

A TECHNICAL INTRODUCTION TO **ZIZO**®

TABLE OF CONTENTS

2	1	Technology Introduction - Database
2	1.1	Zizo® - Database
3	1.2	Pathway - Our Integrated Graphical Development Environment
5	1.3	Pathway - Delivering Complex Analytics
5	2	Conclusion



TECHNOLOGY INTRODUCTION

1.1 zizo® - Database

zizo® is a strategic database platform conceived to meet the challenges of data analytics. zizo® uses proprietary technology to handle vast volumes of data on an incredibly small hardware footprint. The database is managed using a fully integrated graphical development environment named 'Pathway' (described in Section 1.2). The inclusion of this built-in environment allows for a number of key functions to be delivered without the need for any additional tools, such as ETL, cleansing, data analysis and data integration technologies. An in-depth description of the database technology and its capabilities can be found below:

zizo® is a paged in-memory pattern database.

A pattern database differs from either of the commonly used storage techniques for databases, namely row based and column based. A pattern database stores the data neither as rows or columns, but instead decomposes the source data into two parts, a pattern store and a set of indices. The pattern store is a random access optimised structure that allows you to retrieve a pattern based on a reference value/pointer. Patterns may themselves refer to other patterns within the store in which case the pattern is a composition of its children. An index identifies a list of the patterns that are present in the table.

Pattern databases are well suited to in-memory databases. Firstly the pattern store is not accessed linearly, so there is no conventional efficient disk based representation for it. Secondly the pattern store provides a natural de-duplication mechanism whereby a pattern that appears two or more times in the database need only appear once in the store so long as the references to that pattern share a common address. This means that the pattern database is very efficient at using the memory that is available, and represents the smallest possible footprint for the data.

However, all in-memory databases suffer from scalability issues. The amount of RAM is typically a limiting factor and the price/performance curve for machines means that the obvious answer of simply adding more RAM cannot be applied indefinitely. What makes zizo® unique is it combines the performance benefits of an in-memory pattern database with a collection of paging strategies that mean that the database can grow beyond the limits of the RAM available with only minor degradations in performance. zizo® uses special data structures to replace conventional database building blocks like "lists", "sets" and "maps" with implementations that provide the same capabilities but that can also use the unique properties of pattern databases to switch to an efficient disk based storage when the data volumes warrant it. This means that zizo® is able to offer fast access to quantities of data that were not previously available at its price-point.

FOR FURTHER INFORMATION

please call us
01908 366662
or 07894 762766

EMAIL:
matthew.napleton@zizo.co.uk

www.zizo.co.uk



TECHNOLOGY INTRODUCTION

1.2 Pathway – Our Integrated Graphical Development Environment

Working in parallel with zizo® is the other major component to our technology stack, the integrated graphical development environment named 'Pathway'. This unique environment allows the User to interact with the data at every stage of its lifecycle without the need for additional tools for any process such as extract, transform, load, cleanse or integrate. It takes the form of an easy to use graphical 'storyboard', allowing you to add or remove components to create data streams which form the basis of complex queries. These queries are then called by the Web interface to populate the required visual Business Intelligence tool.

Any solutions developed within the 'Pathway' environment can be saved and then re-used during other stages of the development process. This is especially useful when there is a change to the data sources within the development phase and there is a requirement to change the data quickly and effectively to meet new business requirements.

Pathway is the visual programming environment, used for developing solutions in zizo®. However, Pathway is far more than just a visual environment for writing queries.

Pathway is a "computationally complete" programming language in its own right. This means that if a function is computable, it is computable in Pathway. The semantics of Pathway are based on those in the Lucid language originally developed by the Programming Research Group at Oxford University. These semantics describe how computations can be performed over potentially infinite streams of data. The semantics naturally lead to an "on demand" computational model where values are only computed at the point they are required. This makes Pathway ideally suited to the development of software for processing very large data sets. A critical aspect of Pathway is its facility to allow the User to observe the data as it flows through the diagram and to see the effects of changes to the program in real time. This is made possible by the semantics that automatically reduce the computation requirements to the minimum needed to provide the requested information.

The programming model in Pathway is a dataflow/functional one. The developer introduces data sources to the canvas. A database is one example of a data source, but there are many others including files and information pulled from the Internet. Each node/operation in the visual program transforms the data that flows through it. Sometimes nodes merge data flows, or take samples from them according to the rules supplied by the User. Pathway's computational power is built on libraries of components that can be imported into the environment. Libraries are typically built around particular data types. There are libraries for processing text, numbers, dates etc. New libraries can be written in the Java language and added into the development environment without the need for any recompilation and Database support is simply another library that is imported into Pathway.

FOR FURTHER INFORMATION

please call us
01908 366662
or 07894 762766

EMAIL:

matthew.napleton@zizo.co.uk



Pathway's visual programming model is simple and intuitive. Its controls are very similar to those of a computer painting package. It uses context-sensitive menus to avoid overloading the user with options. When the user clicks on an element in the environment, Pathway automatically works out what operations are valid on the data stream at that point and only shows the user those operations. This makes Pathway very easy and quick to learn. It also encourages exploration of data. A typical user starts with one or more data sources and interacts with Pathway to transform this data, working toward the answers they are interested in in easily understandable steps. Along the way new solutions can be discovered and new relationships will emerge. This makes Pathway uniquely suited and capable for the handling of large data sets by the domain experts who best understand the data.'

An explanation of how a combination of zizo® and Pathway can be used to dramatically enhance the delivery of a data oriented project can be found below:

'During a traditional Data Warehouse/Business Intelligence implementation, the extract, transform and load process is one the most important and lengthy elements to the solution. The data must be retrieved from numerous source systems and then transformed into a state whereby it can be integrated into the data warehouse to service the queries from the reporting tool. These processes are applied to the data 'in flight' and have to be constantly monitored to ensure they are effective. If the schema requirements of the data warehouse change, then in turn the ETL process must also change, meaning a huge overhead in both time & complexity when meeting changing business requirements.

In contrast, the process by which data is imported into the zizo® database represents a step change over standard ETL methods. zizo® reduces the complexity of this activity by holding all the data in the most efficient structure possible. This enables all the data to be manipulated and processed using our visual programming tool Pathway. All changes such as data cleansing are made to all the data and the routines created in Pathway are then used in the on-going production system - these only need to be created once and can be applied to all of the required processes. The data stays in the zizo® database from the start of the process and new data sources can be added to the environment without the need for complex schema redesign. This is a major advance over existing ETL methods and greatly reduces the complexity and time and cost of delivering a data warehouse or other similar data-oriented projects.

FOR FURTHER INFORMATION

please call us
01908 366662
or 07894 762766

EMAIL:

matthew.napleton@zizo.co.uk



TECHNOLOGY INTRODUCTION

1.3 Pathway - Delivering Complex Analytics

The analysis of data to fit business rules and populate a data warehouse is not the end of the data processing cycle; often there are additional business questions which will need to be answered. These may not have an impact on the data warehouse or web interface, but may help to define other objectives for the business moving forward.

Domain experts can be provided with either a full replicated dataset or a sub-set of the data to 'mine'. The dataflow model which Pathway uses empowers these Users and encourages them to explore the data to answer the important questions which drive the business agenda. The ability to 'deep dive' into the data regardless of size or complexity to solve critical issues adds another important layer to the solution which would otherwise be supplanted by the use of additional technologies.

The flexibility inherent within the zizo® toolset allows for rapid changes and enhancements to be made to existing datasets, without the need for database redesign. As the technology allows for all of the data to be stored without aggregation, the data held within the database can be used to answer any question, whether it be operational, financial, commercial, customer focussed or insight driven.

CONCLUSION

In conclusion, zizo® have brought a new solution to the data warehouse and management information marketplace. By marrying the latest advances in database technology with a fresh approach to delivering the right information to the business, the capability to deliver actionable insight has been dramatically improved. This in turn means that the complexity and cost of delivering projects of this nature has been massively reduced.

This 'Next Generation' approach affords greater flexibility to business, whilst ensuring that the required strategy is delivered using a single consistent view of data. The ability to adapt to changing requirements, whilst still delivering data analysis and insight 'on time' is key to delivering competitive advantage and improved profitability.

FOR FURTHER INFORMATION

please call us
01908 366662
or 07894 762766

EMAIL:

matthew.napleton@zizo.co.uk