



Business challenge

Bankdata knew that its vast performance data warehouse held the key to further optimizing its IT landscape—but with such a huge volume of data to analyze, it couldn't get the results fast enough.

Transformation

As an IT provider for some of Denmark's largest banks, Bankdata knows that the performance and availability of its systems are paramount. 15 years of 24/7 monitoring have given the company a huge repository of data on system behavior—and today, the IBM® z/OS® Platform for Apache Spark is helping Bankdata fully explore that data for the first time.



Frank Petersen,
IT Architect, Bankdata

Business benefits

Reveals

new ways to boost performance by analyzing data at a more detailed level

Accelerates

data ingestion from hours to minutes, enabling near real-time analytics

Reduces

costs by offloading ETL workload to cost-effective specialty processors

Bankdata

Harnessing mainframe-powered data science to identify and resolve IT issues fast

Bankdata is one of the largest financial IT companies in Denmark. Headquartered in Fredericia, with development teams in Silkeborg, Aarhus and India, the company is owned by 11 Danish banks, who act as its main customers. Bankdata provides complete, end-to-end solutions for the financial sector, including the development of internet and mobile banking, credit and advisory tools, support, and security.

“With Spark on Z, we finally have the tools we need to make advanced, near real-time analysis a reality.”

—Frank Petersen, IT Architect, Bankdata

Share this



Reducing costs, keeping customer satisfaction high by maintaining fast response times

Bankdata is a specialist IT services provider that focuses on running core systems for many of Denmark's leading banks. One of its most important responsibilities is to host and manage online and mobile banking systems, which are required to be online and responsive 24/7, enabling the banks to offer seamless and user-friendly customer service.

For this reason, Bankdata takes the performance and availability of its systems extremely seriously. For the past 15 years, it has continuously collected detailed performance data from each of its managed systems, and loaded this data into a data warehouse for analysis. The total amount of data gathered daily is more than 200 gigabytes.



Frank Petersen, IT Architect at Bankdata, recalls: “Over the years, we have used the performance data warehouse to understand the characteristics and behavior of our systems, and make many optimizations that have reduced our costs or allowed us to make better use of our IT resources.

“However, we wanted to take things to the next level. We were collecting such a large amount of data that it was impossible to analyze it at anything more than an aggregated level, via a set of canned reports. It was also taking three or four hours to ingest the new data into the warehouse, which meant we could only analyze the information retrospectively.

“We wanted a more freestyling, data science-based approach to help us unlock new insights. For example, we wanted to understand fluctuations in demand over time, to help us redistribute workload to balance resource consumption. We also wanted to be able to predict the impact of delays in our batch processes. If a processing job overruns its window, how will that affect the timetable for other jobs? And how quickly can we get back on track?”

“Data is available for analysis almost immediately. Instead of looking at information that is three or four hours old, we have the opportunity to monitor and analyze events across our infrastructure as they happen.”

—Frank Petersen, IT Architect, Bankdata

He adds: “We always want to be able to see issues before our customers do—if we can solve them before they develop into a real problem, that would be perfect.”

Building on Spark

While attending an IBM conference, the Bankdata team saw a demo of the IBM z/OS Platform for Apache Spark, and immediately recognized its potential to help them transform their analytics process.

“Our first thought was that Spark could help us solve our data ingestion problem,” says Petersen. “Instead of waiting three or four hours to get new data into the warehouse, Spark could help us do it in minutes. Moreover, because Spark jobs are written in languages like Java and Scala, and run on the Java Virtual Machine, we could move our extract, transform and load [ETL] processes onto our IBM Z® platform’s zIIP engines, making the whole process much more cost-efficient.”

Once Bankdata had successfully implemented the new Spark-based data ingestion process, the team began considering other ways it could utilize Spark technology.

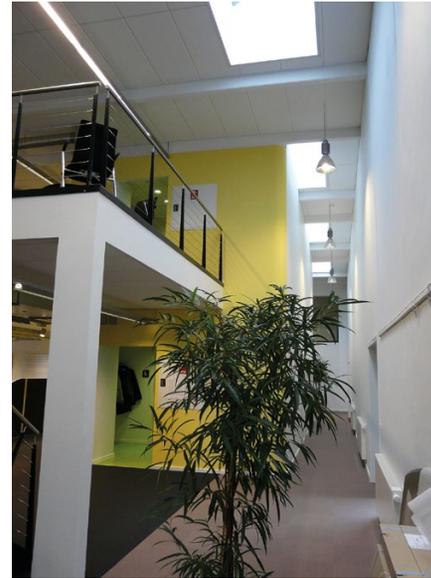
Petersen explains: “We were keen to build a bridge between the ‘old world’ of mainframe data processing and the ‘new world’ of data science, and Spark on Z was the perfect opportunity. With the power of the Spark engine, we would be able to analyze our data directly, instead of having to work with aggregates.

We would also be able to perform analyses much faster, giving us more scope to experiment instead of simply focusing on producing the results.

“Our hope was that with these new capabilities, someone would find a brilliant new way to look at the data. By looking at data in new ways, you will get new truths.”

Since the team at Bankdata had never attempted a data science project before, the team wanted to find an expert partner to help teach them how to build machine learning models and deploy them in an automated way.

“We had previously developed a simple script to look for spikes in behavior of programs over time and create reports when there was an anomaly,” says Petersen. “We used it to analyze over 300,000 programs! It made us realize the value of automation. The only weapon you have against increased workload and complexity is automated analytics.”



Bankdata engaged with IBM to deliver a proof of concept, with the aim of building some useful machine learning models and giving the company’s data science initiative a kick-start on performance data.

“We always have a list of projects that we never have time to do, but we felt that it was vital not to neglect this one,” says Petersen. “We thought it could give us a real competitive advantage. Working with the IBM team has really helped us get up to speed and take advantage of Spark for analytics.”

Realizing the dream of data science

With the IBM z/OS Platform for Apache Spark in place, Bankdata is now in a strong position to capitalize on its dream of optimizing IT performance with near-real time analytics.

Petersen comments: “Today, we ingest 200 gigabytes of data into our performance data warehouse every day, and the data is available for analysis almost immediately. Instead of looking at information that is three or four hours old, we have the opportunity to monitor and analyze events across our infrastructure as they happen.”

The new Spark-powered ETL processes are also much more flexible. “It’s a much easier way to feed the data into our warehouse, and it’s much simpler to make changes. If we want to add a new field to the data we’re capturing, for example, it’s very straightforward. This was much more difficult before.”

Using Spark for automated analytics is also helping Bankdata get much closer to monitoring the actual state of each system in its network. Instead of relying on canned reports generated from summarized data, the team can now take a much deeper dive.

“We’ll be able to see how our workloads fluctuate over different periods—a day, a week, a year, and so on. Perhaps there are key times when we’re under-utilizing our hardware, or maybe there are subtle seasonal patterns in certain workloads that we’ve never noticed before. With a deeper understanding of how our systems behave, we’ll be able to have more fruitful discussions about how to manage workload more effectively and balance our resource consumption.”

He concludes: “The dream is to be able to monitor our batch flow and predict the impact of any delays, so that we can minimize problems and get everything back on track before our clients experience any issues. With Spark on Z, we finally have the tools we need to make this kind of advanced, near real-time analysis a reality.”

Solution components

- IBM® z/OS® Platform for Apache Spark

Connect with us



Take the next step

To learn more about IBM z/OS Platform for Apache Spark, please contact your IBM representative or IBM Business Partner, or visit the following website:

ibm.com/us-en/marketplace/real-time-analysis-on-z-systems

© Copyright IBM Corporation 2017. 1 New Orchard Road, Armonk, New York 10504-1722 United States.

Produced in the United States of America, October 2017. IBM, the IBM logo, ibm.com, and z/OS are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at “Copyright and trademark information” at ibm.com/legal/copytrade.shtml. Java and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates. This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates. The client examples cited are presented for illustrative purposes only. Actual performance results may vary depending on specific configurations and operating conditions. THE INFORMATION IN THIS DOCUMENT IS PROVIDED “AS IS” WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT. IBM products are warranted according to the terms and conditions of the agreements under which they are provided. The client is responsible for ensuring compliance with laws and regulations applicable to it. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the client is in compliance with any law or regulation.



ZSC03359-USEN-00

