



Telcos and Cable Companies Use Hortonworks for Service, Security and Sales

A Modern Data Architecture for Telecommunications and Cable Companies

A Hortonworks White Paper
AUGUST 2014

Contents

| | | |
|------------------------------------|----------------------------------|-----------|
| Introduction | | 3 |
| Network infrastructure | Capacity planning | 5 |
| | Infrastructure upgrades | 6 |
| | Proactive maintenance | 6 |
| | Performance management | 7 |
| | Network traffic shaping | 7 |
| Service and security | Customer experience analytics | 8 |
| | Contact center productivity | 9 |
| | Field service productivity | 10 |
| | Data protection compliance | 11 |
| | End-user device security | 11 |
| Sales and marketing | 360-degree view of the customer | 12 |
| | Personalized marketing campaigns | 13 |
| | Upselling and cross-selling | 13 |
| | Next product to buy (NPTB) | 14 |
| | Churn reduction | 14 |
| New and adjacent businesses | Actionable intelligence for: | |
| | Advertisers | 15 |
| | Merchants | 16 |
| | Payment processors | 16 |
| | Federal governments | 17 |
| | Local governments | 17 |
| | New product development | 18 |
| Enterprise Hadoop | | 19 |

Introduction

Communications and entertainment providers must serve millions of pieces of data per second. Telco subscribers generate more data than ever before with phone calls, text messages and emails. Cable subscribers consume high-bandwidth music, videos and advertisements. Today's subscribers expect nearly limitless choice and instantaneous, personalized delivery. They want the ability to save everything they ever receive, in case they want to consume it later.

All of these changes break the old models of communication and media distribution, so telcos and cable companies are turning to Apache Hadoop for a smarter and more cost-effective way to store and process data.

Hadoop helps these enterprises modernize their data architectures in four important ways:

Multi-use, Multi-workload Data Processing: Hadoop supports multiple access methods (batch, interactive and real-time) for a common, shared data set. Analysts can view and transform data in multiple ways at once. This speeds time-to-insight and strengthens confidence in their conclusions.

New Opportunities for Analytics: Hadoop's schema-on-read architecture lets users store data in its raw format. Analysts can then define unique schemas for the data they need for a particular research question or application.

New Efficiencies for Data Architecture: Hadoop runs on low-cost commodity servers and reduces overall cost of storage. This makes it affordable to retain all source data for much longer periods, which provides applications with far deeper historical context.

Data Warehouse Optimization: ETL workloads also benefit from Hadoop's favorable economics. Data with low per-unit value can be extracted and transformed in Hadoop. This data's value grows in the aggregate when it is joined with other data and stored for longer.

These advantages explain why telcos and cable companies are adopting Hadoop. In fact, Hortonworks Data Platform (HDP) is currently in use at all of the largest US telcos, as well as at US cable companies. A growing number of international providers are also choosing HDP as their Hadoop standard.

Hadoop for Telecommunications

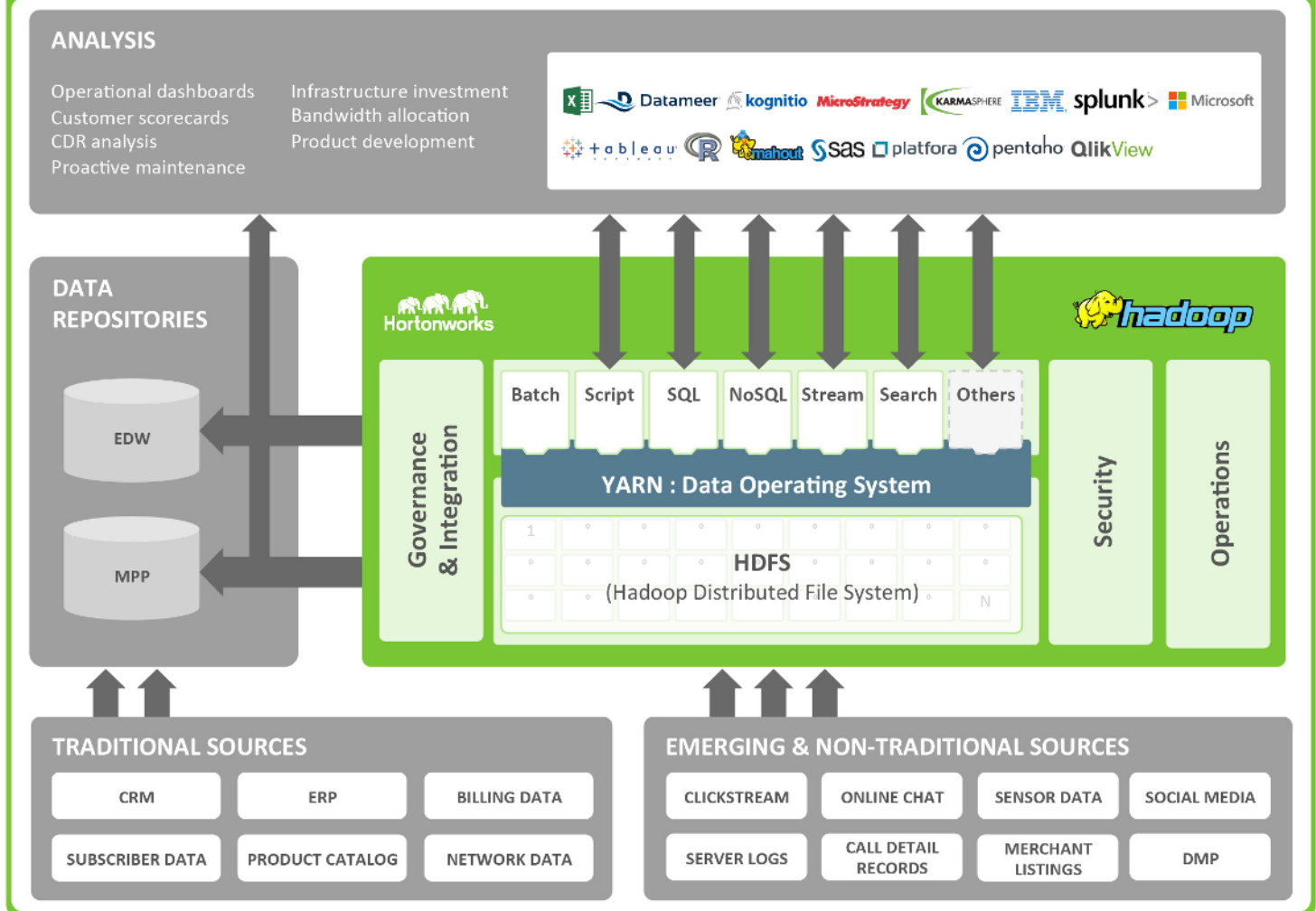


Figure 1: Ingest all the data, store it at scale and analyze it with the tools you already use.

Network infrastructure

Superior data analysis of CAPEX, OPEX and project performance in Hadoop improves business decisions for network infrastructure. The following examples show that Hadoop provides the ability to optimize network capital expenditures. Trends towards network function virtualization, software-defined networking, and unified all-IP networks vastly increase the volume and variety of machine and log data for troubleshooting network performance. Hadoop helps with batch, interactive and real-time network analysis for both emergency response and long-term planning.

CAPACITY PLANNING

Business problem

Investors scrutinize every dollar of network investment for its future value. The consumption of services and bandwidth in a particular neighborhood may be out of sync with a telco's plans to build new towers or transmission lines in that same neighborhood. This leads to a mismatch between expensive infrastructure investments and the revenue that eventually flows from those investments.

Solution

Network capacity planning requires precise information on data flows. By analyzing machine data in the form of call detail records (CDRs) and network loads, telcos can plan infrastructure expansion with greater precision. With this insight, they can then prioritize construction of new transmission towers or fiber optic cables so that the benefit justifies the investment.

Impact

One carrier used Hadoop to optimize rollout of its 4G coverage to match the likely pick-up in service revenue. The company analyzed multiple years of detailed cell tower traffic data in Hadoop. When they compared results from this Hadoop-driven approach with their prior process, they found that they would have spent 10 to 20% more CAPEX for the same outcome.

INFRASTRUCTURE UPGRADES

Business problem

Hadoop is also being used by cable companies for targeted network maintenance and upgrades. One large cable company wanted to understand how cable network congestion affected churn and which network upgrades would produce the most incremental revenue by reducing attrition.

The company needed to analyze customer experience indicators to see how network congestion correlated to customer experience. They wanted to compare dropped packets with calls into the call center and customer attrition. With legacy storage platforms, it was impractical to capture and store huge amounts of multi-form data and analyze it all from a single point of view.

Solution

Only Hadoop was up to the challenge of correlating and analyzing millions of subscriber histories, work orders, call center logs, and IP detail records. The enterprise pooled these data sources—machine, server log, voice-to-text, ERP and CRM data—in one Hadoop data lake.

Impact

By analyzing this previously disparate data in one Hadoop data lake, the company discovered that only a small number of network nodes were responsible for the majority of the negative customer experiences. The cable company prioritized those nodes for upgrades, made improvements and then benefitted from the improved customer experience they could then deliver.

PROACTIVE MAINTENANCE

Business problem

Radio access networks provide the air interface between a mobile provider and end-user mobile devices. Maintenance and repair of radio access networks poses substantial logistical challenges.

In most countries, mobile networks cover more than 95% of a country's surface area. Many transmission towers are in remote and difficult-to-access locations. In urban high-density areas, picocells and femtocells optimize local coverage, but these require maintenance teams to coordinate with the building owner for access. In both rural and urban contexts, it is important to match the cost of maintenance with its impact on the customer experience.

Solution

Telcos can optimize equipment maintenance schedules using Hadoop by comparing real-time streaming data from networks with historical data on maintenance, repairs and overhauls. These historical data sources range from traditional equipment asset databases to newer sources of sensor data. Proactive, data-driven maintenance reduces expenses and service disruptions.

Impact

Hadoop improves a provider's ability to service equipment proactively, which is cheaper and less disruptive than replacement of failed equipment. It also helps a telco's community relationships by allowing its team to service networks in congested areas at times when they will have the lowest impact on traffic, ambient noise, and the safety of work crews and the surrounding neighborhood.

PERFORMANCE MANAGEMENT

Business problem

A US carrier is using Hadoop to improve cell phone service. This operator maintains a sophisticated, end-to-end global network management platform that enables it to understand and fix instances of poor cellular service.

Call Detail Records (CDRs) stream into the network operations center at an average of millions of messages per second. Limited storage capacity for this large data flow limited root cause analysis to a 24-hour time window.

Even though the same customer issue may have generated multiple support calls, the support team could not see relationships between multiple events.

Solution

With Hadoop, this carrier now retains all data affecting service quality for a rolling six-month period. Data sources are machine-based and include CDRs as well as sensor data transmitted from network elements. Now the carrier's team can use a broader range of analytical engines to process a larger, more complex data set. This allows them to see patterns that were invisible in a 24-hour view.

Impact

With more history, the carrier is able to explore root causes of poor service. The new insight helps prioritize large infrastructure investments to optimize network performance (to the benefit of their customers' cellular reception.) With Hadoop, the carrier moved from reactive, incomplete resolution of customer problems to proactive, rational maintenance. Now things just work better.

NETWORK TRAFFIC SHAPING

Business Problem

Hadoop also helps with real-time bandwidth allocation. Certain mobile apps and user activities can hog bandwidth and erode service quality for all other customers. Network operators need to respond to such bandwidth spikes quickly to reallocate virtualized resources and maintain service level agreements (SLAs).

But that rapid response requires that the system administrator know that a spike is imminent. Quite often, the carrier learns of a bandwidth problem only after the phones begin to ring in the support center. By then, the damage is done.

Solution

Apache Hadoop enables both real-time packet inspection and interactive queries over call center transcripts. During a spike, telcos can quickly steer traffic to optimize network quality of service (QoS). With interactive query over historical data, carriers maximize service quality for the largest number of customers and also differentiate tiers of service for those willing to pay a premium.

Impact

Network operators allocate bandwidth more nimbly and manage risk to their SLAs. They can also investigate issues more quickly and suspend devices or services that that degrade other customers' reception. Machine learning algorithms can also predict traffic problems before they occur and generate alerts. Companies that can respond proactively prevent problems that used to be inevitable.

Service and security

An enterprise data lake built with Apache Hadoop makes more data available to more decision-makers over a longer time horizon. With this, telecom companies can drive better service and security outcomes, including: minimizing the total cost to serve, enhancing the customer experience, reducing customer churn, and assuring compliance with security regulations and company policies.

CUSTOMER EXPERIENCE ANALYTICS

Business problem

We discussed how analysis of call detail records (CDRs) helps telcos manage network performance. CDR analysis and archiving is also useful for SLA compliance, billing reconciliation and prevention of network congestion.

A large mobile service provider might generate 1 billion CDRs per day. Most call records are not important, but within massive troves of call data lie records about a particular customer's bad experience. If the same customer suffers recurring problems based on the same underlying root cause, it can be extremely difficult to correlate multiple events across years (and petabytes) of data. Delay causes attrition and harms servicing margins.

Solution

Apache Hadoop helps service personnel find that one important call record among millions and find patterns through the noise. Hadoop can ingest millions of CDRs per second and process the data in real-time with Apache Storm. Because Hadoop economics make it affordable to retain records for years rather than days, telcos can explore their growing data lake for continual improvement.

Impact

One large US carrier captures all information associated with mobile devices. This is tightly integrated with other data assets stored in Teradata, adding up to more than 100 billion records, and growing by 500 terabytes monthly. Hadoop's economies of scale allowed the carrier to extend its data retention from sixty days to three years.

CONTACT CENTER PRODUCTIVITY

Business problem

A major wireless carrier faced a challenge supporting tens of millions of wireless subscribers. When their customers were fewer and using less complex products and services, “batch and blast” communications worked.

But as the number and complexity of customer relationships increased, the company needed a better way to analyze call center interactions to stay close to their customers’ experiences. Why guess what customers need, when they’re calling to tell you thousands of times daily? Because the carrier had difficulty storing and processing their call center logs at scale, it was difficult to extract the message from the noise to identify and respond to each customer experience pattern.

Solution

The company built a modern data architecture with Teradata and Hortonworks Data Platform (HDP). They then analyzed voice-to-text unstructured data and detected that one quarter of the calls to their contact center were requests to waive late fees on monthly bills (even though many callers weren’t actually late). This was an unproductive use of their agents’ time.

Impact

The company was able to off-load these fee waiver requests to online self-service and interactive voice recognition. This freed contact center agents to focus on more valuable customer interactions. The carrier is now extending this solution to focus on resolving other types of issues, and the data lake allows them to do so without any additional IT projects.

FIELD SERVICE PRODUCTIVITY

Business problem

Another telco's contact center agents had insufficient information to diagnose problems with infrastructure equipment, resulting in unnecessary maintenance visits. In particular, the maintenance teams were not able to effectively triage the difference between network and home-based problems. Technicians were dispatched to customer homes for problems that actually resided within the network.

Bottom line? Trucks rolled when they didn't need to, adding up to unnecessary expenses. There was also an opportunity cost, because that same resource might have been used to fix an issue that actually required a visit to another home.

Solution

The provider used Hadoop to correlate network events with increased calls into their contact center, with better insight to distinguish between network issues and in-home issues. They combined voice-to-text unstructured data and machine-based data coming out of their network. Now they are routing calls from affected customers to a different IVR menu, based on better information.

Impact

The provider was able to avoid a large number of "false positive" truck rolls. With each truck roll costing about \$150 fully loaded, the provider was able to save several million dollars within the first year. Although it is difficult to precisely quantify the impact on customer satisfaction and retention, the carrier feels certain that better precision in home visits makes customers happier.

DATA PROTECTION COMPLIANCE

Business problem

The security operations center (SOC) of a large US telco protects the organization's data from misuse by its employees. Unauthorized employees are not allowed to access consumers' billing information, call records or texting histories.

The SOC faced three data security challenges. Investigations of potential breaches were limited to the most recent 90 days of data, so misdeeds older than three months were difficult to investigate. Slow processing speeds meant that the SOC could only maintain a three-day SLA for investigation requests. Finally, regulations mandated the retention of certain data for multiple years.

Solution

The SOC team used Hadoop to analyze various data types from multiple sources, including security event logs, emails, LDAP data, Bluecoat Proxy logs, clickstream, server log and text data. The team could then use several months of historical data to determine (within an hour) if a security breach had occurred. The company also relied on these new capabilities to comply with new data retention regulations.

Impact

The telco now better secures its customer data without dramatic changes to its SOC operations. It identifies and stops more improper activity while it is underway. Longer data retention helps them investigate any employee's activity across various systems over a period of several years. This keeps customer data safe and reduces the litigation and reputation risk posed by data breaches.

END-USER DEVICE SECURITY

Business problem

A mobile operator knew that malware on its customers' mobile devices was harming its customers' experiences and consuming an inordinate amount of bandwidth. The company needed to identify malware threats from non-trusted applications and protect their customers from harm.

This was a needle-in-a-haystack problem. Anecdotally, they suspected multiple needles and haystacks (many malware apps originating from multiple app stores). Even if one piece of malware was identified and blocked, another five could take its place. The company needed a way to continuously detect and respond to threats. They needed their data vigilance to stay one step ahead of the bad guys.

Solution

The company built a Hadoop solution that captured and analyzed web clickstreams on smartphones and then correlated that machine data with known patterns for distributing malware. By unifying these two disparate data sets, the carrier connected the what, where and how. All of this data feeds a real-time dashboard that provides minute-by-minute awareness of new threats.

Impact

The operator identified that 4% of its customers had particular Android malware apps and tied those to suspicious app stores. Hadoop helped detect those instances so operators could take remedial action. This eliminated a disproportionate share of bandwidth consumption. They also notified affected customers and blocked certain malicious URLs from the entire network.

Sales and marketing

Hadoop drives better business outcomes in sales and marketing through its ability to affect conversion rates, cost per acquisition, churn and pricing.

360-DEGREE VIEW OF CUSTOMER VALUE

Business problem

Telcos and cable companies interact with customers across many channels at different points in time. Data about those interactions is often stored in silos maintained by different lines of business. Merging that data in a relational database is slow, expensive and technically difficult. The various parties fear that other groups will misuse the data they contribute to the common repository.

For these reasons, few operators are able to correlate data about store visits, customer purchases, omni-channel marketing campaigns, and online browsing behavior. Telcos need an end-to-end view of customers and households: one data lake with integrated data governance and security.

Solution

Hadoop gives telcos and cable companies a clearer view of their customers' behavior—online and offline. It also allows them to track customer relationships for longer, revealing patterns about how customer behavior changes over time. Providers bring all that information together into a single, multi-channel view with centrally administered security and data governance.

Value realized

One large US provider was challenged to build a unified view of its customers. Multiple acquisitions and a proliferation of customer data obscured the picture. With Hadoop, the company built a cost-effective data lake holding several petabytes of social media, sensor and transactional data. Now the company uses its 360-degree view of its households to improve customer service and cross-sell.

PERSONALIZED MARKETING CAMPAIGNS

Business problem

Marketers have long sought ways to tailor their marketing campaigns to the desires of each individual customer. Telcos are uniquely positioned to help deliver 1:1 marketing, because mobile phones not only follow their owners everywhere, they also reveal data about those owners' interests (through browsing behavior and applications present on their phones).

Telcos are looking for ways to mine that information, both to increase conversion rates for their own campaigns, but also to sell that information to merchants for third-party ads directed at the right target segment, at the right time, in the right location.

Solution

Hadoop enables personalized marketing campaigns that leverage location-based information enriched by social media and other sources. BI tools access data in Hadoop to correlate deep troves of historical data with real-time streaming geospatial data to deliver highly targeted promotional messages.

Impact

One provider risked losing substantial revenue as prepaid customers switched to a competitor. The provider used Hadoop to pinpoint those individual customers most at risk of attrition and then built a targeted campaign to retain their customers in that segment. They established a churn alarm system to minimize revenue leakage.

UPSELLING AND CROSS-SELLING

Business problem

Another provider wanted to upsell smart phones to a large segment of its user base that still had legacy feature phones.

The company needed to determine the right marketing approach for each micro-segment depending on historical usage, recent trends in data use and social profile. It had access to aggregate data from billing systems, network elements and other sources, but it had no cost-effective way to merge that data together in its many forms.

Solution

In Hadoop, data scientists correlated multiple data points for their customers: their rate, device, usage patterns in voice and SMS, web pages visited from fixed lines, and demographic characteristics like age and gender. The analysis produced a tailored, segment-based approach to upselling from feature phones to smart phones.

Impact

The provider was able to push individualized credits and discounts to the target segment at point of sale. The data was enriched by the spend patterns of prepaid customers to estimate worthiness for micro-credits to finance smart phone upgrades. The operator convinced hundreds of thousands of feature phone users to upgrade to smart phones with their associated data plans.

NEXT PRODUCT TO BUY (NPTB) RECOMMENDATIONS

Business problem

As telco product portfolios grow more complex, there are more opportunities to sell additional services to the same customer base. But sales associates can be overwhelmed by the complexity of product and service offerings. They are reluctant to guess about NPTB recommendations, especially since they need to make those recommendations in real-time (face to face or over the phone).

This leads to a paradox. Experienced reps have the knowledge to make confident NPTB recommendations, but relatively few succeed long enough to earn that confidence. NPTB algorithms have the compound effect of improving same-store sales and employee retention. But those algorithms need data to work.

Solution

NPTB algorithms run with data from structured ERP and CRM sources, enriched with clickstream, social media and text data. In one example, Hadoop-based analysis tells a sales person the probability that a particular home cable subscriber will add a cell phone package to their account, and the system makes a recommendation on how best to offer that incremental product.

Impact

Confident NPTB recommendations, based on data from thousands of customers, empower sales associates and improve their interactions with customers. A Hadoop data lake, combined with a machine-learning recommendation application, reduces sales friction and creates a competitive advantage similar to Amazon's NPTB advantage in online commerce.

CHURN REDUCTION

Business problem

A mobile provider faced the following challenge: 50% of new customers terminated their relationship within 6 months of acquisition. The average customer tenure in this segment was 13 months, well short of the 18-month tenure needed to break even.

The company contemplated traditional churn-reduction strategies like targeted discounts or personalized communications to those at risk. But they knew from experience that these could be counter-productive without a robust model to predict attrition probabilities. (Those with no intention of leaving might accept a discount or become annoyed by too many "please don't leave us" messages.)

Solution

The provider took a novel approach to the age-old churn problem. It focused acquisition efforts on the "right" prospects with predicted low propensities to churn. Using Hadoop, the provider built a profit/churn prediction model that could be used in real-time at the point of acquisition. This enabled sales staff to predict the customer lifetime value for each new prospect.

Impact

The provider increased the "right" customer acquisitions by 27% and decreased subsequent churn in this segment by 50%. Moreover, they saved money that they might have spent retaining unprofitable customers. Now the company enjoys a virtuous cycle: fewer bad customers hamper their ability to delight the good customers who recommend the company to other attractive prospects.

New and adjacent businesses

Hadoop can also drive better business outcomes for business partners of cable companies and telcos by uncovering new ways to monetize existing data. Communications and entertainment providers are well positioned to provide big data as a service, helping their business customers better serve their end users.

Hadoop can also be used to develop the provider's new products to meet those secondary needs.

ACTIONABLE INTELLIGENCE FOR ADVERTISERS

Business problem

Data flowing through mobile devices contains critical information for advertisers about what customers want, and where, when and how they search for information about those products and services.

One example of how a telco provides information to advertisers comes from an online real estate marketplace, a telco subsidiary. The site features more than one million properties for rent or sale. The parent telco wanted to drive more marketshare to the site by improving its service to advertisers—typically real estate agents and brokers.

Solution

The company rolled out a data analysis service powered by Hadoop. It analyzes local real estate markets and recommends prices to realtors for renting or selling their properties. Relational data on past transactions is merged with geospatial and clickstream data. The analysis includes sociodemographic data on 19 million properties and 6 million saved searches.

Impact

A senior leader at the telco summed up the impact of the new service: “All told, a small team consisting of a product manager, a data scientist and a few developers was able to make a meaningful contribution to revenue growth.” This new, rich data service helped drive more real estate agents and brokers to the site.

ACTIONABLE INTELLIGENCE FOR MERCHANTS

Business problem

A mobile service provider understood that the geolocation data generated by its customers could be very valuable for retailers looking to deliver real-time promotions and advertisements to customers based on where they tended to shop over time, plus where they were standing at the moment.

They wanted to segment customers using geolocation information from their mobile devices, but such data was semi-structured and flowing in extremely quickly. It didn't make sense to offer the data service with any of its legacy data platforms. Those storage costs were too expensive at scale and the geolocation data had an unpredictable schema: it would never fit nicely in columns and rows.

Solution

The provider uses Hadoop to segment subscribers by locations they frequent. They correlate this geospatial data with their application usage and online search behavior. This combined data informs recommendations on local promotions, campaigns and store staffing levels. Retailers follow their target customer segments, paying particular attention to visits to rival e-commerce sites via their smart phones.

Impact

The retailer customers of the mobile provider were able to increase their reported same-store sales through better campaign management and in-store optimizations. As more data accumulates in Hadoop, the mobile provider has an increasingly predictive data asset that they offer to additional retailers or other clients that aim to improve their service delivery using the geospatial data.

ACTIONABLE INTELLIGENCE FOR PAYMENT PROCESSORS

Business problem

Credit card issuers experience higher rates of fraud when their cardholders travel abroad. Travelers are typically less attentive to their transaction records while traveling, and so they may detect fraud on their accounts only after their monthly statement arrives

This has a two-fold impact on the payment processors. First, they do take losses on actual fraudulent charges, which they are obligated to repay. Second, they have higher customer support costs unlocking inappropriate account holds or resolving "false positive" support calls when a customer feels an actual charge was fraudulent (e.g., "I never would have paid that much for dinner!")

Solution

One issuer used Hadoop to correlate two types of geospatial data: mobile phone location and credit card POS location (from transaction records). If card transactions were made at POS terminals that did not match the location of the customer's mobile phone, those customers would receive a real-time SMS message indicating a suspicious use of their credit card.

Impact

The new solution grew from the carrier's prior experience using Hadoop for real-time CDR analytics to reduce fraud on prepaid mobile services. Ninety-five percent of travelers opted into the SMS alerting service, substantially reducing card fraud. Moreover, the processor earned retention benefits from improved customer confidence that they were staying one step ahead of cyber-criminals.

ACTIONABLE INTELLIGENCE FOR FEDERAL GOVERNMENTS

Business problem

The eastward expansion of the European Union has resulted in a longer and more porous border with states that are not EU members. This makes it more difficult to protect the EU against a stream of illegal goods and refugees, which often travel over land from the EU's eastern neighbors. These flows come through official entry points and also clandestine routes familiar to authorities.

The core challenge is familiar: how does a government distinguish a small amount of data on illegal crossings amidst a much larger set of data on legal crossings, and then take action without abusing the rights of those who have broken no laws.

Solution

One telecommunications provider is using Hadoop to support law enforcement agencies in the implementation of advanced border control measures. Hadoop is used to detect unusual travel patterns in real time by correlating geolocation data with data on travel routes, travel speeds, means of transportation, mobile subscriber phone numbers (MSISDN), and mobile subscriber identities (IMSI).

Impact

With this previously unavailable information, law enforcement agencies are able to allocate their scarce resources much more effectively. Now authorities can choose to intercept suspicious cars traveling in certain areas favored by smugglers at speeds above 130km/h. This radically increases their apprehension rate per mission and helps them better secure their borders.

ACTIONABLE INTELLIGENCE FOR LOCAL GOVERNMENTS

Business problem

In a major city, vehicular traffic to large events caused massive congestion on the city's streets and highways. After these events, citizen complaints would pour in to city offices. Over time, potential planners and sponsors of events shied away from bringing their venues to the city because they feared angering city residents (and potential customers).

City leaders did not want to stifle the cultural and economic benefits brought to the city by these large events, yet they needed actionable data on how best to bring thousands of attendees in and get them home again without angering the millions just trying to move about the city.

Solution

A mobile service provider used mobile phone data in Hadoop to optimize traffic management for the city's mega-events. They tracked and anonymized mobile network location information. This provided granular, real-time visibility into traffic congestion events. Historical reference data from many events identifies trends useful for improving the city's traffic-shaping methods over time.

Impact

The city identified and implemented dozens of specific traffic management measures that relieved congestion around major events. Officials are now exploring how to use these insights for environmental impact studies, city planning and disaster management. A wealth of detailed geolocation data in Hadoop helps the city's planners keep its residents safer, healthier and happier.

NEW PRODUCT DEVELOPMENT

Business problem

Mobile devices produce large amounts of data about where, when, how and why they are used. This data is extremely valuable for product managers, yet much of it is out of reach. It may never be captured or, if captured, stored for only a short period of time. Even if it is captured, storage is typically fragmented across multiple business groups and projects. Because of this fragmentation, the data may never be converted into business insight.

The data volume and variety make it difficult to ingest, store and analyze at scale. As a result, product managers make uncertain guesses when they actually possess empirical data to make confident decisions.

Solution

With Hadoop, product managers get precise visibility into how their customers use their products. Immediate big data feedback on new launches helps product teams respond to actual (not sampled) usage patterns in real time, so they can promote winning features and drop or fix unpopular ones. Data sources include structured ERP and CRM data, and click-stream, social media and machine data.

Impact

One provider logged millions of devices with more than 1 billion events per month. More than 20 projects and pilots used this data within the first 18 months after launch, leading to increased revenue and profitability. As this data trove expands, the manufacturer's rate of innovation also increases. This also provides ideas for net-new products that never would have existed otherwise.

Build a Modern Telecommunications Architecture with Enterprise Hadoop

To realize the value in your investment in big data, use the blueprint for Enterprise Hadoop to integrate with your EDW and related data systems. Building a modern data architecture enables your organization to store and analyze the data most important to your business at massive scale, extract critical business insights from all types of data from any source, and ultimately improve your competitive position in the market and maximize customer loyalty and revenues. Read more at: <http://hortonworks.com/hdp>

Hortonworks Data Platform provides Enterprise Hadoop

Hortonworks Data Platform (HDP) is powered by 100% open source Apache Hadoop. HDP provides all of the Apache Hadoop-related projects necessary to integrate Hadoop alongside an EDW as part of a Modern Data Architecture.

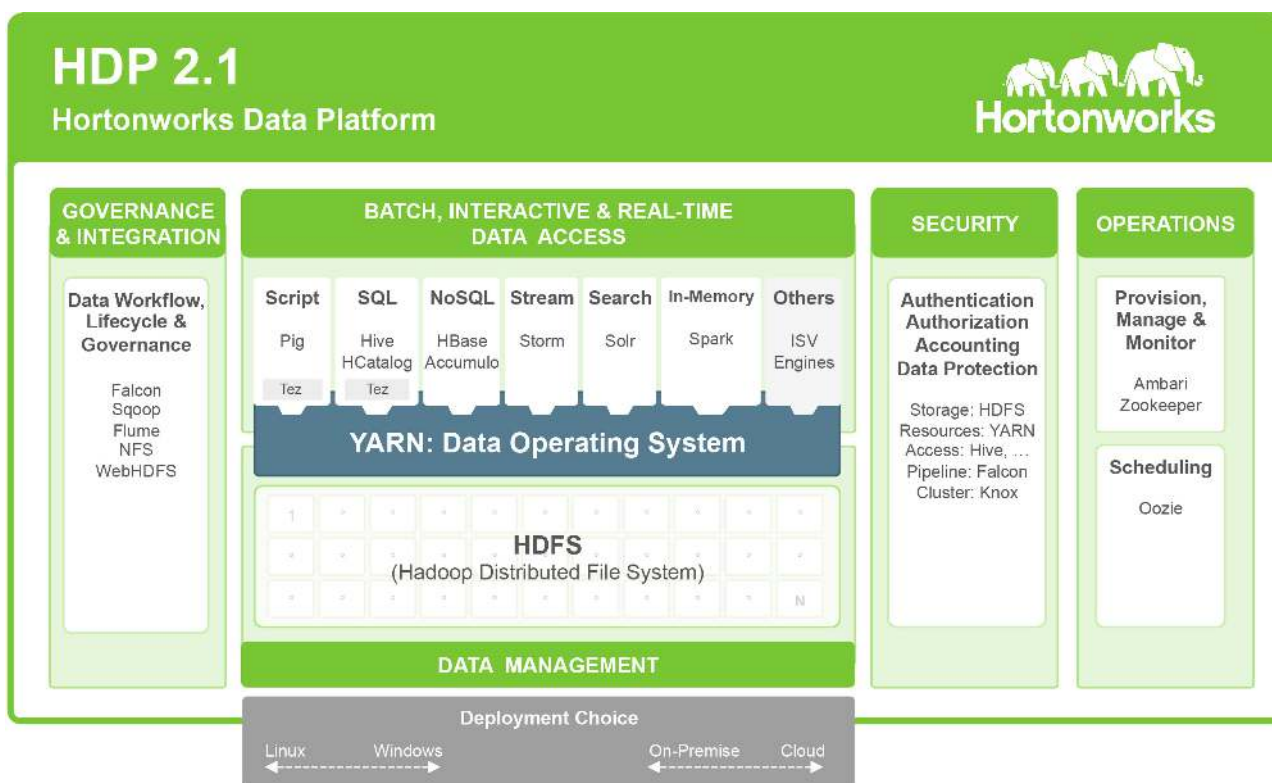


Figure 2: Five core capabilities: data governance & integration, data management, data access, security and operations.

Why Hortonworks for Hadoop?

Founded in 2011 by 24 engineers from the original Yahoo! Hadoop development and operations team, Hortonworks has amassed more Hadoop experience under one roof than any other organization. Our team members are active participants and leaders in Hadoop development, designing, building and testing the core of the Hadoop platform. We have years of experience in Hadoop operations and are best suited to support your mission-critical Hadoop project.

For an independent analysis of Hortonworks Data Platform and its leadership among Apache Hadoop vendors, you can download the [Forrester Wave™: Big Data Hadoop Solutions, Q1 2014](#) report from Forrester Research.

About Hortonworks

Hortonworks develops, distributes and supports the only 100% open source Apache Hadoop data platform. Our team comprises the largest contingent of builders and architects within the Hadoop ecosystem who represent and lead the broader enterprise requirements within these communities. Hortonworks Data Platform deeply integrates with existing IT investments upon which enterprises can build and deploy Hadoop-based applications. Hortonworks has deep relationships with the key strategic data center partners that enable our customers to unlock the broadest opportunities from Hadoop. For more information, visit www.hortonworks.com.