Network Analytics: Turn Big Data into Big Opportunity

Seven Steps for Network Operations, Marketing, Customer Care and IT
Introduction
Mobile communications continues to evolve at breakneck speed within the telecommunications industry. As competition increases, organizations need to focus on ensuring the best possible customer experience while managing costs related to maintenance and improvement of the network, and a burgeoning portfolio of third-party content and interconnected partners. Mobile data traffic congestion, increases in signalling traffic from the explosion of always-on smartphones, and an inability to proactively invest in timely network optimization, can lead to customer churn and missed marketing and revenue opportunities.

Huge volumes of structured and unstructured data are amassed on a daily basis from the network, customer usage, revenue interactions with customers and business partners, and a host of other operational and business support systems. The telecom industry has entered the world of big data in a big way, and there is no more looking back.

Big data represents a potential gold mine for those looking to gain better insight into customer relationships and operations, and innovate in an increasingly competitive industry. But how can organizations manage huge volumes of loosely structured and distributed data from multiple, disparate and dissimilar sources, and fulfill the need for timely and on-demand access to information and insight demanded by users?

The following seven steps are critical for delivering big data projects that can be undertaken to turn the flood of data into an advantage, and truly understand one’s network and one’s customers, while turning up profitability.

Step 1 – Break down big data strategy into manageable objectives
A key challenge of big data-style strategy is to get meaningful information from data that has traditionally been too large and cumbersome to view, manipulate or be handled by traditional technology. By clearly defining measurable objectives for the project, managers can start to define the appropriate scope, internal support, funding and execution for the project.

As an example, a key business strategy may be to improve the customer experience and reduce churn across prepay and postpay subscribers. Understanding customer behavior is key, particularly in the case of prepay where little demographic data is available. This strategy can then be broken down into the following tactical objectives:

- **Customer Care:** The entire customer base should be segmented according to behavior – changes in behavior can be flagged, and more detailed profiling across all customers can be achieved.
- **Marketing:** Profiles of customer behavior should be better understood to help define marketing campaigns targeting a specific audience.
- **Network Operations:** Combinations of handset device failures and service and network location issues should be more closely tracked and correlated to specific customers.

Support for the above tactics will likely depend on the IT or network department defining an appropriate and cost-effective infrastructure that allows timely access to data, and to provide the ability to measure the effectiveness of various initiatives.

Each of the above tactical objectives can be supported by the integration and correlation of between five to ten distinct data source types, including transactional usage data (voice, data, clickstream, etc.); network operational and performance data (including signaling, DPI, etc.); customer information (class of service, demographic information where available, etc.); network inventory and costs (trunk group information, traffic factors, rates, etc.); and reference data such as handset specifications, product/service definitions, etc.

By creating a single data repository from these distinct data sources to support one tactical objective, users can gradually phase-in other data source types to support additional or complementary tactical initiatives. The key is to load and store all transactional and operational data at its most granular, most raw form, to support full access to data from the analytical layer. Also, an architecture must be defined that can be
extended or evolved to support ever growing volumes of data, and users with different requirements to data access, immediacy and access rights.

**Step 2 – Automate as much as possible**

In order to derive the greatest value and agility from analyzing telecom big data, the automation of as many decision-making processes as possible becomes critical. The key is to understand which decisions in an organization can be automated, versus those that are truly best left to manual assessment and intervention.

Exception and issue handling – which is where the bulk of day-to-day decisions are made – must have a configurable workflow framework, and must allow for automated and manual actions and escalation processes. By leveraging the use of automation as much as possible, and where appropriate in exception handling, organizations can evolve towards more proactive management of the network, services and customers. For example, when a customer experiences more than four service failures within a two-hour span, an automatically generated Short Message Service (SMS) can be sent to their handset with an offer based on a predefined retention campaign for this scenario. Additionally, if the exception is determined to be related to the combination of a customer’s handset device and the service they are trying to access – perhaps related to an older version of firmware – an Over-The-Air (OTA) firmware upgrade can be pushed to the handset, and once the customer has activated this, a SMS can notify them that the service should now be accessible.

Outside of exception handling, day-to-day system tasks related to data loading, transformation, processing and trending must be as automated as possible. Power users such as analysts must have the ability to access data sets on-demand without having to revert to one’s IT department. Providing users with this agility significantly improves time-to-value from big data, and also has major implications on the underlying architecture and data storage decisions one makes. Longer, cost-effective storage of online data, and seamless access to offline data, reduce effort and time to analysis, and accelerates decision-making capability across an organization. The use of thresholds and alerts can automate tracking and monitoring of a big data ecosystem. Monitoring behaviors over weeks and months can show organizational analysts how to spot trends and flag them when the system diverges from the expected behavior. It can also pinpoint areas where additional automation can be implemented to increase efficiencies. Managing an organization’s big data can become a self-tuning, self-sustaining process with the judicious use of thresholds and alerts.

**Step 3 – Leverage experts and prototype where needed**

Successfully delivering insight on big data will require organizations to define, implement and empower a cross-functional team made up of people who know the data, people who know the systems architecture and tools, and people who know what output is required. Big data projects must be focused on the delivery of business value in the short-term, not by merely the delivery of a successful project. By focusing their efforts on solving one tactical objective at a time (see Step 1), an environment of collaboration can be created for the intense but short time span of the project.

**High-level roles and responsibilities can be divided into the following:**

- **End user representatives:** Responsible for defining detailed requirements on the output required, including metrics, summaries, aggregations and visualization for the tactical objective. Also responsible for defining the data storage requirements for the analysis, and defining the workflow parameters and exception/issue handling.

- **Data owners:** Responsible for defining the data-sets that are required to support the analysis and tactical objective. They work with systems and tools owners to define how that data is acquired and staged, and define the processes and exception handling for data acquisition.

- **System and tools owners:** Responsible for defining the architecture and tools needed to address the tactical objective based on requirements from the above team members. They define the logical and conceptual integration required to handle the distributed data sources, and define the processes for monitoring the system and exception handling.
By focusing on delivering one objective at a time, project times can be shortened from years to months and time-to-value emphasised and accelerated.

But what if there is no one in an organization who truly knows the data or who can be freed up from their day-to-day jobs to assist? There must then be an agreement to conduct an incremental approach to “run it and see,” that is:

• Load the data and understand the results
• Enrich and report on the data
• Define thresholds and see if they work
• Repeat until satisfied

The above leverages the concept of rapid prototyping: build a model, test it and then tweak it. This is an agile methodology that is well suited to the fast-moving analytics big data requires, and one that delivers ‘usable code’ according to business priorities. The more data that is available and worked by this methodology, the more it leads to answers that can be found to known problems. This results in the uncovering and answering of more new problems. All of this requires a constant effort to focus, fix, automate and find the next new problem. This iterative, focused approach incrementally adds new solutions while improving one’s overall system health and output.

**Step 4 – Get the right equipment and tools**

Due to traditional infrastructure limitations, most organizations have had to choose either big data (huge volumes of a wider and deeper set of data with additional dimensions and a richer history) or big math (analytics without constraints – increased computational intensity supporting complex models). But getting the right insight out of big data requires big math, and the key to delivering big data with big math is the elimination of complexity from the analytics infrastructure, while leveraging an architecture designed from the ground up to enable complex analytics at petascale volumes.

The opportunity in big data is unlocked from near real-time user access to data through efficient data acquisition, interpretation, visualization and storage.

For those interested in capitalizing on big data, data acquisition must move from the traditional ETL (Extract, Transform, Load) approach to an ELT (Extract, Load, Transform) model based on MPP (Massively Parallel Processing) technology, where transformation of data is done in the database. In-database transformation and processing provides two clear benefits in dealing with big data:

• Data is stored in as raw a form as possible, allowing for extreme flexibility and agility in analysis, and providing summary, aggregate and detailed level granularity.
• Data is loaded leveraging the power of advanced MPP database technology designed to handle huge volumes (terabytes/petabytes) of data daily, allowing for analysis to be done in “user-time” as opposed to “system-time.”

The organization of the data in a comprehensive data model is critical to achieving rapid and accessible analytic insight. Data interpretation runs on top of this data model to map the required data elements to the output that is desired. For a truly flexible system, it is recommended that one's underlying architecture be able to provide semantic views to the data model which can support interpretation by third-party analytical applications. This will help to extend the reach of the data as well as provide an internal or external query tool directly running on the big data set to accommodate ad-hoc queries. This hybrid approach facilitates the spread – and accelerates the use – of the same data throughout an enterprise.

Storage of on-line and off-line data must be carefully constructed so as to allow for the most seamless access to data for analysis, balanced against storage costs. One of the advantages of MPP technology is that more data can be retained online. Most vendors have virtualization capabilities that permit on-demand access to off-line data stored in a lower cost, distributed architecture.
Step 5 – Get the right data and get all of it

The variety, velocity and volume of data sources within a telecom environment – network devices, smartphones, tablets – can be daunting when attempting to define a big data project. By scoping the delivery down to one tactical objective, it is easier to identify those data sources which provide the most value for an initial return.

Getting the data right is key, but getting ALL the data is what is required. It is the individual, transactional-level detail that provides the ultimate view to what has happened. Systems that don’t or can’t use details default their users to sampling, inferences, extrapolations, “industry averages” or gut feel. Big data handled correctly drives data-driven decision making.

To further expand on the example objectives outlined earlier, the table above contains suggestions of the minimum detailed data sources required.

The commonality of data sources shared amongst the objectives reinforces the importance of accessibility to network and enterprise data that transcends departments and roles. The “load once, use many” approach using detailed data, reduces costs and drives efficiency by having a single data repository being accessed by all users, regardless of operational group. Everyone has and works from the same numbers – leading to accurate, holistic, timely and data-driven decision making.

<table>
<thead>
<tr>
<th>User Group</th>
<th>Tactical Objective</th>
<th>Minimum Required Data Sources</th>
</tr>
</thead>
</table>
| Customer Care | Segment the entire customer base according to their behavior, flag changes in behavior which could lead to churn, and understand true profitability for improved profiling | • Customer information  
• Transactional usage (voice, messaging and data)  
• Application server data  
• Signalling data  
• Handset specifications  
• Product reference data  
• Cost and revenue data |
| Marketing | Understand root causes of customer behavior to define marketing campaigns with improved uptake | |
| Network | Understand which combinations of handset device, service and network location fail the most often | • Inventory data  
• Transactional usage (voice, messaging and data)  
• Application server data  
• Signalling data  
• Handset specifications  
• Customer reference data  
• Product reference data |
Step 6 – Empower the analysts to do Big Math

True business value in big data can only be derived with the application of Big Math - the complex computational processing that allows near real-time interpretation and output of information from data based on advanced analytics. An organization’s analysts are the power users of applications – they derive the business value out of big data – and therefore require the capability to apply Big Math to:

• Quickly translate information into action
• Have access to all data that is required – on demand
• Be able to ask questions they didn’t think they could ask
• Get the answers in a timeframe previously unattainable

Big Math can be characterized by the following capabilities:

Inform and decide: Multiple options for visualization are needed in order to quickly and efficiently organize big data and make it accessible and useful for users. Visualization should match the objective (see Steps 2 and 4) and seek to reduce the time to pinpoint the problem. The capability to drill-down from within the application from the visual start (e.g. charts, graphs, etc.) to the underlying detailed data (rows on reports) is critical for analytical agility and the identification of patterns necessary to optimize one’s business.

On demand analytics: In these days of big data and the requirement for near real-time analytics, waiting hours or days to load source data and generate the query or report is no longer acceptable. Analysts require pre-packaged analytics and interpretation models in order to quickly become productive. But to gain the most value out of big data, ad hoc, dynamic querying and data loading by the analyst becomes critical.

Flexible interaction with the data: User-driven analytics must also include the ability to go down different drill-down or query paths from what has been previously defined. The use of a flat and comprehensive data model, and the retention of detailed normalized data allows for the most flexibility in supporting previously unknown analytical requirements from a user base. Businesses need query tools that are highly configurable to drive more control into the hands of the users and away from vendor dependence. Big data and users of big data demand flexibility and timeliness in tools.

Rapid response technology: In-database analytics allows for faster time to response and faster time to analytical value by pushing analytical processing closer to the data source. Embedding data integration, analytical modelling and visualization into the data warehouse appliance cuts costs, accelerates time to productivity and business value, and democratizes data across an organization.

Step 7 – Open up the data to the enterprise

The use of leading-edge tools and an architecture capable of handling the volumes and sources required at the level of detail required – and the accessibility of a flexible query layer – all support the potential to drive the use of big data across an entire enterprise. As more users are exposed to the data repository – and begin to use it – the more efficient their organization becomes. This can result in a higher and faster the return on investment, as well as more timely management of customers, services, the network, and ultimately profitability.

Opening up big data across an organization means that decision making becomes data-driven for the whole enterprise. Insights from big data may lead one down unknown paths, but the justification and return for going there will be evident.

In conclusion

Big data represents challenges for any service provider but it also represents great opportunity for users willing to go the distance. Big data is the means by which an organization can get closer to its customers and their experiences on the network. New business models and relationships can be innovated from data coming from across the enterprise. The harnessing of big data combined with the application of Big Math represents a significant competitive advantage to those willing to take it on and willing to execute it correctly. The era of big data is upon us, it is up to you to turn it to your advantage.
About Ventraq

Ventraq® helps telecommunication service providers leverage big data to better profile their customers, define more targeted marketing, better manage quality of experience, network capacity and costs, and understand true profitability.

With domain expertise in telecom network, cost and revenue data and OSS/BSS solutions, Ventraq is uniquely positioned to help service providers get the most from their data. Partnering with market leaders in advanced MPP technology such as IBM® Netezza®, Ventraq provides in-database processing and analytics for Tier 1 service providers, processing petabytes of information daily on a single platform. Through its unique Business Analytics framework – encompassing a data acquisition, interpretation and a query layer in a single platform – Ventraq supports enterprise-wide use of the data by providing semantic views of the data accessible by other third-party analytical applications.

Ventraq Corporate Headquarters
502 Earth City Expressway
Plaza 1 Building, Suite 200
St. Louis, MO 63045

+1 314 770 9700
Americas: +1 856 866 1000
EMEA/APAC: +44 844 500 3900

www.ventraq.com

About IBM® Netezza®

IBM Netezza pioneered the data warehouse appliance space by integrating database, server and storage into a single, easy to manage appliance that requires minimal set-up and ongoing administration while producing faster and more consistent analytic performance. The IBM Netezza family of data warehouse appliances simplifies business analytics dramatically by consolidating all analytic activity in the appliance, right where the data resides, for blisteringly fast performance. Visit neterza.com to see how our family of data warehouse appliances eliminate complexity at every step and lets you drive true business value for your organization. For the latest data warehouse and advanced analytics blogs, videos and more, please visit: thinking.netezza.com.

IBM Data Warehousing and Analytics Solutions

IBM® provides the broadest and most comprehensive portfolio of data warehousing, information management and business analytic software, hardware and solutions to help customers maximize the value of their information assets and discover new insights to make better and faster decisions and optimize their business outcomes.