

White paper

FUJITSU Integrated System PRIMEFLEX for Hadoop & Meteomatics

FUJITSU Integrated System PRIMEFLEX for Hadoop offers end-to-end integration from strategic consulting to use case development and implementation services for specific customer needs. Read how PRIMEFLEX for Hadoop is making an impact on Meteomatics and its analysis of complicated weather systems.

Contents

Introduction	2
Case study: Meteomatics	3
Conclusion	4



Introduction

Improving profitability and revenue is the leading priority of any commercial organization. This requires improving the performance and productivity of employees, and the efficiency, effectiveness and competitiveness of the overall business, while minimizing potential risks. The question is how to achieve this faster, better and to a greater extent than your competitors?

That's where data comes in. Data is one of the most valuable assets for organizations of every kind, enabling smarter decision-making, improved processes, reduced costs and happier customers. However, managing and understanding the ever-increasing volumes of data produced daily is a huge challenge. Business Intelligence solutions can consolidate and transform data into easily analyzed forms but this can be time-consuming and resource-intensive.



Reinventing Business Intelligence

Traditional Business Intelligence considers mainly internal and historical views collected from limited data sources, which is structured and stored in a relational database management system. Business analytics tasks are then designed against a static data model, and happen periodically – every day, week or month in a batch process. As the average end-user isn't trained to do their own complex analysis, the number of direct users initiating queries or dealing with business analytics is limited to a few specialists.

"We wanted to find a way to open the handling and analysis of big data not just to specialists but to other areas of the business such as marketing or engineering," explains Fritz Schinkel, Program Manager, Fujitsu Technology Solutions. "The concept was to agglomerate the most state-of-the-art technology – from the hardware infrastructure to the analytic software to the user interface – using best of breed components to produce a solution that could more effectively manage big data."

PRIMEFLEX for Hadoop

The result is FUJITSU Integrated System PRIMEFLEX for Hadoop, which dramatically reduces the time it takes to gather, process and understand business critical information. Hadoop is the de-facto standard for Big Data and distributed parallel processing, an open source framework primarily for batch operation, written in Java. It is designed to scale up to 1,000s of nodes, to accept server crashes as "normal" in large farms, and to make data storage and analytics robust against failures.

"Distributed parallel processing provides several advantages. Executing a query or any other data operation by many nodes at the same time increases performance and delivers fast results," adds Schinkel. "You can start small, with just a few servers and then add more servers as they are needed. Basically, your infrastructure will linearly scale up without any limits."

Hadoop is integrated with Fujitsu hardware and Datameer on the front end to enable non-specialist users to easily manipulate huge volumes of data from multiple sources. It simplifies the analysis while removing the need for any programming, coding or scripting thereby opening up the potential of big data processing to all areas of the business. Datameer is the only end-to-end big data analytics application purpose-built for Hadoop that enables the fastest time from raw data to new insights.

"PRIMEFLEX for Hadoop is a powerful and scalable platform that provides business users with a more cost-effective way of creating actionable analytics from big data," says Schinkel. "It analyses large volumes of data to extract and make accessible meaningful business-relevant information, combining the convenience of pre-configured and pre-tested hardware and the economic advantages of open source software plus system support and all-round lifecycle management."

Faster, more accurate results

Essentially, PRIMEFLEX shortens the time it takes to crunch data and provides a more direct route to smarter decision-making. This in turn reduces the time it takes to complete projects and speeds ROI. And, because processing can be completed in real-time, it gives more accurate results than traditional batch processing.

PRIMEFLEX for Hadoop comes with preinstalled software including RedHat Enterprise OS, Datameer, Cloudera Manager and Cloudera Distribution for Hadoop. The Entry variant is completely installed and configured in the factory and needs only to be connected to the customer network.

"It's an off-the-shelf solution that comes fully pre-configured so that it makes sense even to non-technical staff," continues Schinkel. "And it has the potential to revolutionize exciting new fields such as predictive maintenance, which could reduce costs for millions of companies that rely on manufacturing."

Fujitsu also offers end-to-end integration and consulting services for PRIMEFLEX for Hadoop, from strategic consulting to use case development and implementation services for specific customer needs. This makes the solution ideal for any organization that needs to understand large volumes of complex information, such as government bodies, manufacturers and oil companies. One area where PRIMEFLEX for Hadoop is already making an appreciable impact is in the analysis of complicated weather systems, which, in the current political and economic climate, is a vitally important subject.

Case study:

Meteomatics predicts weather more accurately with FUJITSU Integrated System PRIMEFLEX for Hadoop

Reliable information about historical and future weather trends is critical to numerous industries. From humankind's oldest endeavor – the cultivation of agriculture – to forward-looking technology such as renewable energy, credible knowledge about long-term climate behavior is invaluable. This knowledge is also vital for insurers to predict possible risks and adjust premiums, however it depends on the accurate and rapid analysis of big data.

Meteomatics was founded in March 2012 and focuses on the development of meteorological software products, processing and providing weather data and consulting work. Its experience covers the whole process chain of a weather service provider: acquisition and processing of weather station data, numerical weather models, radar images, satellite data as well as visualization of meteorological data and the development of customer specific forecasting systems.

Finding a faster way to process data

Meteomatics needs to process weather data on an industrial scale and previously used traditional C++ coding to undertake the task in a sequential fashion. However, this was time-consuming and inefficient. It wanted to find a faster way of crunching the vast amounts of meteorological data that forms the basis of its work.

"We consume massive amounts of data in four dimensions because we look at both geography and the temporal evolution of weather systems," explains Martin Fengler, CEO, Meteomatics. "A numerical forecast on a grid limited to a few kilometers can generate multiple TBs per day so we wanted to find a more efficient way of understanding that data."

The company's work is important because it enables companies to make smarter decisions. For example, it helps locate the best point for wind turbines to be built and informs insurance companies of the inherent risk in particular projects.

"We can estimate the average impact of severe weather which is crucial in many industries, however we wanted to be able to do it more quickly," adds Fengler. "Dealing with 25TB of data is extremely tricky but we knew there had to be a more effective solution."

At Germany's leading technology trade show CeBIT, Fengler and his team came across Fujitsu's booth where the company was demonstrating its own high-performance big data platform: FUJITSU Integrated System PRIMEFLEX for Hadoop.



An open-source, cloud-based approach

"We did some pilot tests using Fujitsu's Cloud platform and 120 CPUs which enabled us to accurately predict real world weather outcomes that other meteorologists had failed to detect," says Fengler. "Because Fujitsu handles the Hadoop programming, it makes the process very easy for the end-user."

Following the success of the trial, Meteoromatics is now using PRIMEFLEX for Hadoop as its core big data processing technology.

Speed and scalability

The principle benefit is the speed at which big data can be analyzed. What once would have taken weeks to process can now be completed within an hour or two. That means faster time to market for Meteoromatics customers as well as the ability to de-risk insurance policies.

"Re-insurance companies use the data to estimate the impact of weather on certain policies. There was a farmer in the US, for example, who wanted to insure his 10,000 cattle against blizzards which is not normally covered," continues Fengler. "We were able to provide accurate probability over a long time series so that a fair value contract could be drawn up. Another example is the construction of wind farms where re-insurance companies can offer investors protection against a lack of wind based on statistics provided by us."

The new Fujitsu solution is not only faster but it also offers more accuracy. The current resolution for meteorological measurement is 2.2km squared in Switzerland where the company is based. Thanks to the speed and capacity offered by PRIMEFLEX for Hadoop, Meteoromatics is hoping to reduce this to a single square kilometer.

"We can process more data at a more granular level which serves to improve our precision and thus enable better decision-making for our customers," comments Fengler. "Fujitsu's unique offering is at the heart of this capability."

Meteoromatics is a growing business with PRIMEFLEX for Hadoop at its core, enabling the rapid analysis of TBs of data on a daily basis. As the company continues its evolution, its partnership with Fujitsu will remain a key part of its business and its ability to predict weather systems come rain or shine.

"We have an excellent ongoing relationship with Fujitsu which has proven instrumental to our development. PRIMEFLEX for Hadoop has transformed how we process our data and enables us to offer much faster and more accurate results for our customers," concludes Fengler.

Conclusion

FUJITSU Integrated System PRIMEFLEX for Hadoop is a powerful and scalable platform that can help analyze big data volumes at high velocity. It combines pre-configured and pre-tested industry-standard hardware with open source software provided by Cloudera and Big Data Analytics from Datameer, making it the most cost-effective, accurate and fast way for companies to best exploit their data for competitive advantage.

Contact

FUJITSU
E-Mail: cic@ts.fujitsu.com
Website: www.fujitsu.com
2015-03-11

© 2015 Fujitsu and the Fujitsu logo are trademarks or registered trademarks of Fujitsu Limited in Japan and other countries. Other company, product and service names may be trademarks or registered trademarks of their respective owners. Technical data subject to modification and delivery subject to availability. Any liability that the data and illustrations are complete, actual or correct is excluded. Designations may be trademarks and/or copyrights of the respective manufacturer, the use of which by third parties for their own purposes may infringe the rights of such owner.