ERICSSON WHITE PAPER
284 23-3211 Uen | August 2013

ERICSSON

BIG DATA ANALYTICS

Successful decision-making will increasingly be driven by analytics-generated insights. From the lowest-level network enablers to high-level business support systems, there will be an opportunity to utilize insights gathered from careful analysis of network data.

Big data is at the core of this opportunity. But the traditional view of big data is not enough. Rather than focusing exclusively on what technology big data brings, we have to look at what value it can create.

This paper explores big data analytics, the opportunities it gives rise to, and how big data should be expanded to support analytics.

INTRODUCTION

People and devices are constantly generating data. While streaming a video, playing the latest game with friends, or making in-app purchases, user activity generates data about their needs and preferences, as well as the quality of their experiences. Even when they put their devices in their pockets, the network is generating location and other data that keeps services running and ready to use.

In 2014, estimates put worldwide data generation at a staggering 7ZB [1], and by 2018 each smartphone is expected to generate 2GB of data every month [2]. At the same time, the big data technology and services market is expected to grow at a 40 percent compound annual growth rate (CAGR) – about seven times the rate of the overall ICT market – with revenues expected to reach USD 16.9 billion in 2015 [3]. Clearly, the age of big data has begun.

Communication service providers (CSPs) can make use of this big data to drive a wide range of important decisions and activities, such as: designing more competitive offers, prices and packages; recommending the most attractive offers to subscribers during the shopping and ordering process; communicating with users about their usage, spending and purchase options; configuring the network to deliver more reliable services; and monitoring QoE to proactively correct any potential problems. All these activities enable improved user experience, increased loyalty, the creation of smarter networks, and extended network functionality to facilitate progress toward the Networked Society.

The profound impact that increased broadband networking will have on society will also create business opportunities in new areas for CSPs. With improved real-time connectivity and data management comes the possibility to create tailored data sets, readily available for analysis and machine learning. This would be the core ingredient in data-driven efficiency improvements in a number of business areas – for example, transport, logistics, energy, agriculture and environmental protection. Furthermore, decision-making in business and society would be facilitated by access to insights based on more accurate and up-to-date data.

In the past, CSPs have been prevented from benefiting from the value of big data on account of its sheer weight. The volume, velocity and variety – or the three Vs – of big data were simply overwhelming. Those data-handling challenges have now largely been met by a variety of easily obtained tools. Distributed databases, complex event-processing frameworks, analytics libraries and so on have been developed in the open-source community and are readily available to CSPs.

But like a blank spreadsheet, these tools are simply a platform for data handling. The real value comes from knowing which combination of the vast array of data elements reveals the desired insights. That is where deep network and operational expertise are of paramount value. Only when these key relationships are understood can the necessary insights – such as user behavior, network performance and causes of experience issues – be gained.

BIG DATA FOR THE CSP

For the CSP, big data not only means a fundamental shift in the way data is stored and managed – it also entails deploying powerful real-time analytics and visualization tools, collaboration platforms, and the ability to automatically create links with existing applications vital to the CSP's business, such as operations and business support systems (OSS/BSS) and customer relationship management (CRM).

Insights create the connection between the CSP's data assets and the desired business results. The first step is to connect data collected from a variety of sources: network and non-network, structured and unstructured. Insights may be derived from properly correlated data. For example, it is possible to create extremely rich subscriber profiles that can be useful for customer care, campaign management, customer intelligence, and so on. Finally, insights can dynamically change network behavior to ensure an optimal subscriber experience.

Figure 1: The big data analytics stack

FROM DATA TO INSIGHT

Successful decision-making will increasingly be driven by analytics-generated insights. And the more accurate and timely these are, the better chance a decision-maker has to anticipate and profit from change.

The keys to effective data-driven decision-making are: the ability to sift through large amounts of data; and the ability to combine data from several sources to gain a more comprehensive view of the business. Seemingly unimportant data can become crucial once combined with other sources to reveal new insights. For example, combining what people write in social media with network data can play a vital role in fully understanding the way users experience a service or knowing whether they are likely to churn.

In practical terms, there are several important factors to be considered to manage the expansion of data, devices and network complexity:

- the growth of mobile broadband
- enhanced connectivity and the adoption of machine-to-machine (M2M)
- the power of cloud computing
- service expansion and the shift of scale and context with relation to such services
- easy and cheaper access to big data technologies.

The complexity of handling this expanded universe of data sources is compounded by the need to link, match and transform data across business heterogeneous entities and systems, while managing scale and timeliness. Organizations need to understand key data relationships, such as complex hierarchies and links between data types and sources.

A NEW VIEW OF BIG DATA

It is clear that we need to revise our thinking when it comes to big data. The focus should not be on storing and processing large amounts of information, but rather on the ability to make decisions in a timely and accurate way. Analytics – guided by domain expertise – allows CSPs to gain more precise and timely customer and network insights, which in turn allow for improved operational effectiveness and open up new opportunities for revenue generation [4].

For CSPs, a flexible and layered analytics platform allows for easy creation of new use cases, and such an approach facilitates insights into their existing business processes and tools. For instance, CSPs will be able to prioritize VIP subscribers, provide premium service in a specific region, and focus on user experience.

MOVING FROM DATA-HANDLING CAPABILITIES TO VALUE-CREATING SERVICES

Today, big data is understood more or less from a technology perspective: the possibility of better storage (volume), the ability to process the information and make it available in real time (velocity) and the ability to deal with various kinds of data sources, including structured, semi-structured and unstructured ones (variety). The technology exists, so the essential issue is how carriers can make sense of the massive volumes of data and deliver value to subscribers and businesses.

Getting value out of big data is no longer a pie-in-the-sky proposition: it is a reality. Still, for the CSP doing network-near analytics, extracting value is a domain-specific task, and off-the-shelf IT components are not good enough. Therefore, CSPs need to look beyond traditional big data techniques and focus on the analytic value that can be gained from transforming huge volumes of complex and high-velocity data into pure business insight. Adequate, accurate and actionable – the three As – insights enable CSPs to improve their existing process and drive better decision-making.

Figure 2: The three Vs and As of big data

CSPs need to shift from analytics application silos to more generic, horizontal analytics environments that take in a wide array of data sources, while supporting a variety of applications and services. The recommended approach involves cutting-edge IT components like data storage, data management and network resources that can work in harmony with domain- specific analytic logic (data models, rules sets, and so on) not only to bring benefits in terms of agility and scalability, but also to maximize capital investments for the CSP.

In order to exploit the full potential of OSS/BSS, the CSP needs to enhance these systems with real-time and predictive analytics functionality, such as continuous updates of network caches based on where a user is expected to move to next, or the offer of dynamic subscriptions based on predicted service usage. The more real-time efficiency an analytics function provides, the more value it can bring to operations. Storing all data in a large database and then running queries on top is a naïve approach that will neither scale nor provide the analytics capabilities needed.

OSS/BSS management functionality, based on analytics, such as customer experience assurance (CEA), end-to-end performance monitoring and device management, allows the CSP to handle more effectively the increased complexity of network management, M2M and ever-increasing data volumes.

OPTIMIZE, ENRICH AND CREATE

To derive value from big data, CSPs must know what data to use, how to process, correlate and limit the data, and how to interpret and apply the resulting insights to different processes (relating to systems, operations, business, customers, and so on). Network data can be either user data transported by the network or data generated by the user device or network element including the core, RAN, transport, terminal, CRM, Home Subscriber Server (HSS), charging and billing, service proxy, policy nodes, self-care application, operator's web portal, and so on.

Figure 3: Creating value from network data

Three core areas have been identified in which CSPs can maximize the value of their data assets: optimize operations, enrich operations and create new business opportunities.

Operations can be optimized using insights gleaned from network data to ensure network resources are used where they generate most value. By connecting the analytics function close to the data source, the insight delay decreases. The network-near data analytics includes real-time distributed processing to collect, correlate and analyze data from all parts of the network, as well as from probes.

To enrich operations, an operator can analyze user behavior and identify new product offerings that will meet their needs more effectively. This type of insight will become more important as we progress toward the Networked Society and our communication involves more new device types.

With new insights, value can be created both for existing and new subscribers. Examples of insights that can be gleaned automatically include mobility patterns that can inform municipal transportation agencies how best to plan their bus routes. In several use cases, the revenue chain is based on a B2B model (where the additional revenues come from a third party, while the subscriber has the benefit of the service).

PRIVACY

Enhanced connectivity, cheap and easy access to big data technologies and massive computing power via cloud computing for extracting insights from the data gives rise to new challenges in terms of security and privacy. Traditionally, CSPs have addressed privacy concerns through enforceable regulations (such as country-specific security and privacy laws relating to CSPs), and this continues to be a differentiator. However, the opportunities to expose insights into third-party providers (3PPs) strengthen the need to protect access to private data and avoid potential misuse of it. A real-time opt-in/opt-out option empowers consumers to decide what data is shared with whom in a transparent manner, and enables the subscriber to gauge the benefits of sharing his or her personal data in return for certain benefits. Approaches that anonymize personal data (such as k-anonymity and hashing) or add noise to exposed insights (such as differential privacy) lead to more coarse-grained information, which will protect the privacy of any particular subscriber while still exposing the correct insight.

Technical approaches to address security and privacy should include secure communication, adjusted data collection, secure data storage, stringent and granular access control, audits and other topics as discussed in [5].

THE EXPANDED BIG DATA SYSTEM

It is analytics that enables us to turn big data into actionable insights. And ultimately, analytics will help make the Networked Society happen by addressing devices, network-connectivity growth, adoption of M2M, growth of the cloud and new business opportunities. To extract meaningful business value using analytics, the big data stack needs to be expanded with a number of components and technologies. In short, these are:

- the management of data assets, with real-time process control for analytics
- scaling up analytics to gain insights from big data, as well as the ability to solve increasingly complex problems using more data
- analytics-as-a-service (AaaS) approach and the flexibility to choose between local and hosted deployment options
- securing customer privacy, while still extracting relevant and timely information from data
- finally, and most importantly, domain-specific analytics tools and access to experts to ensure that true insights are being extracted.

NEW OPPORTUNITIES

Analytics is essential to the creation of cost-efficiencies and business innovation across all CSP business segments. Accordingly, solutions will need to provide cross-domain support with a common architecture to maximize reusability, agility and cross-selling of analytic applications based on facilitated use cases.

Insights can be exposed through an integrated and interactive platform that implements or reuses existing APIs and integrated development environments (IDEs), allowing multiple applications to be supported by the same underlying data infrastructure.

To be clear, underlying technology is not what generates value. It is rather the decisions that are enabled by insights gleaned from newly discovered and defined data relationships that lead to successful and beneficial use cases. The following sections provide a sample of the kind of use cases that can be realized using big data analytics.

SMARTER NETWORK

Smarter networks optimize the use of network resources and management of network traffic, while also enabling the delivery of more compelling service offerings (from both the user's and the service-provider's perspective) and improving the user experience, as envisioned in the Networked Society.

The mix and behavior of services are not static. Every day there are new services available in the app stores. The traffic patterns for individual services also change over time. Thus, it is important to have a solution that is adaptable and can work well in a changing environment, from the network and user perspective alike.

Self-organizing networks (SON) efficiently use network-near analytics in planning, building and optimizing network resources. All SON automation – such as provisioning, configuration and commissioning – can be adapted to changes in the environment and traffic demand based on the insights gained from big data analytics.

Figure 4: Smart and self-optimizing networks

To assess the impact of devices on the signaling network, big data analytics can be utilized to predict parameters based on the characteristics of the traffic, which in turn can be used to reduce battery consumption on the user device, while not increasing network delay.

USER EXPERIENCE

Big data analytics reveals what the CSP needs to know to be able to take timely action to resolve issues that impact the user across devices, subscriptions, services and network resources. For instance, CEA monitors the individual user experience of mobile broadband services [6].

Measuring QoE through system service key performance indicators (S-KPIs) requires the analysis of distributed network data. Analysis of large data sets also enables precise analysis of the user experience and leads to insights into the subscribers' real experience, which could incorporate social data, for example.

CEA collects data down to individual subscriber's sessions, and can automatically pinpoint the root cause of problems affecting customer experience, either reactively (within customer care) or proactively (within the SOC/NOC). In addition, communication between customer care

and the NOC/SOC is simplified by creating a common interface and language, easily linking QoE to the underlying network.

Off-the-shelf data-processing technologies are not enough because root- cause insights are very network-specific. In-depth network knowledge is required to work out what network data must be analyzed – at the user-session level – to understand both the perceived user experience, as well as the directly associated root cause (in real time).

Figure 5: Customer experience assurance

DATA BROKERING AND MULTI-SIDED BUSINESS MODELS

New revenue streams can be captured – monetizing the data that the CSP owns – by leveraging the valuable knowledge extracted from user data. This knowledge is often valuable to a variety of potential business partners.

Data monetization is the process of capturing, storing and managing appropriate data, performing analytics to identify key trends and patterns, and sharing the discovered insights with those who can create value from them. Sharing data in a per-subscriber-based manner is not well suited to monetized data services. Therefore aggregating, summarizing, anonymizing and packaging data in a way that is valuable to partners and third parties – such as advertisers, content publishers and social media – requires infrastructure, systems and interfaces, as well as a clearing house and market for data exchange.

CHURN PREVENTION

Churn prevention is by no means a new concept, but with the introduction of big data analytics it is possible to do far more accurate churn prediction since it enables multifaceted and multisource analysis. The subscriber base is analyzed to identify subscriber micro segments according to a broad range of parameters, in numerous combinations. It takes user data from both network elements and related sources, such as charging and demographic information, to provide a complete picture of the user in order to identify the most valuable and influential ones and pinpoint which of them may be about to churn.

Network elements, such as charging systems and device-management systems provide call data records (CDRs) and other subscriber data, for example, user device type. Applying analytics on such data enables an operator to predict which subscribers are likely to churn in order to target them before they switch operator.

Figure 6: Brokering network data insights

In addition, by understanding the data in social networks (applying insights into the behavior of influential users – or leaders – and followers), it is possible to identify the leaders in those networks who contribute to rotational churn (for instance, in a prepaid charging environment). If a potential churner falls into the leader category, it is important to apply customer-retention schemes immediately.

CONCLUSION

A common, horizontal big data analytics platform is necessary to support a variety of analytics applications. Such a platform analyzes incoming data in real time, makes correlations (guided by domain expertise), produces insights and exposes those insights to various applications. This approach both enhances the performance of each application and leverages the big data investments across multiple applications. Storing and processing huge amounts of information is no longer the issue. The challenge is to know what needs to be done within the big data analytics platform in order to create specific value.

True data-driven insight calls for domain expertise. For the CSP, this means in-depth knowledge of how the network functions, what data to pull from the network's nodes and support systems and an understanding of how to connect – end-to-end – data from multiple sources to yield an enriched set of information sources. This is what ultimately enables a range of service and user-centric applications to be created. Smarter networks, churn prevention and CEA are just a few examples of what is possible.

While big data storage and processing techniques are necessary enablers, the goal must be the creation of the right use cases. The big data tools and technologies deployed have to support the process of finding insights that are adequate, accurate and actionable. So instead of talking about the three Vs of big data, we should focus our attention on these three As of big data.

GLOSSARY

3PP	third-party provider
API	application programming interface
B2B	business-to-business
BSS	business support systems
CAGR	compound annual growth rate
CDR	call detail record
CEA	customer experience assurance
CRM	customer relationship management
CSP	communication service provider
HSS	Home Subscriber Server
IDE	integrated development environment
M2M	machine-to-machine
NOC	network operations center
OSS	operations support systems
S-KPI	system service key performance indicator
SOC	service operation center
SON	self-organizing networks

REFERENCES

1. IDC, 2011. Big Data: What It Is and Why You Should Care. [online] IDC. Available at: http://sites.amd.com/us/Documents/IDC_AMD_Big_Data_Whitepaper.pdf [Accessed July 10, 2013].

2. Ericsson, 2013. Ericsson Mobility Report. [online] Ericsson. Available at: http://www.ericsson.com/res/docs/2013/ericsson-mobility-report-june-2013.pdf [Accessed July 10, 2013].

3. IDC Press Release, 2012. Worldwide Big Data Technology and Services 2012-2016 Forecast. [online] IDC. Available at: http://www.idc.com/getdoc.jsp?containerId=prUS23355112

4. Analysis Mason, 2013. Big Data Analytics: How To Generate Revenue and Customer Loyalty Using Real-time Network Data. [online] Analysis Mason. Available at: http://www.analysysmason.com/Research/Content/Reports/big_data_analytics_Jan2013_RMA05/ [Accessed August 5, 2013].

5. Cloud Security Alliance Big Data Working Group, 2013. Expanded Top Ten Big Data Security and Privacy Challenges. [online] Cloud Security Alliance Big Data Working Group. Available at: https://cloudsecurityalliance.org/research/big-data/#_downloads [Accessed August 5, 2013].

6. Ericsson, white paper, May 2013. Experience-driven CEM, available at: http://www.ericsson.com/res/docs/whitepapers/wp-bss-crm.pdf

© 2013 Ericsson AB – All rights reserved