



## Complementing the traditional data warehouse with Hadoop framework to provide on-time network performance metrics at a lower cost

### The Client

The client is a telecommunications service provider and the fourth-largest wireless network operator in the US serving around 60 million customers.

### Business Challenge

The introduction of 4G LTE technology brought about a massive increase in data volumes. As a direct result, data warehouse size was predicted to grow from 250 TB to 350 TB in a year.

Exploring additional data collection avenues presents opportunities to grow, as well as new challenges to conquer. The addition of over 3 million users, combined with the unprecedented growth of data volume, led to a performance issue in the data warehouse, causing a potential breach of the SLA.

In order to resolve the issue, multiple SQL performance optimization techniques and best practices were applied to the solution, thereby increasing the sustenance of the data warehouse. However, this was not a permanent remedy for the problem at hand.

Options such as introducing a data warehouse appliance and a parallel data warehouse were considered. Technosoft played a crucial role in comparing various data warehouse solutions and appliances such as Teradata, Netezza, MS SQL Parallel Data Warehouse and ExaData. In addition to the ability to handle and process high data volumes, cost played a vital role in the decision about the alternate technology stack.

Around this time, the world had started looking at Hadoop as an application framework. The Hadoop framework's performance, proved itself useful with Google and Yahoo. It's ability to scale elastically with growing data volumes and processing power needs made the platform a preferable option. Additionally, Hadoop was more cost efficient. Compared to alternative methods, Hadoop had the lowest cost per terabyte. Therefore, Hadoop was considered a platform of choice, from both a cost and performance perspective.



Hadoop offered an alternative approach to traditional data warehouses and data warehouse appliances. The framework served as an option that could be used in traditional data warehouses and as an ETL tool. A proof-of-concept was proposed and delivered in 2014. The Hadoop platform feature analysis, gap analysis and workarounds were proposed to implement the data warehouse solution using Hadoop. The proof-of-concept also included reporting tools such as MicroStrategy and Tableau. They were tested with rigor for their ability to connect and work with the data from Hadoop

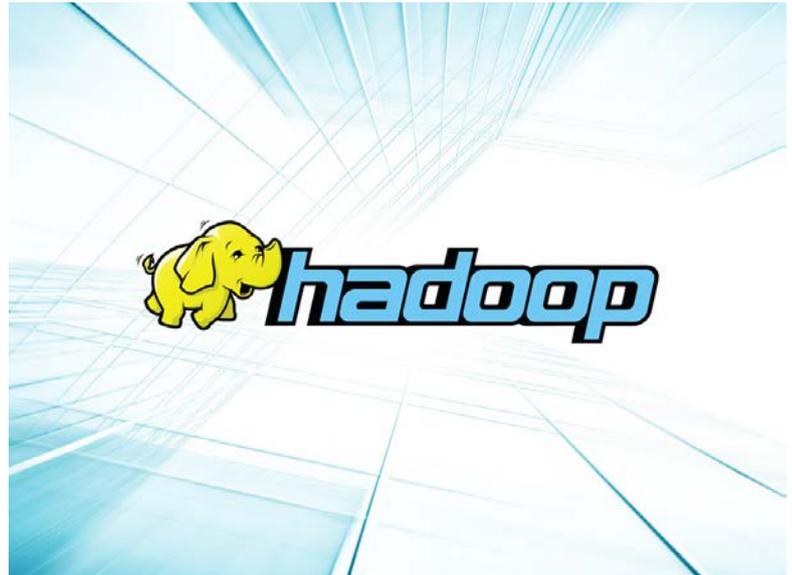
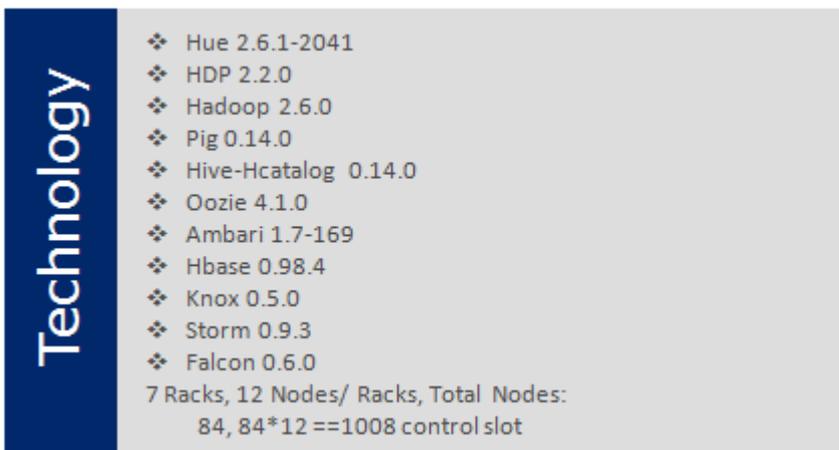
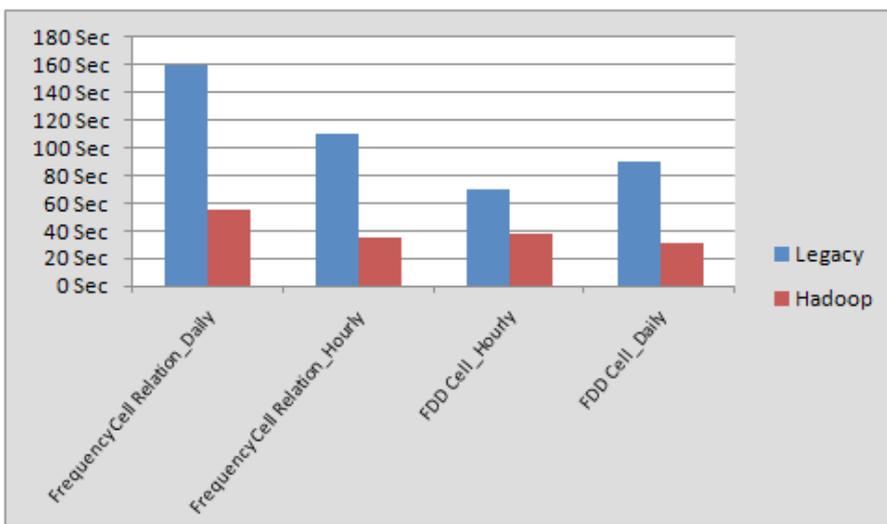


Fig 1 – Technology stack



The 3.5x performance improvement in data delivery brought in confidence about the viability and suitability of the Hadoop platform.

Fig 2 – Comparison of performance between legacy and Hadoop-based solution





Due to its batch processing nature, not all data was moved to Hadoop. Daily operational data was stored in the data warehouse and historical data was stored in HIVE. Data loading and transformation was replaced with HQL and PIG. Teradata was chosen as the platform for storing daily operational metrics.

The proof-of-concept helped finalize the foundational architecture for development in Hadoop. In 2015, the solution for Ericsson and Nokia LTE switches on the Hadoop platform was successfully developed.

With the introduction of Hadoop, the ability to deliver data on time, handle growing data volumes, and bring in cost savings for the client was achieved. Apart from these stated benefits, the client now has the benefit of introducing unstructured data from social media, click streams, web logs etc. in Hadoop.

At Technosoft, we are proud to have helped our client in moving the legacy data warehouse to a Hadoop-based platform that is future-proof and can meet their ever-growing data demands. We would be happy to share this enriching experience in other similar situations.





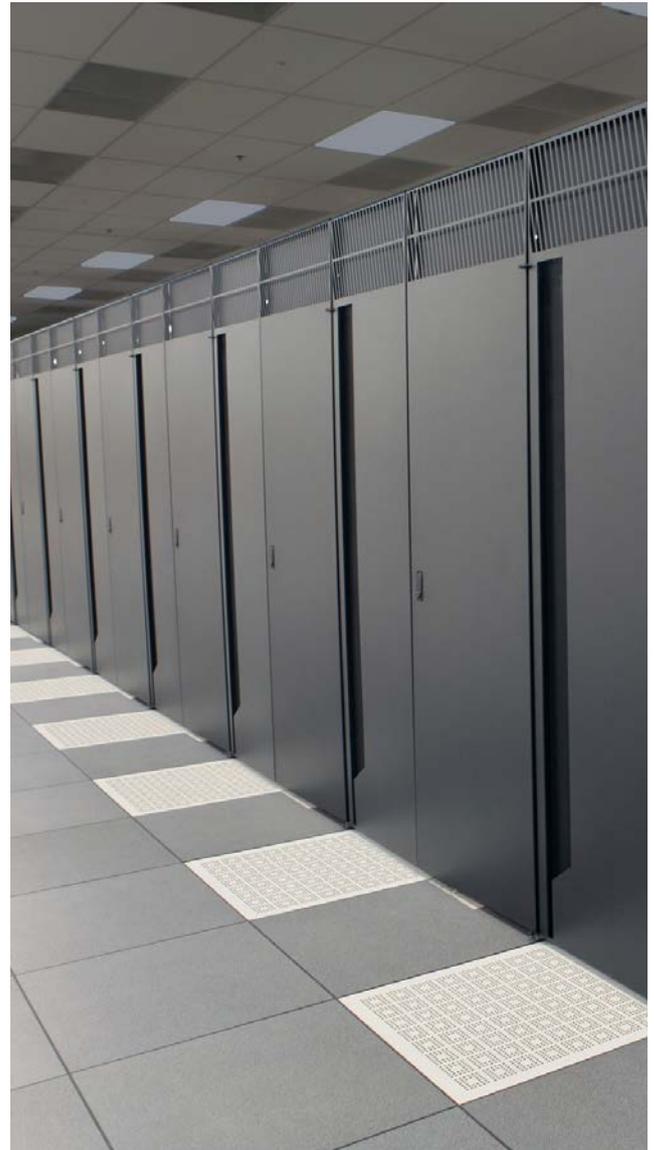
## Technosoft's Solution

Technosoft helped build one of the largest MS SQL Server-based data warehouses for the client.

The 7-year-old solution is used to capture network performance metrics used by Radio Frequency (RF) engineers to analyze and sustain the health of the network. Over 1,000 jobs were involved in the extraction, transformation and subsequent loading of data into the data warehouse. Including switches such as Ericsson, Nokia, Lucent, Nortel, Motorola and Samsung, which run at predefined time intervals (quarter hourly, hourly). Technosoft is involved in providing daily network performance metrics as per the defined service level agreement (SLA) with the client.

Our client works with FQHCs and helps in their efforts to maximize insurance billing revenue and implement sound collection procedures.

In this case, there was very little data on collections. The community centers were reeling under low collection rates and higher denial rates with poor revenue cycle services.



## About Technosoft

Technosoft Corp. (Technosoft) combines strategy, technology and creativity to help companies accelerate their digital transformation journeys. We help our customers gain insights from data that others can't see and we provide bold ideas for innovation. Technosoft offers solutions for digital transformation, data science, robotic process automation, artificial intelligence, Blockchain, cloud computing, application lifecycle management, quality assurance and testing, CRM-ERP (Salesforce, SAP, Oracle), and IT infrastructure management. Technosoft has extensive domain knowledge of the banking and financial services, healthcare, manufacturing, retail, and high tech industries. Headquartered in Southfield, Michigan, Technosoft has 4,400+ global employees and is trusted by more than 35+ Fortune 1000 customers in North America and India. Learn more at [www.technosoftcorp.com](http://www.technosoftcorp.com)

One Towne Square, 6th Floor  
Southfield, MI 48076, Phone: 855.527.3966, [Info@Technosoftcorp.com](mailto:Info@Technosoftcorp.com)

