GIGAOM RESEARCH

How to deliver a comprehensive big data analytics framework to communication service providers

William McKnight

July 31, 2014

This report is underwritten by Guavus.

guavus

TABLE OF CONTENTS

Executive summary

The communications service provider (CSP) industry has undergone a dramatic shift in recent years. The traditional model of competing on subscription plans is no longer an adequate business strategy. Since most internal systems were built with this model in mind, these environments, with non-enriched, non-integrated, and latent data, fit for after-the-fact reporting, are struggling to keep up with the changes.

This research report will explain how CSPs establish a framework for their analytics as well as review the business drivers for telcos and the key benefits that big data analytics provide. It will also address the impact of the business drivers and the advantages of streaming analytics, combined with the ability to harness big data to meet several CSP competitive requirements. It will conclude by summarizing this comprehensive big data analytics framework for CSPs.

Key findings in this analysis include:

- Three main pillars epitomize the competitive environment in telecommunications today: stagnating or declining average revenue per user (ARPU), over-the-top (OTT) competition, and marginalized profits.
- The enriched, integrated, timely, and all-inclusive data known as analytical data forms the basis for telco competition today.
- Key benefits of big data analytics in telecommunications companies include more-accurate capex planning, the creation of new revenue sources, a reduction in opex, and more-precise marketing, upsell, and resell opportunities and the ability to improve the customer experience.
- Telco competitive requirements include customizing marketing messages, analyzing call-detail records, proactive equipment servicing, infrastructure investment protection, context-sensitive bandwidth allocation, and product development.
- Meeting the performance requirements necessitates a comprehensive big data analytics framework, which includes enterprise data warehousing, analytic databases, data integration, scale-out architectures, and real-time stream processing.
- Those telcos transitioning to SDN, SON, and NFV technologies will find that streaming analytics can support and optimize their investments.

Business drivers for CSPs

Three main pillars epitomize the competitive environment in telecommunications today.

Stagnating and declining average revenue per user (ARPU). ARPU has long been an important key performance indicator for a telco. The ability to sell deeper and across services within the customer base is vitally important. There have been recent line extensions into satellite television and into service areas like advanced user access control, content parsing, and voice-response units. Despite these additions, industry ARPUs have stagnated.

Over-the-top (OTT) competition. Users now own multiple devices and consequently use multiple access providers. To secure a position in a disintermediated landscape, a telecommunications company must deliver branded services as a third party.

Marginalized profits. Profits have come under intense pressure. Services with low value per occurrence, such as voice calls, are being commoditized at an alarming rate. In exchange for advertising exposure, text messages and phone calls can now be free to a consumer via third-party services. Telecommunications companies must compete using the insights generated by analytics to maintain profit.

Like other customer-sensitive businesses, telcos have an enormous opportunity to transform their companies into data-driven organizations. They sit on a wealth of data that can lead to innovative services and offerings. While some have launched analytics initiatives, most are still in the infancy stage and have more innovation coming that will involve closed-loop decisioning systems and automated workflows.

CSPs can use big data analytics to plan capex better, create new revenue sources, reduce opex, improve marketing effectiveness, and deliver an enhanced customer experience. Telecommunication companies can also be closer to their customers. Stickiness and farming more revenue from existing customers are vital for survival and profitability.

Examples of big data analytics that support telcos in a modern competitive arena:

- Advertisement catalog and efficacy
- Anticipated impact of network degradation on high-value customers
- Cell-site reach
- Customer categorizations such as promotion shopper
- Customer interests, explicit and imputed
- Customer movements
- Customer network of influence (on other customers)
- Customer price sensitivity
- Customer relationships to other providers and their users
- Customer satisfaction metrics, explicit and imputed
- Customer trouble ticket metrics
- Derived product customizations of interest by customer
- Extensive customer-device information
- High-traffic locations and network-congestion metrics
- Likelihood of uptake of a SIM offer
- Network-performance data
- Package-design efficacy by customer profile
- Predicted customer usage
- Propensity to churn
- Service-level agreement (SLA) conformance by customer
- Usage transactions by customer

Key benefits of big data analytics for the communications service provider industry

Some forward-thinking telco organizations have seized the opportunity with big data analytics. These companies have improved their access to clean and correct customer communication and outcome metrics, which has positively impacted the three pillars mentioned in the preceding section.

Consider the following benefits that big data analytics can support (along with corresponding processes) for a telecommunications organization.

More-accurate capex planning

Soaring capex levels represent an opportunity for making accurate projections and appropriations a priority. LTE deployments are a good example. According to some estimates, global capital spending on LTE infrastructure could exceed $36 billion next year. Operators need LTE because it improves spectral efficiency that enables them to transport the rapidly growing volume of mobile IP data traffic with greater cost-effectiveness than they could with 3G. Cisco estimates that by 2018, LTE network connections will generate six times more traffic on average compared to 3G connections. Those operators that are investing heavily in LTE can use analytics to deploy it with greater ROI.

LTE is already creating extreme traffic trends that are being driven by mobile video. Shortly after introducing LTE in 2012, Vodafone Germany reported that non-revenue-producing mobile video drove 85 percent of its LTE traffic, with moneymaking services using the remaining 15 percent. Vodafone's experience reveals the challenges of monetizing LTE for popular consumer services, especially in an overall environment of stagnating revenues.

LTE engineering KPIs + cost-intelligence analytics = better investment decisions

As these challenges demonstrate, operators must cultivate ROI-consciousness for every LTE decision they make, whether in planning, deployment, or optimization. An operational intelligence platform designed for big data can break down data silos to fuse and correlate key performance indicators (KPIs) for radio access network (RAN) and business conditions. The integrated analytics generate a holistic, cross-sectional view of the network and business in real time, enabling operators to make smarter and expedited decisions based on the ROI impacts of their RAN spending options.

Creation of new revenue sources

Telcos are incorporating current, detailed patterns of utilization and translating the findings into marketable new product development. While revenues from voice services continue to decline, telcos are looking for new growth areas. With huge market potential and demand stimulated by billions of connected devices, a world connected by the internet provides real opportunities for expanding service portfolios and increasing revenues.

Reduction in opex

If telcos are to excel in opex management, they must establish full cost transparency so that they can identify, prioritize, and optimize savings measures. A broad range of new opex savings possibilities can be leveraged through a new generation of technologies, such as software-defined radio (SDR) networks and self-organized networks (SON). Energy efficiency is an important issue when telcos are selecting network elements. The ability to know about a problem while it is happening and to figure out why it is happening can save a company millions of dollars each year.

Additional benefits of big data analytics

Analytics also:

- Creates consistent customer identification and case recognition as well as case management and contact history across all services
- Provides cost-to-serve reporting capabilities across all services for case-load tracking, inbound contact routing, and center efficiency to help agents and managers understand the fully loaded cost to serve a customer
- Creates a common layer allowing a consistent customer experience
- Leverages the core competencies from each service line, allowing the lines to focus on core business processes and systems
- Provides a presentation layer for the agents and customers across the company

Indeed, these are the results of the applications we strive for in telcos, but the underlying foundation of high-quality, integrated, well-performing, and corporately vetted information is usually found by successfully deploying analytics and big data strategic initiatives.

The preponderance of such a variety of data sources — perhaps the most diverse of any industry — and the speed with which they must be addressed can perpetuate the problem of data proliferation. The speed of change in the industry, often driven by regulatory bodies, is an additional factor. Change is almost guaranteed to continue, which is perhaps the most important reason for telecommunications companies to implement a big data analytical foundation as soon as possible.

Big data analytics

As businesses move beyond analyzing basic information for rearview-mirror reporting and move to seeing the business landscape with forward-facing advanced analytics, a game changer has emerged: big data analytics.

With the ability to explore previously unrealized correlations between certain metrics and attributes, streaming analytics data technology — and its combination with operational database technology — greatly increases the effectiveness of big data analytics. When streaming data is correlated, fused, and integrated with data at rest, analytics are enhanced with highly granular data points that open opportunities for taking telecommunications analytics into the realm of minute fine-tuning. Telcos now have an end-to-end understanding of their network at a very precise and granular policy level and with full context. These analytic insights can now be embedded into automated workflows and business processes.

The call-detail record (CDR), which can be utilized in variety of ways across a telecommunications company, is one important resource for a telecommunications organization; the complete record must be stored and made available across the company as part of a sound analytics strategy. Additional value comes from being able to analyze the variety of data and a cross-section of data sources in real time. This use can provide context for better decision-making.

CDRs provide comprehensive information on each and every call occurring on data circuits, including:

- Voice capture for both directions
- Complete signaling information for each direction
- Disruptive alarms and errors occurring during the call, including bipolar violations, loss of frame, cyclic redundancy check errors, and loss of signal
- Detailed voice band event information occurring during the call, including dual tones and more
- Detailed analysis of the call including noise level, speech level, speech activity factor, echo measurements, and more
- Categorization of the call as voice, fax, modem, or data

From data to insights across the CSP network

The telco industry is experiencing a paradigm shift about how networks are operated. Telcos have an opportunity to analyze data to improve network operations and planning, deliver more personalized offerings, enhance customer experience, and improve security offerings. Operations will be able to prioritize field services and maintenance based on where most customers are affected. Customer care agents will have useful and instant information about customer usage and experience to further improve awareness and time to resolution.

A big data analytics strategy can provide additional knowledge and information for better quality and timelier decision-making, empowering the company to be more proactive. Most business units in telcos will benefit from the data that will be available. This includes marketing, product management, sales, enterprise sales, the enterprise network operations center, network planning, and CSO.

Customizing marketing messages

Telco product portfolios are complex. Many cross-sell opportunities exist for the installed customer base. Sales associates use in-person or phone conversations to formulate next-best recommendations, often without the support of analytics.

An analytics approach gives sales and marketing the ability to make confident recommendations and create messages based on data from the entire network. Confident recommendations and messages empower the sales mechanism and improve customer interactions.

Proactive equipment servicing

Transmission towers and their related connections form the network of a telco. Failure of a transmission tower can cause severe service degradation. Since the replacement of equipment is usually more expensive than repair, getting ahead of the replacement cycle with analytics is imperative.

Unstructured, streaming sensor data from the network must be managed. Telcos can derive optimal maintenance schedules by comparing real-time information with learned good patterns from network data. Machine-learning algorithms can reduce both maintenance costs and service disruptions by detecting anomalies and using root-cause analysis both to identify patterns and to determine fixes.

Infrastructure investment protection

Planning ahead saves time and money for a telco. Consumption of bandwidth and services should be in sync with plans for new towers and transmission lines. Telco network operations and capacity planning are correlated as well.

Network-log data helps telcos understand service consumption in a particular location. They can then analyze network loads more intelligently, with data stretching over longer periods of time. This allows for infrastructure investments to be planned with more precision and confidence.

Context-sensitive bandwidth allocation

Telcos must respond quickly to bandwidth spikes — often unforeseen — so that they can reallocate resources and maintain service levels. Popular new applications, such as HBO Go or the latest Hulu, can consume significant bandwidth and reduce service quality for the rest of the network.

A big data analytics approach helps telcos visualize spikes in network data and adjust bandwidth. Text-based sentiment analysis on telecom network unstructured data can also help understand how these spikes impact customer experience. These insights help maintain service quality and customer satisfaction, and they also inform strategic planning to build smarter, more-dynamic networks, especially as telcos move more toward leveraging SDN, SON, and NFV technologies.

Product development

A large part of effective product development is understanding how current products are used in detail. Mobile devices produce data at a high velocity concerning how, why, when, and where they are used. An analytics approach stores all the data without limits and does so economically.

This puts rich product-use data, often tagged by geography or customer segment, in the hands of product managers, which enhances product innovation. Also, immediate feedback on product launches allows product mangers to reel in busts and push harder on successes.

Technology requirements to support a new class of analytically powered applications

In the previous section we discussed several business contexts in which telcos can utilize big data analytics technology to improve competitiveness, profitability, and efficiency. In this section we discuss specific technologies and architectural approaches that should be put in place to ensure a comprehensive big data implementation.

Enterprise data warehousing

For an increasing number of telcos, transformation from legacy voice-only telephony services to converged services — and therefore converged systems — is prompting widespread technological changes to break down silos of data.

There is still a big need for the data-warehousing concept in all telco environments. The idea of sharing the data, platform, model, methods, and tools across different data sets and subject areas brings many benefits. Sharing the data, as long as concurrency is not a technical issue, is highly beneficial.

The data warehouse offers a lowest-common-denominator approach to storing data, which is sufficient for the straightforward reporting workload. Data warehouses will continue to provide clean data to many data marts. Many marts are being built today, so a tremendous expansion of platform features in databases continues as these marts go searching for their best-fit platform.

The data warehouse clearly has efficiency benefits, but by itself, this technology is not sufficient to support this new class of analytically powered apps in a big data world.

Analytic databases

Analytic databases (ADBs) provide fast access to that data and allow the logical progression of telco analysts to drive much deeper into root-cause analysis and deliver more in their analysis window than would be permissible with a row-based 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 ensures that columns not pertaining to a query are excluded from access. This

column elimination results (sometimes dramatically) in improved performance for an important class of query.

Data integration

The new requirements telcos have for their data include unparalleled levels of performance, increased agility, and systems that scale with the workload, the levels of which are more difficult than ever to predict.

Data is seldom confined to one system. Architects may choose to minimize redundancy and utilize data virtualization, but undoubtedly many data elements will flow throughout the architecture, sometimes as-is and sometimes transformed for their new purpose. Being skillful in data integration allows enterprises to take advantage of data-platform innovations by putting workloads in their absolute best platform. Data integration truly holds the telco analytic environment together.

Data integration, clearly, involves moving data. It creates necessary data redundancy. The data integration style can be:

1. Extract, transform, and load (ETL)
2. Extract, load, and transform (ELT)
3. Extract, transform, load, and transform (ETLT)
4. Extract, transform, transform, and load (ETTL)

Where there are two transform steps, one of them is dedicated to data-quality transformations while the other is forming transformations to get the data to work in the destination schema (adding row-level metadata, time variance, history persistence, splitting data, filtering data, calculations, lookups, and summarizations).

Scale-out architecture

A scale-out architecture allows unstructured data to be managed. Scale-out architectures can keep up with the increasing data levels of telcos and do so relatively cheaply, eliminating the need for accurately predicting the data size at the outset. Summaries of the analytics are likely valuable to the telco enterprise data warehouse so interaction with the EDW will likely occur.

Scale-out architectures are premised on a large cluster of commodity-class servers, referred to as nodes. Whenever a node is placed in the cluster, it becomes fair game for data storage in the file system, and its activities will be recorded by a master node tracking the cluster.

Sharding — horizontal partitioning of a database to spread its data across nodes and even geographically across data centers — is important to telcos with global reach.

A scale-out architecture embodies the approach of taking a large problem, breaking it into smaller problems, addressing the smaller problems in parallel, and finally combining the output for a result set. For many telcos, scale-out architectures have enabled the full corpus of CDR data to be retained and analyzed.

The power of performance

Improved query performance allows for revolutionary leaps in disseminating information in multiple ways:

- Enabling query sessions to be truly interactive, not limited by poor performance that causes analysis to stop at 3 interactive queries (instead of 10, 20, or 100) to achieve actionable business insights
- Facilitating rollout of the analytic environment to all knowledge workers of the company as well as customers, supply-chain partners, and broad potential users of the data
- Allowing for years of history to be kept, knowing that high-volume data can be queried
- Adding up the possibilities for CDR, text data clickstream, and other volume-intensive data to be kept at the detail level for analysis
- Permitting analysis of complex data types such as flat files, XML, graphics, and spreadsheets

Improvement in query performance time alone can bring the following benefits:

- Compliance deadlines can be met
- Life cycle analysis can be improved
- Fraud can be prevented

- Appropriate offers can be extended in real time to customers
- Data analysis can be more detailed

The associated technology requirements for telcos must provide this power of performance. This will necessitate different technologies for different workloads.

Real-time stream processing

Real-time processing, in the telco context, refers to sub-millisecond in-memory processing of streaming data. Stream-processing systems are not data stores in and of themselves, since they don't actually store data. However, they are data-processing platforms.

Stream processing is used to process data in real time, handling complex multistream analysis when there are hundreds of thousands of data points (e.g., call segments or network pulses) per second. Stream-processed data is well suited to scale-out systems, given the high-velocity and typically unstructured nature of such data. Stream processing is often the only suitable choice for processing high-velocity and high-volume data effectively.

With workloads appropriately allocated to the enterprise data warehouse, analytical databases and scale-out systems held together by data integration and stream processing handling real-time data, telcos can have a comprehensive analytics framework with the power of performance.

An effective analytic framework must comprise numerous components in order to provide optimal, integrated results. Some solutions combine scale-out architectures, stream processing, and data integration abilities to provide high-performance systems. This is a comprehensive approach to the telcos' challenges.

Many of these challenges require real-time response. For example, if a high-volume subscriber is having a poor video feed, a telco has seconds to remedy the situation. It must be able to identify a real issue and its root cause, then allocate resources in order to address the issue.

Myriad challenges occur in the collection, correlation, fusion, and analysis of big data for a telco that many modern solutions manage. This includes converting formats and fusing multiple streams of data together.

Advanced solutions store data in a variety of formats for their applications, depending on the characteristics of the workload. The primary approach is the archetypal scale-out environment of the Hadoop Distributed File System (HDFS), which is no longer confined to hard disk storage.

Finally, there are in-memory caches with built-in intelligence that store the most recent or most commonly used data for dashboarding or queries.

Telco forensics dictate not only the data appropriate for an in-memory cache but also the time period over which to store the data in the relational database. This complicated formula requires context intelligence to be incorporated into any aspiring telco solution today.

Ultimately, these applications and the availability of data provide value to the telco. With a sound analytic framework, users and data scientists throughout the organization can engage with the data to detect new patterns. The telco can trigger actions based on the data: Do machine learning, detect anomalies, and set up tools to catch specific scenarios and deal with them in real time.

Built on the realities of the modern telco, an analytics approach addresses business drivers and aids telco competitiveness. Starting with the capture of the telco high-volume data stream, it comprises many of the technology requirements, including big data management, to support the new class of analytically powered, domain-specific applications.

Key takeaways

- The three main factors that characterize today's competitive environment in telecommunications are the decline in ARPU, OTT competition, and marginalized profits.
- LTE engineering KPIs plus cost-intelligence analytics equal better investment decisions.
- Big data analytics is the basis for today's telco competition.
- Big data analytics offer CSPs key benefits including: accurate capex planning, the creation of new revenue sources, reduction in opex, more precise marketing, upsell and resell opportunities, and the ability to improve the customer experience.
- Telco competitive requirements include customizing marketing messages, analyzing call-detail records, proactive equipment servicing, infrastructure investment protection, context-sensitive bandwidth allocation, and product development.
- Meeting performance requirements necessitates a comprehensive analytics framework that includes enterprise data warehousing, analytic databases, data integration, scale-out architectures, improved query performance, and real-time stream processing.
- Those telcos transitioning to SDN, SON, and NFV technologies will find that streaming analytics can support and optimize their investments.

About William McKnight

William McKnight takes corporate information and turns it into a bottom-line producing asset. He's worked with companies like Dong Energy, France Telecomm, Pfizer, Samba Bank, ScotiaBank, Teva Pharmaceuticals, and Verizon — 17 of the Global 2000 — and many others. McKnight Consulting Group focuses on delivering business value and solving business problems utilizing proven, streamlined approaches in information management. His teams have won several best practice competitions for their implementations.

McKnight has hundreds of articles and 25 white papers in publication and is a prolific blogger. He is also the author of "Information management: Strategies for Gaining a Competitive Advantage with Data." He is a frequent international keynote speaker and trainer. He provides clients with action plans, architectures, strategies, complete programs, and vendor-neutral tool selection to manage information.

An Ernst & Young Entrepreneur of the Year finalist and frequent best practices judge, McKnight is a former Fortune 50 technology executive and database engineer. He has taught at Santa Clara University, UC Berkelely and UC Santa Cruz.

About Gigaom Research

Gigaom Research gives you insider access to expert industry insights on emerging markets. Focused on delivering highly relevant and timely research to the people who need it most, our analysis, reports, and original research come from the most respected voices in the industry. Whether you're beginning to learn about a new market or are an industry insider, Gigaom Research addresses the need for relevant, illuminating insights into the industry's most dynamic markets.

Visit us at: research.gigaom.com.