



TERADATA® UNIFIED DATA  
ARCHITECTURE

Integrated data.  
Strategic insight.  
Business action.

TERADATA®

## EXECUTIVE SUMMARY

We are at a time in history when vast amounts of data bring tremendous opportunity. The combination of exploding amounts of data paired with advanced technologies for analytics represents an exciting new frontier for businesses.

However, the challenges of capturing and interpreting all this data are significant. Enterprises need flexible, sophisticated, and cost-effective solutions that complement existing technology and scale as the organization grows. The data warehouse continues to be a key component of the architecture to deliver these analytics, when coupled with other technologies for cost-effective storage and fast insight discovery.

Companies need to capture, store, and analyze not only structured data, but new forms of multi-structured data such as Web logs, social media, text, graphics, email, audio, machine-generated data, and much more. Their users need powerful analytics to discover patterns in this data using the skills and tools they already possess.

Until now, these huge amounts of data have been ignored or underutilized simply because the tools didn't exist to make sense of it all. The good news is that increasingly sophisticated software, combined with growing computing horsepower, can be used to create new hybrid analytical ecosystems that can handle all types of data from disparate systems. Using these, organizations can extract intelligence from their data to significantly outperform their rivals on both depth and breadth of valuable business insights.

The Teradata® Unified Data Architecture™ is a hybrid architecture that lets organizations do just that: leverage all their data for new insights and new business opportunities, resulting in growth of revenues, earnings, and corporate productivity.

## BIG DATA IN TODAY'S WORLD

Vast amounts of data are created every day by machines as well as billions of people using computers, smart phones, tablets, and other personal electronics. Companies are now capturing trillions of bytes of information about their customers, suppliers, and operations. Networked sensors in devices such as mobile phones, smart energy meters, automobiles, and industrial machines can sense, create, and communicate data constantly. The sheer magnitude of the numbers around big data is breathtaking. Consider that:

- ~ 2.7 Zetabytes of data exist in the digital universe today.<sup>1</sup>
- ~ According to industry analyst IDC, data volumes are expected to more than double every two years between now and 2020.<sup>2</sup>
- ~ On Facebook alone, 4.75 billion pieces of content are shared daily.<sup>3</sup>
- ~ 422 Terabytes of data have been collected by the U.S. Library of Congress as of May 2013.<sup>4</sup>
- ~ YouTube users upload 100 hours of new video every minute of the day.<sup>5</sup>

## THE CHALLENGE: INTEGRATING AND EXTRACTING VALUE FROM ALL THE DATA

Included in the notion of big data is not just its volume, but also its complexity—known as multi-structured data. In addition to structured data, such as ERP, inventory, and customer purchase data, there are semi-structured data types such as Web logs, XML, JSON, social profiles, and Twitter feeds. Add to that data types such as images, audio, text, and video, and it's evident that it's not just more streams of data companies must contend with, but entirely new ones. In addition, big data is often generated with high velocity and collected at frequent intervals, which is what makes it both valuable but difficult to analyze.

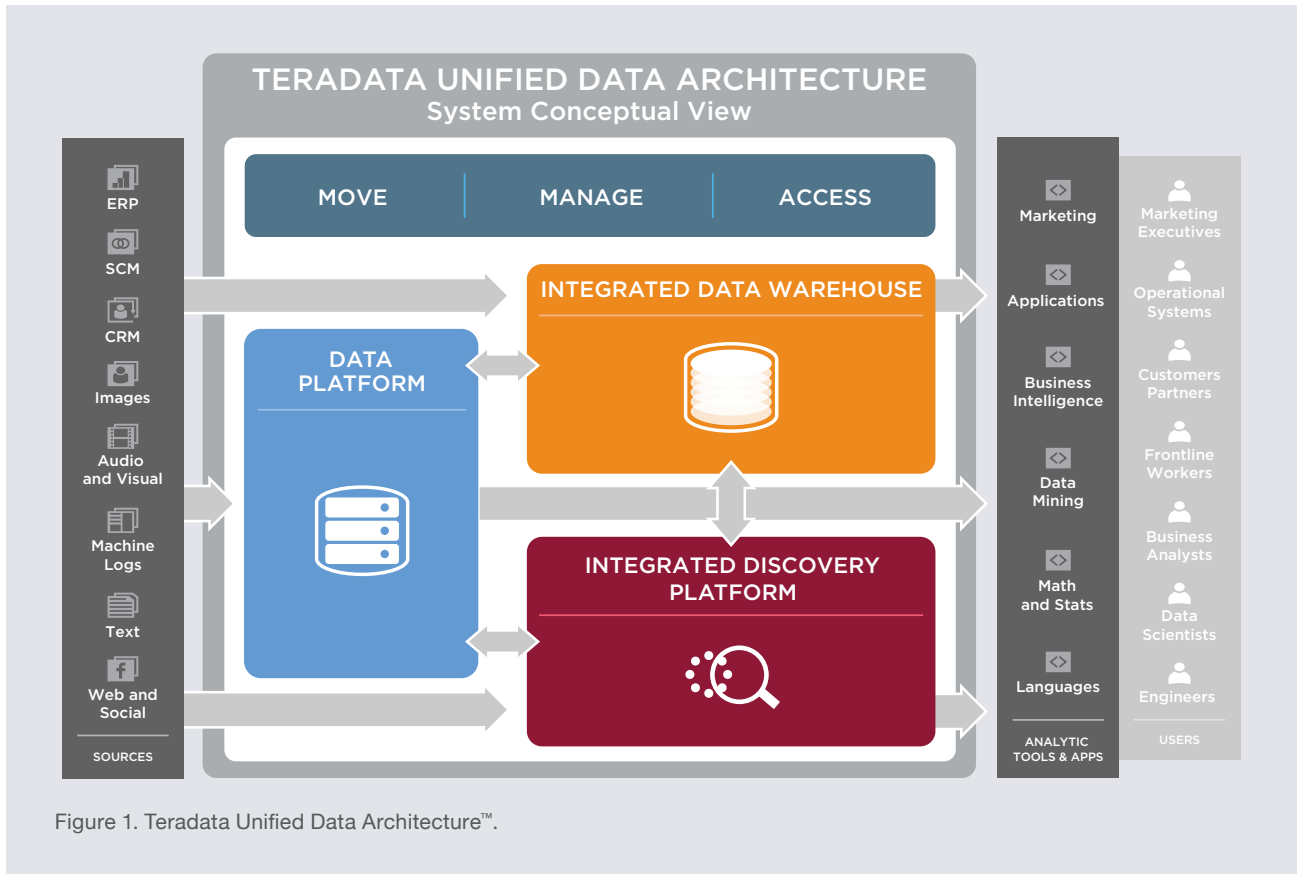


Figure 1. Teradata Unified Data Architecture™.

As a result, there are broad challenges to extracting value from data. Different data has different information densities, carries different worth, comes in different formats, requires different analytics, and serves different business purposes. Enterprises need technologies that can address all these challenges while delivering a holistic view across all of their data to make unique associations and ask the important questions about customers, products, or relationships. So it reasons that employing multiple technologies may be the answer. These fall into three categories: Data Platforms, Data Discovery, and Data Warehousing. Each brings unique value and capabilities. However, integrating these technologies can create additional overhead. Organizations must be careful to choose solutions that maximize business benefit and minimize integration costs.

## THE SOLUTION: TERADATA UNIFIED DATA ARCHITECTURE™

Teradata Unified Data Architecture™ is Teradata’s big data hybrid architecture. It combines best-of-class components— the Teradata Integrated Data Warehouse, the Teradata Aster® Discovery Platform, and two options for the Data Platform—Hortonworks® Hadoop™ or a Teradata Integrated Big Data Platform. These are connected using a high-speed Infiniband interconnect fabric, to provide a scalable, flexible, and robust ecosystem for handling all data and analytic requirements.

Organizations can ask any question of any data at any time to unlock new and valuable insights, resulting in deeper customer predictive models, faster detection of fraud patterns, and better understanding of root causes

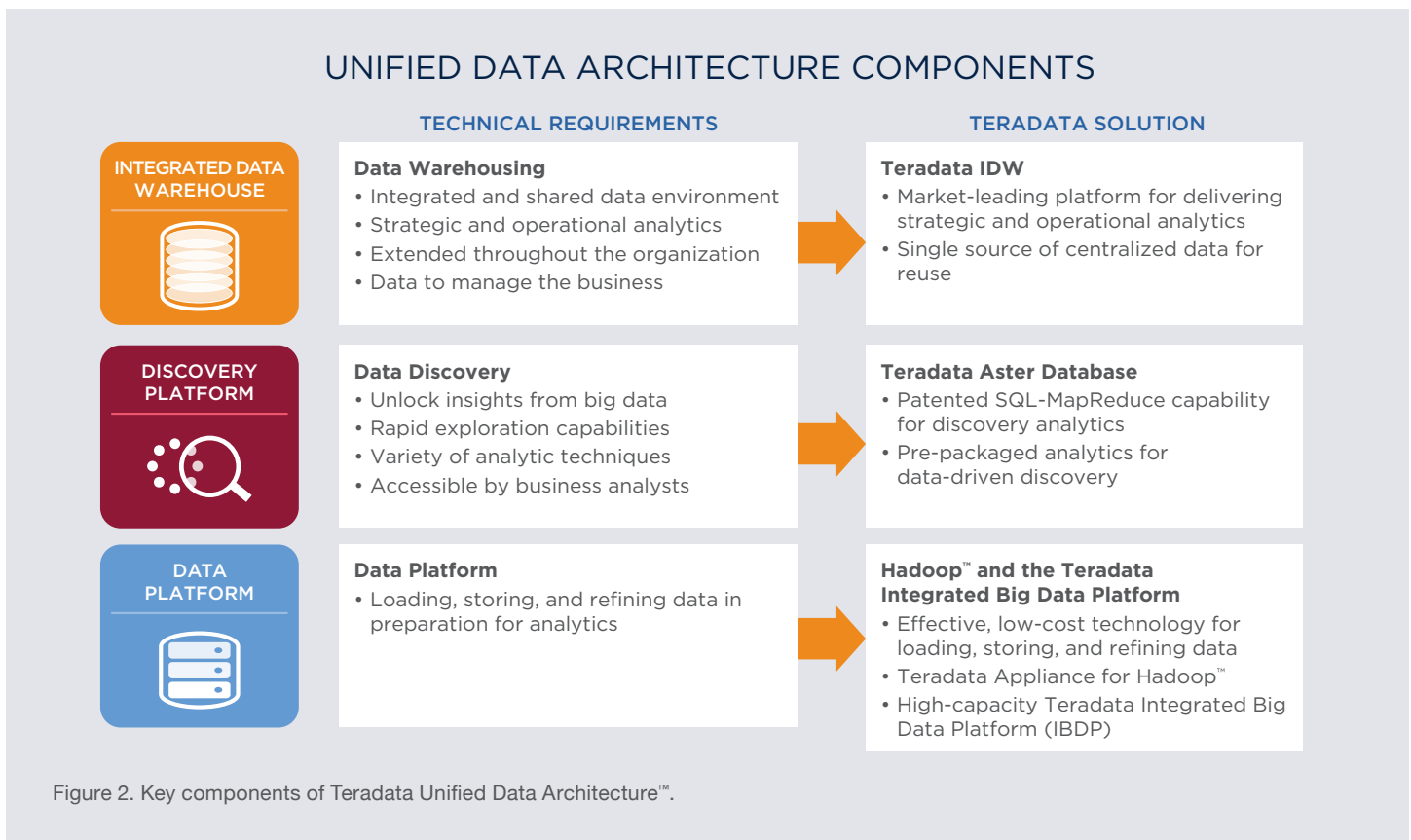
of bad customer service, as a few examples. No data is too big or too complex. Users can ask the most challenging questions and get speedy answers because Teradata Unified Data Architecture™ bridges the gap between the business language of SQL and the emerging popularity of MapReduce.

Key components of the Teradata Unified Data Architecture™ are Data Warehousing, Data Discovery, and Data Platforms. (See Figure 2.)

The Teradata Integrated Data Warehouse (IDW) is the market-leading platform for delivering strategic and operational analytics throughout your organization so users can access a single source of consistent, centralized, integrated traditional data. Teradata provides and encourages data reuse across functional groups, a “store-once, use-many” enterprise-wide advantage.

Users can range from BI experts doing complex trend analytics on seasonality to Web systems that make automated call-outs to Teradata to personalize Web pages for end-customers. You’ll have the flexibility to do it all—all at the same time—with intelligent query processing that supports any physical data model. With more than 30 years of cutting-edge development, Teradata Database is the foundation for the world’s leading data warehouses.

The Teradata Aster Database brings the power of patented, parallel Aster SQL-MapReduce® technology to data scientists and business users so they can do discovery easily and iteratively. A library of more than 800 statistical, manipulation, and visualization functions provides powerful capabilities on both traditional and semi-structured data, speeding up the discovery process. Data scientists can also write their own analytics for use in SQL-MapReduce statements. A suite of





Aster connectors allows data for experimentation to be pulled from a variety of external data sources, including Teradata IDW or IBDP and Hadoop™ systems, as well as flat files and other database sources. This provides the ideal environment for exploration and discovery of new insights across various data types.

For companies that require a large-scale data platform, Teradata has two options—open-source Hortonworks Hadoop,™ and the Teradata Integrated Big Data Platform. Both data platforms utilize Teradata’s robust tools for system management, data access, and one-stop support for all Teradata products. These options in the Unified Data Architecture™ provide extremely cost-effective storage as well as great landing/staging platforms for all kinds of data.

Teradata Unified Data Architecture™ ties these three powerful technologies together and integrates them using value-add connector software, staging, and support.

Teradata Unified Data Architecture™ is the best hybrid solution on the market, leveraging the best-of-breed and complementary values of Teradata Data Warehouse, Teradata Aster Discovery Platform, Teradata Integrated Big Data Platform, as well as open-source Hortonworks Hadoop.™ We align the best technology to the specific analytic need—all engineered, configured, and delivered ready to run. Teradata adds key value-add enabling technologies such as Teradata BYNET® running on Infiniband that unify the solution, allowing companies to focus their efforts on extracting business value from their analytics rather than trying to make them work together. All the solutions are enabled by high performance Intel® Xeon® processors. Other key value-add-enabling technologies are:

#### Transparent Access

- ~ SQL-H™ provides a robust interface for run-time, self-service data access from Aster Database to Hadoop,™ as well as Teradata Database to Hadoop™
- ~ SQL Assistant provides a user-friendly SQL creation front-end for a consistent experience across Teradata Database and Teradata Aster Database
- ~ Unity Director automatically routes users and queries between Teradata systems based on context of the query and system availability

“As a big data practitioner and a data warehouse evangelist, what truly is a forward thinking architecture from my perspective is the ‘unified data architecture.’ This is where the rubber meets the road in my opinion.

“What this architecture does for users is create two platforms at the same time—one for exploration and mining purposes and the other for analytics and management reporting. You can push workloads across the different architectures here and leverage the power of all the pieces of the infrastructure. With the right approach and solution architectures, enterprises can take a giant leap forward for the big data journey on these types of platforms.”

Krish Krishnan, President & CEO,  
Sixth Sense Advisors Inc

### Seamless Data Movement

- ~ Connectors provide easy to use, high-speed data movement between Teradata Database and Teradata Aster Database, Teradata Aster Database and Hadoop™, and Teradata Database and Hadoop™
- ~ Smart Loader for Hadoop™ gives users and administrators a friendly point-and-click, drag-and-drop interface for bi-directional data movement between Teradata Database and Hortonworks Hadoop™
- ~ Unity Data Mover delivers intelligent, high speed data movement between Teradata systems. Its combination of command line or GUI driven interface coupled with its automatic selection of load utility gives users and administrators a powerful tool for data movement

### Single Operational View for Management

- ~ The ease-of-use and anytime, anywhere access of Teradata Viewpoint is now available on the Teradata Aster Big Analytics Appliance, with support for the Teradata Aster Database as well as Hortonworks Hadoop™, giving administrators a single console for managing all the Teradata analytics systems in their environment
- ~ Teradata Vital Infrastructure extends consistent one-stop support for Teradata to the Teradata Aster, and Hortonworks Hadoop™ appliances. Automated monitoring and fault escalation for all three technologies is delivered from a single source
- ~ Unity Ecosystem Manager supplies end-to-end monitoring of process, components, and data across Teradata systems
- ~ Unity Director makes managing multiple systems running the Teradata Database easy by intelligently applying database management commands to all participating Teradata systems

Bringing these technologies together allows organizations to quickly run iterative analytics against a broad, deep set of data using SQL, SQL-MapReduce, non-SQL languages and tools.

As a result, businesses can capture, store, and analyze data from all parts of the organization and turn it into actionable intelligence. By combining analytics capabilities, the Teradata Unified Data Architecture™ allows business to explore vast stores of traditional and new data sources. For example, companies can better understand and predict customer behaviors, and by doing so, improve customer experiences. Transactions, multi-channel interactions, social media, syndicated data through sources such as loyalty cards, and other customer-related information can be combined easily on Teradata Aster from Hadoop™ or Teradata sources. Putting activities in time series order enables data scientists to see a complete picture of customers' preferences and demands. These insights about opportunities for differentiated treatments can be driven into automated Web personalizations and better customer care screens, using Web service callouts to the Teradata Integrated Data Warehouse.

## EXAMPLE #1: FINANCIAL

### Problem

A large global bank was struggling with reducing churn in profitable customer segments. Part of the problem the bank had in tackling this issue was integrating customer interaction data across multiple channels from numerous silo'd repositories. The size of the data—billions of records per month—also made the analysis of this information a very complex exercise on their previous systems.

They sought a solution to proactively help detect and prevent churn of their most profitable customers.

### Solution

Leveraging the powerful analytic capability of Teradata Unified Data Architecture™, this leading financial institution was able to build an enterprise visualization of all customer interactions with the bank and identify most frequent paths to account closure across all interaction channels.

### This solution allowed the bank to:

- ~ Identify the precipitating common causes of churn
- ~ Reduce the causative problems (e.g., excessive fees)
- ~ Decrease churn rate (by an expected 5%)

### This was achieved by using:

- ~ An enterprise data warehouse from Teradata for historical customer transaction, profile and product information
- ~ Teradata Aster Database to analyze and discover patterns through nPath analysis to determine which sequences of customer activities are most likely to lead an account closure
- ~ Hadoop™ for loading, storing and refining data and optimizing storage costs
- ~ Teradata Integrated Marketing Management to make real-time decisions and offers to enhance customer satisfaction and deliver the right offers at the right time preventing account closures and actually growing customer relationships

### MULTI-CHANNEL PATHS TO ACCOUNT CLOSURE



Figure 3. Frequent paths to account closure.

## EXAMPLE #2: RETAIL

### Problem

A global retailer with multiple Web site properties needed to get a better understanding of customer Web activities like to paths to purchase (and non-purchase), as well as bailout analysis.

They wanted to identify which offer personalizations work more effectively for which segments of customers.

### Solution

They used all three components of the Teradata Unified Data Architecture™. The raw Web log records from more than a million visitors per day were captured in Hortonworks Hadoop™. Teradata Aster was used to analyze the time series of Web clicks at the page and product level to build clusters of customers with similar interests. Experiments were done to see which next-best offers had the most appeal for various shopping sequences. Experimental results were then loaded into Teradata, and used to drive real-time Web personalizations.

### The benefits realized include:

- ~ Very fast system creation—it took just few days to create the environment and do discovery work
- ~ Better recommendations, leading to a 15% average increase in the market basket, driving up sales
- ~ A new discovery that some shoppers buy 10-20% more when the Web sequence from

browsing to purchase is shortened. Led to a Web redesign to minimize clicks to purchase.

- ~ An infrastructure is now in place for hundreds of additional experiments, increasing the value of data to the organization.

### This was achieved by using:

- ~ Hadoop™ for raw Web log collection
- ~ Teradata Aster Database for experimental insights, including
  - Fast pattern matching on sequences of Web clicks
  - Measuring path to purchase and bailouts
- ~ Teradata to drive Website personalizations with the CRM tool's Real Time Interaction Manager to measure results.

### BUYING JOURNEYS TERMINATED AT SHIPPING CHARGE PAGE

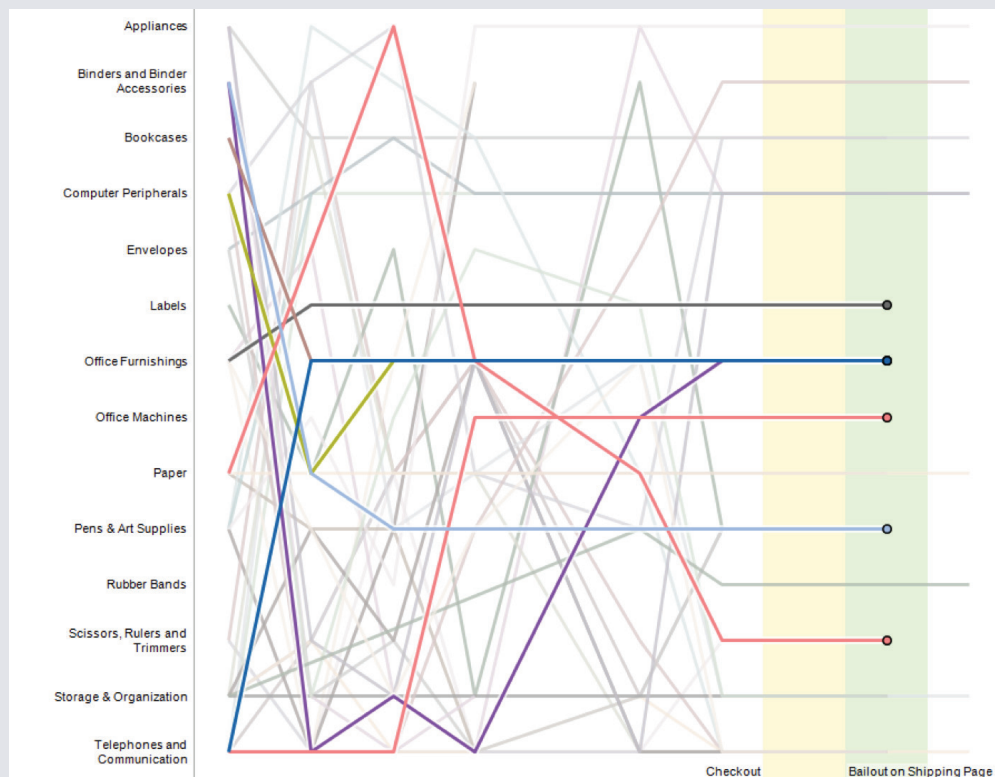


Figure 4. Web visitors resulting in bailouts.



### EXAMPLE #3: TELCO

#### Problem

A Global Communication Service Provider with bundled service (wireline, wireless, cable, and Internet) was interested in improving customer service by analyzing why customers call customer care after they have been on the Web site. They call this “breakage” since self-service for any of these various offers via the Web site isn’t working for the customer. It’s a big problem worth correcting because Web self-service takes substantial cost out of the operations budget for the care centers.

#### Solution

The company collects a large amount of data, ranging from call detail records (CDRs) on each call, switching data from operations, billing information, customer care center call information, truck rolls for installations and repairs. Over time, the company is learning which component of the Teradata Unified Data Architecture™ can most cost-effectively store each type of data (e.g., once CDRs have been used for billing, they can roll onto cheaper storage.) Teradata’s Unified Data Architecture™ was used to first pull all the Web logs from Hadoop™, and interactive voice response (button push sequence) logs, as well as agent interactions and text notes from a variety of other systems including Oracle and MS-SQL into Teradata Aster. Organizing this information from multiple sources into a time sequence for each failed customer episode allowed them to build the database of relevant cases of interest (e.g., one or more Web clicks followed by an immediate call to customer care).

With this database in hand on Teradata Aster, their data scientists could then start analyzing which Web pages were viewed right before a call to customer care and figure out, by working with product managers, how to improve those pages so customers can successfully self-serve.

#### As a result the provider can:

- ~ Identify customers impacted by “breakage”
- ~ Determine problem spots on the Web that are driving dis-satisfaction, even churn
- ~ Work with product managers to fix the guilty Web pages
- ~ Take cost out of the contact center (from fewer calls)
- ~ Provide a more holistic view of customers to both the Web and the customer care group

#### This was achieved by using:

- ~ Teradata Aster Database for time series analytics on the hybrid sequence of Web sessions followed by calls to customer care
- ~ Hadoop™ storage of Web logs and transformation into sessions
- ~ Connectors to contact center IVR logs and agent text notes
- ~ Teradata Integrated Data Warehouse to provide customer valuations and satisfaction scores

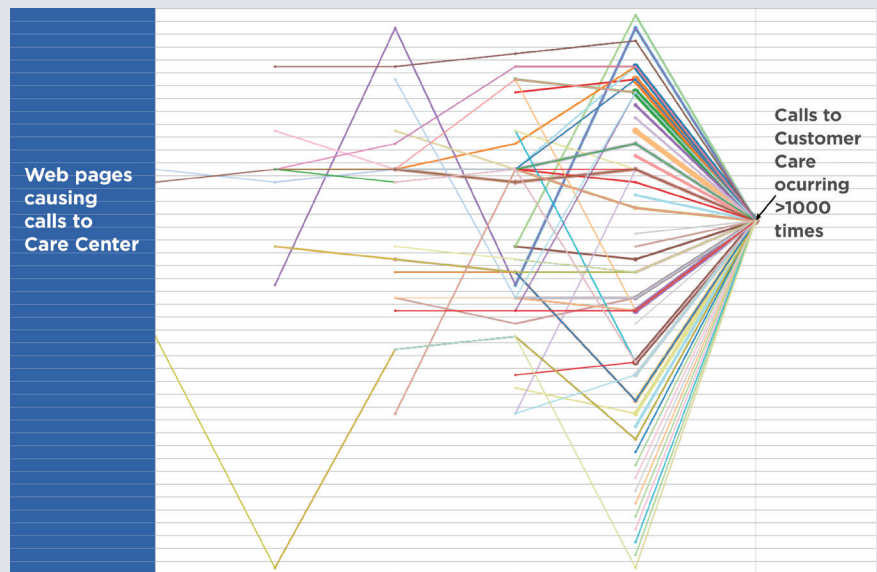


Figure 5. Web pages customers visited before contacting call center.

## CONCLUSION

Teradata is the best choice for companies that are serious about using data to build a competitive edge. The Teradata Unified Data Architecture™ is the first to unify all forms of data into an architecture that enables business insights from across the enterprise.

To learn more about how Teradata can help you architect a custom solution designed to meet your organization's needs and derive the greatest value from traditional and new data types, please visit **Teradata.com**, or call your Teradata representative.

## ENDNOTES

1. Infographic: Big Data Brings Marketing Big Numbers, Marketing Tech Blog, May 9, 2012, <http://www.marketingtechblog.com/ibm-big-data-marketing/>
2. IDC Digital Universe study, "A Universe of Opportunities and Challenges," December 2012 <http://www.emc.com/leadership/digital-universe/iview/executive-summary-a-universe-of.htm>
3. A Focus on Efficiency, A Whitepaper from Facebook, Ericsson and Qualcomm, September 16, 2013, [https://fbcdn-dragon-a.akamaihd.net/hphotos-ak-prn1/851575\\_520797877991079\\_393255490\\_n.pdf](https://fbcdn-dragon-a.akamaihd.net/hphotos-ak-prn1/851575_520797877991079_393255490_n.pdf)
4. Library of Congress, Web Archiving FAQs, December 2013, <http://www.loc.gov/webarchiving/faq.html>
5. YouTube Statistics, December 2013, <http://www.youtube.com/yt/press/statistics.html>



10000 Innovation Drive Dayton, OH 45342 [teradata.com](http://teradata.com)



TERADATA

Unified Data Architecture and SQL-H are trademarks, and SQL-MapReduce, BYNET, Teradata, Aster and the Teradata logo are registered trademarks of Teradata Corporation and/or its affiliates in the U.S. and worldwide. Hadoop is a trademark of the Apache Software Foundation. Hortonworks is a registered trademark of Hortonworks, Inc. Intel and Xeon are trademarks of Intel Corporation in the U.S. and/or other countries. Teradata continually improves products as new technologies and components become available. Teradata, therefore, reserves the right to change specifications without prior notice. All features, functions and operations described herein may not be marketed in all parts of the world. Consult your Teradata representative or [Teradata.com](http://Teradata.com) for more information.