

---

# Almacenamiento de información en XML

---



Ricardo Sánchez

## Utilización de XML para almacenar información

---

Una base de datos XML constituye un sistema software que da persistencia a datos almacenados en formato XML. Estos datos pueden ser interrogados, exportados y serializados. Las bases de datos XML están generalmente asociadas con las bases de datos documentales.

Existen dos grandes clases de bases de datos XML:

- **XML habilitado:** éstas bien pueden almacenar XML en estructuras tradicionales de bases de datos (como las relacionales), aceptando XML como entrada y formateando en XML la salida, o más recientemente soportando tipos XML nativos en la propia base de datos.
- **XML nativo (NXD):** el modelo interno de estas bases de datos usa documentos XML como la unidad elemental de almacenamiento, los cuales no han de almacenarse necesariamente en formato de texto.

Una razón para el uso de XML en bases de datos estriba en el generalizado aumento del transporte de datos en formato XML, lo que significa que los datos se extraen de bases de datos y convertidos a XML o viceversa. Resultaría más eficiente (en términos de costes de conversión) y fácil almacenar los datos en formato XML directamente.

## Bases de datos habilitadas XML

---

Típicamente ofrecen alguna de las siguientes aproximaciones para almacenar XML en la estructura relacional clásica:

1. El XML se almacena en un campo CLOB (*Character large object*)
2. El XML se desgaja en una serie de tablas según un esquema
3. El XML se almacena de forma nativa como tipo XML según define ISO

Entre los SGBD que soportan ISO XML están:

- IBM DB2 (XML puro).
- Microsoft SQL Server.
- Oracle Database.
- PostgreSQL.

Típicamente una base de datos habilitada XML se adapta mejor allí donde la mayoría de los datos no están en XML, para aquellos datos en los que la mayoría están en XML, una base de datos nativa XML se adaptará mejor.

## Bases de datos nativas en XML

---

Las bases de datos nativas definen un modelo lógico para el documento XML (a diferencia de dicho modelo), además, almacena y recupera documentos de la misma manera que los XML. Por lo menos este modelo debe incluir atributos como PCDATA, y documentos en orden. Ejemplos de estos modelos son XPath, XML Infoset y modelos que implican DOM y SAX 1.0.

Una base de datos nativa en XML, no posee campos, ni almacena datos atómicos, lo que ella almacena son documentos XML, por lo tanto a este tipo de bases de datos se les denomina **bases de datos centradas en documentos**.

## Características

Diversos productos brindan diferentes características para las bases de datos nativas en XML, pero generalmente tienen las siguientes características:

### Procesamiento de datos

El procesamiento de datos en este tipo de bases de datos debería ser algo muy beneficioso, pero realmente no es así, debido al formato jerárquico en el que está almacenada la información. Muchas bases de datos necesitan que el usuario recupere todo el documento XML, lo actualice con el XML API de su preferencia y posteriormente vuelva a almacenar el documento en el repositorio. Esto se debe a que aún no existe un lenguaje estándar que permita la actualización, inserción o eliminación de elementos de un documento XML. Existe un lenguaje que permite realizar actualizaciones en un documento XML pero aún no es un estándar y muchos gestores de este tipo de bases de datos no lo soportan, este lenguaje es **Xupdate**.

### Almacenamiento

Por deducción lógica, una base de datos nativa en XML almacena la información en formato XML, pero esto es solamente una deducción lógica, pues este tipo de bases de datos tienen repositorios con un formato "tipo XML", como puede ser DOM o Infoset. En este mismo "repositorio" (paquete de archivos) se almacenan los índices que se generan por cada documento XML almacenado.

### Búsquedas

Este tipo de bases de datos no utiliza SQL como lenguaje de consulta. En lugar de ello utilizan **Xpath**. Algunas bases de datos permiten seleccionar los elementos que deberán tener índices mientras que otras bases de datos indexan todo el contenido del documento. El problema que tienen las búsquedas en este tipo de bases de datos es que no permiten realizar búsquedas muy complicadas, como por ejemplo ordenamiento y consultas en varias tablas, debido a que Xpath no fue creada realmente para búsquedas en bases de datos, sino simplemente para búsquedas en un solo documento.

## Áreas de aplicación

En general, cualquier base de datos nativa en XML es una excelente herramienta para almacenar documentos orientados a datos (Ej. XHTML o DocBook), la información que tiene una compleja estructuración anidada profundamente y la información que esta semi-estructurada en la naturaleza. Básicamente, si la información es representada por XML una base de datos nativa en XML probablemente seria una buena solución. Una base de datos nativa en XML puede almacenar cualquier tipo de información XML, pero no es adecuada para algo como un sistema de contabilidad donde los datos están muy bien definidos y rígidos.

Algunas áreas potenciales de aplicación podrían ser:

- Portales de información corporativa
- Información de catálogos
- BD en partes de manufactura
- Información medica
- BD personalizadas

En estos momentos, este tipo de base de datos es perfecta para las organizaciones que mantienen su información en diferentes formatos, debido a que les permitiría rápidamente pasar su información a un formato XML, utilizando la aplicación de su preferencia, y de esta manera almacenarla en una de estas bases de datos y evitarse el proceso de conversión de un formato a otro. Además, teniendo en cuenta que por ahora aún no existe un lenguaje estándar para el procesamiento de los datos, estas bases de datos son adecuadas para realizar búsquedas sobre los documentos almacenados.

A medida que todas las carencias que tienen estas bases de datos sean superadas, podrán ser una muy buena alternativa para el almacenamiento de la información no solo por que la información estará almacenada en un formato estándar, XML, sino porque facilitará el desarrollo de aplicaciones basadas en XML, como por ejemplo los Web Services.

### Referencias:

[https://es.wikipedia.org/wiki/Base\\_de\\_datos\\_XML](https://es.wikipedia.org/wiki/Base_de_datos_XML)

[https://es.wikipedia.org/wiki/Bases\\_de\\_datos\\_nativas\\_XML](https://es.wikipedia.org/wiki/Bases_de_datos_nativas_XML)