

# Kognitio WX<sub>2</sub> benchmark demonstrates how leading retailer can analyze large volume EPOS transactions

## The Retailer's Challenge

Everyone has stood in line at a supermarket checkout accompanied by the constant beeping as product barcodes are scanned by the tills. In a modern supermarket with its rows of checkouts, even in the relatively short time we are stood in-line, our ears must be subjected to hundreds, probably thousands of individual beeps. Imagine how many beeps that must be a day and then multiply that by the numbers of stores today's large retailers own and you soon realize that there must be billions of beeps every year.

To the retailer each beep is a transaction; a single record of information, detailing what item was bought, for how much and hopefully, if the customer has a loyalty card, who bought it. This record has immense value to the retailer. When analyzed along with all the other transaction records, it can reveal important information about customer buying behavior.

Although the data contains enormous value, its sheer scale means it is very difficult to analyze effectively. To get at the real value locked within this data, inherently complex analysis needs to be performed. Who bought what and when is interesting, but the real value is in understanding what else they bought at the same time and establishing buying patterns across groups of customers. With billions of records to analyze, conventional database technologies, with their reliance on indexing mechanisms, are simply unable to perform this analyze effectively.

## Proof of Concept - Benchmark

To demonstrate WX<sub>2</sub> performance on large volume retail data, a leading retailer asked for a benchmark using HP hardware. Two separate trials were run; the initial trial in HP Grenoble (France) involved 10% of the data and was designed to show that WX<sub>2</sub> had the necessary breadth of functionality for this complex analysis. A second trial involving 100% of the data was then undertaken at HP Böblingen (Germany) to demonstrate scalability.

The benchmark data consisted of 23 billion "point of sale" (EPOS) transactions 24 million customer records and over 660,000 product records. This represented 2 years of transactional data for the retailer. To visualize 23 billion transactions, consider the following: If each transaction were written down as one line on a sheet with 40 lines per page, the stack of paper would be 50 miles high, seven times higher than Mount Everest.

## Fast analysis

The HP blade system in Grenoble consisted of 30 BL35p blades, each with 8GB of memory and 36GB of disk. The system in Böblingen had 125 BL25p blades, each fitted with 16GB of memory and 144GB of disk. In both cases the blades were set-up and configured with SuSE Linux. The blade farm was then made available via a VPN to Kognitio in the UK. No Kognitio employees were required on site. In both trials it took less than two hours for Kognitio to remotely configure the blades as a single WX<sub>2</sub> database server.

For the trial in Grenoble the retailer's data took just 20 minutes to load into WX<sub>2</sub>. At Böblingen, the full volume data took less than an hour to load. No pre-partitioning of the data was required as the data was automatically and randomly distributed across all the blades by WX<sub>2</sub>'s load utilities. Once loaded, the data was immediately available for querying because no tuning or index building is required with WX<sub>2</sub>.

The first simple queries that were run were designed to make the database read every single record in the database and examine it for a match for a given parameter. The Grenoble system was able to read 2.3 billion records in 0.5 seconds and the Böblingen system read 23 billion records in less than 1 second. This ability to read data very quickly is the basis of WX<sub>2</sub>'s power to perform complex analysis on large data volumes.

The customer's retail benchmark consisted of multiple concurrent query streams each running a series of complex queries aimed at discovering the propensity of groups of customers to buy products. Single queries asked the following type of question: "For the set of customers I am interested in, find who, in the given period, bought one of the products I am interested in and then tell me what else they bought in the same product category?"

To satisfy this query the database needed to perform multiple joins between large tables, create interim temporary tables and perform several scans of the data. The results were outstanding. The Grenoble (2.3bn records) system was able to perform this query in 6 seconds and the Böblingen (23bn records) system was able to return results in 10 seconds.

This underpins the flexibility and power of Kognitio WX<sub>2</sub> as a true analytical database powerhouse ideal for retailers who want to maximize their business.

## The Challenge

Leading retailer wants to understand their customers' buying habits on a dynamic level in order to market to them more effectively

## The Solution

Kognitio worked with HP in Grenoble and Böblingen to set up a retail benchmark using WX<sub>2</sub> which sought to demonstrate how the leading retailer could ask any question of the data and obtain comprehensive results on customer spend within seconds

## The Results

Using the analytical power of WX<sub>2</sub>, over 23 billion EPOS records were queried in a matter of seconds instead of hours or days. As a result, the retailer can use the system to conduct train-of-thought analysis in order to maximize their marketing efforts

