant design
- Jumpstart your production workloads with implementation, training and enablement services.

SERVICES

Bundled services include factory integration, implementation and three weeks of immersive Dell Jumpstart training to help your staff learn the system, understand best practices and develop a roadmap for your future state. In addition, Dell's highly qualified service professionals can assist your organization with complex workload development—such as hybrid integration with Microsoft APS and cloud or Hadoop platform migrations.

SUPPORT

The appliance experience is complemented by a unified Microsoft and Dell one-call support model, with Microsoft acting as your single point of contact for all patches, upgrades and issues. Customer support calls go to Microsoft, and Microsoft communicates directly with Dell to enable quick resolution of hardware issues.

One system. Many use cases.

The Dell Analytics Platform System supports a wide range of data analytics use cases. Potential use cases span virtually all industries.

Here are a few of the countless potential examples of how organizations can use the system to unlock valuable insights from diverse sources of data:

- Manufacturers can join structured data from manufacturing systems with unstructured data from engineering tests to analyze defect and fail rates.
- Retailers can merge structured data from point-of-sale systems with unstructured social sentiment data to improved demand forecasting.
- Educational institutions can bring together data from student information systems with unstructured data from diverse sources to identify at-risk students.
- Global enterprises can merge data from dozens or even hundreds of widely distributed database servers into hub-and-spoke data warehouses.
- Healthcare providers can migrate multiple medical claims data stores to a single high-performance data warehouse environment.

The key is to identify your big data challenges and then tailor the configuration of your Dell Analytics Platform System to meet your unique needs.

Key takeaways

In today's digitally driven world, data analytics is the key to competitive advantage. Virtually all enterprises now need to analyze big data to accelerate time to insights. That's the idea behind the Dell Analytics Platform System. With its converged MPP architecture, in-memory analytics and support for virtually all data types, the Dell Analytics Platform System greatly accelerates complex queries that join data from diverse sources.

Compared with conventional approaches, the appliance-based Dell Analytics Platform System makes it much easier and much faster to deploy a complete data warehousing solution that meets your organization's most challenging big data and business intelligence needs.

Unlike today's complex legacy data warehouses, the Dell Analytics Platform System makes scalable, high-performance data analytics available as a turnkey experience. You can now get an out-of-the box solution with high-speed parallel query processing, highly scalable data storage, integration with Hadoop and high-speed data transfer for your most demanding data warehousing workloads.

Best of all, Dell delivers everything you need for an end-to-end big data solution—from the appliance and its hardware and software to services, support and technical guidance.

To learn more, visit Dell.com/microsoft or email APS_Sales@Dell.com.

August 2016 | Version 1.2