



Accelerating Data Velocity by 15x

Overview

With £18 billion (about US\$30 billion) in revenue in 2014, **BT** is one of the largest telecommunications providers in the world. The company serves more than 18 million consumers and nearly three million businesses.

For BT, the key to achieving sustainable, profitable growth in today's competitive landscape is its ability to broaden and deepen customer relationships. To support this goal, BT is using a **Cloudera enterprise data hub** (EDH) to accelerate data velocity and fast-track the delivery of new offerings to its customers. This EDH provides the backbone for an operational data store (ODS) that enables BT to break through data silos to ingest, store, and prepare data for myriad operational and analytical uses.

Within one year, BT has increased data processing velocity by a factor of 15, achieved an ROI between 200 and 250 percent, and is now positioned to take on new projects faster at a lower cost.

The Challenge

"To meet customer expectations, we need to know who our customers are, what services they have, how they're connected to the network, and how those services are performing," explained Phillip Radley, chief data architect at BT. "Maintaining the quality and integrity of those data assets is a constant challenge."

Several years ago, Radley met with colleagues to discuss the next iteration of a critical extract, transform, and load (ETL) "pipeline." In the legacy environment, business client records were spread across multiple databases. They needed to be reconciled and updated daily with Dun & Bradstreet data in order to provide business units with the most relevant and up-to-date information. With nearly one billion records being compared and reconciled daily, BT's legacy ETL platform, built on a traditional relational database, couldn't keep up any more. At any given point in time, its business units were working with day-old data.

"It wasn't processing all the data we wanted to process, and it was taking more than 24 hours to process 24 hours of the data," said Radley.

Key Highlights

Industry

- Telecommunications

Location

- Headquarters: London, England
- Operates in 170+ countries

Business Applications Supported

- An operational data store deployed on an enterprise data hub
- Supports service delivery, customer care, performance management, and more

Impact

- Processes 5x more customer data
- Increased data velocity by 15x
- Delivered ROI of 200-250% in one year

Technologies in Use

- Hadoop Platform: Cloudera Enterprise, Data Hub Edition
- Hadoop Components: Apache Hive, Apache Pig, Apache Sentry, Apache Spark, Cloudera Manager, Cloudera Navigator, Impala

The Solution

“We had a proposal to re-platform the system to a new relational database,” said Radley. “But as we sat down, our discussion turned to [\[Apache\] Hadoop](#). We realized we basically had a data velocity problem. We had to process the data faster and increase the volume that we could ingest—both of which Hadoop excels at.”

BT engaged Cloudera to install a production-ready Hadoop cluster that replaced the batch ETL application with [MapReduce](#) routines. For Radley and his colleagues, this new enterprise data hub approach would not only solve BT’s immediate ETL problem, but also tackle a host of big data challenges to help BT fast-track the delivery of new offerings.

“We selected Cloudera as they offered a mature enterprise Hadoop platform that met our needs at a very competitive price,” said Radley. “We worked with Cloudera and went from PowerPoint to production in nine months, which is pretty good by our standards.”

Rather than hiring new talent, the company wanted its Linux administrators to manage the Hadoop platform, and Cloudera provided the training they needed.

“We were able to save time and money on head hunting for Hadoop experts by training the skilled teams that we already had,” said Radley. “The [Cloudera University](#) training course was not only high quality, but also the trainers were able to understand what we were trying to accomplish and helped ramp up the team quickly. The same people who run our 30,000 Linux servers also now run Hadoop, and they can do that on top of their other responsibilities.”

Today, this multi-tenant Hadoop environment provides a single, cost-effective infrastructure that enables BT to gain unified views of its data across its multiple business units. Additionally, the environment will enable the company to extend data retention from one year to more than 10 years when needed and implement innovative knowledge management use cases.

“We have 1,900 operational systems and several of the world’s largest data warehouses,” said Radley. “The EDH runs below our operational systems and these systems extend their data into the EDH. We can then expose and share that data as we need to.”

As BT moves to more sophisticated big data applications, it will use [Apache Spark](#) to combine batch, streaming, and interactive analytics, and [Impala](#) enables the business intelligence (BI) teams to perform SQL queries on the data.

For Radley, [security](#) and stability were vital to the platform’s success. “Security has to be as good as our business-as-usual security, and the operational SLAs have to be as good as our existing IT systems,” said Radley. “[Cloudera Manager](#) rolling upgrades have allowed us to keep our Hadoop platform on the latest release so that we can get quick access to new features without service interruptions, and [Cloudera Navigator](#) saves us time auditing the platform and tracking data lineage.”

Impact: Increasing Data Velocity by 15x

Moving its ETL platform to Hadoop enabled BT to accelerate data velocity. “We were able to increase data velocity by a factor of 15,” said Radley. “We’re processing five times the data in a third of the time. The business sponsors don’t know that we moved to Hadoop and they don’t care. All they know is that they’re now working with today’s data instead of yesterday’s.”

The move also delivered substantial cost savings for BT.

“Putting the data on Hadoop was much cheaper than putting it on a standalone system,” said Radley. “Our one-year return on investment from the Hadoop deployment is in the 200 to 250 percent range. Additionally, now that the enterprise data hub is implemented, BT can undertake new projects quickly and at a much lower incremental cost.”

Impact: Providing Better Broadband Service to Customers

Following the success of its ETL initiative, BT is now utilizing the Cloudera EDH to help deliver its broadband services.

The speed of an individual line is dominated by its length (the distance from network equipment to a customer’s premises), but many other factors can have a significant impact on customer experience.

“We use our Cloudera EDH to join network topology (GIS) data with terabytes of DSL performance (time series) and electrical line test data to grade the quality of every line in the network,” said Radley. “Using this network analysis, the probability of a successful outcome of an engineer dispatch can be predicted. This reduces wasted engineer visits and truck rolls.”

He added, “Better network analytics helps us understand how to deliver better network performance, which is good for customers. Hadoop’s power and flexibility, with scale-out and schema-on-read capabilities, makes network analytics better because we can use all the raw data, processing it faster than we have ever been able to achieve and at a much lower cost. Cloudera’s EDH enables our IT teams to deliver Hadoop to the business as a secure and reliable IT service on a low-cost shared infrastructure.”

Impact: Supporting Urban Planning with IoT Data

BT’s work with Cloudera is also helping position the company to take advantage of the Internet of Things (IoT). Take its work with Milton Keynes (MK), a fast-growing town in Buckinghamshire, England.

BT is part of the MK:Smart initiative, which includes early IoT solutions such as sensors in car parking spaces that broadcast if the spots are vacant or occupied. Citizens and visitors can then use a smartphone app that guides them to the nearest free parking space based on the sensor data.

According to Radley, the same data ultimately will be used to better inform multi-million pound infrastructure decisions, such as the location and size of future car parks.

Radley said, “Early projects, such as the MK Data Hub, a key component of MK:Smart, show us that the data from things, such as car parking spaces, recycling bins, and street lights, can provide valuable insights and needs to be captured, analyzed, and made available. When you scale this up for a large town or city, the volumes of data can become large and meaningful, providing significant insights and value to the community and business, and therein lies the potential for our Hadoop application.”

“We’re processing five times the data in a third of the time. The business sponsors don’t know that we moved to Hadoop and they don’t care. All they know is that they’re now working with today’s data instead of yesterday’s.”

— Phillip Radley, Chief Data Architect, BT

About Cloudera

Cloudera is revolutionizing enterprise data management by offering the first unified Platform for big data, an enterprise data hub built on Apache Hadoop. Cloudera offers enterprises one place to store, access, process, secure, and analyze all their data, empowering them to extend the value of existing investments while enabling fundamental new ways to derive value from their data. Cloudera's open source big data platform is the most widely adopted in the world, and Cloudera is the most prolific contributor to the open source Hadoop ecosystem. As the leading educator of Hadoop professionals, Cloudera has trained over 40,000 individuals worldwide. Over 1,700 partners and a seasoned professional services team help deliver greater time to value. Leading organizations in every industry plus top public sector organizations globally run Cloudera in production. www.cloudera.com



cloudera.com

1-888-789-1488 or 1-650-362-0488

Cloudera, Inc. 1001 Page Mill Road, Palo Alto, CA 94304, USA

© 2015 Cloudera, Inc. All rights reserved. Cloudera and the Cloudera logo are trademarks or registered trademarks of Cloudera Inc. in the USA and other countries. All other trademarks are the property of their respective companies. Information is subject to change without notice.

Cloudera_CaseStudy_BT_104