StretchDB®

Making Business Flow™

Case Study

Country / Region: Colombia / Latinamerica
Industry: Financial Sector

Customer Profile:

As the second largest company in the trust sector in Colombia, ***Fiduciaría Bogotá,*** which mission is to determine the financial needs and provide quality specialized fiduciary services, generating economic growth in target markets.

Current Situation:

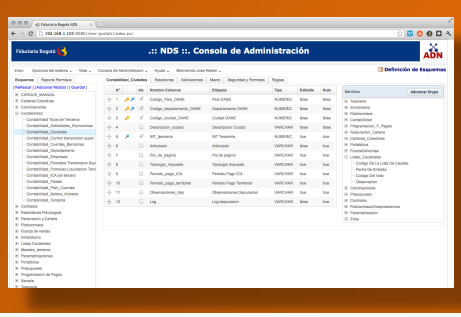
The “Fiduciaria Bogotá” is now implementing an ERP-Like solution in the trust sector, and needs to shutdown 33 applications and integrate them all in this ERP-Like solution, and there is a need for a solution to migrate data from all those applications and generate a portal for data quality and assurance.

Solution Goals

Create an ETL process from which data mining, from all 33 applications can be taken from, as well as to create a portal solution from which users can scrub data for quality assurance. The idea is to use riverETL® Studio for creating Data Maps and automate them for the mining, and port the extracted data to a StretchDB® enabled server, using riverPortals® as web application solution, from which users will use to depurate records in order to load up clean data to the new ERP-Like system.

Expected Benefits

- Have a portal solution from which data can be depurated and later moved to the new system.
- Have a workflow for data approval using security features from riverPortals®
- Aggregate data from 33 applications, even data stored in Microsoft Excel or Plain Text Files, in a single structure from which data can be depurated.



Stretching Relational Databases Towards the Limits of NoSQL and Schema-Free

Today's business needs are pushing the boundaries for database manipulation, in a way that IT divisions cannot cope with the fact of the many changes in the database structures. What to do then? Move to new fascinating technologies? Cassandra? MongoDB? Illuminate? Hadoop? Or just rely on the current infrastructure and know-how? StretchDB® is here as a concept to bridge that gap.

Nowadays industries and standards on data storage and manipulation, strives towards and elastic mode of growing databases and its containing schemas, being able to adapt without having to shake the whole application, requiring costly adaptations to the new model. For this the NoSQL databases and Schema-Free are becoming the standard in the industry, databases like Casandra, MongoDB, Hadoop, HBase among others are gaining momentum.

But what if you really know good relational databases and you know well XML? What if the company you work for cannot afford moving to these new trends on databases? What if the company wants to keep existing know-how up to date and use it to the max in order to adapt the business to the growing data demands?

Most relational databases support storing XML as part of their bound-schema tables, and it’s possible to loosely reach the boundaries for a Schema-Free Database, and using XML as backbone to accomplish what NoSQL databases aim for, on relational databases becomes a real possibility.

This approach is that we want to present with a real world example, running on a top-notch company in the Financial and Banking sector in Colombia.

This paper describes the full set of features for StretchDB® from a technical perspective, out of which a systems architect and a solution engineer can benefit from, in order to plan his next big application deployment.

StretchDB® Bridge the Gap Between Traditional RDBMS and NoSQL.

Solution Foundations

Modern Traditional Relational Databases

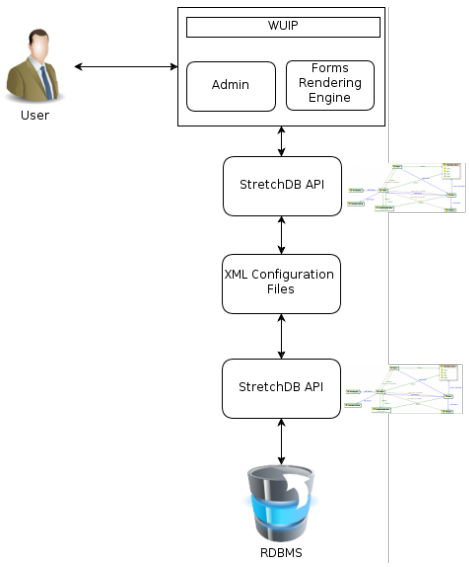
- Just One Single Table!
- Must have support for XML types.
- Must support strong XPATH functions for retrieving data
- Ideally must support Function Based Indexes based on XPATH

Strong API Handling / Administration (StretchDB®)

- Build all in JAVA, Plain Old Java Objects (POJO) based solution
- Build on top of riverSDK® open source framework
- Web User Interface Plugins (WUIP) based on riverPortals®

Traditional Know-How Exploited to the Max!

- Standard SQL Statements combined with XPATH functions
- Standard JAVA Programming
- XML Manipulation through marshalling and un-marshalling.



- **StretchDB**® is not a technology by itself, is a functional framework built on top of XML and RDMBS for a better use of data storage technology.
- Tables can be created and be ready to be filled out with information by end-users with the built-in “on-the-fly” forms rendering engine.
- Columns can be added and removed at runtime with ***almost*** no side effect to business logic. **StretchDB**® logic is no affected by this. Specific project logic can be affected if developers aren’t able to look for either removed or added newly columns.
- Enriched column metadata descriptors, allows to have real control of schema and data at the same time, unlike to traditional relational databases, opening a world of new possibilities for data-driven applications.
- Ensured ACID compliance through the underlying RDBMS that supports **StretchDB**® concepts.
- Strong API that encapsulates the hurdles of manipulating XML and Database calls.
- Most of the time, it is either an administrator or a data-entry end user, which interacts with riverPortals® and the StretchDB® Web User Interface POJO (WUIP) bundle, either to create or modify existing data structures, as well as entering data.

1. The **StretchDB**® API encapsulates all the hurdles to manipulate XML configuration files, as well as takes care of all security implications during the process.
2. At the end, it’s again the **StretchDB**® API in charge to communicate all the changes to the RDBMS where the underlying table structure is located into.
3. What you end up knowing? From the API standpoint, just a Table that has Column Schema, and a collection of Rows containing Row objects, which at the end contains a collection of Columns, composed of Column objects.
4. Very much like when you manipulate tables in HTML but additionally having a full blown up schema for you to benefit from.

StretchDB® management console is deployed on riverPortals® as a WUIP bundle, nevertheless, if someone doesn’t have this facility, or do not want to use riverPortals®, it can always be used the underlying XML configuration files to create all what is required, and since this is an open source project, it can always be customized to be accommodated to any other platform.

The following is the representation of a column-schema stored in the RDBMS as part of the whole metadata manipulation done by StretchDB®.

```
<?xml version="1.0" encoding="UTF-8"?>
<row>
    <Empresa dirty="false" mutable="false">13194</Empresa>
    <Banco dirty="false" mutable="false">52</Banco>
    <Sucursal dirty="false" mutable="false">AV VILLAS</Sucursal>
    <Numero_Cuenta dirty="false" mutable="false">059010967</Numero_Cuenta>
    <Clase_Cuenta dirty="false" mutable="false">CORRIENTE</Clase_Cuenta>
    <Nombre_Cuenta dirty="false" mutable="false">FBOGOTA S.A. - SOLIDDA GROUP</Nombre_Cuenta>
    <Proceso_Cuenta dirty="false" mutable="false">RECAUDADORA, DISPERSORA</Proceso_Cuenta>
    <Tipo_Cuenta dirty="false" mutable="false">RECAUDADORA, DISPERSORA</Tipo_Cuenta>
    <Es_Cuenta_Nacional dirty="false" mutable="false">S</Es_Cuenta_Nacional>
    <Manejo_de_Chequera dirty="false" mutable="false">I</Manejo_de_Chequera>
    <Nombre_Abreviado_Cuenta dirty="false" mutable="false"></Nombre_Abreviado_Cuenta>
    <Cuenta_Mayor dirty="false" mutable="false">PENDIENTE HOMOLOGACION</Cuenta_Mayor>
</row>
```

*The idea behind **StretchDB**® is to use existent know-how and technology to broaden the scope of traditional-relational way of thinking.*

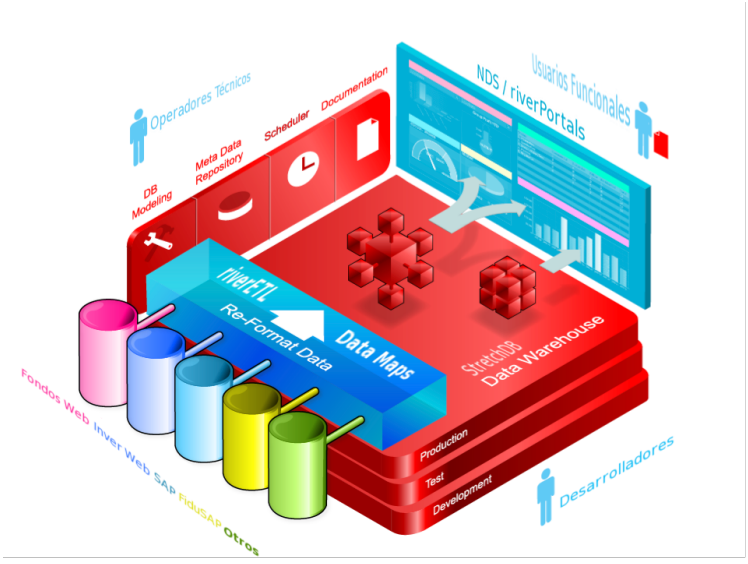
StretchDB®

Making Business Flow™

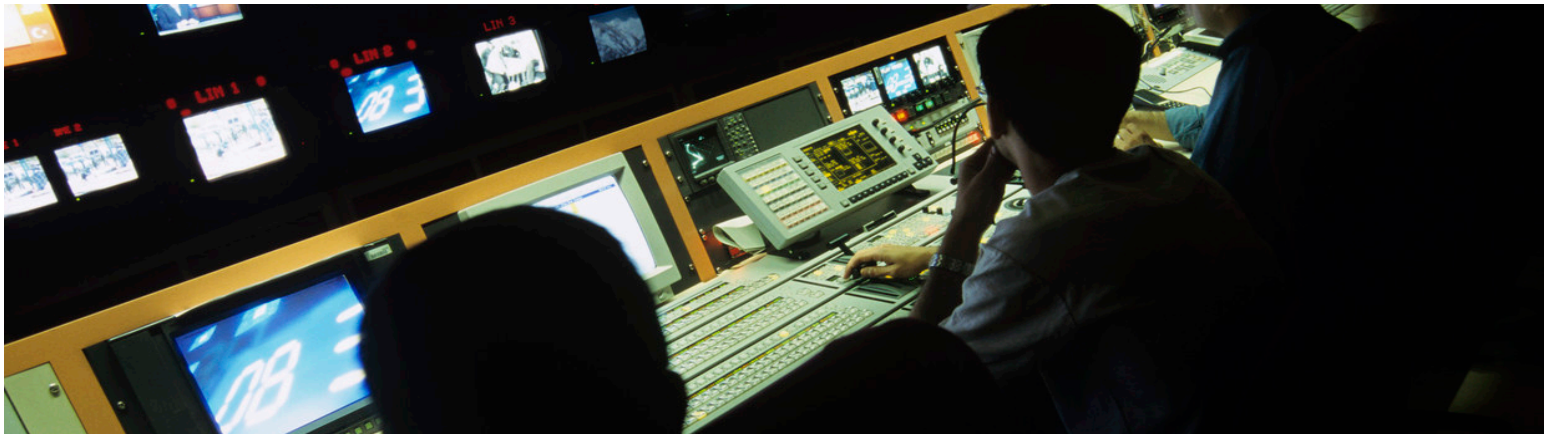
Features List

- Completely integrated with riverPortals® as a UI-POJO bundle.
- Strong API allowing programmers to benefit from.
- Flexible schema creation via its management console, storing those schemas into XML, allowing:
  - i. Creation of schema taxonomy.
  - ii. Creation of attributes for column creation via drag-n-drop from the attribute sandbox.
  - iii. Visual column sorting, allowing controlling the correct position of a column in a table.
  - iv. Local and foreign keys and indexes creation and administration.
  - v. Use of metadata at column level, to categorize the data stored within the column.
  - vi. Relation management from which data can be updated in cascade or a dependency relations from which data related within a column referenced in another column will pull data from that table just because the relation is defined.
  - vii. Validation rules created at column level, in order to control the behaviour of the auto-generated out-of-the-box formulary by StretchDB®.
  - viii. Ability to enable macros at the column level, to be executed at the "save" event. One can thing of them as triggers executed on the column on which is defined to.
  - ix. Security enforcement at all levels of the schema/table (provided by riverPortals®), setting rules of what groups or users can do.
- Ability to create POJO based controllers for each schema, called ***TableHandlers***, in order to control how the data is stored in the RDBMS. One can thing of them as being triggers at the table level.
- Ability to create EDA rules, stored as JSON, at the column level to control how StretchDB® Formulary rendering engine creates the user form, in a way of "***When-this-Then-That***".
- Allows exporting data on Microsoft Excel format. Here, every 50K rows a new sheet is created until data can be stored in the file.
- Security reports, can be created by table at user level in order to identify what a user can do over a set of schemas/tables.
- Automatic formulary creation by StretchDB® Formulary Rendering Engine to be used by riverPortals®, out of which a user can:
  - i. Taxonomic schema visualization, depending on the security grants given to a user.
  - ii. Exporting data if the security level allows.
  - iii. Do actions on data like: Approve, Review, Reject, Assign, Create, Edit or Erase.
  - iv. Powerful search engine that allows the user to filter data upon his requirements.
  - v. Grouping and sorting at grid level where the data is displayed onto.
  - vi. Format error visualization at the grid when a value within a column doesn't match the defined schema.
  - vii. Colour categorization of workflows that can be defined on the data.
  - viii. Data quality assurance by means of categorizing data as "dirty" if later needs to be cleaned up.

Our vision encloses a full data cycle that can be started out an ETL process, as described in the picture below, that we have put in place in the financial sector in Colombia with success.



StretchDB® is not a technology by itself, but rather a framework that runs on top of proven technology, therefore all the burden of performance lies on the database of choice and its support for manipulating XML, and mostly its management for large sets of XML chunks.



Coordinated as a Control Center We Work in Order to Help You Reach Your Goals!

Sistem Requirements

StretchDB® Minimum Server Hardware Requirements

- 2 CPUs 1Ghz+
- 2+ GB RAM
- 60+ GB available hard drive space
- Network connectivity via TCP/IP

StretchDB® Minimum RDBMS Required

- Oracle 11g or up.
- PostgreSQL 9 or up

River Software Technologies S.A.S

Making Business Flow ™

Headquarters
Bogotá – Colombia
Email: info@river-ss.com / salarcon@river-ss.com / jwalker@river-ss.com
http://www.river-ss.com