

From Big Data to Knowledge: Value Chain for Communication Service Providers

Big Data and advanced analytics create an opportunity for communications service providers (CSPs) to cultivate intelligence for operating networks more efficiently, analyze the success of their services, and create a better personal experience for customers. This white paper analyzes the complete value chain that can transform CSPs' data to knowledge. It covers the sources of information, data collection tools, analytics, and finally, the most prominent business use cases emerging among CSPs and telecommunications companies.

Table of Contents

page

Introduction

1

From Data To Knowledge—Better Business Decisions

2

Data Sources

3

Data Collections

4

Vertica: An Advanced Analytics Platform for CSPs.....

7

Information collected from data sources: caller, called party, jitter, packet loss, latency, call duration, setup time, codec, throughput, mobile ID (IMSI, IMEI), phone number, user login, IP address, MAC address, date and time of login/logoff, subject of email/chat/webmail, sender, receiver, attached documents, response time, data transfer sessions, visited website, time spent on visit, basket share, referent, and so on.

Introduction

Society becomes more connected daily; SMS, MMS, video on demand, email, and basic mobile conversations are a few examples of the countless ways people interact electronically.

Over the years, your large communication network and their associated switches, billing systems, and service departments may have generated hundreds of millions of individual call detail records (CDRs) daily. Terabytes of dynamic customer data will continue to grow exponentially as carriers add new services and as IP-based traffic increases.

Why Now?

What we have seen in the last five years is an evolution of advanced analytics technologies that can now turn the network and communication data explosion into an opportunity for business transformation and monetization. The analytics equation described in figure 1 has made possible solutions and approaches that were impossibly expensive and complex just a few years ago.

- The cost of processing data and doing analytics on it in real time has gone down, making it increasingly possible for live data to trigger system actions, rather than just generate static reports for human consumption.
- The commoditized access of the information from multiple systems, network elements, sensor data that could outstrip all others, and Internet data that have equal value to internal data, has facilitated the explosion of the data you can analyze.
- The cost of data storage has fallen dramatically. This, coupled with the application of new data processing techniques, is enabling far larger datasets to be captured, stored, and analyzed.
- New sources of user data and the urge to share from smartphones and social networks have come on-stream, enabling a multidimensional view of the customer.
- Data can explain behavior interest—many operators are able to apply statistical techniques, complex relationships across data and graphs, social network analysis, and their own business problems.

From a strategic point of view, you want to be able to connect with customers with a full understanding of their needs, detect business trends and opportunities, timely detect fraud, and comply with regulatory requirements.

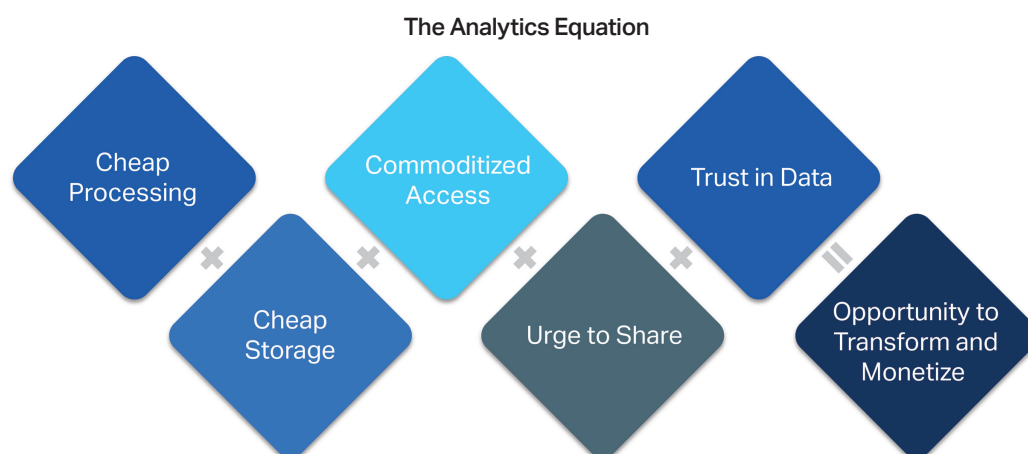


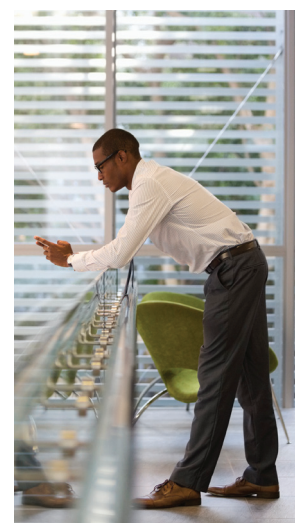
Figure 1. New advancements create opportunity for business transformation and data monetization

From Data To Knowledge— Better Business Decisions

To achieve better business decision, CSPs need to consider the complete value chain that can transform their data to knowledge, bringing all components together in one end-to-end solution. It includes (see *Figure 2 on the following page*):

- **Data sources:** From network information, billing systems, subscriber profile, devices, or social networks
- **Data collections:** Including different technologies such as a network probe that captures the data
- **Data management and structuring:** The heart of your business knowledge, analytic databases (ADB) provide fast access to that data
- **Data access:** Enabling query sessions to be truly interactive and available to various internal teams and users
- **Business intelligence:** The analytics process applied in a number of CSPs' specific use cases
- **Presentation and visualization:** Proposing predictions and displaying results to staff in a usable format

Let's analyze each of these components in more detail.



You can meet compliance deadlines, improve lifecycle analysis, prevent fraud, extend appropriate offers to customers in real time, provide detailed data analysis, and yield pattern-improving results to the customers—all of this, just from an improvement in query performance time.

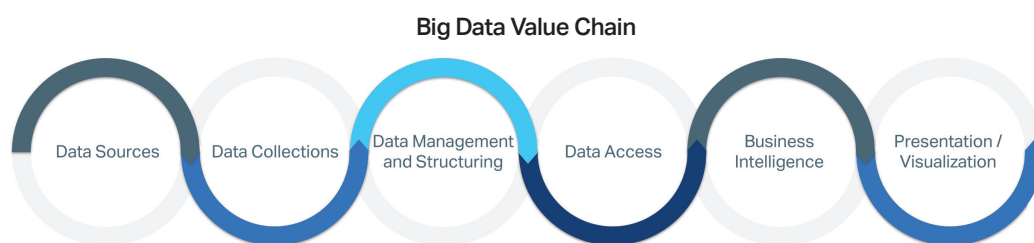


Figure 2. Complete value chain needed to turn Big Data into knowledge

Data Sources

Success for a Big Data industry solution strategy lies in recognizing the different types of Big Data sources, using the proper mining technologies to find valuable information within each type, and then integrating and presenting those new insights appropriately according to your unique goals. Understanding these steps is necessary to enable your organization to make more effective business decisions. The different sources of information for CSPs include:

- **Network usage:** Such as CDR or Internet Protocol Detail Record (IPDR) and information from business and operational support systems.
- **Sensors:** The global market for wireless sensor devices used in end vertical applications reached nearly \$1.2 billion USD in 2013. This market is expected to increase at a 23.7 percent compound annual growth rate (CAGR) and reach an estimated \$4.3 billion USD by 2019.¹
- **Connected devices:** According to Gartner, IoT devices will encompass more than 6.4 billion connected objects in use by 2016, and further explode to 20.8 billion things by 2020.²
- **Apps:** Information from the number of applications available along with the marketing around those apps is expected to rise. The number of apps downloaded worldwide is expected to reach over 268 billion in 2017, dramatically increasing the amount of information and marketing associated with them.³ Subscriber profiles come from network systems such as home location register/home subscriber server (HLR/HSS) or provisioning or CRM.
- **Services:** Where we are engaging with a service to buy or sell something; ultimately creating an opportunity to analyze behavior, target a pattern, and so on.
- **Social network profiles:** Tapping user profiles from Facebook, LinkedIn, Yahoo, Google, and interest-specific social or travel sites, to cull individuals' profiles and demographic information, and extend that to capture their hopefully like-minded networks.

¹ *Wireless Sensors: Technologies and Global Markets,* report code: IAS019B, BCC Research, August 2014.

² "Gartner: 21 Billion IoT Devices To Invade By 2020," *InformationWeek*, November, 2015.

³ "Forecast: Number of mobile app downloads worldwide from 2009 to 2017 (in millions)," *Statista*, 2016.

Data Collections

Network probe is a technology that decodes protocols, and extracts information embedded in the traffic or transmitting over the traffic. Then it delivers this information in the form of metadata and content feeds to an application developed by the user that leverages the information provided by the network probe.

The probe makes an acquire specifying what information is required. The network probe delivers this information in a tabular format, just like in a database to a storage engine. This technology can also deliver packets and packets contents. The process of extracting and delivering the information from the network is done in real time and scale up to 10 Gbps. Different protocols are supported such as: network protocols, and application protocols such as webmail, email database, or any kind of network application. For each protocol, tens of metadata are delivered, which make at least thousands of metadata in your application. These protocols are regularly updated and new protocols are added to the protocol plug-in library.

Network probe intelligence is designed to be embedded into your application, so you can rely on the real-time visibility provided by the network probe to develop application, processing the traffic information, or storing this information, for some reporting or traffic shaping.

Data Management and Structuring

Analytic databases provide fast access to large sets of data, allowing business analysts to drive much deeper into root cause analysis and delivery greater insights than would not be permissible with a traditional database. Many analytic databases differ in the way the data is stored on disk. In those that have adapted a “columnar” orientation, disk files are occupied by the values of a single column, instead of complete rows. This physical division allows more frequently used data to be assigned to faster storage tiers. It also warrants that columns not pertaining to a query are excluded from access. This “column elimination” results in improved (sometimes dramatically improved) performance for an important class of query in an enterprise. The clustering of similar values in a column also makes columnar databases much more disposed to a new class of compression capabilities. Column elimination and compression help to alleviate the I/O problems that have been plaguing analytic queries for years. Techniques include:

- Run-length encoding, whereby column values that repeat across consecutive rows can be stored once
- Dictionary compression, which abstracts the real values and stores only tokens in the record
- Delta compression, which stores only offsets from a set value
- In addition, some analytic databases, such as the one Vertica offers, are “hybrid” in the way they store data, allowing for multiple columns to be stored in a single disk file. When you access columns together, you could place them in the same disk file. This would streamline the process at the end of the query where the result set columns are brought together for presentation. When a table has a large number of columns, like CDRs do, it is frequently a candidate for effective multiple-column disk files.



The number of subscribers to mobile, fixed-line, and cable communications services is growing by millions of people every year.

Data Access

The vendor's usual statement for analytic query performance seems to be one of indifference or that the users are not fully exploiting their existing technology. The only real answer to the problem, if you are sticking with the existing stack, is more hardware. However, the reality is that most systems are already fully optimized for the hardware in place. If there were a way to increase query performance in business intelligence without resorting to more hardware, would allow revolutionary leaps in information dissemination in multiple ways:

- Enable analytics to be truly interactive and responsive, not limited by poor performance that causes analysis to stop at three interactive queries instead of the 10, 20, or 100 needed to achieve actionable insight into business problems
- Facilitate rollout of the analytic environment to all knowledge workers of the company as well as customers, supply chain partners, and other potential users of the data
- Allow for years of history to be kept, knowing that high-volume data can be queried without significant performance constraints
- Add the possibilities for CDR, text data clickstream, and other volume-intensive data to be kept for a more detailed level of analysis
- Add the possibilities for analysis of complex data types such as flat files, XML, graphics, and spreadsheets

Business Intelligence

Analytics brings together statistics, operational research, and computational knowledge to solve business challenges by analyzing data held within different business environments. From patterns in data, you can use statistical methods to create models and algorithms that forecast future events. The analytics process can be described as two distinct activities: modeling and prediction.

The modeling starts with the process of mining data to identify patterns and relationships, with the intention of explaining a set of actions or of looking for anomalies that identify a sector or cluster. After patterns and relationships have been identified, a model is created that describes the behavior, for example the likelihood of customer churn, the impact of video production on network bandwidth, the likelihood of product adoption through new bundles, and so on.

The frequency with which the model is run depends on the application, computational power, the amount of data, and the complexity of the model. There may also be limitations on how quickly new data is fed from source data points. With the advent of more powerful database technologies such as Vertica, the time required to run an analytics model is decreasing.

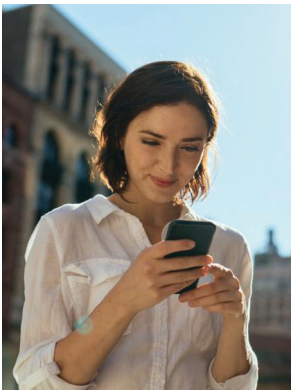
Analytics tools have a generic capability that can be applied in a number of different use cases (see figure 3). Some of these use cases are already widely understood, but in the majority of them the unique nature of the data, requirements, and processes demand customized implementation.

CSP Analytics Use Cases			
Sales and Marketing	Network	Products and Services	Supplier
<ul style="list-style-type: none">■ Partner analysis■ Channel analysis■ Campaign analysis■ Customer insight analytics■ Sales analytics■ Social network analytics■ Customer segmentation■ Customer network analysis■ Customer churn analysis■ Customer cross-sell and upsell■ Customer lifetime value■ Customer profitability■ Customer acquisition	<ul style="list-style-type: none">■ Capacity management■ Performance management	<ul style="list-style-type: none">■ Performance analysis■ Margin analysis■ Pricing impact and stimulation■ Cannibalization impact of new services■ What-if analysis for new product launch■ Impact of supply chain	<ul style="list-style-type: none">■ Cost and contribution analysis■ Quality control■ Regulatory requirements■ Fulfillment analysis■ Quality analysis■ Roaming analytics■ Settlements
Customer Management	Operational and Resource Management	Enterprise	
<ul style="list-style-type: none">■ Customer churn (retention)■ Customer experience■ Service level agreements■ Credit scoring■ Customer retention analysis■ Customer problem case analysis	<ul style="list-style-type: none">■ Operational analysis of key processes:<ul style="list-style-type: none">– Mean time from order to cash– Trouble ticketing resolution– Performance management■ Facility profitability analytics	<ul style="list-style-type: none">■ Fraud management and prevention■ Revenue assurance	

Figure 3. Big Data drives value across a range of CSP use cases

Visualization

Presentation and visualization functions allow your teams and automated systems that require predictions and results to receive them in a usable format. Traditionally, delivery has been to a staff member who then acts on it manually. But this is increasingly being automated, as actions are required in near-real time, including the use of customer data for real-time personalized on-device customer interaction. In this case, the CSP and subscriber jointly manage their relationship through direct engagement on a mobile application to strengthen customer intimacy, enhance experience, and open new revenue streams.



Vertica: An Advanced Analytics Platform for CSPs

The time for database innovation in telecom is now. The number of subscribers to mobile, fixed-line, and cable communications services is growing by millions of people every year. And the volume of CDR, IPDR, subscriber profile information, network probe, and machine-to-machine data that communications companies must store and analyze is growing year over year.

A Platform Architecture Built for Telecom Analytics

Together, the key features of the Vertica Analytics Platform create an elegant architecture for high-speed, cost-effective analysis of large volumes of CDR, SNMP, IPDR, network probe information, subscriber profile information, and machine-to-machine data. Partitioning by column, extensive use of compression, and hybrid storage model—all these Vertica features reduce the I/O required to execute queries. Telecom records such as CDRs contain many columns per customer, but individual reports cull only a few of these. The types of analytic applications that use a small, arbitrary subset of columns per query are both common, and ideally suited for the vertical partitioning provided by the Vertica Analytics Platform. At the same time, Vertica data partitioning, which divides work across multiple nodes in a computer cluster, supports the very large data volumes to which these applications often grow. These analytic applications also typically support users from many disciplines, interrogating the database from multiple individual perspectives. Here, Vertica's ability to keep multiple physical projections of the data is a natural fit for such usage patterns.

Enabling CSP Data Analytics

The Vertica Analytics Platform enables communications service providers to analyze and make informed decisions in near-real time with unparalleled efficiency, performance, and scalability.

Our communication service provider customers and partners routinely address the following challenges:

- Determine behaviors that may ultimately lead to either customer or product churn
- Understand customer experiences at a transactional level and determine investment criteria
- Improve offerings and portfolio in a highly competitive market targeting high-value, high-margin infrastructure and applications based upon empirical data
- Analyze volumes of data in near-real time that eclipse the capabilities of legacy infrastructures

Leveraging the Vertica Analytics Platform, our CSPs' customers and partners derive benefits relating to capacity management, performance, scalability, and availability. A few examples include:

- Deliver significantly higher customer satisfaction, retention, and profitability
- Optimize portfolio by focusing on alleviating high-cost, low-value products and services
- Manage and scale portfolio dynamically, without sacrificing details of any individual customer, transaction, or product
- Store, access, analyze, and monetize the vast amounts of customer and network data without sacrificing time, scale, or detail



Vertica Customer Case Study: Anritsu

Anritsu provides products and services for the development, manufacturing and maintenance of a range of communication systems for mobile phones and internet connectivity. Anritsu is now leveraging its experience with service assurance technologies to develop solutions for 5G, M2M, IoT, and other wireline and wireless communication markets.

CUSTOMER CHALLENGE

Anritsu's customers now consider waiting several minutes to receive critical diagnostic analysis as unacceptable. This meant that Anritsu needed a new analytical database that enabled the same data-intensive advanced analytics solutions the company currently provides, but much more quickly and at a comparable price to their legacy solution.

Anritsu is expanding its service offerings to address multiple IoT verticals with advanced analytics that focus on applications related to mobile asset connectivity, security, and asset behavior, and required an advanced data management and analytics solution to enable these offerings.

VERTICA SOLUTION

As the amount of data being produced by networks and assets increases and customers demand insight more quickly, open source solutions such as Hadoop or commercial appliance based solutions such as Oracle were no longer viable for Anritsu. Vertica allowed Anritsu to implement a scalable, cost-effective analytics solution that offers fast data access.

A mix of streaming and historical data is critical to Anritsu's development of predictive algorithms that train on historical data. Vertica has provided Anritsu with the technology necessary for the company to implement predictive analytics solutions that have only been theoretical until now.

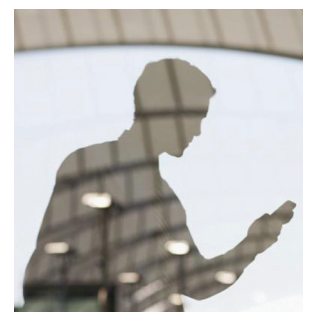
Anritsu has already realized rapid ROI after implementing Vertica in place of a legacy Oracle solution: 351% ROI with a payback of just 4 months.

[Read the ROI Case Study](#)

[Watch the Customer Video](#)

Learn More At
www.vertica.com

Try Vertica At
www.vertica.com/try



Additional contact information and office locations:
www.vertica.com

www.vertica.com